

Witnessing Wigner Negativity

Ulysse Chabaud^{*1,2}, Pierre-Emmanuel Emeriau^{*3}, and Frédéric Grosshans³

¹Institute for Quantum Information and Matter, Caltech

²Université de Paris, IRIF, CNRS, France

³Sorbonne Université, CNRS, LIP6, F-75005 Paris, France

Negativity of the Wigner function is arguably one of the most striking non-classical features of quantum states. Beyond its fundamental relevance, it is also a necessary resource for quantum speedup with continuous variables. As quantum technologies emerge, the need to identify and characterize the resources which provide an advantage over existing classical technologies becomes more pressing. Here we derive witnesses for Wigner negativity of single mode and multimode quantum states, based on fidelities with Fock states, which can be reliably measured using standard detection setups. They possess a threshold expectation value indicating whether the measured state has a negative Wigner function. Moreover, the amount of violation provides an operational quantification of Wigner negativity. We phrase the problem of finding the threshold values for our witnesses as an infinite-dimensional linear optimisation. By relaxing and restricting the corresponding linear programs, we derive two hierarchies of semidefinite programs, which provide numerical sequences of increasingly tighter upper and lower bounds for the threshold values. We further show that both sequences converge to the threshold value. Moreover, our witnesses form a complete family—each Wigner negative state is detected by at least one witness—thus providing a reliable method for experimentally witnessing Wigner negativity of quantum states from few measurements. From a foundational perspective,

our findings provide insights on the set of positive Wigner functions which still lacks a proper characterisation.

1 Introduction

Quantum information with continuous variables [1]—where information is encoded in continuous degrees of freedom of quantum systems—is one of the promising directions for the future of quantum technologies. For example, continuous-variable quantum optics enables the deterministic experimental preparation of entangled states over millions of modes [2] and also offers reliable and efficient detection methods, such as homodyne or heterodyne detection [3]. From a theoretical point of view, quantum information with continuous variables provides different perspectives from quantum information with discrete variables and is described via the formalism of infinite-dimensional Hilbert spaces.

To handily manipulate states in those infinite spaces, mathematical tools initially inspired by physics have been developed such as phase-space formalism [4]. In this framework, quantum states are represented by a quasi-probability distribution over phase space, like the Wigner function [5]. These representations provide a geometric intuition of quantum states [6]: quantum states are separated into two categories, Gaussian and non-Gaussian, depending on whether their Wigner function is a Gaussian function or not.

Non-Gaussian quantum states are essential to a variety of quantum information processing tasks such as quantum state distillation [7–9], quantum error-correction [10], universal quantum computing [1, 11] or quantum computational speedup [12, 13]. Within those, an important subclass of non-Gaussian states are the states which display negativity in the Wigner function. These two classes of states coincide for pure

uchabaud@caltech.edu

pierre-emmanuel.emeriau@lip6.fr

frederic.grosshans@lip6.fr

* These authors contributed equally.

states—namely, non-Gaussian pure states have a negative Wigner function—as pure states with a positive Wigner function are necessarily Gaussian states by Hudson theorem [14, 15]. However, this is not the case for mixed states and the (convex) set of states with a positive Wigner function becomes much harder to characterise [16, 17].

Negativity of other phase-space quasiprobabilities have been used to define different notions of quantumness [18]. In particular, the negativity of the Glauber–Sudarshan P quasiprobability distribution of a quantum state is known as its non-classicality [19]. The Wigner function can be obtained from the more singular P function by a Gaussian convolution, thus positivity of the latter implies positivity of the former. In particular, negativity of the Wigner function implies non-classicality, although there are non-classical states with positive Wigner function, such as squeezed states.

In addition to its fundamental relevance as a non-classical property of physical systems [20], Wigner negativity is also essential for quantum computing, since continuous-variable quantum computations described by positive Wigner functions can be simulated efficiently classically [21]. Wigner negativity is thus a necessary resource, though not sufficient [22], for quantum computational speedup with continuous variables.

With the rapid development of quantum technologies [23], finding efficient methods for assessing the correct functioning of quantum devices is of timely importance [24]. Detecting key properties of a quantum state, such as entanglement or Wigner negativity, can be done by a full tomographic reconstruction [25]. However, such reconstructions are very costly in terms of the number of measurements needed, and require performing a tomographically complete set of measurements—thus usually involving multiple measurement settings. In the continuous-variable setting the task is even more daunting, since the Hilbert space of quantum states is infinite-dimensional [26, 27].

Instead, one may introduce witnesses for specific properties of quantum states [28–32] that are easier to measure experimentally. These witnesses should be observables that possess a threshold expectation value indicating whether the measured state exhibits the desired property or not. Intuitively, a witness for a given property

can be thought of as a separating hyperplane in the set of quantum states, such that any state on one side of this hyperplane has this property (see Fig. 1). In particular, some states with the sought property may remain unnoticed by the witness. In that regard, one may use a complete set of witnesses such that for each state exhibiting the desired property, there exists at least one witness in the set that captures it.

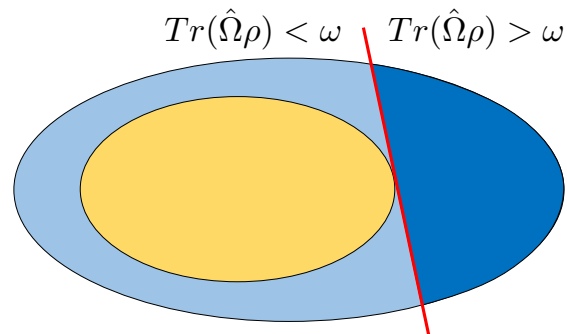


Figure 1: Pictorial representation of a witness $\hat{\Omega}$ with threshold value ω for a given property. In yellow: states without the property. In blue: states with the property. In red: witness threshold value. In light blue: states with the property undetected by the witness.

Here, we introduce and study a complete family of witnesses for Wigner negativity of single-mode quantum states, expressed using Fock states projectors. The expectation values of our witnesses are linear functions of the state which may be efficiently estimated experimentally using standard homodyne or heterodyne detection, thus providing a reliable method for detecting Wigner negativity with certifiable bounds. Additionally, we show that the amount by which the measured expectation value exceeds the threshold value of the witness provides an operational measure of Wigner negativity: it directly lower bounds the distance between the measured state and the set of states with positive Wigner function.

We cast the computation of the threshold values of the witnesses as infinite-dimensional linear programs, which can be either relaxed or restricted. Upper and lower bounds for the threshold values of our witnesses are then given by two hierarchies of finite-dimensional semidefinite programs, similar in spirit to the Lasserre–Parrilo hierarchy [33, 34] and the subsequent Lasserre hierarchy [35]. In particular, we show that both

hierarchies converge to the threshold values.

Finally, we discuss the generalisation to multi-mode quantum states and show that most of our results are also applicable in this case.

Our work is thought to be interesting for physicists interested in characterising Wigner negativity of quantum states—either theoretically or experimentally—and mathematicians interested in infinite-dimensional convex optimisation theory. To that end, the rest of the paper is structured as follows: we give some notations and background in the next section 2 before a detailed exposition of our witnesses in section 3. Section 4 describes the experimental procedure for witnessing Wigner negativity of a quantum state using these witnesses, together with use-case examples. The following section 5—which deals with infinite-dimensional optimisation techniques of independent interest—is devoted to estimating the threshold values of our witnesses: after some technical background in section 5.2, section 5.3 reformulates the problem of finding the threshold value of a witness as an infinite-dimensional linear optimisation, while section 5.4 derives two hierarchies of semidefinite relaxations and restrictions for this linear program, yielding numerical upper and lower bounds for the threshold value. Section 5.5 establishes the proof of convergence of these hierarchies of upper and lower bounds towards the threshold value. We introduce the generalisation to the multimode case in section 6 and conclude with a few open questions in section 7.

2 Notations and background

2.1 Preliminary notations

For all $m \in \mathbb{N}^*$, Sym_m denotes the space of $m \times m$ real symmetric matrices. An exponent T denotes the transpose while an exponent \dagger denotes the conjugate transpose.

\mathcal{H} denotes a separable infinite-dimensional Hilbert space equipped with a countable orthonormal single-mode Fock basis $\{|n\rangle\}_{n \in \mathbb{N}}$. We write $\mathcal{D}(\mathcal{H})$ the set of quantum states (positive semidefinite operators with unit trace) over \mathcal{H} . A single-mode quantum state ρ can be expanded in Fock basis as $\rho = \sum_{k,l=0}^{+\infty} \rho_{kl} |k\rangle\langle l|$.

The fidelity between two quantum states ρ and σ is denoted $F(\rho, \sigma) = \text{Tr}(\sqrt{\sqrt{\rho}\sigma\sqrt{\rho}})^2$.

When one of the states is pure, it reduces to $F(\rho, \sigma) = \text{Tr}(\rho\sigma)$. The trace distance between two quantum states ρ and σ is denoted $D(\rho, \sigma) = \frac{1}{2}\text{Tr}(\sqrt{(\rho - \sigma)^2})$. The trace distance can be related to the maximum probability of distinguishing between two quantum states. The fidelity and trace distance are related by $1 - F \leq D \leq \sqrt{1 - F}$ [36].

We denote by \hat{a} and \hat{a}^\dagger the single-mode annihilation and creation operators, respectively, defined by their action on the Fock basis:

$$\begin{aligned} \hat{a}|n\rangle &= \sqrt{n}|n-1\rangle, & \text{for } n \in \mathbb{N}^*, \\ \hat{a}|0\rangle &= 0, \\ \hat{a}^\dagger|n\rangle &= \sqrt{n+1}|n+1\rangle, & \text{for } n \in \mathbb{N}. \end{aligned} \quad (1)$$

These operators satisfy the canonical commutation relation $[\hat{a}, \hat{a}^\dagger] = \mathbb{1}$. For all $\alpha \in \mathbb{C}$, we write

$$\hat{D}(\alpha) = e^{\alpha\hat{a}^\dagger - \alpha^*\hat{a}} \quad (2)$$

the displacement operator of amplitude α [37]. The coefficients of the displacement operator in Fock basis are given by [38]:

$$\begin{aligned} \langle k|\hat{D}(\alpha)|l\rangle &= e^{-\frac{1}{2}|\alpha|^2} \\ &\times \sum_{p=0}^{\min k,l} \frac{\sqrt{k!l!}(-1)^{l-p}}{p!(k-p)!(l-p)!} \alpha^{k-p} \alpha^{*l-p}, \end{aligned} \quad (3)$$

for all $k, l \in \mathbb{N}$ and all $\alpha \in \mathbb{C}$.

2.2 Wigner function

The Wigner function of a single-mode quantum state ρ is an equivalent representation of the state in phase space which can be expressed as [39, 40]:

$$W_\rho(\alpha) = \frac{2}{\pi} \text{Tr} \left[\hat{D}(\alpha) \hat{\Pi} \hat{D}^\dagger(\alpha) \rho \right], \quad (4)$$

for all $\alpha \in \mathbb{C}$, where \hat{D} is defined above and $\hat{\Pi} = (-1)^{\hat{a}^\dagger \hat{a}} = \sum_{n \geq 0} (-1)^n |n\rangle\langle n|$ is the parity operator. In particular, the Wigner function of a quantum state is related to the expectation value of displaced parity operators.

The Wigner function is a real-valued quasiprobability distribution [41], i.e., a normalised distribution which can take negative values. Hence, it cannot be sampled directly experimentally. However, its marginals are probability distributions which can be sampled using homodyne detection [26]. Alternatively, heterodyne detection (also called double homodyne detection) allows for sampling from a smoothed version of the

Wigner function [42, 43]. In both cases, applying a displacement before the detection is equivalent to measuring directly with homodyne or heterodyne detection and applying a classical post-processing procedure—namely, a translation of the classical outcome according to the displacement amplitude [32, 44].

As mentioned in the introduction, continuous-variable quantum states are classified in two categories, Gaussian and non-Gaussian, depending on the shape of their Wigner function. The set of Gaussian states is well-understood [45] but has a limited power, while characterising the set of non-Gaussian states is an active research topic [46–49].

The negativity of the Wigner function can only decrease under Gaussian operations [46], i.e., operations that map Gaussian states to Gaussian states. In particular, it is invariant under displacements. It is also a robust property, since two almost indistinguishable quantum states have similar Wigner functions. An operational measure of Wigner negativity for a quantum state $\rho \in \mathcal{D}(\mathcal{H})$ is given by its distance with the set of states having a positive Wigner function [30]:

$$\eta_\rho = \inf_{\substack{\sigma \in \mathcal{D}(\mathcal{H}) \\ W_\sigma \geq 0}} D(\rho, \sigma), \quad (5)$$

where D denotes the trace distance, thus quantifying the operational distinguishability between the state ρ and any state having a positive Wigner function [36].

A natural choice for a witness of Wigner negativity is the fidelity with a pure state having a Wigner function with negative values, since it is a quantity that can be accessed experimentally by direct fidelity estimation [25, 27]. Building on this intuition, and given that all Fock states—with the exception of the (Gaussian) vacuum state $|0\rangle$ —have a negative Wigner function, we introduce in the following section a broad family of Wigner negativity witnesses for single-mode continuous-variable quantum states based on fidelities with Fock states.

3 Wigner negativity witnesses

We introduce the following Wigner negativity witnesses:

$$\hat{\Omega}_{\mathbf{a},\alpha} := \sum_{k=1}^n a_k \hat{D}(\alpha) |k\rangle\langle k| \hat{D}^\dagger(\alpha), \quad (6)$$

for $n \in \mathbb{N}^*$, $\mathbf{a} = (a_1, \dots, a_n) \in [0, 1]^n$, with $\max_k a_k = 1$, and $\alpha \in \mathbb{C}$. These operators are weighted sums of displaced Fock states projectors. They can be thought of as Positive Operator-Valued Measure (POVM) elements, and their expectation value for a quantum state ρ is given by

$$\text{Tr}(\hat{\Omega}_{\mathbf{a},\alpha} \rho) = \sum_{k=1}^n a_k F(\hat{D}^\dagger(\alpha) \rho \hat{D}(\alpha), |k\rangle), \quad (7)$$

where F is the fidelity. This quantity can be directly estimated from homodyne or heterodyne detection of multiple copies of the state ρ by translating the samples obtained by the amplitude α in the classical postprocessing and performing fidelity estimation with the Fock states $|1\rangle, \dots, |n\rangle$ [25, 27].

For $n \in \mathbb{N}^*$, each choice of $(\mathbf{a}, \alpha) \in [0, 1]^n \times \mathbb{C}$ yields a different Wigner negativity witness. In particular, when $\alpha = 0$ and one entry of the vector \mathbf{a} is equal to 1 and all the other entries are 0, the expectation value of the witness is given by the fidelity with a single Fock state.

To each witness $\hat{\Omega}_{\mathbf{a},\alpha}$ is associated a threshold value defined as:

$$\omega_{\mathbf{a}} := \sup_{\substack{\rho \in \mathcal{D}(\mathcal{H}) \\ W_\rho \geq 0}} \text{Tr}(\hat{\Omega}_{\mathbf{a},\alpha} \rho). \quad (8)$$

Since negativity of the Wigner function is invariant under displacements, the threshold values do not depend on the value of the displacement amplitude α and we thus write $\omega_{\mathbf{a}}$ (rather than $\omega_{\mathbf{a},\alpha}$) for the threshold value associated to the witness $\hat{\Omega}_{\mathbf{a},\alpha}$. This is sensible, given that the threshold value asks for non-negativity anywhere in phase space, so a displacement in phase space should not change its value. Combining (i) that the threshold value associated to a witness does not depend on the displacement parameter α and (ii) that we can always take into account displacement via classical post-processing if one uses homodyne or heterodyne detection associated to $\hat{\Omega}_{\mathbf{a},0}$ [32, 44], we can restrict the analysis to witnesses of the form $\hat{\Omega}_{\mathbf{a},0}$ that will generate the family $\{\hat{\Omega}_{\mathbf{a},\alpha}\}_{\alpha \in \mathbb{C}}$. Note however that the choice of displacement amplitude can play an important role for certifying negativity of certain quantum states.

If the measured expectation value for an experimental state is higher than the threshold value

given by Eq. (8), this implies by definition that its Wigner function takes negative values. Moreover, the following result shows that the amount by which the expectation value exceeds the threshold value directly provides an operational quantification of Wigner negativity for that state:

Lemma 1. *Let $\rho \in \mathcal{D}(\mathcal{H})$ Wigner negative, and fix a witness $\hat{\Omega}_{\mathbf{a},\alpha}$ defined in Eq. (6), for $n \in \mathbb{N}^*$, $\mathbf{a} = (a_1, \dots, a_n) \in [0, 1]^n$, and $\alpha \in \mathbb{C}$, with threshold value $\omega_{\mathbf{a}}$. Let us further assume that that it violates the threshold value of the witness i.e. $\text{Tr}(\hat{\Omega}_{\mathbf{a},\alpha} \rho) > \omega_{\mathbf{a}}$ and denote the amount of violation as*

$$\delta_{\mathbf{a},\alpha}(\rho) := \text{Tr}(\hat{\Omega}_{\mathbf{a},\alpha} \rho) - \omega_{\mathbf{a}}. \quad (9)$$

Then,

$$\eta_{\rho} \geq \delta_{\mathbf{a},\alpha}(\rho), \quad (10)$$

where η_{ρ} is the distance between ρ and the set of states having a positive Wigner function, defined in Eq. (5).

Proof. We use the notations of the lemma. Let us consider the binary POVM $\{\hat{\Omega}_{\mathbf{a},\alpha}, \mathbb{1} - \hat{\Omega}_{\mathbf{a},\alpha}\}$. For all $\sigma \in \mathcal{D}(\mathcal{H})$, we write $P_{\mathbf{a},\alpha}^{\sigma}$ the associated probability distribution: $P_{\mathbf{a},\alpha}^{\sigma}(0) = 1 - P_{\mathbf{a},\alpha}^{\sigma}(1) = \text{Tr}(\hat{\Omega}_{\mathbf{a},\alpha} \sigma)$.

Let σ be a state with a positive Wigner function, so that $\text{Tr}(\hat{\Omega}_{\mathbf{a},\alpha} \sigma) \leq \omega_{\mathbf{a}}$, by definition of the threshold value. We have:

$$\begin{aligned} \delta_{\mathbf{a},\alpha}(\rho) &= \text{Tr}(\hat{\Omega}_{\mathbf{a},\alpha} \rho) - \omega_{\mathbf{a}} \\ &\leq |\text{Tr}(\hat{\Omega}_{\mathbf{a},\alpha} \rho) - \text{Tr}(\hat{\Omega}_{\mathbf{a},\alpha} \sigma)| \\ &= |P_{\mathbf{a},\alpha}^{\rho}(0) - P_{\mathbf{a},\alpha}^{\sigma}(0)| \\ &= \|P_{\mathbf{a},\alpha}^{\rho} - P_{\mathbf{a},\alpha}^{\sigma}\|_{\text{tvd}} \\ &\leq D(\rho, \sigma), \end{aligned} \quad (11)$$

where we used $\delta_{\mathbf{a},\alpha}(\rho) \geq 0$ in the second line, $\|P - Q\|_{\text{tvd}} = \frac{1}{2} \sum_x |P(x) - Q(x)|$ denotes the total variation distance, and we used the operational property of the trace distance in the last line [36]. With Eq. (5), taking the infimum over σ concludes the proof. \square

This results directly extends to the case where only an upper bound of the threshold value is known: the amount by which the expectation value exceeds this upper bound is also a lower bound of the distance to the set of states having a positive Wigner function.

Importantly, the family of Wigner negativity witnesses $\{\hat{\Omega}_{\mathbf{a},\alpha}\}$ is complete, i.e., for any quantum state with negative Wigner function there exists a choice of witness (\mathbf{a}, α) such that the expectation value of $\hat{\Omega}_{\mathbf{a},\alpha}$ for this state is higher than the threshold value. Indeed, this family includes as a subclass the complete family of witnesses from [32], by taking $\mathbf{a} = (1, 0, 1, 0, 1, \dots)$.

The threshold value in Eq. (8) is given by an optimisation problem over quantum states having a positive Wigner function, which is a convex subset of an infinite-dimensional space that does not possess a well-characterised structure. While solving this optimisation problem thus seems unfeasible in general, it turns out that we can obtain increasingly good numerical upper and lower bounds for the threshold value using semidefinite programming.

Semidefinite programming is a particular case of conic programming—a subfield of convex optimisation—where one optimises linear functions within the convex cone of positive semidefinite matrices [50]. This is a powerful optimisation technique as semidefinite programs (SDP) can be solved efficiently using interior point methods.

The relevant programs are derived in section 5.4.2 and 5.4.3 where their convergence is proven. Since these proofs introduces several intermediate forms of the programs, we explicitly give them below to avoid confusion on which programs to implement numerically. For $n \in \mathbb{N}^*$, $\mathbf{a} = (a_1, \dots, a_n) \in [0, 1]^n$, and $m \geq n$, the hierarchies of semidefinite programs that respectively provide lower bounds and upper bounds for the threshold value $\omega_{\mathbf{a}}$ associated to the witnesses $\{\hat{\Omega}_{\mathbf{a},\alpha}\}_{\alpha \in \mathbb{C}}$ are:

$$\begin{aligned} &\text{Find } Q \in \text{Sym}_{m+1} \text{ and } \mathbf{F} \in \mathbb{R}^{m+1} \\ &\text{maximising } \sum_{k=1}^n a_k F_k \\ &\text{subject to} \\ &\left\{ \begin{array}{l} \sum_{k=0}^m F_k = 1 \\ \forall k \in \llbracket 0, m \rrbracket, \quad F_k \geq 0 \\ \forall l \in \llbracket 1, m \rrbracket, \quad \sum_{i+j=2l-1} Q_{ij} = 0 \\ \forall l \in \llbracket 0, m \rrbracket, \quad \sum_{i+j=2l} Q_{ij} = \sum_{k=l}^m \frac{(-1)^{k+l}}{l!} \binom{k}{l} F_k \\ Q \succeq 0, \end{array} \right. \quad (\text{SDP}_{\mathbf{a}}^{m, \leq}) \end{aligned}$$

and

Find $Q \in \text{Sym}_{m+1}$, $\boldsymbol{\mu} \in \mathbb{R}^{m+1}$ and $y \in \mathbb{R}$
 minimising y
 subject to

$$\begin{cases} \forall k \in \llbracket 0, n \rrbracket, & y \geq a_k + \mu_k \\ \forall k \in \llbracket n+1, m \rrbracket, & y \geq \mu_k \\ \forall l \in \llbracket 0, m \rrbracket, & \sum_{i+j=2l} Q_{ij} = \sum_{k=l}^m \frac{(-1)^{k+l}}{l!} \binom{k}{l} \mu_k \\ Q \succeq 0. \end{cases} \quad (\text{SDP}_{\mathbf{a}}^{m, \geq})$$

Let $\omega_{\mathbf{a}}^{m, \geq}$ be the optimal value of $(\text{SDP}_{\mathbf{a}}^{m, \geq})$. We show in section 5 that the sequence $\{\omega_{\mathbf{a}}^{m, \geq}\}_{m \geq n}$ is a decreasing sequence of upper bounds of $\omega_{\mathbf{a}}$, which converges to $\omega_{\mathbf{a}}$. Similarly, let $\omega_{\mathbf{a}}^{m, \leq}$ be the optimal value of $(\text{SDP}_{\mathbf{a}}^{m, \leq})$. We show that the sequence $\{\omega_{\mathbf{a}}^{m, \leq}\}_{m \geq n}$ is an increasing sequence of lower bounds of $\omega_{\mathbf{a}}$, which also converges to $\omega_{\mathbf{a}}$.

In particular, the numerical upper bounds $\omega_{\mathbf{a}}^{m, \geq}$ can be used instead of the threshold value $\omega_{\mathbf{a}}$ to witness Wigner negativity, while the numerical lower bounds $\omega_{\mathbf{a}}^{m, \leq}$ may be used to control how much the upper bounds differ from the threshold value. We give a detailed procedure in the following section, together with use-case examples and details on the numerical implementation.

State-of-the-art Our Wigner negativity witnesses outperform existing ones [30, 51] in terms of generality and practicality, since they form a complete family and provide much more flexibility with the choice of $n \in \mathbb{N}^*$, $\mathbf{a} \in [0, 1]^n$ and $\alpha \in \mathbb{C}$. They are accessible with optical homodyne or heterodyne measurement, and do not require making any assumption on the measured state, unlike other existing methods [52]. Moreover, our witnesses generalise those of [32], and may provide simpler alternatives to detect Wigner negativity. We also provide two converging hierarchies to approximate the threshold values associated to these witnesses. Proving convergence is crucial and was not considered in the other approaches mentioned above. Finally, our approach also generalises to the multimode setting, as discussed in section 6.

4 Witnessing Wigner negativity

4.1 Procedure

In this section, mainly devoted to experimentalists, we describe a procedure to check whether a continuous-variable quantum state exhibits Wigner negativity using our witnesses.

The main subroutine of this procedure is to estimate fidelities with displaced Fock states using classical samples from homodyne or heterodyne detection¹, in order to compute the experimental value for a Wigner negativity witness. Displacement can be achieved with classical post-processing by translating the classical samples according to the displacement amplitude, and performing direct fidelity estimation with Fock states (see section 2.2 and [25–27]).

Upper and lower bounds on the threshold value of the witness are then obtained using semidefinite programming, and comparing the experimental witness value to these bounds gives insight about the Wigner negativity of the measured quantum state.

We give a detailed procedure for using our witnesses for detecting Wigner negativity in the framed box below. This procedure starts by the choice of a specific witness, and we explain here a heuristic method for picking a good witness.

If the experimental state is anticipated to have negativity at α , then one may use the witness with parameters (n, \mathbf{a}, α) with $\mathbf{a} = (1, 0, 1, 0, \dots)$, which will detect negativity for n large enough [32]. However, this may imply having to estimate fidelities with Fock states having a high photon number with homodyne or heterodyne detection, which requires a lot of samples, while simpler witnesses can suffice for the task and be more efficient, as we show in the next section. Moreover, there are cases where the state to be characterised is fully unknown.

Instead, a simple heuristic for picking a good witness for Wigner negativity is the following:

- From samples of homodyne or heterodyne detection of multiple copies of an experimental state, estimate the expected values of witnesses in Eq. (6) for a small value of n and a large set of values \mathbf{a} and α , using the same samples for all witnesses.

¹Actually, using a fidelity witness rather than a fidelity estimate is sufficient for our purpose.

Witnessing Wigner negativity:

1. Choose a fidelity-based witness $\hat{\Omega}_{\mathbf{a},\alpha}$ defined in Eq. (6) by picking $n \in \mathbb{N}^*$, $\mathbf{a} \in [0, 1]^n$ and $\alpha \in \mathbb{C}$.
2. Run the upper bound semidefinite program ($\text{SDP}_{\mathbf{a}}^{m,\geq}$) for $m \geq n$, get a numerical estimate $\omega_{\mathbf{a}}^{m,\geq}$. These values are already computed for $\mathbf{a} = (0, \dots, 0, 1)$ and $n \leq 10$ in Table 1.
3. Run the lower bound semidefinite program ($\text{SDP}_{\mathbf{a}}^{m,\leq}$) for $m \geq n$, get a numerical estimate $\omega_{\mathbf{a}}^{m,\leq}$. These values are already computed for $\mathbf{a} = (0, \dots, 0, 1)$ and $n \leq 10$ in Table 1.
4. Estimate the expectation value for that witness of the experimental state from samples of homodyne or heterodyne detection by translating the samples by α and performing fidelity estimation with the corresponding Fock states. This yields an experimental witness value denoted ω_{exp} .
5. Compare the value obtained experimentally with the numerical bounds: if it is greater than the numerical upper bound, i.e., $\omega_{\text{exp}} \geq \omega_{\mathbf{a}}^{m,\geq}$ then the state displays Wigner negativity, and its distance to the set of Wigner positive states is lower bounded by $\omega_{\text{exp}} - \omega_{\mathbf{a}}^{m,\geq}$. Otherwise, if it is lower than the numerical lower bound, i.e., $\omega_{\text{exp}} \leq \omega_{\mathbf{a}}^{m,\leq}$ then the witness cannot detect Wigner negativity for this state. Finally, if the experimental value is between the numerical bounds, i.e., $\omega_{\mathbf{a}}^{m,\leq} < \omega_{\text{exp}} < \omega_{\mathbf{a}}^{m,\geq}$, compute tighter upper and lower bounds by running steps 3 and 4 with higher m .

- Based on these values, pick the simplest witness possible—with the smallest value of n —that is able to witness Wigner negativity with a reasonable violation. This is done by comparing the estimated expected values with the upper bounds on the corresponding

threshold values. These bounds depend only the choice of the witness parameters n , \mathbf{a} and can be precomputed using ($\text{SDP}_{\mathbf{a}}^{m,\geq}$). To facilitate the use of our methods, we have collected such bounds for $\mathbf{a} = (0, \dots, 0, 1)$ and $n \leq 10$ in Table 1. We also precomputed these bounds for $n = 3$ and a large number of values of \mathbf{a} in Appendix I and [53].

- Then, estimate the expected value for that witness using a new collection of samples—thus obtaining proper error bars and avoiding the accumulation of statistical errors.

In what follows, we give a few theoretical examples for using our witnesses to detect negativity of the Wigner function of single-mode quantum states.

4.2 Examples

We identify three levels of generality within our family of witnesses in Eq. (6): (i) fidelities with single Fock states, (ii) linear combinations of fidelities with Fock states, and (iii) displaced linear combinations of fidelities with Fock states.

Fidelities with Fock states are the most practical of our witnesses, since they require the estimation of only one diagonal element of the density matrix of the measured state. The corresponding values in Table 1 can be used directly by experimentalists: if an estimate of $\langle n|\rho|n\rangle$ for appropriate n is above one of these numerical upper bounds then it ensures that ρ has a Wigner function with negative values. Moreover, by Lemma 1, the amount by which the estimate of $\langle n|\rho|n\rangle$ exceeds the numerical upper bound directly provides a lower bound on the distance between ρ and the set of states having a positive Wigner function.

For instance, if we focus on $n = 3$ in Table 1, the threshold value ω_3 satisfies $0.378 \leq \omega_3 \leq 0.427$. Having a state ρ such that $\langle 3|\rho|3\rangle > 0.427$ guarantees that ρ has Wigner negativity. If $\langle 3|\rho|3\rangle < 0.378$ then we conclude that the witness cannot detect negativity for this state. When the experimental state is close to a Fock state (different from the vacuum), a natural choice for the witness thus is the fidelity with the corresponding Fock state. For instance, consider a photon-subtracted squeezed vacuum state [54]

$$|p\text{-ssvs}(r)\rangle = \frac{1}{\sinh r} \hat{a} \hat{S}(r) |0\rangle, \quad (12)$$

n	Lower bound	Upper bound
1	0.5	0.5
2	0.5	0.5
3	0.378	0.427
4	0.375	0.441
5	0.314	0.385
6	0.314	0.378
7	0.277	0.344
8	0.280	0.348
9	0.256	0.341
10	0.262	0.334

Table 1: Table of numerical upper and lower bounds for the threshold value ω_n of various Wigner negativity witnesses obtained using our hierarchies of semidefinite programs at rank up to around 30. The witnesses considered here are Fock states projectors (photon-number states $|n\rangle$) from 1 to 10. Note that the gap between the lower and upper bounds never exceeds 0.1. Additionally, the bounds in the first two lines are analytical (see section 5.3) and for the upper bounds, the corresponding numerical values are 0.528 and 0.551, respectively. See section 4.3 and [53] for the numerical implementation.

where $\hat{S}(r) = e^{\frac{r}{2}(\hat{a}^2 - \hat{a}^{\dagger 2})}$ is a squeezing operator with parameter $r \in \mathbb{R}$. Its fidelity with the single-photon Fock state $|1\rangle$ is given by:

$$\frac{1}{(\sinh r)^2} \left| \langle 1 | \hat{a} \hat{S}(r) | 0 \rangle \right|^2 = \frac{1}{(\cosh r)^3}. \quad (13)$$

When the squeezing parameter is small, this state is close to a single-photon Fock state. In particular, for $0 < r < 0.70$, the fidelity $F(\text{p-ssvs}(r), 1)$ in Eq. (13) is greater than $\omega_1^> = \frac{1}{2}$ and our witness can be used to detect Wigner negativity of this state (see Fig. 2 (a)).

Another example is given by superpositions of coherent states: we consider the cat state [55]

$$|\text{cat}_2(\alpha)\rangle = \frac{|\alpha\rangle + |-\alpha\rangle}{\sqrt{2(1 + e^{-2|\alpha|^2})}}, \quad (14)$$

and the compass state [56]

$$|\text{cat}_4(\alpha)\rangle = \frac{|\alpha\rangle + |-\alpha\rangle + |i\alpha\rangle + |-i\alpha\rangle}{2\sqrt{1 + e^{-|\alpha|^2}(2\cos(|\alpha|^2) + 1)}}, \quad (15)$$

where $|\alpha\rangle = e^{-\frac{1}{2}|\alpha|^2} \sum_{k \geq 0} \frac{\alpha^k}{\sqrt{k!}} |k\rangle$ is the coherent state of amplitude $\alpha \in \mathbb{C}$. We have

$$|\langle 2 | \text{cat}_2(\alpha) \rangle|^2 = \frac{|\alpha|^4}{2 \cosh(|\alpha|^2)}, \quad (16)$$

and

$$|\langle 4 | \text{cat}_4(\alpha) \rangle|^2 = \frac{|\alpha|^8/12}{\cosh(|\alpha|^2) + \cos(|\alpha|^2)}. \quad (17)$$

For $1.63 \leq |\alpha|^2 \leq 2.59$, the fidelity $F(\text{cat}_2(\alpha), 2)$ in Eq. (16) is greater than $\omega_2^> = \frac{1}{2}$ and our witness corresponding to $n = 2$ can be used to detect Wigner negativity of this state. Similarly, for $2.10 \leq |\alpha|^2 \leq 6.70$, the fidelity $F(\text{cat}_4(\alpha), 4)$ in Eq. (17) is greater than $\omega_4^> = 0.441$ and our witness corresponding to $n = 4$ can be used to detect Wigner negativity of this state (see Fig. 2 (b) and (c)).

Some quantum states will remain unnoticed by all single Fock state negativity witnesses. For example, the state $\rho_{0,1,2} := \frac{1}{9} |0\rangle\langle 0| + \frac{4}{9} |1\rangle\langle 1| + \frac{4}{9} |2\rangle\langle 2|$ has a negative Wigner function but is not detected by any of the single Fock state negativity witnesses, since the lower bounds for $n = 1, 2$ in Table 1 are higher than $\frac{4}{9}$, and this state has fidelity 0 with higher Fock states. However, it is detected by simple witnesses based on linear combinations of fidelities. For example, with $n = 2$ and $\mathbf{a} = (1, 1)$ we find numerically that the threshold value of the witness $|1\rangle\langle 1| + |2\rangle\langle 2|$ is less than 0.875 when running the corresponding program ($\text{SDP}_{\mathbf{a}}^{m, \geq}$) for $m = 7$. And since $\text{Tr}(\rho_{0,1,2}(|1\rangle\langle 1| + |2\rangle\langle 2|)) = \frac{8}{9} > 0.875$, this linear combination of Fock state fidelities can indeed detect Wigner negativity for this state.

However, some quantum states with a negative Wigner function will always go unnoticed by the previous witnesses because these witnesses are invariant under phase-space rotations while the Wigner function of those states becomes positive under random dephasing: consider for instance the superposition $\sqrt{1 - \frac{1}{s}} |0\rangle + \frac{1}{\sqrt{s}} |1\rangle$, for $s > 2$ (which under random dephasing is mapped to $(1 - \frac{1}{s}) |0\rangle\langle 0| + \frac{1}{s} |1\rangle\langle 1|$). In that case, the Wigner negativity of such states can still be witnessed by using the displaced version of our witnesses. In particular, if any single-mode quantum state has a Wigner function negative at $\alpha \in \mathbb{C}$, then there is a choice of $n \in \mathbb{N}^*$ such that the witness in Eq. (6) defined by $\mathbf{a} = (1, 0, 1, 0, 1, \dots)$ and

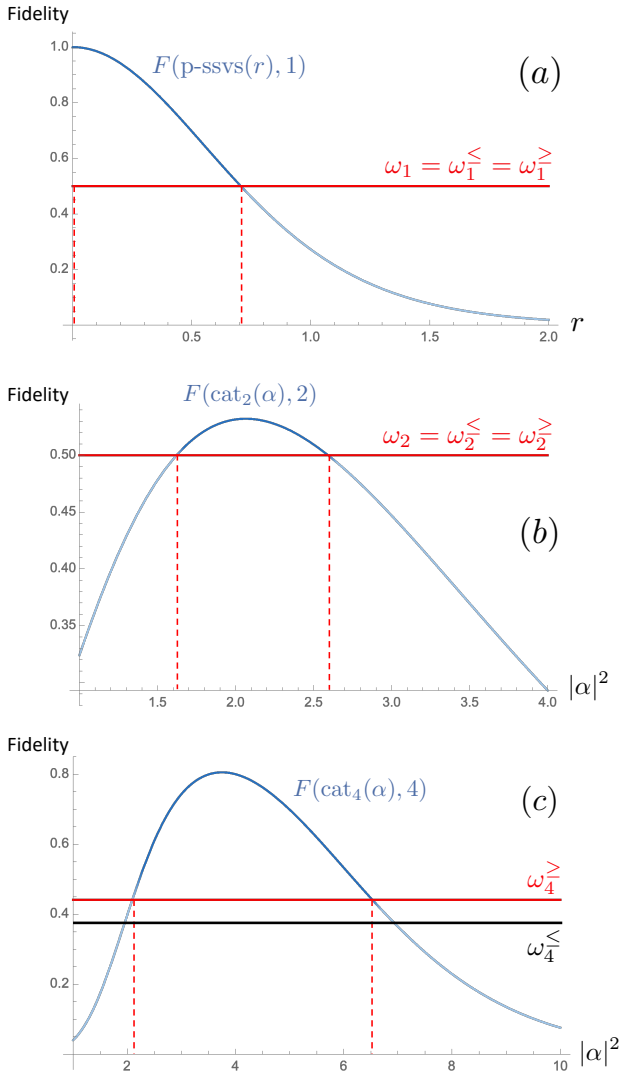


Figure 2: (a) Fidelities of photon-subtracted squeezed vacuum states $|p\text{-ssvs}(r)\rangle$ with squeezing parameter $r \in \mathbb{R}$ with the Fock state $|1\rangle$. (b) Fidelities of cat states $|\text{cat}_2(\alpha)\rangle$ with amplitude $\alpha \in \mathbb{C}$ with the Fock state $|2\rangle$. (c) Fidelities of compass states $|\text{cat}_4(\alpha)\rangle$ with amplitude $\alpha \in \mathbb{C}$ with the Fock state $|4\rangle$. The dashed red lines delimit the intervals of parameter values where our witnesses from Table 1 can be used to detect Wigner negativity of the corresponding state, i.e., when the fidelity (blue curve) is above the witness upper bound (red line). When it is below the witness lower bound (black line), we are guaranteed that the witness (here $|4\rangle\langle 4|$) cannot be used to detect Wigner negativity of the state. In all cases, the height difference when the blue curve is above the red line directly provides a lower bound on the distance between the corresponding state and the set of states having a positive Wigner function.

the displacement amplitude α detects its negativity [32]. In practice, simpler witnesses may suffice to detect negativity, and the choice of witness will ultimately depend on the experimental state at hand. Hereafter we discuss the heuristics

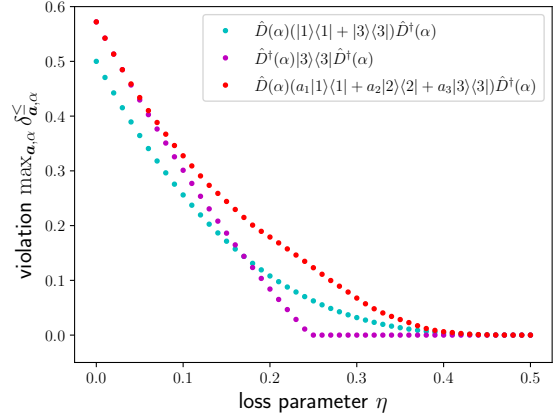


Figure 3: Lower bounds on the violation $\delta_{\mathbf{a},\alpha} = \text{Tr}(\hat{\Omega}_{\mathbf{a},\alpha} \rho_{3,\eta}) - \omega_{\mathbf{a}}$ for a lossy 3-photon Fock state $\rho_{3,\eta} = \eta^3 |3\rangle\langle 3| + 3\eta^2(1-\eta) |2\rangle\langle 2| + 3\eta(1-\eta)^2 |1\rangle\langle 1| + (1-\eta)^3 |0\rangle\langle 0|$ for $\mathbf{a} = (a_1, a_2, a_3)$ with the loss parameter η . Precomputed bounds on threshold values for witnesses of the form $\hat{\Omega}_{\mathbf{a}} = a_1 |1\rangle\langle 1| + a_2 |2\rangle\langle 2| + a_3 |3\rangle\langle 3|$ can be found in Appendix I. We use these values to find the witness $\hat{\Omega}_{\mathbf{a},\alpha}$ giving the maximum lower bound $\delta_{\mathbf{a},\alpha}^{\leq} = \text{Tr}(\hat{\Omega}_{\mathbf{a},\alpha} \rho_{3,\eta}) - \omega_{\mathbf{a}}^{\geq}$ on the violation $\delta_{\mathbf{a},\alpha}$, for different values of the loss parameter η . These lower bounds are represented in red. In blue is the maximal violation that can be detected using the witnesses $\hat{\Omega}_{(1,0,1),\alpha} = \hat{D}(\alpha)(|1\rangle\langle 1| + |3\rangle\langle 3|)\hat{D}^\dagger(\alpha)$ [32]. In violet is the maximal violation that can be detected using the more naive witness $\hat{\Omega}_{(0,0,1),\alpha} = D(\alpha) |3\rangle\langle 3| D^\dagger(\alpha)$. Note that $\rho_{3,\eta}$ has a non-negative Wigner function for $\eta \geq 0.5$.

for picking a good witness, with the theoretical example of the lossy 3-photon Fock state:

$$\rho_{3,\eta} := (1-\eta)^3 |3\rangle\langle 3| + 3\eta(1-\eta)^2 |2\rangle\langle 2| + 3\eta^2(1-\eta) |1\rangle\langle 1| + \eta^3 |0\rangle\langle 0|, \quad (18)$$

where $0 \leq \eta \leq 1$ is the loss parameter. Setting $\eta = 0$ gives $\rho_{3,\eta} = |3\rangle\langle 3|$ while setting $\eta = 1$ gives $\rho_{3,\eta} = |0\rangle\langle 0|$. This state has a non-negative Wigner function for $\eta \geq \frac{1}{2}$. The fidelities of $\rho_{3,\eta}$

with displaced Fock states $\hat{D}(\alpha)|l\rangle$ are given by:

$$\begin{aligned}
F(\rho_{3,\eta}, \hat{D}(\alpha)|l\rangle) &= (1-\eta)^3 |\langle 3|\hat{D}(\alpha)|l\rangle|^2 \\
&+ 3\eta(1-\eta)^2 |\langle 2|\hat{D}(\alpha)|l\rangle|^2 \quad (19) \\
&+ 3\eta^2(1-\eta) |\langle 1|\hat{D}(\alpha)|l\rangle|^2 \\
&+ \eta^3 |\langle 0|\hat{D}(\alpha)|l\rangle|^2.
\end{aligned}$$

where the coefficients of the displacement operator in Fock basis are given in Eq. (3). In an experimental scenario, the state would be unknown and these fidelities should be estimated using samples from a homodyne or heterodyne detection of the state translated by α , and estimating the fidelities with Fock states [32, 44].

Following the heuristic detailed in the previous section, we have determined good Wigner negativity witnesses for 50 values of the loss parameter η between 0 and 0.5 as follows: for each value η , we have computed numerically the values of the fidelities in Eq. (19) for $l = 1, 2, 3$, and for displacement parameters $\alpha = q/10 + ip/10$ for all $q, p \in \llbracket 0, 10 \rrbracket$. Using these values, we have computed the expectation value of the witnesses $\hat{\Omega}_{\mathbf{a},\alpha}$ for multiple choices of $\mathbf{a} = (a_1, a_2, a_3)$ with $\max_i a_i = 1$. We have used the corresponding precomputed bounds $\omega_{\mathbf{a}}^{\geq}$ on the threshold values of $\hat{\Omega}_{\mathbf{a},\alpha}$ in Appendix I to determine the witness leading to the maximal lower bound $\delta_{\mathbf{a},\alpha}^{\leq} := \text{Tr}(\hat{\Omega}_{\mathbf{a},\alpha} \rho_{3,\eta}) - \omega_{\mathbf{a}}^{\geq}$ on the violation $\delta_{\mathbf{a},\alpha} = \text{Tr}(\hat{\Omega}_{\mathbf{a},\alpha} \rho_{3,\eta}) - \omega_{\mathbf{a}}$ over the choice of (\mathbf{a}, α) .

We have represented these violations for each value of the loss parameter η in Fig. 3. For all values of η , we find that the optimal displacement parameter is $\alpha = 0$. On the other hand, we find different optimal choices of \mathbf{a} for different values of η . To illustrate the usefulness of the optimisation over the choice of witnesses parametrised by (\mathbf{a}, α) , we have also represented the violations obtained when using the witnesses $\hat{\Omega}_{(1,0,1),\alpha} = \hat{D}(\alpha)(|1\rangle\langle 1| + |3\rangle\langle 3|)\hat{D}^\dagger(\alpha)$ from [32] for all values of η . In that setting, the violations obtained quantify how hard it is to detect the Wigner negativity of the state: a larger violation implies that a less precise estimate of the witness expectation value is needed to witness Wigner negativity. In particular, we obtain that our optimised witnesses always provide a greater violation to detect negativity than the previous witnesses which will result in an easier experimental detection. We also represented the vio-

lation obtained when using the more naive witnesses $\hat{\Omega}_{(0,0,1),\alpha} = \hat{D}(\alpha)|3\rangle\langle 3|\hat{D}^\dagger(\alpha)$ and we see that it is only useful when the loss parameter is smaller than 0.25, while the optimised witnesses may detect negativity of the state $\rho_{3,\eta}$ up to $\eta = 0.5$ —when the Wigner functions becomes non-negative—provided the estimates of the fidelities are precise enough.

Overall, this procedure only amounts to a simple classical post-processing of samples from homodyne or heterodyne detection and yields a good witness for detecting Wigner negativity.

4.3 Numerical implementation

Here we discuss numerical implementations of the semidefinite programs ($\text{SDP}_{\mathbf{a}}^{m,\leq}$) and ($\text{SDP}_{\mathbf{a}}^{m,\geq}$). All codes are available [here](#) [53].

We implemented the semidefinite programs with Python through the interface provided by PICOS [57]. We first used the solver Mosek [58] to solve these problems but, while the size of the semidefinite programs remains relatively low for small values of n and m , binomial terms grow rapidly and numerical precision issues arise quickly (usually for $m = 12, n \leq m$). The linear constraints involving Q_{ij} in the semidefinite programs come from a polynomial equality (see Lemma 3). While polynomial equalities are usually written in the canonical basis, a first trick is to express them in a different basis—for instance the basis $(1, \frac{X}{1!}, \frac{X^2}{2!}, \dots)$ —to counterbalance the binomial terms.

However, this may not be sufficient to probe larger values of m . Instead, we used the solver SDPA-GMP [59, 60] which allows arbitrary precision arithmetic. While much slower, this solver is dedicated to solve problems requiring a lot of precision. Because our problems remain rather small, time efficiency is not an issue and this solver is particularly well-suited. All problems were initially solved on a regular laptop as warning flags on optimality were raised before the problems were too large. A high-performance computer²—handling floating point arithmetic more accurately—was later used to compute further ranks in the hierarchy.

Using the semidefinite programs ($\text{SDP}_{\mathbf{a}}^{m,\geq}$) and ($\text{SDP}_{\mathbf{a}}^{m,\leq}$) for values of m up to around 30 and

²DELL PowerEdge R440, 384 Gb RAM, Intel Xeon Silver 4216 processor, 64 threads from LIP6.

$\mathbf{a} = (0, 0, \dots, 0, 1)$ (where the size n of the vector \mathbf{a} is ranging from 1 to 10), we have obtained upper and lower bounds for the threshold values of Wigner negativity witnesses corresponding to fidelities with Fock states from 1 to 10, reported in Table 1.

We also computed upper and lower bounds on the threshold values of witnesses of the form:

$$\hat{\Omega}_{(a_1, a_2, a_3)} = \sum_{k=1}^3 a_k |k\rangle\langle k|, \quad (20)$$

where $\forall i \in \{1, 2, 3\}$, $0 \leq a_i \leq 1$ and $\max_i a_i = 1$. We focused on these particular witnesses for experimental considerations as it is challenging to obtain fidelities with higher Fock states. We fix one coefficient equal to 1 and vary each other a_i from 0 to 1 with a step of 0.1. The resulting bounds on the threshold values can be found in appendix I.

We now turn to the mathematical proofs of our results, i.e., that the threshold values in Eq. (8) can be upper bounded and lower bounded by the optimal values of the converging hierarchies of semidefinite programs $(\text{SDP}_{\mathbf{a}}^{m, \geq})_{m \geq n}$ and $(\text{SDP}_{\mathbf{a}}^{m, \leq})_{m \geq n}$, respectively.

The following section is rather technical as we dive into infinite-dimensional optimisation techniques to prove the convergence of the two hierarchies of semidefinite programs. Some readers may want to skip directly to section 6.

5 Infinite-dimensional optimisation

In this section we use infinite-dimensional optimisation techniques: (i) to phrase the computation of the witness threshold value introduced in Eq. (8) as an infinite-dimensional linear program in section 5.3, (ii) to derive two hierarchies of finite-dimensional semidefinite programs that upper bound and lower bound the threshold value in section 5.4, and (iii) to show that these sequences of upper and lower bounds both converge to the threshold value in section 5.5 (see Fig. 4). Given the technicalities of the proofs above, we sketch them in section 5.1 before detailing them in the following sections.

For clarity, we treat the case where the witnesses are given by the fidelity with a single Fock state, corresponding to the case where one entry of the vector \mathbf{a} is equal to 1 and all the other

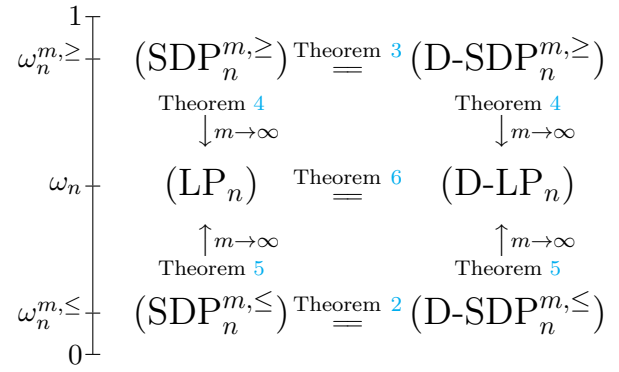


Figure 4: Hierarchies of semidefinite relaxations and restrictions converging to the linear program (LP_n) , together with their dual programs. The upper index m denotes the level of the relaxation or restriction. On the left are the associated optimal values. The equal sign denotes strong duality, i.e., equality of optimal values, and the arrows denote convergence of the corresponding sequences of optimal values.

entries are 0. The generalisation to linear combinations of fidelities with Fock states is straightforward by linearity.

A linear program is an optimisation problem where variables are linearly constrained [61]. This can be expressed as:

$$\sup_{\mathbf{x} \in K} f(\mathbf{x}), \quad (21)$$

for some linear real-valued function f and some set $K \subset E$ where E is the optimisation space which has the structure of a locally convex topological vector space and K is specified by a set of linear constraints.

An element $\mathbf{x} \in E$ which belongs to K , i.e., which satisfies all the linear constraints, is called a *feasible plan* or *feasible solution*, and *strictly feasible solution* when it strictly satisfies the constraints. The set of feasible plans is called the *feasible set*. If the supremum in Eq. (21) is attained, a feasible plan that reaches it is called an *optimal plan* or *optimal solution*.

5.1 Sketch of the proofs

This section aims at giving an overview of the rather technical proofs that follow. Our reasoning can be split into the following steps:

Obtaining an infinite linear program and its dual

- We exploit the phase-space rotational invariance of Fock states in Lemma 2 to express the computation of threshold values Eq. (8) on states that are diagonal in the Fock basis.
- This provides a maximisation problem which can be rephrased as an infinite-dimensional linear program (LP_n) on a space of functions.
- By duality, we recast this as an optimisation problem on measures which is given by the dual infinite-dimensional linear program (D-LP_n).

Hierarchy of relaxations

See upper part of figure 4.

- We relax the program (LP_n): instead of asking for a positive function, we require that this function has a positive inner product with positive polynomials of fixed degree m . This degree fixes a level within a hierarchy of relaxations. These constraints can be cast as a positive semidefinite constraint and we thus obtain a hierarchy of semidefinite programs (see ($\text{SDP}_n^{m,\geq}$)).
- For each level m , we derive the dual program ($\text{D-SDP}_n^{m,\geq}$) and show that strong duality holds in Theorem 3 by finding a strictly feasible solution of ($\text{SDP}_n^{m,\geq}$).
- We prove the convergence of the hierarchy of semidefinite relaxations towards (D-LP_n) in Theorem 4 by first showing that the feasible set of ($\text{SDP}_n^{m,\geq}$) is compact. Then, we perform a diagonal extraction on a sequence of optimal solutions of ($\text{SDP}_n^{m,\geq}$) and we finally show that this provides a feasible solution for (LP_n).

Convergence of hierarchy of restrictions

See bottom part of figure 4.

- We restrict the program (LP_n): instead of searching for a positive function, we look for a positive polynomial of fixed degree m . Again, this degree fixes a level within a hierarchy of restrictions. Using the fact that univariate positive polynomials are sum-of-squares, we obtain a hierarchy of semidefinite programs (see ($\text{SDP}_n^{m,\leq}$)).

- For each level m , we derive the dual program ($\text{D-SDP}_n^{m,\leq}$) and show that strong duality holds in Theorem 2 by finding a strictly feasible solution of ($\text{SDP}_n^{m,\leq}$).
- We prove the convergence of the hierarchy in Theorem 5 by first showing that the feasible set of ($\text{D-SDP}_n^{m,\leq}$) is compact. This is nontrivial since it requires exhibiting an analytical feasible solution. Then, we perform a diagonal extraction on a sequence of optimal solutions of ($\text{D-SDP}_n^{m,\leq}$). A technicality arises as it does not necessarily provide a feasible solution of (D-LP_n). We thus extend the program to a larger feasible set (the space of tempered distributions over \mathbb{R}_+ rather than $L^2(\mathbb{R}^+)$), obtaining the program ($\text{D-LP}_n^{\mathcal{S}}$). We then show that a diagonal extraction provides a feasible solution of ($\text{D-LP}_n^{\mathcal{S}}$). Since ($\text{D-LP}_n^{\mathcal{S}}$) is itself a relaxation of (D-LP_n), weak duality inequalities, together with the strong duality between ($\text{SDP}_n^{m,\leq}$) and ($\text{D-SDP}_n^{m,\leq}$) allow us to conclude the proof.

5.2 Function spaces

We now review some function spaces which appear in the following sections, together with a few notations.

Hereafter, the half line of non-negative real numbers is denoted \mathbb{R}_+ . The space of real square-integrable functions over \mathbb{R}_+ is denoted $L^2(\mathbb{R}_+)$ and is equipped with the usual scalar product:

$$\langle f, g \rangle = \int_{\mathbb{R}_+} f(x)g(x)dx, \quad (22)$$

for $f, g \in L^2(\mathbb{R}_+)$. This space is isomorphic to the space of square-summable real sequences l^2 , by considering the expansion in a countable basis. Such a basis is given, e.g., by the Laguerre functions [62], modified here by a $(-1)^k$ prefactor to correspond to Fock state Wigner functions:

$$\mathcal{L}_k(x) := (-1)^k L_k(x) e^{-\frac{x}{2}}, \quad (23)$$

for all $k \in \mathbb{N}$ and all $x \in \mathbb{R}_+$, where $L_k(x) = \sum_{l=0}^k \frac{(-1)^l}{l!} \binom{k}{l} x^l$ is the k^{th} Laguerre polynomial. These functions form an orthonormal basis: for all $p, q \in \mathbb{N}$, $\langle \mathcal{L}_p, \mathcal{L}_q \rangle = \delta_{pq}$, where δ_{pq} is the Kronecker symbol.

The space $L^2(\mathbb{R}_+)$ is also isomorphic to its dual space $L^{2'}(\mathbb{R}_+)$ via the Radon–Nikodym theorem

[63]: elements of $L^2(\mathbb{R}_+)$ can be identified by the Lebesgue measure on \mathbb{R}_+ times the corresponding function in $L^2(\mathbb{R}_+)$.

We write $\mathcal{S}(\mathbb{R}_+)$ the space of Schwartz functions over \mathbb{R}_+ , i.e., the space of C^∞ functions that go to 0 at infinity faster than any inverse polynomial, as do their derivatives. $\mathcal{S}'(\mathbb{R}_+)$ is its dual space, the space of tempered distributions over \mathbb{R}_+ . We denote the space of rapidly decreasing real sequences by $\mathcal{S}(\mathbb{N})$ (sequences that go to 0 at infinity faster than any inverse polynomial), together with its dual space of slowly increasing real sequences $\mathcal{S}'(\mathbb{N})$ (sequences that are upper bounded by a polynomial). The spaces $\mathcal{S}(\mathbb{R}_+)$ and $\mathcal{S}(\mathbb{N})$ are isomorphic: any Schwartz function over \mathbb{R}_+ can be expanded uniquely in the basis of Laguerre functions with a rapidly decreasing sequence of coefficients. Similarly, the spaces $\mathcal{S}'(\mathbb{R}_+)$ and $\mathcal{S}'(\mathbb{N})$ are also isomorphic: any tempered distribution over \mathbb{R}_+ can be written uniquely as a formal series of Laguerre functions with a slowly increasing sequence of coefficients [64]. We extend the definition of the duality $\langle -, - \rangle$ in Eq. (22) to these spaces.

In order to denote non-negative elements of these spaces, we will use the notations $L_+^2(\mathbb{R}_+)$, $L_+^{2'}(\mathbb{R}_+)$, $\mathcal{S}_+(\mathbb{R}_+)$ and $\mathcal{S}'_+(\mathbb{R}_+)$. A distribution μ in $L_+^{2'}(\mathbb{R}_+)$ (resp. in $\mathcal{S}'_+(\mathbb{R}_+)$) satisfies: $\forall f \in L_+^2(\mathbb{R}_+)$ (resp. $\forall f \in \mathcal{S}_+(\mathbb{R}_+)$), $\langle \mu, f \rangle \geq 0$.

For all $m \in \mathbb{N}$, we define the following space of truncated series of Laguerre functions over \mathbb{R}_+ :

$$\mathcal{R}_m(\mathbb{R}_+) := \text{span}_{\mathbb{R}}\{\mathcal{L}_k\}_{0 \leq k \leq m}, \quad (24)$$

which is equal to the set of real polynomials over \mathbb{R}_+ of degree less or equal to m multiplied by the function $x \mapsto e^{-\frac{x}{2}}$. We denote by $\mathcal{R}_{m,+}(\mathbb{R}_+)$ its subset of non-negative elements.

For all $\mathbf{s} \in \mathbb{R}^{\mathbb{N}}$, we define the associated formal series of Laguerre functions:

$$f_{\mathbf{s}} := \sum_{k \geq 0} s_k \mathcal{L}_k, \quad (25)$$

with the (formal) relation:

$$s_k = \langle f_{\mathbf{s}}, \mathcal{L}_k \rangle, \quad (26)$$

for all $k \in \mathbb{N}$. We refer to \mathbf{s} as the sequence of Laguerre moments of $f_{\mathbf{s}}$. We extend this definition to finite sequences by completing these sequences with zeros. For $m \in \mathbb{N}$, we also define the matrix

$A_{\mathbf{s}}$ (thus omitting the dependence in m) by

$$(A_{\mathbf{s}})_{0 \leq i, j \leq m} := \begin{cases} \sum_{k=0}^l s_k \binom{l}{k}! & \text{if } i + j = 2l, \\ 0 & \text{otherwise.} \end{cases} \quad (27)$$

$A_{\mathbf{s}}$ can be seen as the Laguerre moment matrix of the measure $f_{\mathbf{s}}$. In what follows, we use standard techniques relating to the Stieltjes moment problem [65], which seeks conditions for a real sequence $\nu = (\nu_k)_{k \in \mathbb{N}} \in \mathbb{R}^{\mathbb{N}}$ to be the sequence of moments $\int_{\mathbb{R}_+} x^k d\nu(x)$ of a non-negative distribution ν over \mathbb{R}_+ . We adapt these techniques to the basis of Laguerre functions, rather than the canonical basis. In particular, we make use of the following result:

Theorem 1. *Let $\mu = (\mu_k)_{k \in \mathbb{N}} \in \mathbb{R}^{\mathbb{N}}$. The sequence μ is the sequence of Laguerre moments $\int_{\mathbb{R}_+} \mathcal{L}_k(x) d\mu(x)$ of a non-negative distribution μ supported on \mathbb{R}_+ if and only if*

$$\forall m \in \mathbb{N}, \forall g \in \mathcal{R}_{m,+}(\mathbb{R}_+), \langle f_{\mu}, g \rangle \geq 0. \quad (28)$$

We give a proof in Appendix A for this result, which is based on the classical Riesz–Haviland theorem [66, 67].

5.3 Linear program

In this section, we phrase the computation of the witness threshold value introduced in Eq. (8) as an infinite-dimensional linear program, in the case where one entry of the vector \mathbf{a} is equal to 1 and all the other entries are 0, the generalisation being straightforward by linearity.

Formally, we fix hereafter $n \in \mathbb{N}^*$ and we look for the witnesses threshold value ω_n defined as

$$\omega_n := \sup_{\substack{\rho \in \mathcal{D}(\mathcal{H}) \\ W_{\rho} \geq 0}} \langle n | \rho | n \rangle. \quad (29)$$

This is the maximal values such that for all states $\rho \in \mathcal{D}(\mathcal{H})$:

$$\langle n | \rho | n \rangle > \omega_n \quad \Rightarrow \quad \exists \alpha \in \mathbb{C}, W_{\rho}(\alpha) < 0. \quad (30)$$

Let $\mathcal{C}(\mathcal{H})$ be the set of states that are invariant under phase-space rotations:

$$\mathcal{C}(\mathcal{H}) := \{ \sigma \in \mathcal{D}(\mathcal{H}) \text{ such that: } \forall \varphi \in [0, 2\pi], e^{i\varphi \hat{n}} \sigma e^{-i\varphi \hat{n}} = \sigma \}, \quad (31)$$

where $\hat{n} = \hat{a}^\dagger \hat{a}$ is the number operator. The witnesses corresponding to the fidelity with a single Fock state feature a rotational symmetry in phase space, which we exploit in the following lemma.

Lemma 2. *The threshold value in Eq. (29) can be expressed as*

$$\omega_n = \sup_{\substack{\sigma \in \mathcal{C}(\mathcal{H}) \\ W_\sigma \geq 0}} \langle n | \sigma | n \rangle. \quad (32)$$

Proof. Let $\rho \in \mathcal{D}(\mathcal{H})$. We start by applying a random dephasing to the state ρ :

$$\sigma = \int_0^{2\pi} \frac{d\varphi}{2\pi} e^{i\varphi \hat{n}} \rho e^{-i\varphi \hat{n}} \in \mathcal{C}(\mathcal{H}). \quad (33)$$

The random dephasing does not change the fidelity with any Fock state because of the rotational symmetry in phase space of the latter, i.e.,

$$\forall n \in \mathbb{N}, \langle n | \sigma | n \rangle = \langle n | \rho | n \rangle. \quad (34)$$

Moreover, it can only decrease the maximum negativity of the Wigner function: for all $\alpha \in \mathbb{C}$,

$$\begin{aligned} W_\sigma(\alpha) &= \int_0^{2\pi} \frac{d\varphi}{2\pi} W_{e^{i\varphi \hat{n}} \rho e^{-i\varphi \hat{n}}}(\alpha) \\ &= \int_0^{2\pi} \frac{d\varphi}{2\pi} W_\rho(\alpha e^{i\varphi}) \\ &\geq \min_{\varphi \in [0, 2\pi]} W_\rho(\alpha e^{i\varphi}) \\ &\geq \min_{\beta \in \mathbb{C}} W_\rho(\beta), \end{aligned} \quad (35)$$

and taking the minimum over all $\alpha \in \mathbb{C}$ then gives

$$\min_{\alpha \in \mathbb{C}} W_\sigma(\alpha) \geq \min_{\beta \in \mathbb{C}} W_\rho(\beta). \quad (36)$$

In particular, applying a random dephasing to a Wigner positive state yields a Wigner positive mixtures of Fock states, which is invariant under phase-space rotations. Hence, we can restrict without loss of generality to states that are invariant under phase-space rotations when looking for the maximum fidelity of Wigner positive states with a given Fock state $|n\rangle$. \square

Lemma 2 ensures that the supremum in Eq. (29) can be computed over states that have a rotational symmetry in phase space. Such states σ can be expanded diagonally in the Fock basis:

$$\sigma = \sum_{k=0}^{\infty} F_k |k\rangle\langle k|, \quad (37)$$

with $\sum_k F_k = 1$ and $0 \leq F_k \leq 1$ for all $k \in \mathbb{N}$. By linearity of the Wigner function:

$$\forall \alpha \in \mathbb{C}, W_\sigma(\alpha) = \sum_k F_k W_k(\alpha), \quad (38)$$

where W_k is the Wigner function of the k^{th} Fock state [20]:

$$\forall \alpha \in \mathbb{C}, W_k(\alpha) = \frac{2}{\pi} \mathcal{L}_k(4|\alpha|^2), \quad (39)$$

with \mathcal{L}_k the k^{th} Laguerre function, defined in Eq. (23). As noted before, Fock states are invariant under phase-space rotations: their Wigner function only depends on the amplitude of the phase-space point considered. We fix $x = 4|\alpha|^2 \in \mathbb{R}^+$ hereafter.

With Lemma 2, the computation of the threshold value for the witness ω_n can thus be expressed as the following infinite-dimensional linear program:

$$\begin{aligned} &\text{Find } (F_k)_{k \in \mathbb{N}} \in \ell^2 \\ &\text{maximising } F_n \\ &\text{subject to} \\ &\left\{ \begin{array}{l} \sum_k F_k = 1 \\ \forall k \in \mathbb{N}, F_k \geq 0 \\ \forall x \in \mathbb{R}_+, \sum_k F_k \mathcal{L}_k(x) \geq 0. \end{array} \right. \quad (\text{LP}_n) \end{aligned}$$

The first constraint ensures unit trace of the corresponding state σ , the second one ensures that its fidelity with each Fock state is non-negative, and the last one ensures that its Wigner function W_σ is non-negative. Note that $\omega_n > 0$ for all $n \in \mathbb{N}^*$, by considering a mixture of $|0\rangle$ and $|n\rangle$ with the vacuum component close enough to 1.

We refer the interested reader to Appendix B where this program is expressed in the canonical form of infinite-dimensional linear programs [61] and its dual is derived. This dual linear program reads:

$$\begin{aligned} &\text{Find } y \in \mathbb{R} \text{ and } \mu \in L^2(\mathbb{R}_+) \\ &\text{minimising } y \\ &\text{subject to} \\ &\left\{ \begin{array}{l} \forall k \neq n \in \mathbb{N}, y \geq \int_{\mathbb{R}_+} \mathcal{L}_k d\mu \\ y \geq 1 + \int_{\mathbb{R}_+} \mathcal{L}_n d\mu \\ \forall f \in L^2_+(\mathbb{R}_+), \langle \mu, f \rangle \geq 0. \end{array} \right. \quad (\text{D-LP}_n) \end{aligned}$$

Weak duality of linear programming ensures that the optimal value ω_n of (LP_n) is upper bounded by the optimal value of (D-LP_n) . Hence, a possible way of solving the optimisation (LP_n) is to

exhibit a feasible solution for (LP_n) and a feasible solution for (D-LP_n) that have the same value.

For $n = 1$, choosing $(F_k)_{k \in \mathbb{N}} = (\frac{1}{2}, \frac{1}{2}, 0, 0, \dots)$ gives a feasible solution for (LP_1) with the value $\frac{1}{2}$, while choosing $(y, \mu) = (\frac{1}{2}, \frac{1}{2}\delta(x))$, where δ is the Dirac delta function³ over \mathbb{R}_+ , gives a feasible solution for (D-LP_1) with the value $\frac{1}{2}$. This shows that $\omega_1 = \frac{1}{2}$.

Similarly, for $n = 2$, choosing $(F_k)_{k \in \mathbb{N}} = (\frac{1}{2}, 0, \frac{1}{2}, 0, 0, \dots)$ gives a feasible solution for (LP_2) with the value $\frac{1}{2}$, while choosing $(y, \mu) = (\frac{1}{2}, \frac{\varepsilon}{2}\delta(x - 2))$ gives a feasible solution for (D-LP_2) , up to a conjecture⁴, with the value $\frac{1}{2}$. This shows that $\omega_2 = \frac{1}{2}$.

While this approach is sensible for small values of n , finding optimal analytical solutions for higher values of n seems highly nontrivial. Moreover, the infinite number of variables prevents us from performing the optimisation (LP_n) numerically. A natural workaround is to find finite-dimensional relaxations or restrictions of the original problem—thus providing upper and lower bounds for the optimal value ω_n , respectively. This is the approach we follow in the next section.

5.4 Hierarchies of semidefinite programs

Semidefinite programming is a convex optimisation technique in the cone of positive semidefinite matrices. It comes with a duality theory: if a primal semidefinite program is a maximisation problem, then one may deduce a dual minimisation problem which is again a semidefinite program. Like linear programs, these programs satisfy a weak duality condition: the optimal value of the primal problem is upper bounded by the optimal value of the dual program. The difference between the optimal values is called the duality gap. When there is no duality gap, we say that there is strong duality between the programs.

5.4.1 Preliminaries

In this section, we introduce preliminary technical lemmas and we refer the reader to Appendix C

³Technically $\delta \notin L^2(\mathbb{R}_+)$, but the result holds by considering a sequence of functions converging to a Dirac delta.

⁴We checked numerically the corresponding constraints $|L_k(2)| \leq 1$ for k up to 10^3 and, considering asymptotic behaviors, we conjecture that these hold for all $k \geq 0$.

for their proofs.

We recall the following standard result, which comes from the fact that any univariate polynomial non-negative over \mathbb{R} can be written as a sum of squares:

Lemma 3 ([68]). *Let $p \in \mathbb{N}$ and let P be a univariate polynomial of degree $2p$. Let $X = (1, x, \dots, x^p)$ be the vector of monomials. Then, P is non-negative over \mathbb{R} if and only if there exists a real $(p + 1) \times (p + 1)$ positive semidefinite matrix Q such that for all $x \in \mathbb{R}$,*

$$P(x) = X^T Q X. \quad (40)$$

From this lemma we deduce the following characterisation of non-negative polynomials over \mathbb{R}_+ :

Lemma 4. *Non-negative polynomials on \mathbb{R}_+ can be written as sums of polynomials of the form $\sum_{l=0}^p x^l \sum_{i+j=2l} y_i y_j$, where $p \in \mathbb{N}$ and $y_i \in \mathbb{R}$, for all $0 \leq i \leq p$.*

Note that the characterisation in Lemma 4 differs from that of Stieltjes [65], which expresses non-negative polynomials over \mathbb{R}_+ as $x \mapsto A_1(x) + xA_2(x)$, where A_1 and A_2 are sums of squares. This slightly slows down the numerical resolution, which does not matter given the size of the programs considered. At the same time, this allows us to obtain more compact expressions for the semidefinite programs.

We use the characterisation of Lemma 4 to obtain the following result: for $\mathbf{s} \in \mathbb{R}^{\mathbb{N}}$, the fact that the series $f_{\mathbf{s}}$ defined in Eq. (25) has non-negative scalar product with non-negative truncated Laguerre series up to degree m can be expressed as a positive semidefinite constraint involving the matrix $A_{\mathbf{s}}$ defined in Eq. (27). Formally:

Lemma 5. *Let $m \geq n$ and let $\mathbf{s} \in \mathbb{R}^{\mathbb{N}}$. The following propositions are equivalent:*

$$(i) \quad \forall g \in \mathcal{R}_{m,+}(\mathbb{R}_+), \langle f_{\mathbf{s}}, g \rangle \geq 0,$$

$$(ii) \quad A_{\mathbf{s}} \succeq 0.$$

Using these results, we derive hierarchies of semidefinite relaxations and restrictions for the infinite-dimensional linear program (LP_n) in the following sections.

5.4.2 Semidefinite relaxations

One way to obtain a relaxation of (LP_n) is to relax the constraint:

$$\forall x \in \mathbb{R}_+, f_{\mathbf{F}}(x) = \sum_k F_k \mathcal{L}_k(x) \geq 0. \quad (41)$$

Instead, one may impose the weaker constraint:

$$\forall g \in \mathcal{R}_{m,+}(\mathbb{R}_+), \langle f_{\mathbf{F}}, g \rangle \geq 0, \quad (42)$$

for some fixed $m \geq n$. By Lemma 5, this constraint may in turn be expressed as a positive semidefinite constraint on the $(m+1) \times (m+1)$ matrix $A_{\mathbf{F}}$ defined in Eq. (27).

Each choice of m thus leads to a different semidefinite program, whose optimal value gets closer to ω_n as m increases (since the constraint (42) gets stronger when m increases). Moreover, when replacing the constraint (41) by the constraint (42) in (LP_n) , for $l > m$ the variables F_l do not appear in the matrix $A_{\mathbf{F}}$, and are constrained only by $\sum_k F_k = 1$ and $F_l \geq 0$. Since $m \geq n$ and the optimal value is F_n , we may thus set $F_l = 0$ for $l > m$ without loss of generality. This gives a hierarchy of finite-dimensional semidefinite relaxations for (LP_n) , and the semidefinite relaxation of order m is given by:

Find $A \in \text{Sym}_{m+1}$ and $\mathbf{F} \in \mathbb{R}^{m+1}$

maximising F_n

subject to

$$\begin{cases} \sum_{k=0}^m F_k = 1 \\ \forall k \leq m, F_k \geq 0 \\ \forall l \leq m, \forall i+j=2l, A_{ij} = \sum_{k=0}^l F_k \binom{l}{k} l! \\ \forall l \in \llbracket 1, m \rrbracket, \forall i+j=2l-1, A_{ij} = 0 \\ A \succeq 0. \end{cases} \quad (\text{SDP}_n^{m,\geq})$$

Let us denote its optimal value by $\omega_n^{m,\geq}$. The sequence $\{\omega_n^{m,\geq}\}_{m \geq n}$ is a decreasing sequence and for all $m \geq n$, we have $\omega_n \leq \omega_n^{m,\geq}$.

For each $m \geq n$, the program $(\text{SDP}_n^{m,\geq})$ has a dual semidefinite program which is given by (see

Appendix D for a detailed derivation):

Find $Q \in \text{Sym}_{m+1}$, $\boldsymbol{\mu} \in \mathbb{R}^{m+1}$ and $y \in \mathbb{R}$

minimising y

subject to

$$\begin{cases} y \geq 1 + \mu_n \\ \forall k \leq m, y \geq \mu_k \\ \forall l \leq m, \sum_{i+j=2l} Q_{ij} = \sum_{k=l}^m \frac{(-1)^{k+l}}{l!} \binom{k}{l} \mu_k \\ Q \succeq 0. \end{cases} \quad (\text{D-SDP}_n^{m,\geq})$$

Note that the semidefinite program $(\text{SDP}_n^{m,\geq})$ presented in section 3 actually corresponds to this dual semidefinite program $(\text{D-SDP}_n^{m,\geq})$. We show in Theorem 3 that strong duality holds between the primal and the dual versions of this semidefinite program. In particular, numerical computations with either of these programs will yield the same optimal value.

5.4.3 Semidefinite restrictions

A trivial way to obtain a restriction of (LP_n) is to impose $F_l = 0$ for $l > m$, for some $m \geq n$. What is less trivial is that this yields a finite-dimensional semidefinite program. Indeed, the constraint (41) becomes

$$\forall x \in \mathbb{R}_+, \sum_{k=0}^m F_k \mathcal{L}_k(x) \geq 0, \quad (43)$$

or equivalently:

$$\forall x \in \mathbb{R}, \sum_{k=0}^m (-1)^k F_k L_k(x^2) \geq 0, \quad (44)$$

where we used Eq. (23). By Lemma 3, writing $X = (1, x, \dots, x^m)$, this is equivalent to the existence of a positive semidefinite matrix $Q = (Q_{ij})_{0 \leq i, j \leq m}$ such that for all $x \in \mathbb{R}$,

$$\begin{aligned} \sum_{k=0}^m (-1)^k F_k L_k(x^2) &= X^T Q X \\ &= \sum_{l=0}^m x^l \sum_{i+j=l} Q_{ij}. \end{aligned} \quad (45)$$

This is in turn equivalent to the linear constraints:

$$\forall l \in \llbracket 1, m \rrbracket, \sum_{i+j=2l-1} Q_{ij} = 0,$$

and

$$\forall l \leq m, \sum_{i+j=2l} Q_{ij} = \frac{(-1)^l}{l!} \sum_{k=l}^m (-1)^k \binom{k}{l} F_k, \quad (46)$$

by identifying the coefficients in front of each monomial in Eq. (45).

Hence, the restriction of (LP_n) obtained by imposing $F_l = 0$ for $l > m$, for a fixed $m \geq n$, is a semidefinite program given by:

$$\begin{aligned} & \text{Find } Q \in \text{Sym}_{m+1} \text{ and } \mathbf{F} \in \mathbb{R}^{m+1} \\ & \text{maximising } F_n \\ & \text{subject to} \\ & \left\{ \begin{array}{l} \sum_{k=0}^m F_k = 1 \\ \forall k \leq m, F_k \geq 0 \\ \forall l \in \llbracket 1, m \rrbracket, \sum_{i+j=2l-1} Q_{ij} = 0 \\ \forall l \leq m, \sum_{i+j=2l} Q_{ij} = \sum_{k=l}^m \frac{(-1)^{k+l}}{l!} \binom{k}{l} F_k \\ Q \succeq 0. \end{array} \right. \quad (\text{SDP}_n^{m,\leq}) \end{aligned}$$

Let us denote its optimal value by $\omega_n^{m,\leq}$. Each choice of m leads to a different semidefinite restriction of (LP_n) , whose optimal value gets closer to ω_n as m increases (since the constraint $F_l = 0$ for $l > m$ gets weaker when m increases). The sequence $\{\omega_n^{m,\leq}\}_{m \geq n}$ is thus an increasing sequence and for all $m \geq n$, we have $\omega_n^{m,\leq} \leq \omega_n$.

For each $m \geq n$, the program $(\text{SDP}_n^{m,\leq})$ has a dual semidefinite program which is given by (see Appendix D for a detailed derivation):

$$\begin{aligned} & \text{Find } A \in \text{Sym}_{m+1}, \boldsymbol{\mu} \in \mathbb{R}^{m+1} \text{ and } y \in \mathbb{R} \\ & \text{minimising } y \\ & \text{subject to} \\ & \left\{ \begin{array}{l} y \geq 1 + \mu_n \\ \forall k \leq m, y \geq \mu_k \\ \forall l \leq m, \forall i+j=2l, A_{ij} = \sum_{k=0}^l \mu_k \binom{l}{k} l! \\ A \succeq 0. \end{array} \right. \quad (\text{D-SDP}_n^{m,\leq}) \end{aligned}$$

We show in Theorem 2 that strong duality holds between the primal and the dual versions of this semidefinite program.

5.4.4 Strong duality of semidefinite programs

We fix $m \geq n$ and we show that strong duality holds both for the semidefinite restriction $(\text{SDP}_n^{m,\leq})$ and the semidefinite relaxation $(\text{SDP}_n^{m,\geq})$.

We first consider the semidefinite restrictions:

Theorem 2. *Strong duality holds between the programs $(\text{SDP}_n^{m,\leq})$ and $(\text{D-SDP}_n^{m,\leq})$.*

Proof. We make use of Slater condition for (finite-dimensional) semidefinite programs: strict feasibility of $(\text{SDP}_n^{m,\leq})$ implies strong duality between $(\text{SDP}_n^{m,\leq})$ and $(\text{D-SDP}_n^{m,\leq})$.

In order to obtain a strictly feasible solution, we define $Q := \frac{1}{2^{m+1}-1} \text{Diag}_{k=0,\dots,m}(\frac{1}{k!}) \in \text{Sym}_{m+1}$ and $\mathbf{F} = (F_0, \dots, F_m) \in \mathbb{R}^{m+1}$, where for all $k \in \llbracket 0, m \rrbracket$, $F_k := \frac{1}{2^{m+1}-1} \binom{m+1}{k+1}$. Then $Q \succ 0$ and $F_k > 0$ for all $k \in \llbracket 0, m \rrbracket$. Moreover, we have

$$\begin{aligned} \sum_{k=0}^m F_k &= \frac{1}{2^{m+1}-1} \sum_{k=0}^m \binom{m+1}{k+1} \\ &= \frac{1}{2^{m+1}-1} \left(\sum_{k=0}^{m+1} \binom{m+1}{k} - 1 \right) \\ &= 1. \end{aligned} \quad (47)$$

We also have $\sum_{i+j=2l-1} Q_{ij} = 0$ for all $l \in \llbracket 1, m \rrbracket$, since Q is diagonal. Furthermore, for all $l \leq m$,

$$\sum_{i+j=2l} Q_{ij} = Q_{ll} = \frac{1}{2^{m+1}-1} \frac{1}{l!}, \quad (48)$$

and

$$\begin{aligned} & \frac{(-1)^l}{l!} \sum_{k=l}^m (-1)^k \binom{k}{l} F_k \\ &= \frac{1}{2^{m+1}-1} \frac{1}{l!} \sum_{k=l}^m (-1)^{k-l} \binom{k}{m} \binom{m+1}{k+1} \\ &= \frac{1}{2^{m+1}-1} \frac{1}{l!} \binom{m}{l} \sum_{q=0}^{m-l} (-1)^q \frac{m+1}{q+l+1} \binom{m-l}{q} \\ &= \frac{1}{2^{m+1}-1} \frac{1}{l!}, \end{aligned} \quad (49)$$

where we used [69, (1.41)] in the last line. Therefore, (Q, \mathbf{F}) is a strictly feasible solution of $(\text{SDP}_n^{m,\leq})$, which implies strong duality. \square

As a consequence of the proof of Theorem 2, we also obtain strong duality for the semidefinite relaxations:

Theorem 3. *Strong duality holds between the programs $(\text{SDP}_n^{m,\geq})$ and $(\text{D-SDP}_n^{m,\geq})$.*

Proof. The program $(\text{SDP}_n^{m,\geq})$ is a relaxation of (LP_n) and $(\text{SDP}_n^{m,\leq})$ is a restriction of (LP_n) , so $(\text{SDP}_n^{m,\geq})$ is a relaxation of $(\text{SDP}_n^{m,\leq})$. Hence, the strictly feasible solution of $(\text{SDP}_n^{m,\leq})$ derived in the proof of Theorem 2 yields a strictly feasible solution (A, \mathbf{F}) for $(\text{SDP}_n^{m,\geq})$: we set $\mathbf{F} = (F_0, \dots, F_m) \in \mathbb{R}^{m+1}$, where for all $k \in \llbracket 0, m \rrbracket$,

$F_k := \frac{1}{2^{m+1}-1} \binom{m+1}{k+1}$ and $A = A_{\mathbf{F}} \in \text{Sym}_{m+1}$, where $A_{\mathbf{F}}$ is defined in Eq. (27).

With Slater condition, this shows again that strong duality holds between the programs $(\text{SDP}_n^{m,\geq})$ and $(\text{D-SDP}_n^{m,\geq})$. \square

5.5 Convergence of the hierarchies of semidefinite programs

From the previous sections, for $m \geq n$ the optimal values $\omega_n^{m,\geq}$ and $\omega_n^{m,\leq}$ of $(\text{SDP}_n^{m,\geq})$ and $(\text{SDP}_n^{m,\leq})$ form decreasing and increasing sequences, respectively, which satisfy

$$0 \leq \omega_n^{m,\leq} \leq \omega_n \leq \omega_n^{m,\geq} \leq 1, \quad (50)$$

where ω_n is the optimal value of (LP_n) . These sequences thus both converge, and the remaining question is whether they converge to ω_n . In this section, we show that this is indeed the case, for both sequences.

5.5.1 Convergence of the sequence of upper bounds

Theorem 4. *The decreasing sequence of optimal values of $(\text{SDP}_n^{m,\geq})$ converges to the optimal value of (LP_n) :*

$$\lim_{m \rightarrow +\infty} \omega_n^{m,\geq} = \omega_n. \quad (51)$$

In order to prove this theorem, we extract a limit from a sequence of optimal solutions of $(\text{SDP}_n^{m,\geq})$, for $m \geq n$, and we show using Theorem 1 that it provides a feasible solution of (LP_n) .

Proof. For all $m \geq n$, the feasible set of $(\text{SDP}_n^{m,\geq})$ is non-empty (consider, e.g., $\mathbf{F} = (1, 0, 0, \dots, 0) \in \mathbb{R}^{m+1}$). Moreover, due to the constraints $\sum_{k=0}^m F_k = 1$ and $F_k \geq 0$ for all $k \leq m$, the feasible set of $(\text{SDP}_n^{m,\geq})$ is compact. Hence, the program $(\text{SDP}_n^{m,\geq})$ has feasible optimal solutions, for all $m \geq n$, by diagonal extraction.

The matrix A in $(\text{SDP}_n^{m,\geq})$ is entirely fixed by the choice of \mathbf{F} . Let $(\mathbf{F}^m)_{m \geq n}$ be a sequence of optimal solutions of $(\text{SDP}_n^{m,\geq})$, for $m \geq n$. For each $m \geq n$, we have by optimality that $F_n^m = \omega_n^{m,\geq}$, and the sequence $(F_n^m)_{m \geq n}$ converges. We complete each tuple $\mathbf{F}^m = (F_0^m, F_1^m, \dots, F_m^m) \in \mathbb{R}^{m+1}$ with zeros to obtain a sequence in $\mathbb{R}^{\mathbb{N}}$, which we still denote $\mathbf{F}^m = (F_0^m, F_1^m, \dots, F_m^m, 0, 0, \dots) \in \mathbb{R}^{\mathbb{N}}$.

Performing a diagonal extraction ϕ on the sequence of optimal solutions $(\mathbf{F}^m)_{m \geq n}$, we obtain a sequence of sequences $(\mathbf{F}^{\phi(m)})_{m \geq n}$ such that each sequence $(F_k^{\phi(m)})_{m \geq n}$ converges when $m \rightarrow +\infty$, for all $k \in \mathbb{N}$. Let F_k denote its limit, for each $k \in \mathbb{N}$. We write $\mathbf{F} = (F_k)_{k \in \mathbb{N}} \in \mathbb{R}^{\mathbb{N}}$ the sequence of limits.

For all $m \geq n$, $F_k^{\phi(m)} \geq 0$ for all $k \in \mathbb{N}$ and $\sum_k F_k^{\phi(m)} = 1$, so taking $m \rightarrow +\infty$ we obtain $F_k \geq 0$ for all $k \in \mathbb{N}$, and $\sum_k F_k \leq 1$. Moreover,

$$F_n = \lim_{m \rightarrow +\infty} F_n^m = \lim_{m \rightarrow +\infty} \omega_n^{m,\geq}. \quad (52)$$

For all $m \geq n$, we have $\omega_n^{m,\geq} \geq \omega_n$, so $F_n \geq \omega_n > 0$. In particular, $\sum_k F_k > 0$, so without loss of generality we may assume that $\sum_k F_k = 1$ (otherwise we can always replace F_k by $\frac{F_k}{\sum_l F_l}$).

Let $f_{\mathbf{F}} = \sum_k F_k \mathcal{L}_k \in L^2(\mathbb{R}_+)$. By construction we have:

$$\forall m \geq n, \forall g \in \mathcal{R}_{m,+}(\mathbb{R}_+), \langle f_{\mathbf{F}}, g \rangle \geq 0. \quad (53)$$

Hence, by Theorem 1, \mathbf{F} is the sequence of Laguerre moments of a non-negative distribution over \mathbb{R}_+ (the Lebesgue measure times the function $f_{\mathbf{F}}$). In particular,

$$\forall x \in \mathbb{R}_+, f_{\mathbf{F}}(x) = \sum_k F_k \mathcal{L}_k(x) \geq 0. \quad (54)$$

With the constraints $F_k \geq 0$ for all $k \in \mathbb{N}$, and $\sum_k F_k \leq 1$, this implies that \mathbf{F} is a feasible solution of (LP_n) , and in particular $F_n \leq \omega_n$, since (LP_n) is a maximisation problem. Since we already had $F_n \geq \omega_n$ we obtain with Eq. (52):

$$\lim_{m \rightarrow +\infty} \omega_n^{m,\geq} = \omega_n, \quad (55)$$

which concludes the proof. \square

5.5.2 Convergence of the sequence of lower bounds

Theorem 5. *The increasing sequence of optimal values of $(\text{SDP}_n^{m,\leq})$ converges to the optimal value of (LP_n) :*

$$\lim_{m \rightarrow +\infty} \omega_n^{m,\leq} = \omega_n. \quad (56)$$

The proof is similar to that of Theorem 4 using the dual programs: we attempt to construct a feasible optimal solution of (D-LP_n) by extracting a limit from a sequence of optimal solutions

of (D-SDP $^{m,\leq}_n$), for $m \geq n$, in order to conclude using the strong duality between (SDP $^{m,\leq}_n$) and (D-SDP $^{m,\leq}_n$), which we proved in Theorem 2. However, it turns out that (D-LP $_n$) may not have feasible optimal solutions in $L^2(\mathbb{R}_+)$ (as anticipated with the analytical optimal solutions for $n = 1$ and $n = 2$ from section 5.3).

To deal with this issue, we extend the formulation of (D-LP $_n$) to a larger space where it has feasible optimal solutions, namely the space of tempered distributions $\mathcal{S}'(\mathbb{R}_+)$. We thus introduce the linear program over Schwartz space:

$$\begin{aligned} & \text{Find } (F_k)_{k \in \mathbb{N}} \in \mathcal{S}(\mathbb{N}) \\ & \text{maximising } F_n \\ & \text{subject to} \\ & \left\{ \begin{array}{l} \sum_k F_k = 1 \\ \forall k \in \mathbb{N}, F_k \geq 0 \\ \forall x \in \mathbb{R}_+, \sum_k F_k \mathcal{L}_k(x) \geq 0, \end{array} \right. \quad (\text{LP}_n^{\mathcal{S}}) \end{aligned}$$

together with its dual linear program:

$$\begin{aligned} & \text{Find } y \in \mathbb{R} \text{ and } \mu \in \mathcal{S}'(\mathbb{R}_+) \\ & \text{minimising } y \\ & \text{subject to} \\ & \left\{ \begin{array}{l} \forall k \neq n \in \mathbb{N}, y \geq \int_{\mathbb{R}_+} \mathcal{L}_k d\mu \\ y \geq 1 + \int_{\mathbb{R}_+} \mathcal{L}_n d\mu \\ \forall f \in \mathcal{S}_+(\mathbb{R}_+), \langle \mu, f \rangle \geq 0. \end{array} \right. \quad (\text{D-LP}_n^{\mathcal{S}}) \end{aligned}$$

Note that the semidefinite restrictions (SDP $^{m,\leq}_n$) of (LP $_n$) are also restrictions of (LP $_n^{\mathcal{S}}$) for all $m \geq n$, while (LP $_n^{\mathcal{S}}$) is itself a restriction of (LP $_n$), and (D-LP $_n^{\mathcal{S}}$) is a relaxation of (D-LP $_n$). We denote by $\omega_n^{\mathcal{S}}$ the optimal value of (LP $_n^{\mathcal{S}}$) and by $\omega_n^{\prime\mathcal{S}}$ the optimal value of (D-LP $_n^{\mathcal{S}}$). Recall that the optimal value of (LP $_n$) is denoted ω_n . We denote by ω'_n the optimal value of its dual program (D-LP $_n$). By weak duality of linear programming we thus have

$$\omega_n^{m,\leq} \leq \omega_n^{\mathcal{S}} \leq \omega_n \leq \omega'_n \leq \omega_n^{\prime\mathcal{S}}, \quad (57)$$

for all $m \geq n$.

Before proving Theorem 5, we introduce two intermediate technical results. The first one is a reformulation of (D-SDP $^{m,\leq}_n$) over Schwartz space:

Lemma 6. *For all $m \geq n$, the program (D-SDP $^{m,\leq}_n$) is equivalent to the following program:*

$$\begin{aligned} & \text{Find } y \in \mathbb{R} \text{ and } \mu \in \mathcal{S}'(\mathbb{N}) \\ & \text{minimising } y \\ & \text{subject to} \\ & \left\{ \begin{array}{l} \forall k \neq n \in \mathbb{N}, y \geq \mu_k \\ y \geq 1 + \mu_n \\ \forall g \in \mathcal{R}_{m,+}(\mathbb{R}_+), \langle f_\mu, g \rangle \geq 0, \end{array} \right. \quad (\text{D-SDP}_n^{m,\leq}) \end{aligned}$$

where $f_\mu = \sum_k \mu_k \mathcal{L}_k$.

This reformulation uses Stieltjes characterisation of non-negative polynomials over \mathbb{R}_+ [65] and is detailed in Appendix E.

The second result provides a nontrivial analytical solution to the primal program (SDP $^{m,\leq}_n$): let us define $\mathbf{F}^n = (F_k^n)_{k \in \mathbb{N}} \in \mathbb{R}^{\mathbb{N}}$ by

- if n is even:

$$F_k^n := \begin{cases} \frac{1}{2^n} \binom{k}{\frac{k}{2}} \binom{n-k}{\frac{n-k}{2}} & \text{when } k \leq n, k \text{ even,} \\ 0 & \text{otherwise,} \end{cases} \quad (58)$$

- if n is odd:

$$F_k^n := \begin{cases} \frac{1}{2^n} \frac{\binom{\lfloor \frac{n}{2} \rfloor}{\lfloor \frac{k}{2} \rfloor}^2}{\binom{n}{k}}, & \text{when } k \leq n, \\ 0 & \text{otherwise.} \end{cases} \quad (59)$$

In both cases,

$$\begin{aligned} F_n^n &= \frac{1}{2^n} \binom{n}{\lfloor \frac{n}{2} \rfloor} \\ &\geq \frac{1}{n+1} \end{aligned} \quad (60)$$

where we used $\binom{n}{j} \leq \binom{n}{\lfloor \frac{n}{2} \rfloor}$ for all $j \in \llbracket 0, n \rrbracket$, summed over j from 0 to n .

Lemma 7. *For all $m \geq n$, \mathbf{F}^n is a feasible solution of (SDP $^{m,\leq}_n$). Moreover, it is optimal when $m = n$.*

The proof of feasibility consists in checking that the constraints of (SDP $^{m,\leq}_n$) are satisfied by \mathbf{F}^n . To do so, we make use of Zeilberger's algorithm [70], a powerful algorithm for proving binomial identities. The proof of optimality for $m = n$ is obtained by deriving an analytical feasible solution of (D-SDP $^{n,\leq}_n$) with the same optimal value. Note that the optimality of this solution does not

play a role in the proof of convergence. We refer to Appendix F for a detailed proof.

As a consequence, we obtain the following analytical lower bound for the optimal value of (LP_n) :

$$\omega_n \geq \omega_n^{m,\leq} = \frac{1}{2^n} \binom{n}{\lfloor \frac{n}{2} \rfloor} \underset{n \rightarrow +\infty}{\sim} \sqrt{\frac{2}{\pi n}}, \quad (61)$$

which is superseded by numerical bounds when $n \geq 3$ (see Table 1).

We now combine Lemma 6 and Lemma 7 to prove Theorem 5.

Proof. The feasible set of $(\text{D-SDP}_n^{m,\leq})$ is non-empty, by considering the null sequence, which achieves value 1. Without loss of generality, we add the constraint $y \leq 1$ in (D-LP_n^S) and $(\text{D-SDP}_n^{m,\leq})$.

Let $m \geq n$ and let $(y, \boldsymbol{\mu}) \in \mathbb{R} \times \mathcal{S}'(\mathbb{N})$ be a feasible solution of $(\text{D-SDP}_n^{m,\leq})$ expressed in the form given by Lemma 6. The constraint $\langle f_{\boldsymbol{\mu}}, x \mapsto e^{-\frac{x}{2}} \rangle \geq 0$ implies $\mu_0 \geq 0$ and thus $y \geq 0$. Without loss of generality, we may set $\mu_k = 0$ for $k > m$, since these coefficients are only constrained by $\mu_k \leq y \leq 1$. We also have $\mu_k \leq 1$ for all $k \in \mathbb{N}$.

By Lemma 7, \mathbf{F}^l is a feasible solution of $(\text{SDP}_l^{m,\leq})$ for all $l \leq m$, so in particular $f_{\mathbf{F}^l} = \sum_{k=0}^l F_k^l \mathcal{L}_k \in \mathcal{R}_{m,+}(\mathbb{R}_+)$. Hence, $\boldsymbol{\mu}$ satisfies the constraint $\langle f_{\boldsymbol{\mu}}, f_{\mathbf{F}^l} \rangle \geq 0$, which gives

$$\sum_{k=0}^l \mu_k F_k^l \geq 0, \quad (62)$$

for all $l \leq m$. Thus we have, for all $l \in \llbracket 1, m \rrbracket$,

$$\begin{aligned} \mu_l &\geq -\frac{1}{F_l^l} \sum_{k=0}^{l-1} \mu_k F_k^l \\ &\geq -\frac{1}{F_l^l} \sum_{k=0}^{l-1} F_k^l \\ &= 1 - \frac{1}{F_l^l} \\ &\geq -l, \end{aligned} \quad (63)$$

where we used $F_l^l > 0$ in the first line, $\mu_k \leq 1$ and $F_k^l \geq 0$ in the second line, $\sum_{k=0}^l F_k^l = 1$ in the third line and Eq. (60) in the last line. With $\mu_k \leq 1$, we obtain $|\mu_k| \leq k$ for $k \in \mathbb{N}^*$, and thus $|\mu_k| \leq k + 1$ for all $k \in \mathbb{N}$. Hence, the feasible set of $(\text{D-SDP}_n^{m,\leq})$ is compact and the program

$(\text{D-SDP}_n^{m,\leq})$ has feasible optimal solutions for all $m \geq n$, by diagonal extraction.

Let $(y^m, \boldsymbol{\mu}^m)_{m \geq n}$ be a sequence of optimal solutions of $(\text{D-SDP}_n^{m,\leq})$, for $m \geq n$. By Theorem 2, we have strong duality between the programs $(\text{SDP}_n^{m,\leq})$ and $(\text{D-SDP}_n^{m,\leq})$, so the optimal value of $(\text{D-SDP}_n^{m,\leq})$ is given by $\omega_n^{m,\leq}$, for all $m \geq n$. By optimality $y^m = \omega_n^{m,\leq}$, for all $m \geq n$, and the sequence $(y^m)_{m \geq n}$ converges.

Performing a diagonal extraction ϕ on the sequence $(\boldsymbol{\mu}^m)_{m \geq n}$, we obtain a sequence of sequences $(\boldsymbol{\mu}^{\phi(m)})_{m \geq n}$ such that each sequence $(\mu_k^{\phi(m)})_{m \geq n}$ converges when $m \rightarrow +\infty$, for all $k \in \mathbb{N}$. Let μ_k denote its limit, for each $k \in \mathbb{N}$. We write $\boldsymbol{\mu} = (\mu_k)_{k \in \mathbb{N}} \in \mathbb{R}^{\mathbb{N}}$ the sequence of limits. We also write

$$y := \lim_{m \rightarrow +\infty} y^m = \lim_{m \rightarrow +\infty} \omega_n^{m,\leq}. \quad (64)$$

For all $m \geq n$, we have $\omega_n^{\phi(m),\leq} \geq \mu_k^{\phi(m)}$ for all $k \in \mathbb{N}$ and $\omega_n^{\phi(m),\leq} \geq 1 + \mu_n^{\phi(m)}$, so taking $m \rightarrow +\infty$ we obtain $y \geq \mu_k$ for all $k \in \mathbb{N}$ and $y \geq 1 + \mu_n$. By Eq. (57), we have $\omega_n^{m,\leq} \leq \omega_n^S$ for all $m \geq n$, so $y \leq \omega_n^S$.

Moreover, $|\mu_k^{\phi(m)}| \leq k + 1$ for all $k \in \mathbb{N}$, so taking $m \rightarrow +\infty$ we obtain $|\mu_k| \leq k + 1$ for all $k \in \mathbb{N}$, which implies that $\boldsymbol{\mu} \in \mathcal{S}'(\mathbb{N})$ [64]. Let $f_{\boldsymbol{\mu}} = \sum_k \mu_k \mathcal{L}_k \in \mathcal{S}'(\mathbb{R}_+)$. We have

$$\mu_k = \langle f_{\boldsymbol{\mu}}, \mathcal{L}_k \rangle. \quad (65)$$

By construction we also have:

$$\forall m \geq n, \forall g \in \mathcal{R}_{m,+}(\mathbb{R}_+), \langle f_{\boldsymbol{\mu}}, g \rangle \geq 0. \quad (66)$$

By Theorem 1, this implies that the distribution $\mu := f_{\boldsymbol{\mu}}(x) \in \mathcal{S}'(\mathbb{R}_+)$ is non-negative, i.e.,

$$\forall f \in \mathcal{S}_+(\mathbb{R}_+), \langle f_{\boldsymbol{\mu}}, f \rangle \geq 0. \quad (67)$$

With the constraints $y \geq \mu_k$ for all $k \in \mathbb{N}$ and $y \geq 1 + \mu_n$, this implies that $(y, \boldsymbol{\mu})$ is a feasible solution of (D-LP_n^S) , and in particular $y \geq \omega_n'^S$, since (D-LP_n^S) is a minimisation problem. Since $y \leq \omega_n^S$ we obtain with Eq. (57) and Eq. (64):

$$\lim_{m \rightarrow +\infty} \omega_n^{m,\leq} = \omega_n^S = \omega_n = \omega_n' = \omega_n'^S. \quad (68)$$

□

As a direct corollary of the proof of Theorem 5, we obtain the following strong duality result:

Theorem 6. *Strong duality holds between the programs (LP_n^S) and $(D-LP_n^S)$ and between the programs (LP_n) and $(D-LP_n)$.*

For completeness, we give a different and more direct proof of the strong duality between (LP_n) and $(D-LP_n)$ in Appendix G.

We have shown the convergence of the semidefinite hierarchies of upper and lower bounds $(SDP_n^{m,\geq})_{m\geq n}$ and $(SDP_n^{m,\leq})_{m\geq n}$ towards the optimal value of (LP_n) . By linearity, these results generalise straightforwardly to the case of witnesses corresponding to linear combinations of fidelities with displaced Fock states:

$$\lim_{m\rightarrow+\infty} \omega_{\mathbf{a}}^{m,\leq} = \omega_{\mathbf{a}} = \lim_{m\rightarrow+\infty} \omega_{\mathbf{a}}^{m,\geq}, \quad (69)$$

for all $n \in \mathbb{N}^*$ and all $\mathbf{a} \in [0, 1]^n$.

6 Witnessing multimode Wigner negativity

In this section we discuss the generalisation of our Wigner negativity witnesses to the more challenging multimode setting. Hereafter, M denotes the number of modes.

6.1 Multimode Wigner negativity witnesses

Using multi-index notations (see Appendix H.1), the single-mode Wigner negativity witnesses defined in Eq. (6) are naturally generalised to

$$\hat{\Omega}_{\mathbf{a},\alpha} := \sum_{1\leq\mathbf{k}\leq\mathbf{n}} a_{\mathbf{k}} \hat{D}(\alpha) |\mathbf{k}\rangle\langle\mathbf{k}| \hat{D}^\dagger(\alpha), \quad (70)$$

for $\mathbf{n} = (n_1, \dots, n_M) \in \mathbb{N}^M \setminus \{\mathbf{0}\}$, $\mathbf{a} = (a_{\mathbf{k}})_{1\leq\mathbf{k}\leq\mathbf{n}} \in [0, 1]^{n_1 \cdots n_M}$, with $\max_{\mathbf{k}} a_{\mathbf{k}} = 1$, and $\alpha \in \mathbb{C}^M$. Similar to the single-mode case, these POVM elements are weighted sums of multimode displaced Fock states projectors, and their expectation value for a quantum state $\rho \in \mathcal{D}(\mathcal{H}^{\otimes M})$ is given by

$$\text{Tr}(\hat{\Omega}_{\mathbf{a},\alpha}\rho) = \sum_{1\leq\mathbf{k}\leq\mathbf{n}} a_{\mathbf{k}} F(\hat{D}^\dagger(\alpha)\rho\hat{D}(\alpha), |\mathbf{k}\rangle) \quad (71)$$

where F is the fidelity. Unlike in the single-mode case however, estimating this quantity with homodyne or heterodyne detection by direct fidelity estimation is no longer efficient when the number of modes becomes large. Instead, one may use robust lower bounds on the multimode fidelity from [44] which can be obtained efficiently

with homodyne or heterodyne detection. A lower bound on the estimated experimental multimode fidelity will allow to detect Wigner negativity if it is larger than an upper bound on the threshold value associated to a given witness.

These lower bounds are obtained as follows: given a target multimode Fock state $|\mathbf{n}\rangle = |n_1\rangle \otimes \cdots \otimes |n_M\rangle$ and multiple copies of an M -mode experimental state ρ , measure all single-mode subsystems of ρ and perform fidelity estimation with each corresponding single-mode target Fock state. That is, the samples obtained from the detection of the i^{th} mode of ρ are used for single-mode fidelity estimation with the Fock state $|n_i\rangle$. Let F_1, \dots, F_M be the single-mode fidelity estimates obtained and let $\tilde{F}(\rho, |\mathbf{n}\rangle) := 1 - \sum_{i=1}^M (1 - F_i)$. Then [44],

$$1 - M(1 - F(\rho, |\mathbf{n}\rangle)) \leq \tilde{F}(\rho, |\mathbf{n}\rangle) \leq F(\rho, |\mathbf{n}\rangle). \quad (72)$$

In particular, \tilde{F} provides a good estimate of the multimode fidelity F whenever F is not too small: for instance, when the fidelity is equal to 0.5 the equation above yields $0.3 < \tilde{F} \leq 0.5$, and the bound gets tighter as the fidelity increases.

The same procedure is followed in the case of target displaced Fock states, with classical translations of the samples in order to account for the displacement parameters.

To each witness $\hat{\Omega}_{\mathbf{a},\alpha}$ is associated its threshold value:

$$\omega_{\mathbf{a}} := \sup_{\substack{\rho \in \mathcal{D}(\mathcal{H}^{\otimes M}) \\ W_\rho \geq 0}} \text{Tr}(\hat{\Omega}_{\mathbf{a},\alpha}\rho). \quad (73)$$

With Eq. (72), if the value of the bound \tilde{F} obtained experimentally is greater than $\omega_{\mathbf{a}}$, then the state ρ has a negative Wigner function.

With the same arguments as in the single-mode case, the multimode Wigner negativity witnesses in Eq. (70) form a complete family and retain the interpretation from Lemma 1: the violation of the threshold value provides a lower bound on the distance to the set of multimode states with non-negative Wigner function. However, the limited robustness of the bound \tilde{F} may affect the performance of the witnesses in practical scenarios, in particular for witnesses that are sums of different projectors. Still, we show in section 6.3 the applicability of the method with a genuinely multimode example.

We first generalise the single-mode semidefinite programming approach for approximating the threshold values to the multimode case.

6.2 Approximating the multimode threshold values

By linearity, we restrict our analysis to the case of Wigner negativity witnesses that are projectors onto a single multimode Fock state $|\mathbf{n}\rangle$, for $\mathbf{n} \in \mathbb{N}^M \setminus \{\mathbf{0}\}$. We thus consider the computation of

$$\omega_{\mathbf{n}} = \sup_{\substack{\rho \in \mathcal{D}(\mathcal{H}^{\otimes M}) \\ W_{\rho} \geq 0}} \langle \mathbf{n} | \rho | \mathbf{n} \rangle. \quad (74)$$

A similar reasoning to the single-mode case shows that the computation of the corresponding threshold value in Eq. (73) may be rephrased as the following infinite-dimensional linear program:

$$\begin{aligned} & \text{Find } (F_{\mathbf{k}})_{\mathbf{k} \in \mathbb{N}^M} \in \ell^2 \\ & \text{maximising } F_{\mathbf{n}} \\ & \text{subject to} \\ & \left\{ \begin{array}{l} \sum_{\mathbf{k}} F_{\mathbf{k}} = 1 \\ \forall \mathbf{k} \in \mathbb{N}^M, F_{\mathbf{k}} \geq 0 \\ \forall \mathbf{x} \in \mathbb{R}_+^M, \sum_{\mathbf{k}} F_{\mathbf{k}} \mathcal{L}_{\mathbf{k}}(\mathbf{x}) \geq 0, \end{array} \right. \quad (\text{LP}_{\mathbf{n}}) \end{aligned}$$

where the optimisation is over square-summable real sequences indexed by elements of \mathbb{N}^M . Its dual linear program reads

$$\begin{aligned} & \text{Find } y \in \mathbb{R} \text{ and } \mu \in L^2(\mathbb{R}_+^M) \\ & \text{minimising } y \\ & \text{subject to} \\ & \left\{ \begin{array}{l} \forall \mathbf{k} \neq \mathbf{n} \in \mathbb{N}^M, y \geq \int_{\mathbb{R}_+^M} \mathcal{L}_{\mathbf{k}} d\mu \\ y \geq 1 + \int_{\mathbb{R}_+^M} \mathcal{L}_{\mathbf{n}} d\mu \\ \forall f \in L^2(\mathbb{R}_+^M), \langle \mu, f \rangle \geq 0. \end{array} \right. \quad (\text{D-LP}_{\mathbf{n}}) \end{aligned}$$

In the single-mode case, we obtained hierarchies of SDP relaxations and restrictions for $(\text{LP}_{\mathbf{n}})$ by replacing constraints involving non-negative functions by constraints involving non-negative polynomials P of fixed degree. We then exploited the existence of a sum-of-squares decomposition for non-negative monovariate polynomials. In the multimode setting, the polynomials involved are multivariate, so that the set of non-negative polynomials over \mathbb{R} of a given degree may be strictly

larger than the set of sum-of-square polynomials [68]. Instead, we replace directly constraints involving non-negative functions over \mathbb{R}_+ by constraints involving non-negative polynomials P of fixed degree such that $\mathbf{x} \mapsto P(\mathbf{x}^2)$ ⁵ has a sum-of-squares decomposition, implying that the multimode semidefinite relaxations and restrictions are possibly looser than their single-mode counterparts.

Moreover, the dimension of the semidefinite programs increases exponentially with the level of the hierarchy m , as the number of M -variate monomials of degree less or equal to m is given by $\binom{M+m}{m}$. This implies that the semidefinite programs remain tractable only for a constant number of levels.

In spite of these observations, and following similar steps to the single-mode case (see Appendix H.2), the SDP relaxations providing upper bounds for the threshold value are given by

$$\begin{aligned} & \text{Find } A = (A_{\mathbf{ij}})_{|\mathbf{i}|, |\mathbf{j}| \leq m} \in \text{Sym}_{\binom{M+m}{m}} \\ & \text{and } \mathbf{F} = (F_{\mathbf{k}})_{|\mathbf{k}| \leq m} \in \mathbb{R}^{\binom{M+m}{m}} \\ & \text{maximising } F_{\mathbf{n}} \\ & \text{subject to} \\ & \left\{ \begin{array}{l} \sum_{|\mathbf{k}| \leq m} F_{\mathbf{k}} = 1 \\ \forall |\mathbf{k}| \leq m, F_{\mathbf{k}} \geq 0 \\ \forall |\mathbf{l}| \leq m, \forall \mathbf{i} + \mathbf{j} = 2\mathbf{l}, A_{\mathbf{ij}} = \sum_{\mathbf{k} \leq \mathbf{l}} F_{\mathbf{k}} \binom{\mathbf{l}}{\mathbf{k}}! \\ \forall |\mathbf{r}| \leq 2m, \mathbf{r} \neq 2\mathbf{l}, \forall |\mathbf{l}| \leq m, \forall \mathbf{i} + \mathbf{j} = \mathbf{r}, A_{\mathbf{ij}} = 0 \\ A \succeq 0, \end{array} \right. \quad (\text{SDP}_{\mathbf{n}}^{m, \geq}) \end{aligned}$$

for all $m \geq |\mathbf{n}|$. Similarly, the semidefinite restrictions providing lower bounds for the threshold value are given by

$$\begin{aligned} & \text{Find } Q = (Q_{\mathbf{ij}})_{|\mathbf{i}|, |\mathbf{j}| \leq m} \in \text{Sym}_{\binom{M+m}{m}} \\ & \text{and } \mathbf{F} = (F_{\mathbf{k}})_{|\mathbf{k}| \leq m} \in \mathbb{R}^{\binom{M+m}{m}} \\ & \text{maximising } F_{\mathbf{n}} \\ & \text{subject to} \\ & \left\{ \begin{array}{l} \sum_{|\mathbf{k}| \leq m} F_{\mathbf{k}} = 1 \\ \forall |\mathbf{k}| \leq m, F_{\mathbf{k}} \geq 0 \\ \forall |\mathbf{l}| \leq m, \sum_{\mathbf{i} + \mathbf{j} = 2\mathbf{l}} Q_{\mathbf{ij}} = \sum_{\mathbf{k} \geq \mathbf{l}} \frac{(-1)^{|\mathbf{k}| + |\mathbf{l}|}}{\mathbf{l}!} \binom{\mathbf{k}}{\mathbf{l}} F_{\mathbf{k}} \\ \forall |\mathbf{r}| \leq 2m, \mathbf{r} \neq 2\mathbf{l}, \forall |\mathbf{l}| \leq m, \sum_{\mathbf{i} + \mathbf{j} = \mathbf{r}} Q_{\mathbf{ij}} = 0 \\ Q \succeq 0, \end{array} \right. \quad (\text{SDP}_{\mathbf{n}}^{m, \leq}) \end{aligned}$$

⁵We write \mathbf{x}^2 in short for $\mathbf{x}^{2\mathbf{1}} = (x_1^2, \dots, x_M^2)$.

for all $m \geq |\mathbf{n}|$. In these semidefinite programs, the optimisations are over matrices and vectors indexed by elements of \mathbb{N}^m with sum of coefficients lower than m .

While our proof of convergence of the single-mode hierarchy of upper bounds transfers easily to the multimode setting, the proof of convergence of the hierarchy of lower bounds requires the analytical expression of feasible solutions for each level of the hierarchy. We show how to construct such solutions in the multimode case using products of single-mode feasible solutions—this requires introducing an equivalent hierarchy of restrictions, where constraints are expressed on polynomials of M variables with the degree in each individual variable being less or equal to m , rather than on polynomials of degree m (that is, constraints of the form $\mathbf{k} \leq \mathbf{m}$ where $\mathbf{m} = (m, \dots, m) \in \mathbb{N}^M$ rather than $|\mathbf{k}| \leq m$). Along the way, we also prove strong duality of the programs involved. We refer to Appendix H.3 for the proofs.

Summarising our results, we find that the semidefinite programs ($\text{SDP}_{\mathbf{n}}^{m, \geq}$) and ($\text{SDP}_{\mathbf{n}}^{m, \leq}$) respectively provide converging sequences of upper and lower bounds to the sought threshold value $\omega_{\mathbf{n}}$ given by ($\text{LP}_{\mathbf{n}}$). We study a concrete application in the next section.

6.3 Multimode example

To illustrate the usefulness of our Wigner negativity witnesses in the multimode setting, we consider a lossy Fock state over two-modes:

$$\begin{aligned} \rho_{1,1,\eta} := & (1 - \eta)^2 |1\rangle\langle 1| \otimes |1\rangle\langle 1| \\ & + \eta(1 - \eta) |1\rangle\langle 1| \otimes |0\rangle\langle 0| \\ & + \eta(1 - \eta) |0\rangle\langle 0| \otimes |1\rangle\langle 1| \\ & + \eta^2 |0\rangle\langle 0| \otimes |0\rangle\langle 0|, \end{aligned} \quad (75)$$

with loss parameter $0 \leq \eta \leq 1$. Setting $\eta = 0$ gives $\rho_{1,1,\eta} = |1\rangle\langle 1| \otimes |1\rangle\langle 1|$ while setting $\eta = 1$ gives $\rho_{1,1,\eta} = |0\rangle\langle 0| \otimes |0\rangle\langle 0|$. This state has a non-negative Wigner function for $\eta \geq \frac{1}{2}$.

We also consider the multimode Wigner negativity witness $|1\rangle\langle 1| \otimes |1\rangle\langle 1|$, which is a projector onto the Fock state $|1\rangle \otimes |1\rangle$. Solving numerically the corresponding hierarchy ($\text{SDP}_{\mathbf{n}}^{m, \leq}$) up to $m = 3$, we obtain the lower bound 0.266 and solving the hierarchy ($\text{SDP}_{\mathbf{n}}^{m, \geq}$) up to $m = 10$, we obtain the upper bound 0.320.

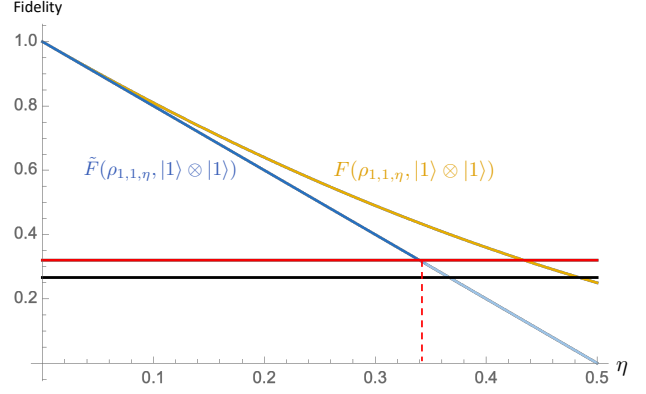


Figure 5: Witnessing Wigner negativity of the lossy Fock state $\rho_{1,1,\eta}$ over two modes using the witness $|1\rangle\langle 1| \otimes |1\rangle\langle 1|$. The threshold value for that witness is upper bounded by 0.320 and lower bounded by 0.266. The dashed red line delimits the interval of loss parameter values where the witness can be used to detect Wigner negativity of $\rho_{1,1,\eta}$ efficiently, i.e., when the robust bound $\tilde{F}(\rho_{1,1,\eta}, |1\rangle \otimes |1\rangle)$ (blue curve) on the fidelity from Eq. (77) is above the witness upper bound (red line). When it is below the witness lower bound (black line), we are guaranteed that the witness cannot be used to detect Wigner negativity of the state. The fidelity $F(\rho_{1,1,\eta}, |1\rangle \otimes |1\rangle)$ is also depicted above (yellow curve). Note that $\rho_{1,1,\eta}$ has a non-negative Wigner function for $\eta \geq 0.5$.

A direct consequence of the numerical lower bound is that tensor product states are not the closest among Wigner positive states to tensor product states with a negative Wigner function. Indeed, the maximum achievable fidelity with the state $|1\rangle \otimes |1\rangle$ using Wigner positive tensor product states is equal to the square of the maximum achievable fidelity with the state $|1\rangle$ using single-mode Wigner positive states, that is $0.5^2 = 0.25 < 0.266$.

We now use the upper bound to witness the Wigner negativity of the state $\rho_{1,1,\eta}$ (see Fig. 5). The fidelity between $\rho_{1,1,\eta}$ and $|1\rangle \otimes |1\rangle$ is given by $F(\rho_{1,1,\eta}, |1\rangle \otimes |1\rangle) = (1 - \eta)^2$, for all $0 \leq \eta \leq 1$. This fidelity is above the upper bound 0.320 on the threshold value of the witness $|1\rangle\langle 1| \otimes |1\rangle\langle 1|$ when $\eta \leq 0.434$.

However, in practice one would not obtain a precise estimate of the fidelity efficiently, but rather a robust lower bound on the fidelity computed from single-mode fidelities, which satisfies

Eq. (72). In the worst case, the estimate obtained is closer to $1 - 2(1 - F(\rho_{1,1,\eta}, |1\rangle \otimes |1\rangle))$ than to $F(\rho_{1,1,\eta}, |1\rangle \otimes |1\rangle)$. When the value of this robust lower bound is greater than the threshold value of the witness, this implies that the state has a negative Wigner function.

In the present case, the two single-mode reduced states of $\rho_{1,1,\eta}$ are the same, given by

$$\text{Tr}_2(\rho_{1,1,\eta}) = (1 - \eta) |1\rangle\langle 1| + \eta |0\rangle\langle 0|, \quad (76)$$

so the single-mode fidelities with $|1\rangle$ are equal for each mode and given by $1 - \eta$. Hence, the robust lower bound on $F(\rho_{1,1,\eta}, |1\rangle \otimes |1\rangle)$ is given by

$$\tilde{F}(\rho_{1,1,\eta}, |1\rangle \otimes |1\rangle) = 1 - 2\eta. \quad (77)$$

It is above the upper bound 0.320 on the threshold value of the witness $|1\rangle\langle 1| \otimes |1\rangle\langle 1|$ when $\eta \leq 0.340$.

This example highlights the use of efficient and robust lower bounds on multimode fidelities rather than fidelity estimates [44], in conjunction with our family of multimode witnesses to detect Wigner negativity of realistic experimental states.

7 Conclusion and open problems

Characterising quantum properties of physical systems is an important step in the development of quantum technologies and negativity of the Wigner function, a necessary resource for any quantum computational speedup, is no exception. In this work, we have derived a complete family of Wigner negativity witnesses which provide an operational quantification of Wigner negativity, both in the single-mode and multimode settings. In the context of quantum optical information processing, the main application of our method is in experimental scenarios, where it leads to robust and efficient certification of negativity of the Wigner function. Witnesses of Wigner negativity also provide witnesses of non-classicality although they are no longer complete in this case.

What is more, our witnesses also delineate the set of quantum states with positive Wigner function, and it would be interesting to understand whether additional insights on this set can be obtained using these witnesses.

The Wigner function has been extended to the discrete-variable setting [71, 72], where it

has been linked to contextuality [73–75], a necessary resource for discrete-variable quantum computing [76, 77]. A framework for treating contextuality and computing the amount of contextuality in continuous-variable settings has recently been developed [78]. As we obtained reliable Wigner negativity witnesses, it would be interesting to investigate the link between Wigner negativity and continuous-variable contextuality.

Hierarchies of semidefinite programs (in particular with non-commutative variables [79]) have found many recent applications in quantum information theory. From an infinite-dimensional linear program, we were able to use numerically both a hierarchy of upper bounds and a hierarchy of lower bounds—thus obtaining a certificate for the optimality of these bounds by looking at their difference—whereas this only works in specific cases for the Lasserre hierarchy of upper bounds [35]. Can we find other interesting cases where we can exploit both hierarchies? Moreover, we obtained an analytical sequence of lower bounds for the threshold value of the program (LP_n). Can we also get an analytical sequence of upper bounds? In particular, we anticipate that Fock states $|n\rangle$ get further away from the set of states having a positive Wigner function as n increases and that $\omega_n = \mathcal{O}(\frac{1}{\sqrt{n}})$ as $n \rightarrow +\infty$.

Finally, using our multimode Wigner negativity witnesses for studying the interplay between Wigner negativity and entanglement [52] is a very interesting prospect which we leave for future work.

Acknowledgments

U. Chabaud acknowledges stimulating discussions with S. Gribling, T. Freiman and T. Vidick. P.-E. Emeriau acknowledges interesting discussions with A. Oustry, E. Galvão and R. Soares Barbosa. We thank J. Eisert for his valuable comments on a previous version of this work and P. Paule for providing access to the Mathematica package for implementing Zeilberger’s algorithm. U. Chabaud acknowledges funding provided by the Institute for Quantum Information and Matter, an NSF Physics Frontiers Center (NSF Grant PHY-1733907). F. Grosshans acknowledges funding from the ANR through the ANR-17-CE24-0035 VanQuTe project.

References

- [1] S. Lloyd and S. L. Braunstein, “Quantum computation over continuous variables,” in *Quantum Information with Continuous Variables*, pp. 9–17. Springer, 1999.
- [2] S. Yokoyama, R. Ukai, S. C. Armstrong, C. Sornphiphatphong, T. Kaji, S. Suzuki, J.-i. Yoshikawa, H. Yonezawa, N. C. Menicucci, and A. Furusawa, “Ultra-large-scale continuous-variable cluster states multiplexed in the time domain,” *Nature Photonics* **7**, 982 (2013).
- [3] U. Leonhardt, “Essential Quantum Optics,” Cambridge University Press, Cambridge, UK, 1st ed., 2010.
- [4] J. E. Moyal, “Quantum mechanics as a statistical theory,” in *Mathematical Proceedings of the Cambridge Philosophical Society*, vol. 45, pp. 99–124, Cambridge University Press. 1949.
- [5] E. P. Wigner, “On the quantum correction for thermodynamic equilibrium,” in *Part I: Physical Chemistry. Part II: Solid State Physics*, pp. 110–120. Springer, 1997.
- [6] C. T. Lee, “Measure of the nonclassicality of nonclassical states,” *Physical Review A* **44**, R2775 (1991).
- [7] G. Giedke and J. I. Cirac, “Characterization of Gaussian operations and distillation of Gaussian states,” *Physical Review A* **66**, 032316 (2002).
- [8] J. Eisert, S. Scheel, and M. B. Plenio, “Distilling Gaussian states with Gaussian operations is impossible,” *Physical Review Letters* **89**, 137903 (2002).
- [9] J. Fiurášek, “Gaussian transformations and distillation of entangled Gaussian states,” *Physical Review Letters* **89**, 137904 (2002).
- [10] J. Niset, J. Fiurášek, and N. J. Cerf, “No-go theorem for Gaussian quantum error correction,” *Physical Review Letters* **102**, 120501 (2009).
- [11] S. Ghose and B. C. Sanders, “Non-Gaussian ancilla states for continuous variable quantum computation via Gaussian maps,” *Journal of Modern Optics* **54**, 855–869 (2007).
- [12] S. D. Bartlett, B. C. Sanders, S. L. Braunstein, and K. Nemoto, “Efficient Classical Simulation of Continuous Variable Quantum Information Processes,” *Physical Review Letters* **88**, 097904 (2002).
- [13] U. Chabaud, G. Ferrini, F. Grosshans, and D. Markham, “Classical simulation of Gaussian quantum circuits with non-Gaussian input states,” [arXiv:2010.14363](https://arxiv.org/abs/2010.14363).
- [14] R. L. Hudson, “When is the Wigner quasi-probability density non-negative?,” *Reports on Mathematical Physics* **6**, 249–252 (1974).
- [15] F. Soto and P. Claverie, “When is the Wigner function of multidimensional systems nonnegative?,” *Journal of Mathematical Physics* **24**, 97–100 (1983).
- [16] A. Mandilara, E. Karpov, and N. Cerf, “Extending Hudson’s theorem to mixed quantum states,” *Physical Review A* **79**, 062302 (2009).
- [17] R. Filip and L. Mišta Jr, “Detecting quantum states with a positive Wigner function beyond mixtures of Gaussian states,” *Physical Review Letters* **106**, 200401 (2011).
- [18] K. C. Tan, S. Choi, and H. Jeong, “Negativity of quasiprobability distributions as a measure of nonclassicality,” *Physical review letters* **124**, 110404 (2020).
- [19] U. Titulaer and R. Glauber, “Correlation functions for coherent fields,” *Physical Review* **140**, B676 (1965).
- [20] A. Kenfack and K. Życzkowski, “Negativity of the Wigner function as an indicator of non-classicality,” *Journal of Optics B: Quantum and Semiclassical Optics* **6**, 396 (2004).
- [21] A. Mari and J. Eisert, “Positive Wigner Functions Render Classical Simulation of Quantum Computation Efficient,” *Physical Review Lett.* **109**, 230503 (2012).
- [22] L. García-Álvarez, C. Calcluth, A. Ferraro, and G. Ferrini, “Efficient simulatability of continuous-variable circuits with large Wigner negativity,” [arXiv:2005.12026](https://arxiv.org/abs/2005.12026).
- [23] J. Preskill, “Quantum Computing in the NISQ era and beyond,” *Quantum* **2**, 79 (2018).
- [24] J. Eisert, D. Hangleiter, N. Walk, I. Roth, D. Markham, R. Parekh, U. Chabaud, and E. Kashefi, “Quantum certification and

- benchmarking,” *Nature Reviews Physics* **2**, 382–390 (2020).
- [25] G. M. D’Ariano, M. G. Paris, and M. F. Sacchi, “Quantum tomography,” *Advances in Imaging and Electron Physics* **128**, 206–309 (2003), [arXiv:quant-ph/0302028](#).
- [26] A. I. Lvovsky and M. G. Raymer, “Continuous-variable optical quantum-state tomography,” *Reviews of Modern Physics* **81**, 299 (2009).
- [27] U. Chabaud, T. Douce, F. Grosshans, E. Kashefi, and D. Markham, “Building Trust for Continuous Variable Quantum States,” in *15th Conference on the Theory of Quantum Computation, Communication and Cryptography*. 2020.
- [28] B. M. Terhal, “A family of indecomposable positive linear maps based on entangled quantum states,” *Linear Algebra and its Applications* **323**, 61–73 (2001).
- [29] M. Lewenstein, B. Kraus, J. I. Cirac, and P. Horodecki, “Optimization of entanglement witnesses,” *Physical Review A* **62**, 052310 (2000).
- [30] A. Mari, K. Kieling, B. M. Nielsen, E. Polzik, and J. Eisert, “Directly estimating nonclassicality,” *Physical Review Letters* **106**, 010403 (2011).
- [31] T. Kiesel and W. Vogel, “Universal nonclassicality witnesses for harmonic oscillators,” *Physical Review A* **85**, 062106 (2012).
- [32] U. Chabaud, G. Roeland, M. Walschaers, F. Grosshans, V. Parigi, D. Markham, and N. Treps, “Certification of non-Gaussian states with operational measurements,” [arXiv:2011.04320](#).
- [33] J.-B. Lasserre, “Global optimization with polynomials and the problem of moments,” *SIAM Journal on optimization* **11**, 796–817 (2001).
- [34] P. A. Parrilo, *Structured semidefinite programs and semialgebraic geometry methods in robustness and optimization*. PhD thesis, California Institute of Technology, 2000.
- [35] J. B. Lasserre, “A new look at nonnegativity on closed sets and polynomial optimization,” *SIAM Journal on Optimization* **21**, 864–885 (2011).
- [36] M. A. Nielsen and I. L. Chuang, “Quantum Computation and Quantum Information: 10th Anniversary Edition,”. Cambridge University Press, New York, NY, USA, 10th ed., 2011.
- [37] C. Weedbrook, S. Pirandola, R. García-Patrón, N. J. Cerf, T. C. Ralph, J. H. Shapiro, and S. Lloyd, “Gaussian quantum information,” *Reviews of Modern Physics* **84**, 621 (2012).
- [38] A. Wünsche, “Laguerre 2D-functions and their application in quantum optics,” *Journal of Physics A: Mathematical and General* **31**, 8267 (1998).
- [39] A. Royer, “Wigner function as the expectation value of a parity operator,” *Physical Review A* **15**, 449 (1977).
- [40] K. Banaszek, C. Radzewicz, K. Wódkiewicz, and J. Krasieński, “Direct measurement of the Wigner function by photon counting,” *Physical Review A* **60**, 674 (1999).
- [41] K. E. Cahill and R. J. Glauber, “Density operators and quasiprobability distributions,” *Physical Review* **177**, 1882 (1969).
- [42] K. Husimi, “Some formal properties of the density matrix,” *Proceedings of the Physico-Mathematical Society of Japan. 3rd Series* **22**, 264–314 (1940).
- [43] T. Richter, “Determination of photon statistics and density matrix from double homodyne detection measurements,” *Journal of Modern Optics* **45**, 1735–1749 (1998).
- [44] U. Chabaud, F. Grosshans, E. Kashefi, and D. Markham, “Efficient verification of Boson Sampling,” [arXiv:2006.03520](#).
- [45] A. Ferraro, S. Olivares, and M. G. Paris, “Gaussian states in continuous variable quantum information,” [arXiv:quant-ph/0503237](#).
- [46] F. Albarelli, M. G. Genoni, M. G. Paris, and A. Ferraro, “Resource theory of quantum non-Gaussianity and Wigner negativity,” *Physical Review A* **98**, 052350 (2018).
- [47] R. Takagi and Q. Zhuang, “Convex resource theory of non-Gaussianity,” *Physical Review A* **97**, 062337 (2018).
- [48] Q. Zhuang, P. W. Shor, and J. H. Shapiro, “Resource theory of non-Gaussian

- operations,” *Physical Review A* **97**, 052317 (2018).
- [49] U. Chabaud, D. Markham, and F. Grosshans, “Stellar representation of non-Gaussian quantum states,” *Physical Review Letters* **124**, 063605 (2020).
- [50] L. Vandenberghe and S. Boyd, “Semidefinite programming,” *SIAM review* **38**, 49–95 (1996).
- [51] J. Fiurášek and M. Ježek, “Witnessing negativity of Wigner function by estimating fidelities of catlike states from homodyne measurements,” *Physical Review A* **87**, 062115 (2013).
- [52] M. Walschaers, C. Fabre, V. Parigi, and N. Treps, “Entanglement and Wigner Function Negativity of Multimode Non-Gaussian States,” *Physical Review Letters* **119**, 183601 (2017).
- [53] U. Chabaud and P.-E. Emeriau, “Zeilberger’s algorithm and Hierarchy of semidefinite programs.” *Software Heritage repository* `swh:1:dir:d98f70e386783ef69bf8c2ecafdb7b328b19b7ec` containing the numerical tools developed for this article.
- [54] A. Ourjoumtsev, R. Tualle-Brouri, J. Laurat, and P. Grangier, “Generating optical Schrödinger kittens for quantum information processing,” *Science* **312**, 83–86 (2006).
- [55] B. C. Sanders, “Entangled coherent states,” *Physical Review A* **45**, 6811 (1992).
- [56] W. H. Zurek, “Sub-Planck structure in phase space and its relevance for quantum decoherence,” *Nature* **412**, 712–717 (2001).
- [57] G. Sagnol and M. Stahlberg, “Picos, a python interface to conic optimization solvers,” in *Proceedings of the in 21st International Symposium on Mathematical Programming*. 2012.
- [58] M. ApS, *MOSEK Optimizer API for Python 9.2.36*, 2019. <https://docs.mosek.com/9.2/pythonapi/index.html>.
- [59] M. Nakata, “A numerical evaluation of highly accurate multiple-precision arithmetic version of semidefinite programming solver: SDPA-GMP,-QD and-DD.,” in *2010 IEEE International Symposium on Computer-Aided Control System Design*, pp. 29–34, IEEE. 2010.
- [60] K. Fujisawa, M. Kojima, K. Nakata, and M. Yamashita, *SDPA (SemiDefinite Programming Algorithm) User’s Manual—Version 6.2. 0*, 2002.
- [61] A. Barvinok, “A course in convexity,” vol. 54 of *Graduate Studies in Mathematics*. American Mathematical Society, 2002.
- [62] G. Szegő, “Orthogonal Polynomials, revised ed,” in *American Mathematical Society Colloquium Publications*, vol. 23. 1959.
- [63] O. Nikodym, “Sur une généralisation des intégrales de M. J. Radon,” *Fundamenta Mathematicae* **15**, 131–179 (1930).
- [64] M. Guillemot-Teissier, “Développements des distributions en séries de fonctions orthogonales. Séries de Legendre et de Laguerre,” *Annali della Scuola Normale Superiore di Pisa-Classe di Scienze* **25**, 519–573 (1971).
- [65] M. Reed and B. Simon, “II: Fourier Analysis, Self-Adjointness,” vol. 2. Elsevier, 1975.
- [66] M. Riesz, “Sur le problème des moments, Troisième Note,” *Ark. Mat. Fys* **16**, 1–52 (1923).
- [67] E. Haviland, “On the momentum problem for distribution functions in more than one dimension. II,” *American Journal of Mathematics* **58**, 164–168 (1936).
- [68] D. Hilbert, “Über die darstellung definiter formen als summe von formenquadraten,” *Mathematische Annalen* **32**, 342–350 (1888).
- [69] H. W. Gould, “Combinatorial Identities: A standardized set of tables listing 500 binomial coefficient summations,”. Morgantown, W Va, 1972.
- [70] D. Zeilberger, “The method of creative telescoping,” *Journal of Symbolic Computation* **11**, 195–204 (1991).
- [71] U. Leonhardt, “Quantum-state tomography and discrete Wigner function,” *Physical Review Letters* **74**, 4101 (1995).
- [72] D. Gross, “Hudson’s theorem for finite-dimensional quantum systems,” *Journal of mathematical physics* **47**, 122107 (2006).
- [73] R. W. Spekkens, “Negativity and contextuality are equivalent notions of nonclassicality,” *Physical Review Letters* **101**, 020401 (2008).
- [74] N. Delfosse, C. Okay, J. Bermejo-Vega, D. E. Browne, and R. Raussendorf,

- “Equivalence between contextuality and negativity of the Wigner function for qudits,” *New Journal of Physics* **19**, 123024 (2017).
- [75] R. Raussendorf, D. E. Browne, N. Delfosse, C. Okay, and J. Bermejo-Vega, “Contextuality and Wigner-function negativity in qubit quantum computation,” *Physical Review A* **95**, 052334 (2017).
- [76] M. Howard, J. Wallman, V. Veitch, and J. Emerson, “Contextuality supplies the ‘magic’ for quantum computation,” *Nature* **510**, 351 (2014).
- [77] J. Bermejo-Vega, N. Delfosse, D. E. Browne, C. Okay, and R. Raussendorf, “Contextuality as a resource for models of quantum computation with qubits,” *Physical Review Letters* **119**, 120505 (2017).
- [78] R. S. Barbosa, T. Douce, P.-E. Emeriau, E. Kashefi, and S. Mansfield, “Continuous-variable nonlocality and contextuality,” [arXiv:1905.08267](https://arxiv.org/abs/1905.08267).
- [79] M. Navascués, S. Pironio, and A. Acín, “A convergent hierarchy of semidefinite programs characterizing the set of quantum correlations,” *New Journal of Physics* **10**, 073013 (2008).
- [80] R. E. Curto and L. A. Fialkow, “An analogue of the Riesz–Haviland theorem for the truncated moment problem,” *Journal of Functional Analysis* **255**, 2709–2731 (2008).
- [81] D. Henrion and M. Korda, “Convex computation of the region of attraction of polynomial control systems,” *IEEE Transactions on Automatic Control* **59**, 297–312 (2014).
- [82] J.-B. Lasserre, “Moments, positive polynomials and their applications,” in *Series on Optimization and its Applications*, vol. 1. Imperial College Press, 2009.

A Riesz–Haviland theorem in Laguerre basis

Let $\nu = (\nu_l)_{l \in \mathbb{N}} \in \mathbb{R}^{\mathbb{N}}$. Let us introduce the Riesz functional

$$\begin{aligned} L_\nu : \mathbb{R}[x] &\longrightarrow \mathbb{R} \\ P(x) = \sum_{k=0}^p p_k x^k &\longmapsto \sum_{k=0}^p p_k \nu_k, \end{aligned} \quad (78)$$

which maps real polynomials to real numbers. Let K be a closed subset of \mathbb{R} . We say that L_ν is K -non-negative if $L_\nu(P) \geq 0$ for all $P \in \mathbb{R}[x]$ non-negative on K . We recall the classical Riesz–Haviland theorem [66, 67] (see, e.g., [80] for a recent formulation):

Theorem 7 (Riesz–Haviland). *The sequence $\nu = (\nu_k)_{k \in \mathbb{N}} \in \mathbb{R}^{\mathbb{N}}$ is the sequence of moments $\int_K x^k d\nu(x)$ of a non-negative distribution ν supported on K if and only if L_ν is K -non-negative.*

We prove a modified version of this result in the basis of Laguerre functions. To that end, we introduce the following change of basis:

Lemma 8. *Let $\mu, \nu \in \mathbb{R}^{\mathbb{N}}$. For all $m \in \mathbb{N}$, the following conditions are equivalent:*

$$\begin{aligned} (i) \quad \forall k \in \llbracket 0, m \rrbracket, \quad \mu_k &= \sum_{l=0}^k \nu_l \frac{(-1)^{k+l}}{l!} \binom{k}{l}, \\ (ii) \quad \forall l \in \llbracket 0, m \rrbracket, \quad \nu_l &= \sum_{k=0}^l \mu_k \binom{l}{k} l!. \end{aligned}$$

As a direct consequence, we retrieve the formula:

$$x^l = \sum_{k=0}^l (-1)^k \binom{l}{k} l! L_k(x), \quad (79)$$

for all $l \in \mathbb{N}$ and all $x \in \mathbb{R}_+$.

Proof. (i) \Rightarrow (ii): suppose that

$$\forall k \in \llbracket 0, m \rrbracket, \quad \mu_k = \sum_{p=0}^k \nu_p \frac{(-1)^{k+p}}{p!} \binom{k}{p}. \quad (80)$$

Then, for all $l \in \llbracket 0, m \rrbracket$,

$$\begin{aligned} \sum_{k=0}^l \mu_k \binom{l}{k} l! &= \sum_{k=0}^l \sum_{p=0}^k \nu_p \frac{(-1)^{k+p}}{p!} \binom{k}{p} \binom{l}{k} l! \\ &= \sum_{p=0}^l \nu_p \frac{l!}{p!} \binom{l}{p} \sum_{k=p}^l (-1)^{k-p} \binom{l-p}{k-p} \\ &= \sum_{p=0}^l \nu_p \frac{l!}{p!} \binom{l}{p} \sum_{q=0}^{l-p} (-1)^q \binom{l-p}{q} \\ &= \nu_l, \end{aligned} \quad (81)$$

where we used Eq. (80) in the first line and the binomial theorem in the last line which imposes $l = p$.

(ii) \Rightarrow (i): suppose that

$$\forall l \in \llbracket 0, m \rrbracket, \quad \nu_l = \sum_{p=0}^l \mu_p \binom{l}{p} l!. \quad (82)$$

Then, for all $k \in \llbracket 0, m \rrbracket$,

$$\begin{aligned}
\sum_{l=0}^k \nu_l \frac{(-1)^{k+l}}{l!} \binom{k}{l} &= \sum_{l=0}^k \sum_{p=0}^l \mu_p (-1)^{k+l} \binom{l}{p} \binom{k}{l} \\
&= \sum_{p=0}^k \mu_p (-1)^{k+p} \binom{k}{p} \sum_{l=p}^k (-1)^{l-p} \binom{k-p}{l-p} \\
&= \sum_{p=0}^k \mu_p (-1)^{k+p} \binom{k}{p} \sum_{q=0}^{k-p} (-1)^q \binom{k-p}{q} \\
&= \mu_k,
\end{aligned} \tag{83}$$

where we used Eq. (82) in the first line and the binomial theorem in the last line which imposes $k = p$. \square

We may now prove the Riesz–Haviland theorem in Laguerre basis (Theorem 1 from the main text):

Theorem 1. *Let $\boldsymbol{\mu} = (\mu_k)_{k \in \mathbb{N}} \in \mathbb{R}^{\mathbb{N}}$. The sequence $\boldsymbol{\mu}$ is the sequence of Laguerre moments $\int_{\mathbb{R}_+} \mathcal{L}_k(x) d\mu(x)$ of a non-negative distribution μ supported on \mathbb{R}_+ if and only if*

$$\forall m \in \mathbb{N}, \forall g \in \mathcal{R}_{m,+}(\mathbb{R}_+), \langle f_{\boldsymbol{\mu}}, g \rangle \geq 0. \tag{84}$$

Proof. Let $\boldsymbol{\mu} = (\mu_k)_{k \in \mathbb{N}} \in \mathbb{R}^{\mathbb{N}}$, and suppose that the sequence $\boldsymbol{\mu}$ is the sequence of Laguerre moments $\int_{\mathbb{R}_+} \mathcal{L}_k(x) d\mu(x)$ of a non-negative distribution μ supported on \mathbb{R}_+ .

Let $m \geq 0$ and let $g = \sum_{k=0}^m g_k \mathcal{L}_k \in \mathcal{R}_{m,+}(\mathbb{R}_+)$. The distribution μ is non-negative, so $\langle \mu, g \rangle \geq 0$. Moreover,

$$\begin{aligned}
\langle f_{\boldsymbol{\mu}}, g \rangle &= \sum_{k=0}^m \mu_k g_k \\
&= \int_{\mathbb{R}_+} \sum_{k=0}^m g_k \mathcal{L}_k d\mu \\
&= \langle \mu, g \rangle.
\end{aligned} \tag{85}$$

Hence, for all $m \in \mathbb{N}$ and all $g \in \mathcal{R}_{m,+}(\mathbb{R}_+)$, $\langle f_{\boldsymbol{\mu}}, g \rangle \geq 0$.

Conversely, let $\boldsymbol{\mu} = (\mu_k)_{k \in \mathbb{N}} \in \mathbb{R}^{\mathbb{N}}$, and suppose that for all $m \in \mathbb{N}$ and all $g \in \mathcal{R}_{m,+}(\mathbb{R}_+)$, $\langle f_{\boldsymbol{\mu}}, g \rangle \geq 0$. We define the sequence $\boldsymbol{\nu} = (\nu_l)_{l \in \mathbb{N}} \in \mathbb{R}^{\mathbb{N}}$ by

$$\nu_l := \sum_{k=0}^l \mu_k \binom{l}{k} l!, \tag{86}$$

for all $l \in \mathbb{N}$.

Let $m \in \mathbb{N}$ and let $P(x) = \sum_{l=0}^m p_l x^l$ be a non-negative polynomial over \mathbb{R}_+ . By Eq. (79), for all $x \in \mathbb{R}_+$,

$$\begin{aligned}
P(x) &= \sum_{l=0}^m p_l \sum_{k=0}^l (-1)^k \binom{l}{k} l! L_k(x) \\
&= \sum_{k=0}^m (-1)^k L_k(x) \left(\sum_{l=k}^m p_l \binom{l}{k} l! \right).
\end{aligned} \tag{87}$$

Let $g_P(x) := P(x)e^{-\frac{x}{2}}$, for $x \in \mathbb{R}_+$. We have $g_P \in \mathcal{R}_{m,+}(\mathbb{R}_+)$, so $\langle f_{\boldsymbol{\mu}}, g_P \rangle \geq 0$. Moreover, with Eq. (87)

$$\begin{aligned}
\langle f_{\boldsymbol{\mu}}, g_P \rangle &= \sum_{k=0}^m \mu_k \left(\sum_{l=k}^m p_l \binom{l}{k} l! \right) \\
&= \sum_{l=0}^m \left(\sum_{k=0}^l \mu_k \binom{l}{k} l! \right) p_l \\
&= L_{\boldsymbol{\nu}}(P)
\end{aligned} \tag{88}$$

where we used Eq. (86) and the definition of the Riesz functional from Eq. (78) in the last line. In particular, $L_\nu(P) \geq 0$, and this holds for all non-negative polynomials P over \mathbb{R}_+ . By the Riesz–Haviland theorem (Theorem 7), this implies that ν is the sequence of moments of a non-negative distribution ν supported on \mathbb{R}_+ .

Furthermore, we have that for all $k \in \mathbb{N}$:

$$\begin{aligned}
\mu_k &= \sum_{l=0}^k \nu_l \frac{(-1)^{k+l}}{l!} \binom{k}{l} \\
&= \sum_{l=0}^k \frac{(-1)^{k+l}}{l!} \binom{k}{l} \int_{\mathbb{R}_+} x^l d\nu(x) \\
&= \int_{\mathbb{R}_+} (-1)^k \sum_{l=0}^k \frac{(-1)^l}{l!} \binom{k}{l} x^l d\nu(x) \\
&= \int_{\mathbb{R}_+} (-1)^k L_k(x) d\nu(x) \\
&= \int_{\mathbb{R}_+} \mathcal{L}_k(x) e^{\frac{x}{2}} d\nu(x)
\end{aligned} \tag{89}$$

where we used Lemma 8 in the first line. Hence, μ is the sequence of Laguerre moments of the distribution $\mu(x) := e^{\frac{x}{2}} \nu(x)$ supported on \mathbb{R}_+ , which is non-negative since ν is non-negative. \square

B Theory for infinite-dimensional linear programs

This appendix is dedicated to expressing formally our linear program as presented in [61, IV–(6.1)] so that readers unfamiliar with global optimisation may better understand why (LP_n) is indeed an infinite-dimensional linear program and how to derive its dual program. We recall our initial program (LP_n) :

$$(LP_n) \quad \left\{ \begin{array}{l} \sup_{(F_k)_{k \in \mathbb{N}} \in \mathbb{R}^{\mathbb{N}}} F_n \\ \text{subject to} \quad \sum_k F_k = 1 \\ \text{and} \quad \forall k \in \mathbb{N}, \quad F_k \geq 0 \\ \text{and} \quad \forall x \in \mathbb{R}_+, \quad \sum_k F_k \mathcal{L}_k(x) \geq 0. \end{array} \right.$$

Let us introduce the spaces:

- $E_1 = \ell^2 \times L^2(\mathbb{R}_+)$ ⁶.
- $F_1 = \ell^2 \times L^2(\mathbb{R}_+)$ ⁷ the dual space of E_1 .
- $E_2 = \mathbb{R} \times L^2(\mathbb{R}_+)$.
- $F_2 = \mathbb{R} \times L^2(\mathbb{R}_+)$ the dual space of E_2 .

We also define the dualities $\langle -, - \rangle_1 : E_1 \times F_1 \rightarrow \mathbb{R}$ and $\langle -, - \rangle_2 : E_2 \times F_2 \rightarrow \mathbb{R}$ as follows:

$$\begin{aligned}
\forall e_1 = ((u_k), f) \in E_1, \forall f_1 = ((v_k), \mu) \in F_1, \quad \langle e_1, f_1 \rangle_1 &:= \sum_k u_k v_k + \int_{\mathbb{R}_+} f d\mu, \\
\forall e_2 = (x, f) \in E_2, \forall f_2 = (y, \mu) \in F_2, \quad \langle e_2, f_2 \rangle_2 &:= xy + \int_{\mathbb{R}_+} f d\mu.
\end{aligned} \tag{90}$$

⁶Recall that via expansion on a basis of $L^2(\mathbb{R}_+)$, $L^2(\mathbb{R}_+)$ and ℓ^2 are isomorphic and that the family of Laguerre functions forms a basis of $L^2(\mathbb{R}_+)$.

⁷The spaces $L^2(\mathbb{R}_+)$ and $L^{2'}(\mathbb{R}_+)$ are isomorphic by the Radon–Nikodym theorem. We associate a measure in the dual of $L^2(\mathbb{R}_+)$ with the Lebesgue measure on \mathbb{R}_+ times the corresponding function in $L^2(\mathbb{R}_+)$.

Let $A: E_1 \rightarrow E_2$ be the following linear transformation:

$$\forall e_1 = ((u_k), f) \in E_1, \quad A(e_1) := \left(\sum_k u_k, x \in \mathbb{R}_+ \mapsto f(x) - \sum_k u_k \mathcal{L}_k(x) \right), \quad (91)$$

and $A^*: F_2 \rightarrow F_1$ be defined as:

$$\forall f_2 = (y, \mu) \in F_2, \quad A^*(f_2) := \left(y - \int_{\mathbb{R}_+} \mathcal{L}_k d\mu, \mu \right). \quad (92)$$

We can easily verify that A^* is the dual transformation of A , i.e., $\forall e_1 \in E_1, \forall f_2 \in F_2$ we have $\langle A(e_1), f_2 \rangle_2 = \langle e_1, A^*(f_2) \rangle_1$.

Recall that $L_+^2(\mathbb{R}_+)$ is the cone of non-negative functions in $L^2(\mathbb{R}_+)$ and ℓ_+^2 the cone of sequences in ℓ^2 with non-negative coefficients. We will optimise in the convex cones $K_1 = \ell_+^2 \times L_+^2(\mathbb{R}_+) \subset E_1$ and $K_2 = \{0\}$. The dual cones are then respectively: $K_1^* = \{f_1 \in F_1 : \forall e_1 \in K_1, \langle e_1, f_1 \rangle \geq 0\}$ and $K_2^* = F_2$.

We can now rewrite the problem (LP_n) as a standard linear program in convex cones. We choose the vector function in the objective to be $c_n = ((\delta_{kn})_k, \mathbf{0}) \in F_1$ and we also set $b = (1, \mathbf{0}) \in E_2$ for the constraints. The standard form of (LP_n) in the sense of [61] can be written as follows:

$$(LP_n) \quad \begin{cases} \sup_{e_1 \in E_1} \langle e_1, c_n \rangle_1 \\ \text{subject to} & A(e_1) = b \\ \text{and} & e_1 \geq_{K_1} \mathbf{0}. \end{cases}$$

The standard form of the dual $(D-LP_n)$ of problem (LP_n) can be expressed as follows:

$$(D-LP_n) \quad \begin{cases} \inf_{f_2 \in F_2} \langle b, f_2 \rangle_2 \\ \text{subject to} & A^*(f_2) \geq_{K_1^*} c_n, \end{cases}$$

which can be expanded as:

$$(D-LP_n) \quad \begin{cases} \inf_{\substack{y \in \mathbb{R} \\ \mu \in L^{2'}(\mathbb{R}_+)}} y \\ \text{subject to} & \forall k \in \mathbb{N}, \quad y \geq \int_{\mathbb{R}_+} \mathcal{L}_k d\mu \\ \text{and} & y \geq 1 + \int_{\mathbb{R}_+} \mathcal{L}_n d\mu \\ \text{and} & \forall f \in L_+^2(\mathbb{R}_+), \quad \langle \mu, f \rangle = \int_{\mathbb{R}_+} f d\mu \geq 0. \end{cases}$$

Note that a similar derivation holds for the more general form where one uses a linear combination of fidelities with Fock states. The displaced Fock states version can be obtained by classical post-processing as detailed in section 2.2. For $n \in \mathbb{N}^*$ and some vector $\mathbf{a} = (a_1, a_2, \dots, a_n) \in [0, 1]^n$, the computation of:

$$\omega_{\mathbf{a}} := \sup_{\substack{\rho \in \mathcal{D}(\mathcal{H}) \\ W_\rho \geq 0}} \text{Tr} \left(\hat{\Omega}_{\mathbf{a}, 0} \rho \right), \quad (93)$$

can be expressed as:

$$(\text{LP}_a) \quad \left\{ \begin{array}{l} \sup_{(F_k)_{k \in \mathbb{N}} \in \mathbb{R}^{\mathbb{N}}} \sum_{k=1}^n a_k F_k \\ \text{subject to} \quad \sum_k F_k = 1 \\ \text{and} \quad \forall k \in \mathbb{N}, \quad F_k \geq 0 \\ \text{and} \quad \forall x \in \mathbb{R}_+, \quad \sum_k F_k \mathcal{L}(x) \geq 0. \end{array} \right.$$

Its dual reads:

$$(\text{D-LP}_a) \quad \left\{ \begin{array}{l} \inf_{\substack{y \in \mathbb{R} \\ \mu \in L^2(\mathbb{R}_+)}} y \\ \text{subject to} \quad \forall k \leq n \in \mathbb{N}, \quad y \geq a_k + \int_{\mathbb{R}_+} \mathcal{L}_k d\mu \\ \text{and} \quad \forall k > n \in \mathbb{N}, \quad y \geq \int_{\mathbb{R}_+} \mathcal{L}_k d\mu \\ \text{and} \quad \forall f \in L^2_+(\mathbb{R}_+), \quad \langle \mu, f \rangle = \int_{\mathbb{R}_+} f d\mu \geq 0. \end{array} \right.$$

C Proof of technical lemmas

In this section we prove the technical lemmas from section 5.4. For completeness, we include the proof of Lemma 3 below:

Lemma 3 ([68]). *Let $p \in \mathbb{N}$ and let P be a univariate polynomial of degree $2p$. Let $X = (1, x, \dots, x^p)$ be the vector of monomials. Then, P is non-negative over \mathbb{R} if and only if there exists a sum of squares decomposition for P , i.e., a real $(p+1) \times (p+1)$ positive semidefinite matrix Q such that for all $x \in \mathbb{R}$,*

$$P(x) = X^T Q X. \quad (94)$$

Proof. If for all $x \in \mathbb{R}$, $P(x) = X^T Q X$ with $X = (1, x, \dots, x^p)$ for Q positive semidefinite, then P is clearly non-negative over \mathbb{R} .

Conversely, suppose the univariate polynomial P of degree $2p$ is non-negative over \mathbb{R} . It can thus be written as a sum of squares of polynomials of degree at most p (e.g., by considering the factorisation of P and writing each term in the product as a sum of squares, given that its zeros on the real line have even multiplicity and that each complex zero is associated to a conjugate zero with same multiplicity): for all $x \in \mathbb{R}$,

$$P(x) = \sum_i S_i^2(x), \quad (95)$$

for some real polynomials S_i of degree at most p . Then, for some vectors of coefficients $\mathbf{s}_i \in \mathbb{R}^{p+1}$ we have

$$\begin{aligned} P(x) &= \sum_i (\mathbf{s}_i^T X)^2 \\ &= \sum_i (X^T \mathbf{s}_i)(\mathbf{s}_i^T X) \\ &= X^T \left(\sum_i \mathbf{s}_i \mathbf{s}_i^T \right) X. \end{aligned} \quad (96)$$

Setting $Q := \sum_i \mathbf{s}_i \mathbf{s}_i^T \succeq 0$ completes the proof. \square

We now turn to the proof of Lemma 4:

Lemma 4. *Non-negative polynomials on \mathbb{R}_+ can be written as sums of polynomials of the form $\sum_{l=0}^p x^l \sum_{i+j=2l} y_i y_j$, where $p \in \mathbb{N}$ and $y_i \in \mathbb{R}$, for all $0 \leq i \leq p$.*

Proof. Let P be a univariate polynomial of degree p which is non-negative on \mathbb{R}_+ . Writing $X = (1, x, \dots, x^p)$, the polynomial $x \mapsto P(x^2)$ of degree $2p$ is non-negative on \mathbb{R} , so by Lemma 3 there exists a real positive semidefinite matrix $Q = (Q_{ij})_{0 \leq i, j \leq p}$ such that for all $x \in \mathbb{R}$.

$$\begin{aligned} P(x^2) &= X^T Q X \\ &= \sum_{k=0}^{2p} x^k \sum_{i+j=k} Q_{ij} \\ &= \sum_{l=0}^p x^{2l} \sum_{i+j=2l} Q_{ij}, \end{aligned} \tag{97}$$

where the last line comes from the fact that $x \mapsto P(x^2)$ has no monomial of odd degree. Hence, for all $x \in \mathbb{R}_+$,

$$P(x) = \sum_{l=0}^p x^l \sum_{i+j=2l} Q_{ij}. \tag{98}$$

Q is a real $(p+1) \times (p+1)$ positive semidefinite matrix, so via Cholesky decomposition

$$Q = \sum_{k=0}^p \mathbf{y}^{(k)} \mathbf{y}^{(k)T}, \tag{99}$$

where $\mathbf{y}^{(k)} \in \mathbb{R}^{p+1}$ for all $k \in \llbracket 0, p \rrbracket$. We finally obtain, for all $x \in \mathbb{R}_+$,

$$\begin{aligned} P(x) &= \sum_{l=0}^p x^l \sum_{i+j=2l} \sum_{k=0}^p \left(\mathbf{y}^{(k)} \mathbf{y}^{(k)T} \right)_{ij} \\ &= \sum_{k=0}^p \left(\sum_{l=0}^p x^l \sum_{i+j=2l} \mathbf{y}_i^{(k)} \mathbf{y}_j^{(k)} \right). \end{aligned} \tag{100}$$

□

We recall a few definitions from the main text. For $\mathbf{s} \in \mathbb{R}^{\mathbb{N}}$, we define the associated formal series of Laguerre functions:

$$f_{\mathbf{s}} := \sum_{k \geq 0} s_k \mathcal{L}_k, \tag{101}$$

where for all $x \in \mathbb{R}_+$, $\mathcal{L}_k(x) = (-1)^k L_k(x) e^{-\frac{x}{2}}$, with $L_k(x) = \sum_{l=0}^k \frac{(-1)^l}{l!} \binom{k}{l} x^l$ the k^{th} Laguerre polynomial. For $m \in \mathbb{N}$, we also define the associated matrix $A_{\mathbf{s}}$ by

$$(A_{\mathbf{s}})_{0 \leq i, j \leq m} = \begin{cases} \sum_{k=0}^l s_k \binom{l}{k} l! & \text{when } i+j=2l, \\ 0 & \text{otherwise.} \end{cases} \tag{102}$$

For all $m \in \mathbb{N}$, the set of series of Laguerre functions over \mathbb{R}_+ truncated at m is denoted $\mathcal{R}_m(\mathbb{R}_+)$, and $\mathcal{R}_{m,+}(\mathbb{R}_+)$ denotes its subset of non-negative elements.

Lemma 5. *Let $m \in \mathbb{N}$ and let $\mathbf{s} \in \mathbb{R}^{\mathbb{N}}$. The following propositions are equivalent:*

- (i) $\forall g \in \mathcal{R}_{m,+}(\mathbb{R}_+)$, $\langle f_{\mathbf{s}}, g \rangle \geq 0$,
- (ii) $A_{\mathbf{s}} \succeq 0$.

Proof. By Lemma 4, any non-negative polynomial over \mathbb{R}_+ of degree less or equal to m can be expressed as a sum of polynomials of the form $\sum_{l=0}^m x^l \sum_{i+j=2l} y_i y_j$, where $Y = (y_0, \dots, y_m) \in \mathbb{R}^{m+1}$. Hence, any non-negative truncated Laguerre series (the elements of $\mathcal{R}_{m,+}(\mathbb{R}_+)$) can be expressed as a sum of terms of the form $e^{-\frac{x}{2}} \sum_{l=0}^m x^l \sum_{i+j=2l} y_i y_j$. By linearity, it is sufficient to check that the scalar products with these expressions are non-negative.

For all $k \in \mathbb{N}$ we have

$$s_k = \int_{\mathbb{R}_+} \mathcal{L}_k(x) f_{\mathbf{s}}(x) dx. \quad (103)$$

Thus,

$$\begin{aligned} A_{\mathbf{s}} \succeq 0 &\Leftrightarrow \forall Y \in \mathbb{R}^{m+1}, Y^T A_{\mathbf{s}} Y \geq 0 \\ &\Leftrightarrow \forall Y \in \mathbb{R}^{m+1}, \sum_{i,j=0}^m y_i y_j (A_{\mathbf{s}})_{ij} \geq 0 \\ &\Leftrightarrow \forall Y \in \mathbb{R}^{m+1}, \sum_{l=0}^m \sum_{i+j=2l}^m y_i y_j \sum_{k=0}^l s_k \binom{l}{k} l! \geq 0 \\ &\Leftrightarrow \forall Y \in \mathbb{R}^{m+1}, \int_{\mathbb{R}_+} \sum_{l=0}^m \sum_{i+j=2l}^m y_i y_j \sum_{k=0}^l \binom{l}{k} l! \mathcal{L}_k(x) f_{\mathbf{s}}(x) dx \geq 0 \\ &\Leftrightarrow \forall Y \in \mathbb{R}^{m+1}, \int_{\mathbb{R}_+} \sum_{l=0}^m \sum_{i+j=2l}^m y_i y_j \sum_{k=0}^l (-1)^k \binom{l}{k} l! L_k(x) e^{-\frac{x}{2}} f_{\mathbf{s}}(x) dx \geq 0 \\ &\Leftrightarrow \forall Y \in \mathbb{R}^{m+1}, \int_{\mathbb{R}_+} \left(e^{-\frac{x}{2}} \sum_{l=0}^m x^l \sum_{i+j=2l}^m y_i y_j \right) f_{\mathbf{s}}(x) dx \geq 0 \\ &\Leftrightarrow \forall Y \in \mathbb{R}^{m+1}, \left\langle f_{\mathbf{s}}, x \mapsto e^{-\frac{x}{2}} \sum_{l=0}^m x^l \sum_{i+j=2l}^m y_i y_j \right\rangle \geq 0 \\ &\Leftrightarrow \forall g \in \mathcal{R}_{m,+}(\mathbb{R}_+), \langle f_{\mathbf{s}}, g \rangle \geq 0, \end{aligned} \quad (104)$$

where we used Eq. (103) in the fourth line and Eq. (79) in the sixth line. \square

D Dual semidefinite programs

In this section, we detail the derivation of the dual semidefinite programs (D-SDP $_{\mathbf{n}}^{m,\geq}$) and (D-SDP $_{\mathbf{n}}^{m,\leq}$). The generalisations for (SDP $_{\mathbf{a}}^{m,\geq}$) and (SDP $_{\mathbf{a}}^{m,\leq}$) are straightforward.

A standard form for a semidefinite program is given by [50]:

$$(SDP) \quad \begin{cases} \sup & \text{Tr}(C^T X) \\ & X \in \text{Sym}_N \\ \text{subject to} & \forall i \in \llbracket 1, M \rrbracket, \quad \text{Tr}(B^{(i)} X) = b_i \\ \text{and} & X \succeq 0, \end{cases}$$

where $M, N \in \mathbb{N}$, $\mathbf{b} = (b_1, \dots, b_M) \in \mathbb{R}^M$, $C \in \text{Sym}_N$, and $B^{(i)} \in \text{Sym}_N$ for all $i \in \llbracket 1, M \rrbracket$. Its dual semidefinite program reads:

$$(D-SDP) \quad \begin{cases} \inf_{\mathbf{y} \in \mathbb{R}^M} & \mathbf{b}^T \mathbf{y} \\ \text{subject to} & \sum_{i=1}^M y_i B^{(i)} \succeq C. \end{cases}$$

D.1 Dual program for the semidefinite relaxations

We fix $m \geq n$ and we recall below the expression of $(\text{SDP}_n^{m,\geq})$:

$$(\text{SDP}_n^{m,\geq}) \quad \left\{ \begin{array}{l} \sup_{\substack{A \in \text{Sym}_{m+1} \\ \mathbf{F} \in \mathbb{R}^{m+1}}} F_n \\ \text{subject to} \quad \sum_{k=0}^m F_k = 1 \\ \text{and} \quad \forall k \leq m, F_k \geq 0 \\ \text{and} \quad \forall l \leq m, \forall i+j=2l, \quad A_{ij} = \sum_{k=0}^l F_k \binom{l}{k} l! \\ \text{and} \quad \forall l \in \llbracket 1, m \rrbracket, \forall i+j=2l-1, \quad A_{ij} = 0 \\ \text{and} \quad A \succeq 0. \end{array} \right.$$

To put $(\text{SDP}_n^{m,\geq})$ in the standard form (SDP) we set $N = 2 \times (m+1)$ and $M = 1 + (m+1)^2$. For all $r \in \mathbb{N}^*$ and all $i, j \in \llbracket 1, r \rrbracket$, let $E_r^{(i,j)}$ be the $r \times r$ matrix whose (i, j) entry is 1 and all other entries are 0. We set

$$\begin{aligned} X &= \text{Diag}_{k=0,\dots,m}(F_k) \oplus A \in \text{Sym}_N, \\ C &= E_N^{(n,n)} = E_{m+1}^{(n,n)} \oplus \mathbf{0}_{m+1} \in \text{Sym}_N, \\ \mathbf{b} &= (1, 0, 0, \dots, 0) \in \mathbb{R}^M, \\ B^{(0)} &= \mathbf{1}_{m+1} \oplus \mathbf{0}_{m+1} \in \text{Sym}_N, \\ \forall i, j \in \llbracket 0, m \rrbracket, \quad B^{(i,j)} &= \begin{cases} \text{Diag}_{k=0,\dots,m} \left(-\binom{l}{k} l! \right) \oplus \left(\frac{1}{2} E_{m+1}^{(i,j)} + \frac{1}{2} E_{m+1}^{(j,i)} \right) & \text{when } i+j=2l, \\ \mathbf{0}_{m+1} \oplus \left(\frac{1}{2} E_{m+1}^{(i,j)} + \frac{1}{2} E_{m+1}^{(j,i)} \right) & \text{otherwise,} \end{cases} \end{aligned} \quad (105)$$

with the convention $\binom{l}{k} = 0$ when $k > l$. The matrix $B^{(0)}$ corresponds to the constraint $\sum_{k=0}^m F_k = 1$, and we denote the corresponding dual variable $y \in \mathbb{R}$. Similarly, the matrices $B^{(i,j)}$ correspond to the $(m+1)^2$ constraints defining the symmetric matrix A , and we denote the corresponding dual variables $Q_{ij} \in \mathbb{R}$, with $Q_{ij} = Q_{ji}$ for all $i, j \in \llbracket 0, m \rrbracket$. We write $Q = (Q_{ij})_{0 \leq i, j \leq m}$. The standard form (D-SDP) of the dual program $(\text{D-SDP}_n^{m,\geq})$ thus reads:

$$(\text{D-SDP}_n^{m,\geq}) \quad \left\{ \begin{array}{l} \inf_{\substack{Q \in \text{Sym}_{m+1} \\ y \in \mathbb{R}}} y \\ \text{subject to} \quad \text{Diag}_{k=0,\dots,m} \left[y - \sum_{l=0}^m \sum_{i+j=2l} Q_{ij} \binom{l}{k} l! \right] \oplus \frac{1}{2} Q \succeq E_{m+1}^{(n,n)} \oplus \mathbf{0}_{m+1}. \end{array} \right.$$

Due to the block-diagonal structure of the matrices involved, the positive semidefinite constraint above is equivalent to the following constraints:

$$\left\{ \begin{array}{l} y \geq 1 + \sum_{l=0}^m \sum_{i+j=2l} Q_{ij} \binom{l}{n} l!, \\ \forall k \in \llbracket 0, m \rrbracket \setminus \{n\}, \quad y \geq \sum_{l=0}^m \sum_{i+j=2l} Q_{ij} \binom{l}{k} l!, \\ Q \succeq 0. \end{array} \right. \quad (106)$$

For $k \in \llbracket 0, m \rrbracket$, we define

$$\mu_k := \sum_{l=0}^m \sum_{i+j=2l} Q_{ij} \binom{l}{k} l! \in \mathbb{R}. \quad (107)$$

We obtain the program:

$$(D\text{-SDP}_n^{m,\geq}) \quad \begin{cases} \inf_{\substack{Q \in \text{Sym}_{m+1} \\ y, \mu \in \mathbb{R} \times \mathbb{R}^{m+1}}} y \\ \text{subject to } y \geq 1 + \mu_n \\ \text{and } \forall k \in \llbracket 0, m \rrbracket \setminus \{n\}, \quad y \geq \mu_k \\ \text{and } \forall k \in \llbracket 0, m \rrbracket, \quad \mu_k = \sum_{l=0}^m \sum_{i+j=2l} Q_{ij} \binom{l}{k} l! \\ \text{and } Q \succeq 0. \end{cases}$$

Finally, in order to obtain the form of $(D\text{-SDP}_n^{m,\geq})$ from the main text we prove the following result:

Lemma 9. *Let $u, v \in \mathbb{R}^{m+1}$. The following propositions are equivalent:*

$$(i) \quad \forall k \in \llbracket 0, m \rrbracket, \quad u_k = \sum_{l=0}^m v_l \binom{l}{k} l!,$$

$$(ii) \quad \forall l \in \llbracket 0, m \rrbracket, \quad v_l = \sum_{k=l}^m \frac{(-1)^{l+k}}{l!} \binom{k}{l} u_k.$$

Proof. The proof is similar to that of Lemma 8.

(i) \Rightarrow (ii): suppose that

$$\forall k \in \llbracket 0, m \rrbracket, \quad u_k = \sum_{p=0}^m v_p \binom{p}{k} p!. \quad (108)$$

Then, for all $l \in \llbracket 0, m \rrbracket$,

$$\begin{aligned} \frac{(-1)^l}{l!} \sum_{k=l}^m (-1)^k \binom{k}{l} u_k &= \frac{(-1)^l}{l!} \sum_{k=l}^m (-1)^k \binom{k}{l} \sum_{p=0}^m v_p \binom{p}{k} p! \\ &= \sum_{p=0}^m v_p \frac{p!}{l!} \sum_{k=l}^m (-1)^{k+l} \binom{p}{k} \binom{k}{l} \\ &= \sum_{p=l}^m v_p \frac{p!}{l!} \sum_{k=l}^p (-1)^{k+l} \binom{p}{k} \binom{k}{l} \\ &= \sum_{p=l}^m v_p \frac{p!}{l!} \sum_{k=l}^p (-1)^{k+l} \frac{p!k!}{k!(p-k)!l!(k-l)!} \\ &= \sum_{p=l}^m v_p \frac{p!}{l!} \binom{p}{l} \sum_{q=0}^{p-l} (-1)^q \binom{p-l}{q} \\ &= v_l, \end{aligned} \quad (109)$$

where we used Eq. (108) in the first line, the fact that $\binom{p}{k} = 0$ if $k > p$ in the third line, $q := k - l$ in the fifth line, and the binomial theorem in the last line which imposes $p = l$.

(ii) \Rightarrow (i): suppose that

$$\forall l \in \llbracket 0, m \rrbracket, \quad v_l = \frac{(-1)^l}{l!} \sum_{p=l}^m (-1)^p \binom{p}{l} u_p. \quad (110)$$

Then, for all $k \in \llbracket 0, m \rrbracket$,

$$\begin{aligned}
\sum_{l=0}^m v_l \binom{l}{k} l! &= \sum_{l=0}^m \frac{(-1)^l}{l!} \sum_{p=l}^m (-1)^p \binom{p}{l} u_p \binom{l}{k} l! \\
&= \sum_{p=0}^m u_p (-1)^p \sum_{l=0}^p (-1)^l \binom{p}{l} \binom{l}{k} \\
&= \sum_{p=k}^m u_p (-1)^p \sum_{l=k}^p (-1)^l \binom{p}{l} \binom{l}{k} \\
&= \sum_{p=k}^m u_p (-1)^p \sum_{l=k}^p (-1)^l \frac{p! l!}{l! (p-l)! k! (l-k)!} \\
&= \sum_{p=k}^m u_p (-1)^{p-k} \binom{p}{k} \sum_{q=0}^{p-k} (-1)^q \binom{p-k}{q} \\
&= u_k,
\end{aligned} \tag{111}$$

where we used Eq. (110) in the first line, the fact that $\binom{l}{k} = 0$ if $k > l$ in the third line, $q := l - k$ in the fifth line, and the binomial theorem in the last line which imposes $p = k$. \square

Combining Lemma 9 for $u_k = \mu_k$ and $v_l = \sum_{i+j=2l} Q_{ij}$ for all $k, l \in \llbracket 0, m \rrbracket$ with the previous expression of (D-SDP $_n^{m, \geq}$) we finally obtain:

$$\text{(D-SDP}_n^{m, \geq}) \quad \left\{ \begin{array}{l} \inf_{\substack{Q \in \text{Sym}_{m+1} \\ y, \mu \in \mathbb{R} \times \mathbb{R}^{m+1}}} y \\ \text{subject to } y \geq 1 + \mu_n \\ \text{and } \forall k \in \llbracket 0, m \rrbracket \setminus \{n\}, \quad y \geq \mu_k \\ \text{and } \forall l \in \llbracket 0, m \rrbracket, \quad \sum_{i+j=2l} Q_{ij} = \frac{(-1)^l}{l!} \sum_{k=l}^m (-1)^k \binom{k}{l} \mu_k \\ \text{and } Q \succeq 0. \end{array} \right.$$

Note that the constraint $y \geq 1 + \mu_n$ implies the constraint $y \geq \mu_n$.

D.2 Dual program for the semidefinite restrictions

The derivation is analogous to that of the previous section. We fix $m \geq n$ and we recall below the expression of (SDP $_n^{m, \leq}$):

$$\text{(SDP}_n^{m, \leq}) \quad \left\{ \begin{array}{l} \sup_{\substack{Q \in \text{Sym}_{m+1} \\ \mathbf{F} \in \mathbb{R}^{m+1}}} F_n \\ \text{subject to } \sum_{k=0}^m F_k = 1 \\ \text{and } \forall k \leq m, \quad F_k \geq 0 \\ \text{and } \forall l \in \llbracket 1, m \rrbracket, \quad \sum_{i+j=2l-1} Q_{ij} = 0 \\ \text{and } \forall l \leq m, \quad \sum_{i+j=2l} Q_{ij} = \frac{(-1)^l}{l!} \sum_{k=l}^m (-1)^k \binom{k}{l} F_k \\ \text{and } Q \succeq 0. \end{array} \right.$$

To put $(\text{SDP}_n^{m,\leq})$ in the standard form (SDP) we set $N = 2 \times (m + 1)$ and $M = 1 + m + (m + 1)$. For all $r \in \mathbb{N}^*$ and all $i, j \in \llbracket 1, r \rrbracket$, recall that $E_r^{(i,j)}$ denotes the $r \times r$ matrix whose (i, j) entry is 1 and all other entries are 0. We set

$$\begin{aligned}
X &= \text{Diag}_{k=0,\dots,m}(F_k) \oplus Q \in \text{Sym}_N, \\
C &= E_N^{(n,n)} = E_{m+1}^{(n,n)} \oplus \mathbf{0}_{m+1} \in \text{Sym}_N, \\
\mathbf{b} &= (1, 0, 0, \dots, 0) \in \mathbb{R}^M, \\
B'^{(0)} &= \mathbf{1}_{m+1} \oplus \mathbf{0}_{m+1} \in \text{Sym}_N, \\
\forall l \in \llbracket 1, m \rrbracket, \quad B'^{(l)} &= \mathbf{0}_{m+1} \oplus \left(\sum_{i+j=2l-1} E_{m+1}^{(i,j)} \right), \\
\forall l \in \llbracket 0, m \rrbracket, \quad B^{(l)} &= \text{Diag}_{k=0,\dots,m} \left(-\frac{(-1)^{k+l}}{l!} \binom{k}{l} \right) \oplus \left(\sum_{i+j=2l} E_{m+1}^{(i,j)} \right),
\end{aligned} \tag{112}$$

with the convention $\binom{k}{l} = 0$ when $l > k$. The matrix $B'^{(0)}$ corresponds to the constraint $\sum_{k=0}^m F_k = 1$ and we denote the corresponding dual variable $y \in \mathbb{R}$. Similarly, the matrices $B'^{(l)}$ correspond to the m constraints $\sum_{i+j=2l-1} Q_{ij} = 0$, and we denote the corresponding dual variables $\nu'_l \in \mathbb{R}$. Finally, the matrices $B^{(l)}$ correspond to the $m + 1$ constraints $\sum_{i+j=2l} Q_{ij} = \frac{(-1)^l}{l!} \sum_{k=l}^m (-1)^k \binom{k}{l} F_k$, and we denote the corresponding dual variables $\nu_l \in \mathbb{R}$.

The standard form (D-SDP) of the dual program $(\text{D-SDP}_n^{m,\leq})$ thus reads:

$$(\text{D-SDP}_n^{m,\leq}) \quad \left\{ \begin{array}{l} \inf_{\substack{\nu, \nu' \in \mathbb{R}^{m+1} \times \mathbb{R}^m \\ y \in \mathbb{R}}} y \\ \text{subject to} \quad \text{Diag}_{k=0,\dots,m} \left[y - \sum_{l=0}^k \nu_l \frac{(-1)^{k+l}}{l!} \binom{k}{l} \right] \\ \oplus \left(\sum_{l=0}^m \sum_{i+j=2l} \nu_l E_{m+1}^{(i,j)} + \sum_{l=1}^m \sum_{i+j=2l-1} \nu'_l E_{m+1}^{(i,j)} \right) \succeq E_{m+1}^{(n,n)} \oplus \mathbf{0}_{m+1}. \end{array} \right.$$

Due to the block-diagonal structure of the matrices involved, the positive semidefinite constraint above is equivalent to the following constraints:

$$\left\{ \begin{array}{l} y \geq 1 + \sum_{l=0}^n \nu_l \frac{(-1)^{n+l}}{l!} \binom{n}{l}, \\ \forall k \in \llbracket 0, m \rrbracket \setminus \{n\}, \quad y \geq \sum_{l=0}^k \nu_l \frac{(-1)^{k+l}}{l!} \binom{k}{l}, \\ \left(\sum_{l=0}^m \sum_{i+j=2l} \nu_l E_{m+1}^{(i,j)} + \sum_{l=1}^m \sum_{i+j=2l-1} \nu'_l E_{m+1}^{(i,j)} \right) \succeq 0. \end{array} \right. \tag{113}$$

Let us define $A = (A_{ij})_{0 \leq i, j \leq m}$ by

$$A := \sum_{l=0}^m \sum_{i+j=2l} \nu_l E_{m+1}^{(i,j)} + \sum_{l=1}^m \sum_{i+j=2l-1} \nu'_l E_{m+1}^{(i,j)}, \tag{114}$$

or equivalently

$$A_{ij} = \begin{cases} \nu_l & \text{when } i + j = 2l, \\ \nu'_l & \text{when } i + j = 2l - 1, \end{cases} \tag{115}$$

for all $i, j \in \llbracket 0, m \rrbracket$. For $k \in \llbracket 0, m \rrbracket$, we also define

$$\mu_k := \sum_{l=0}^k \nu_l \frac{(-1)^{k+l}}{l!} \binom{k}{l} \in \mathbb{R}. \quad (116)$$

By Lemma 8, the following conditions are equivalent:

- (i) $\forall k \in \llbracket 0, m \rrbracket, \quad \mu_k = \sum_{l=0}^k \nu_l \frac{(-1)^{k+l}}{l!} \binom{k}{l},$
- (ii) $\forall l \in \llbracket 0, m \rrbracket, \quad \nu_l = \sum_{k=0}^l \mu_k \binom{l}{k} l!.$

With Eq. (115) we thus have

$$A_{ij} = \sum_{k=0}^l \mu_k \binom{l}{k} l! \quad \text{when } i + j = 2l, \quad (117)$$

and we obtain the following expression for (D-SDP $_n^{m, \leq}$):

$$(D\text{-SDP}_n^{m, \leq}) \quad \left\{ \begin{array}{l} \inf_{\substack{A \in \text{Sym}_{m+1} \\ y, \mu \in \mathbb{R} \times \mathbb{R}^{m+1}}} y \\ \text{subject to } y \geq 1 + \mu_n \\ \text{and } \forall k \in \llbracket 0, m \rrbracket \setminus \{n\}, \quad y \geq \mu_k \\ \text{and } \forall l \leq m, \forall i + j = 2l, \quad A_{ij} = \sum_{k=0}^l \mu_k \binom{l}{k} l! \\ \text{and } A \succeq 0. \end{array} \right.$$

Note that the constraint $y \geq 1 + \mu_n$ implies the constraint $y \geq \mu_n$.

E Proof of Lemma 6

In this section, we prove the following result:

Lemma 6. *For all $m \geq n$, the program (D-SDP $_n^{m, \leq}$) is equivalent to the following program:*

$$(D\text{-SDP}_n^{m, \leq}) \quad \left\{ \begin{array}{l} \inf_{\substack{y \in \mathbb{R} \\ \mu \in \mathcal{S}'(\mathbb{N})}} y \\ \text{subject to } y \geq 1 + \mu_n \\ \text{and } \forall k \neq n \in \mathbb{N}, \quad y \geq \mu_k \\ \text{and } \forall g \in \mathcal{R}_{m,+}(\mathbb{R}_+), \quad \langle f_\mu, g \rangle \geq 0, \end{array} \right.$$

where $f_\mu = \sum_k \mu_k \mathcal{L}_k$.

Proof. We first obtain a reformulation of (SDP $_n^{m, \leq}$) and we derive its dual program. This reformulation is obtained using Stieltjes characterisation of non-negative polynomials over \mathbb{R}_+ rather than Lemma 4:

Lemma 10 ([65]). *Let $m \in \mathbb{N}$ and let P be a univariate polynomial of degree m . Let $a_1 = \lfloor \frac{m}{2} \rfloor$ and $a_2 = \lfloor \frac{m-1}{2} \rfloor$. For all $q \in \mathbb{N}$, let $X_q = (1, x, \dots, x^q)$ be the vector of univariate monomials up to degree q . Then, P is non-negative over \mathbb{R}_+ if and only if there exist sum of squares polynomials A_1 and A_2 of degree $2a_1$ and $2a_2$, respectively, such that $P(x) = A_1(x) + xA_2(x)$ for all $x \in \mathbb{R}_+$, or equivalently, if and only if there exist real positive semidefinite matrices A_1 and A_2 of size $(a_1 + 1) \times (a_1 + 1)$ and $(a_2 + 1) \times (a_2 + 1)$, respectively, such that for all $x \in \mathbb{R}_+$,*

$$P(x) = X_{a_1}^T A_1 X_{a_1} + x X_{a_2}^T A_2 X_{a_2}. \quad (118)$$

By Lemma 10, the program $(\text{SDP}_n^{m,\leq})$, obtained by imposing $F_k = 0$ for $k > m$ in (LP_n) , is equivalent to the following program:

$$(\text{SDP}_n^{m,\leq}) \left\{ \begin{array}{l} \sup_{\substack{A_1, A_2 \in \text{Sym}_{a_1+1} \times \text{Sym}_{a_2+1} \\ \mathbf{F} \in \mathbb{R}^{m+1}}} F_n \\ \text{subject to } \sum_{k=0}^m F_k = 1 \\ \text{and } \forall k \leq m, F_k \geq 0 \\ \text{and } \forall l \leq m, \frac{(-1)^l}{l!} \sum_{k=l}^m (-1)^k \binom{k}{l} F_k = \sum_{\substack{i+j=l \\ 0 \leq i, j \leq a_1}} (A_1)_{ij} + \sum_{\substack{i+j=l-1 \\ 0 \leq i, j \leq a_2}} (A_2)_{ij} \\ \text{and } A_1 \succeq 0 \\ \text{and } A_2 \succeq 0. \end{array} \right.$$

To put $(\text{SDP}_n^{m,\leq})$ in the standard form (SDP) we set $N = (m+1) + (a_1+1) + (a_2+1)$ and $M = 1 + (m+1)$. For all $r \in \mathbb{N}^*$ and all $i, j \in \llbracket 1, r \rrbracket$, recall that $E_r^{(i,j)}$ denotes the $r \times r$ matrix whose (i, j) entry is 1 and all other entries are 0. We set

$$\begin{aligned} X &= \text{Diag}_{k=0,\dots,m}(F_k) \oplus A_1 \oplus A_2 \in \text{Sym}_N, \\ C &= E_N^{(n,n)} = E_{m+1}^{(n,n)} \oplus \mathbb{0}_{a_1+1} \oplus \mathbb{0}_{a_2+1} \in \text{Sym}_N, \\ \mathbf{b} &= (1, 0, 0, \dots, 0) \in \mathbb{R}^M, \\ B^{(-1)} &= \mathbb{1}_{m+1} \oplus \mathbb{0}_{a_1+1} \oplus \mathbb{0}_{a_2+1} \in \text{Sym}_N, \end{aligned} \tag{119}$$

$$\forall l \in \llbracket 0, m \rrbracket, \quad B^{(l)} = \text{Diag}_{k=0,\dots,m} \left(-\frac{(-1)^{k+l}}{l!} \binom{k}{l} \right) \oplus \left(\sum_{\substack{i+j=l \\ 0 \leq i, j \leq a_1}} E_{a_1+1}^{(i,j)} \right) \oplus \left(\sum_{\substack{i+j=l-1 \\ 0 \leq i, j \leq a_2}} E_{a_2+1}^{(i,j)} \right),$$

with the convention $\binom{k}{l} = 0$ when $l > k$. The matrix $B^{(-1)}$ corresponds to the constraint $\sum_{k=0}^m F_k = 1$ and we denote the corresponding dual variable $y \in \mathbb{R}$. Similarly, the matrices $B^{(l)}$ correspond to the $m+1$ other linear constraints, and we denote the corresponding dual variable $\nu \in \mathbb{R}^{m+1}$. The standard form (D-SDP) of the dual program $(\text{D-SDP}_n^{m,\leq})$ thus reads:

$$(\text{D-SDP}_n^{m,\leq}) \left\{ \begin{array}{l} \inf_{\substack{\nu \in \mathbb{R}^{m+1} \\ y \in \mathbb{R}}} y \\ \text{subject to } \text{Diag}_{k=0,\dots,m} \left[y - \sum_{l=0}^k \nu_l \frac{(-1)^{k+l}}{l!} \binom{k}{l} \right] \\ \oplus \left(\sum_{l=0}^m \sum_{\substack{i+j=l \\ 0 \leq i, j \leq a_1}} \nu_l E_{a_1+1}^{(i,j)} \right) \\ \oplus \left(\sum_{l=1}^m \sum_{\substack{i+j=l-1 \\ 0 \leq i, j \leq a_2}} \nu_l E_{a_2+1}^{(i,j)} \right) \succeq E_{m+1}^{(n,n)} \oplus \mathbb{0}_{a_1} \oplus \mathbb{0}_{a_2}. \end{array} \right.$$

Due to the block-diagonal structure of the matrices involved, the positive semidefinite constraint above

is equivalent to the following constraints:

$$\begin{cases} y \geq 1 + \sum_{l=0}^n \nu_l \frac{(-1)^{n+l}}{l!} \binom{n}{l}, \\ \forall k \in \llbracket 0, m \rrbracket \setminus \{n\}, \quad y \geq \sum_{l=0}^k \nu_l \frac{(-1)^{k+l}}{l!} \binom{k}{l}, \\ \sum_{l=0}^m \sum_{\substack{i+j=l \\ 0 \leq i, j \leq a_1}} \nu_l E_{a_1+1}^{(i,j)} \succeq 0 \\ \sum_{l=1}^m \sum_{\substack{i+j=l-1 \\ 0 \leq i, j \leq a_2}} \nu_l E_{a_2+1}^{(i,j)} \succeq 0. \end{cases} \quad (120)$$

For $k \in \llbracket 0, m \rrbracket$, we define

$$\mu_k := \sum_{l=0}^k \nu_l \frac{(-1)^{k+l}}{l!} \binom{k}{l} \in \mathbb{R}. \quad (121)$$

By Lemma 8, the following conditions are equivalent:

- (i) $\forall k \in \llbracket 0, m \rrbracket, \quad \mu_k = \sum_{l=0}^k \nu_l \frac{(-1)^{k+l}}{l!} \binom{k}{l},$
- (ii) $\forall l \in \llbracket 0, m \rrbracket, \quad \nu_l = \sum_{k=0}^l \mu_k \binom{l}{k} l!.$

We thus have

$$\forall l \in \llbracket 0, m \rrbracket, \quad \nu_l = \sum_{k=0}^l \mu_k \binom{l}{k} l! \quad (122)$$

Let us introduce the moment matrices $M_\nu \in \text{Sym}_{a_1+1}$ and $M_\nu^{(1)} \in \text{Sym}_{a_2+1}$:

$$\forall i, j \in \llbracket 0, a_1 \rrbracket, (M_\nu)_{i,j} := \nu_{i+j}, \quad (123)$$

$$\forall i, j \in \llbracket 0, a_2 \rrbracket, (M_\nu^{(1)})_{ij} := \nu_{i+j+1}. \quad (124)$$

The constraints (120) are equivalent to:

$$\begin{cases} y \geq 1 + \mu_n, \\ \forall k \in \llbracket 0, m \rrbracket \setminus \{n\}, \quad y \geq \mu_k, \\ M_\nu \succeq 0 \\ M_\nu^{(1)} \succeq 0. \end{cases} \quad (125)$$

We complete the vector $\boldsymbol{\mu}$ with zeros to obtain an element of $\mathcal{S}'(\mathbb{N})$. We have $f_\boldsymbol{\mu} = \sum_k \mu_k \mathcal{L}_k \in \mathcal{S}'(\mathbb{R}_+)$. We prove the following result, analogous to Lemma 5:

Lemma 11. *The following propositions are equivalent:*

- (i) $\forall g \in \mathcal{R}_{m,+}(\mathbb{R}_+), \langle f_\boldsymbol{\mu}, g \rangle \geq 0,$
- (ii) $M_\nu \succeq 0$ and $M_\nu^{(1)} \succeq 0.$

Proof. The proof is similar to that of Lemma 5: for all $k \in \mathbb{N}$ we have

$$\mu_k = \int_{\mathbb{R}_+} \mathcal{L}_k(x) f_\boldsymbol{\mu}(x) dx. \quad (126)$$

Thus,

$$\begin{aligned}
M_\nu \succeq 0 &\Leftrightarrow \forall Y \in \mathbb{R}^{a_1+1}, Y^T M_\nu Y \geq 0 \\
&\Leftrightarrow \forall Y \in \mathbb{R}^{a_1+1}, \sum_{i,j=0}^{a_1} y_i y_j (M_\nu)_{ij} \geq 0 \\
&\Leftrightarrow \forall Y \in \mathbb{R}^{a_1+1}, \sum_{l=0}^{2a_1} \sum_{i+j=l} y_i y_j \nu_l \geq 0 \\
&\Leftrightarrow \forall Y \in \mathbb{R}^{a_1+1}, \sum_{l=0}^{2a_1} \sum_{i+j=l} y_i y_j \sum_{k=0}^l \mu_k \binom{l}{k} l! \geq 0 \\
&\Leftrightarrow \forall Y \in \mathbb{R}^{a_1+1}, \int_{\mathbb{R}_+} \sum_{l=0}^{2a_1} \sum_{i+j=l} y_i y_j \sum_{k=0}^l \binom{l}{k} l! \mathcal{L}_k(x) f_\mu(x) dx \geq 0 \\
&\Leftrightarrow \forall Y \in \mathbb{R}^{a_1+1}, \int_{\mathbb{R}_+} \sum_{l=0}^{2a_1} \sum_{i+j=l} y_i y_j \sum_{k=0}^l (-1)^k \binom{l}{k} l! L_k(x) e^{-\frac{x}{2}} f_\mu(x) dx \geq 0 \\
&\Leftrightarrow \forall Y \in \mathbb{R}^{a_1+1}, \int_{\mathbb{R}_+} e^{-\frac{x}{2}} \sum_{l=0}^{2a_1} x^l \sum_{i+j=l} y_i y_j f_\mu(x) dx \geq 0 \\
&\Leftrightarrow \forall Y \in \mathbb{R}^{a_1+1}, \int_{\mathbb{R}_+} e^{-\frac{x}{2}} \left(\sum_{k=0}^{a_1} y_k x^k \right)^2 f_\mu(x) dx \geq 0 \\
&\Leftrightarrow \forall Y \in \mathbb{R}^{a_1+1}, \left\langle f_\mu, x \mapsto e^{-\frac{x}{2}} \left(\sum_{k=0}^{a_1} y_k x^k \right)^2 \right\rangle \geq 0,
\end{aligned} \tag{127}$$

where we used Eq. (123) in the third line, Eq. (122) in the fourth line, Eq. (126) in the fifth line and Eq. (79) in the seventh line. Similarly,

$$M_\nu^{(1)} \succeq 0 \Leftrightarrow \forall Y \in \mathbb{R}^{a_2+1}, \left\langle f_\mu, x \mapsto e^{-\frac{x}{2}} x \left(\sum_{k=0}^{a_2} y_k x^k \right)^2 \right\rangle \geq 0. \tag{128}$$

Combining Eq. (127) and Eq. (128) with Lemma 10 we obtain

$$M_\nu \succeq 0 \text{ and } M_\nu^{(1)} \succeq 0 \Leftrightarrow \forall g \in \mathcal{R}_{m,+}(\mathbb{R}_+), \langle f_\mu, g \rangle \geq 0, \tag{129}$$

by linearity. □

Combining Lemma 11 with the constraints (125) finally yields:

$$(\text{D-SDP}_n^{m,\leq}) \begin{cases} \inf_{\substack{y \in \mathbb{R} \\ \mu \in \mathcal{S}'(\mathbb{N})}} y \\ \text{subject to } \forall k \neq n \in \mathbb{N}, y \geq \mu_k \\ \text{and } y \geq 1 + \mu_n \\ \text{and } \forall g \in \mathcal{R}_{m,+}(\mathbb{R}_+), \langle f_\mu, g \rangle \geq 0. \end{cases}$$
□

F Proof of Lemma 7

We recall the definition of $\mathbf{F}^n = (F_k^n)_{k \in \mathbb{N}} \in \mathbb{R}^{\mathbb{N}}$:

- if n is even:

$$F_k^n := \begin{cases} \frac{1}{2^n} \binom{k}{\frac{k}{2}} \binom{n-k}{\frac{n-k}{2}} & \text{when } k \leq n, k \text{ even,} \\ 0 & \text{otherwise,} \end{cases} \quad (130)$$

- if n is odd:

$$F_k^n := \begin{cases} \frac{1}{2^n} \frac{\binom{\lfloor \frac{n}{2} \rfloor}{\lfloor \frac{k}{2} \rfloor} \binom{\lfloor \frac{n}{2} \rfloor}{\lfloor \frac{k}{2} \rfloor}}{\binom{n}{k}}, & \text{when } k \leq n, \\ 0 & \text{otherwise.} \end{cases} \quad (131)$$

In both cases,

$$F_n^n = \frac{1}{2^n} \binom{n}{\lfloor \frac{n}{2} \rfloor}. \quad (132)$$

We extrapolated these analytical expressions from numerical values. Running $(\text{SDP}_n^{m, \leq})$ for several values of n and m allowed us to deduce these sequences (we acknowledge here the great help from oeis.org).

We start by showing two results, corresponding to n even and n odd, respectively, where we make use of Zeilberger's algorithm, a powerful algorithm for proving binomial identities [70]. Given a holonomic function, this algorithm outputs a recurrence relation that it satisfies, thus reducing the proof of identity between binomial expressions to the verification that the initialisation is correct. A Mathematica notebook is available for the implementation of Zeilberger's algorithm [53].

Lemma 12. For $n \in \mathbb{N}$ even:

$$\sum_{k=0}^n (-1)^k F_k^n L_k(x) = \sum_{l=0}^n x^l \sum_{i+j=2n-2l} p_{n-i} p_{n-j}, \quad (133)$$

where:

$$p_{n-k} := \begin{cases} \sqrt{\frac{1}{2^n n!} \binom{n}{\frac{n}{2}}}, & \text{when } k = 0, \\ (-1)^{\frac{k}{2}} 2^{\frac{k}{2}} \left(\frac{k}{2}\right)! \left(\frac{n}{2}\right)^2 p_n, & \text{when } 0 < k \leq n, k \text{ even,} \\ 0, & \text{otherwise.} \end{cases} \quad (134)$$

The coefficients p_{n-k} (and q_{n-k} later on) were found by hand when looking for an analytical sum of squares decomposition.

Proof. To prove the polynomial equality (133), we start by equating the coefficients in x^l for $l \in \llbracket 0, n \rrbracket$ which gives:

$$\begin{aligned} \frac{(-1)^l}{l!} \sum_{k=l}^n (-1)^k \binom{k}{l} F_k^n &= \sum_{i+j=2n-2l} p_{n-i} p_{n-j} \\ &= \sum_{i=0}^{2n-2l} p_{n-i} p_{n-(2n-2l-i)} \\ &= \sum_{i=0}^{n-l} p_{n-2i} p_{n-(2n-2l-2i)}, \end{aligned} \quad (135)$$

where we used $p_{n-i} = 0$ for $n-i < 0$ in the second line. These are equalities between holonomic functions of parameters n and l that are trivial for $l > n$.

- For $l \leq n$ even, because $F_k^n = 0$ for k odd, Eq. (135) becomes:

$$\sum_{k=\frac{l}{2}}^{\frac{n}{2}} \frac{1}{l!} \binom{2k}{l} F_{2k}^n = \sum_{i=0}^{n-l} p_{n-2i} p_{n-(2n-2l-2i)}, \quad (136)$$

that is, taking into account the parity of $l = 2s$ and $n = 2t$,

$$\sum_{k=s}^t \frac{1}{(2s)!} \binom{2k}{2s} F_{2k}^{2t} = \sum_{i=0}^{2t-2s} p_{2t-2i} p_{2t-(4t-4s-2i)}. \quad (137)$$

Inserting the expressions from Eq. (130) and Eq. (134), we thus have to check the identity:

$$\sum_{k=s}^t \frac{1}{2^{2t}(2s)!} \binom{2k}{2s} \binom{2k}{k} \binom{2t-2k}{t-k} = \sum_{i=0}^{2t-2s} \frac{i!(2t-2s-i)!}{2^{2s}(2t)!} \binom{2t}{t} \binom{t}{i}^2 \binom{t}{2t-2s-i}^2, \quad (138)$$

for all $t \in \mathbb{N}$ and all $s \leq t$ (with the convention $\binom{k}{j} = 0$ for $j > k$). We ran Zeilberger's algorithm to show that the right-hand side and the left-hand side of Eq. (137) both satisfy the following recurrence relation, for all $s, t \in \mathbb{N}$:

$$2(t+1)^2 S(s, t) + (-2s^2 - 4t^2 + 4st + 5s - 11t - 8)S(s, t+1) + (s-t-2)(2s-2t-3)S(s, t+2) = 0. \quad (139)$$

It remains to check that the initialisation is correct. For all $s \in \mathbb{N}$, this recurrence relation in t is of order 2. Since the identities in Eq. (138) are trivial when $l > n$, i.e., $s > t$, we thus only need to check Eq. (138) for $(s, t) = (0, 0)$, $(s, t) = (0, 1)$, and $(s, t) = (1, 1)$, which is straightforward: we obtain the values 1, 1 and $\frac{1}{4}$ respectively, for both sides of Eq. (138).

- For $l \leq n$ odd, Eq. (135) becomes:

$$- \sum_{k=\frac{l+1}{2}}^{\frac{n}{2}} \frac{1}{l!} \binom{2k}{l} F_{2k}^n = \sum_{i=0}^{n-l} p_{n-2i} p_{n-(2n-2l-2i)}, \quad (140)$$

that is, taking into account the parity of $l = 2s + 1$ and $n = 2t$:

$$- \sum_{k=s+1}^t \frac{1}{(2s+1)!} \binom{2k}{2s+1} F_{2k}^{2t} = \sum_{i=0}^{2t-2s-1} p_{2t-2i} p_{2t-(4t-4s-2i-2)}. \quad (141)$$

Inserting the expressions from Eq. (130) and Eq. (134), we thus have to check the identity:

$$\begin{aligned} \sum_{k=s+1}^t \frac{1}{2^{2t}(2s+1)!} \binom{2k}{2s+1} \binom{2k}{k} \binom{2t-2k}{t-k} \\ = \sum_{i=0}^{2t-2s-1} \frac{i!(2t-2s-i-1)!}{2^{2s+1}(2t)!} \binom{2t}{t} \binom{t}{i}^2 \binom{t}{2t-2s-i-1}^2, \end{aligned} \quad (142)$$

for all $t \in \mathbb{N}$ and all $s \leq t$ (with the convention $\binom{k}{j} = 0$ for $j > k$). Likewise, we ran Zeilberger's algorithm to show that the right-hand side and the left-hand side of Eq. (141) both satisfy the following recurrence relation, for all $s, t \in \mathbb{N}$:

$$2(t+1)^2 S(s, t) + (-2s^2 - 4t^2 + 4st + 3s - 9t - 6)S(s, t+1) + (s-t-1)(2s-2t-3)S(s, t+2) = 0. \quad (143)$$

It remains to check that the initialisation is correct. For all $s \in \mathbb{N}$, this recurrence relation in t is of order 2. Since the identities in Eq. (142) are trivial when $l > n$, i.e., $s \geq t$, we thus only need to check Eq. (142) for $(s, t) = (0, 1)$, which is straightforward: we obtain the value 1 for both sides of Eq. (142).

Lemma 13. For $n \in \mathbb{N}$ odd:

$$\sum_{k=0}^n (-1)^k F_k^n L_k(x) = \sum_{l=0}^n x^l \sum_{i+j=2l} q_i q_j \quad (144)$$

where:

$$q_{n-k} := \begin{cases} \sqrt{\frac{1}{2^n n!} \binom{n}{\lfloor \frac{n}{2} \rfloor}}, & \text{when } k = 0, \\ (-1)^{\frac{k}{2}} 2^{\frac{k}{2}} \left(\frac{k}{2}\right)! \frac{n+1}{n-k+1} \left(\frac{\lfloor \frac{n}{2} \rfloor}{\frac{k}{2}}\right)^2 q_n, & \text{when } 0 < k < n, k \text{ even,} \\ 0. & \text{otherwise.} \end{cases} \quad (145)$$

Proof. Unlike the case where n is even, F^n is non-zero for $k \leq n$ odd, and the expression of F_k^n depends on the parity of k . Thus we cannot use directly Eq. (131) in Zeilberger's algorithm as we did in the previous lemma, hence the development below in order to obtain expressions that the algorithm can take as inputs. We fix n odd and $l \leq n$.

We start by equating coefficients in x^l in Eq. (144):

$$\begin{aligned} \frac{(-1)^l}{l!} \sum_{k=l}^n (-1)^k \binom{k}{l} F_k^n &= \sum_{i+j=2n-2l} q_{n-i} q_{n-j} \\ &= \sum_{i=0}^{n-l} q_{n-2i} q_{n-(2n-2l-2i)}. \end{aligned} \quad (146)$$

- For l even, writing $l = 2s$ and $n = 2t + 1$, the left-hand side of Eq. (146) becomes:

$$\begin{aligned} &\frac{1}{(2s)!} \sum_{k=2s}^{2t+1} (-1)^k \binom{k}{2s} F_k^{2t+1} \\ &= \frac{1}{(2s)!} \sum_{k=0}^{2t+1-2s} (-1)^k \binom{k+2s}{2s} F_{k+2s}^{2t+1} \\ &= \frac{1}{(2s)!} \sum_{k=0}^{t-s} \left(\binom{2k+2s}{2s} F_{2k+2s}^{2t+1} - \binom{2k+2s+1}{2s} F_{2k+2s+1}^{2t+1} \right) \\ &= \frac{\binom{2t+1}{t}}{2^{2t+1} (2s)!} \sum_{k=0}^{t-s} \left(\frac{\binom{2k+2s}{2s} \binom{t}{k+s}^2}{\binom{2t+1}{2k+2s}} - \frac{\binom{2k+2s+1}{2s} \binom{t}{k+s}^2}{\binom{2t+1}{2k+2s+1}} \right) \\ &= q_{2t+1}^2 \frac{(2t+1)!}{(2s)!} \sum_{k=0}^{t-s} \frac{\binom{t}{k+s}^2 \binom{2k+2s}{2s}}{\binom{2t+1}{2k+2s}} \left(1 - \frac{(2k+2s+1)^2}{(2k+1)(2t-2s-2k+1)} \right). \end{aligned} \quad (147)$$

With Eq. (145) and Eq. (146), we thus have to check the identity:

$$\begin{aligned} &\frac{(2t+1)!}{(2s)!} \sum_{k=0}^{t-s} \frac{\binom{t}{k+s}^2 \binom{2k+2s}{2s}}{\binom{2t+1}{2k+2s}} \left(1 - \frac{(2k+2s+1)^2}{(2k+1)(2t-2s-2k+1)} \right) \\ &= \sum_{i=0}^{2t+1-2s} \frac{q_{2t+1-2i} q_{2t+1-(4t+2-4s-2i)}}{q_{2t+1}^2} \\ &= - \sum_{i=0}^{2t+1-2s} 2^{2t-2s+1} \frac{(2t+2)^2 i! (2t+1-2s-i)!}{(2t-2i+2)(4s+2i-2t)} \binom{t}{i}^2 \binom{t}{2t-2s-i+1}^2, \end{aligned} \quad (148)$$

for all $t \in \mathbb{N}$ and all $s \leq t$ (with the convention $\binom{k}{j} = 0$ for $j > k$). Zeilberger's algorithm certifies that the right-hand side and the left-hand side of Eq. (148) both satisfy for all $s, t \in \mathbb{N}$:

$$\begin{aligned} & -32(t+2)^3(t+1)^2(t+3)S(s,t) \\ & + 4(t+3)(t+2)(2s^2+4t^2-4st-7s+15t+14)S(s,t+1) \\ & + (-2s+2t+5)(s-t-2)S(s,t+2) = 0. \end{aligned} \quad (149)$$

It remains to check that the initialisation is correct. For all $s \in \mathbb{N}$, this recurrence relation in t is of order 2. Since the identities in Eq. (148) are trivial when $l > n$, i.e., $s > t$, we thus only need to check Eq. (148) for $(s,t) = (0,0)$, $(s,t) = (0,1)$, and $(s,t) = (1,1)$, which is straightforward: we obtain the values 0, 0 and -8 respectively, for both sides of Eq. (148).

- For l odd, writing $l = 2s + 1$ and $n = 2t + 1$, the left-hand side of Eq. (146) becomes:

$$\begin{aligned} & \frac{-1}{(2s+1)!} \sum_{k=2s+1}^{2t+1} (-1)^k \binom{k}{2s+1} F_k^{2t+1} \\ & = \frac{-1}{(2s+1)!} \sum_{k=0}^{2t-2s} (-1)^{k+2s+1} \binom{k+2s+1}{2s+1} F_{k+2s+1}^{2t+1} \\ & = \frac{1}{(2s+1)!} \sum_{k=0}^{t-s} \left(-\binom{2k+2s}{2s+1} F_{2k+2s}^{2t+1} + \binom{2k+2s+1}{2s+1} F_{2k+2s+1}^{2t+1} \right) \\ & = q_{2t+1}^2 \frac{(2t+1)!}{(2s+1)!} \sum_{k=0}^{t-s} \left(-\frac{\binom{2k+2s}{2s+1} \binom{t}{k+s}^2}{\binom{2t+1}{2k+2s}} + \frac{\binom{2k+2s+1}{2s+1} \binom{t}{k+s}^2}{\binom{2t+1}{2k+2s+1}} \right) \\ & = q_{2t+1}^2 \frac{(2t+1)!}{(2s+1)!} \sum_{k=0}^{t-s} \frac{\binom{2k+2s}{2s+1} \binom{t}{k+s}^2}{\binom{2t+1}{2k+2s}} \left(-1 + \frac{(2k+2s+1)^2}{2k(2t-2k-2s+1)} \right), \end{aligned} \quad (150)$$

where we used that $\binom{2k+2s}{2s+1} = 0$ for $k = 0$ in the third line. Note that when factorising we introduced an indeterminate form in the last line that Zeilberger's algorithm can resolve. This is necessary since the algorithm cannot deal with differences of binomial terms. With Eq. (145) and Eq. (146), we thus have to check the identity:

$$\begin{aligned} & \frac{(2t+1)!}{(2s+1)!} \sum_{k=0}^{t-s} \frac{\binom{2k+2s}{2s+1} \binom{t}{k+s}^2}{\binom{2t+1}{2k+2s}} \left(-1 + \frac{(2k+2s+1)^2}{2k(2t-2k-2s+1)} \right) \\ & = \sum_{i=0}^{2t-2s} \frac{q_{2t+1-2i} q_{2t+1-(4t-4s-2i)}}{q_{2t+1}^2} \\ & = \sum_{i=0}^{2t-2s} 2^{2t-2s} \frac{(2t+2)^2 i! (2t-2s-i)!}{(2t-2i+2)(4s+2i-2t+2)} \binom{t}{i}^2 \binom{t}{2t-2s-i}, \end{aligned} \quad (151)$$

for all $t \in \mathbb{N}$ and all $s \leq t$ (with the convention $\binom{k}{j} = 0$ for $j > k$). Zeilberger's algorithm certifies that both the right-hand side and the left-hand side of Eq. (151) satisfy for all $s \leq t$:

$$\begin{aligned} & -32(t+2)^3(t+1)^2(t+3)S(s,t) \\ & + 4(t+3)(t+2)(2s^2+4t^2-4st-5s+13t+11)S(s,t+1) \\ & + (-2s+2t+3)(s-t-2)S(s,t+2) = 0. \end{aligned} \quad (152)$$

It remains to check that the initialisation is correct. For all $s \in \mathbb{N}$, this recurrence relation in t is of order 2. Since the identities in Eq. (151) are trivial when $l > n$, i.e., $s > t$, we thus only need to check Eq. (151) for $(s,t) = (0,0)$, $(s,t) = (0,1)$, and $(s,t) = (1,1)$, which is straightforward: we obtain the values 1, 16 and 1 respectively, for both sides of Eq. (151).

□

Having derived these identities, we now recall the lemma from the main text we wish to prove:

Lemma 7. *For all $m \geq n$, \mathbf{F}^n is a feasible solution of $(\text{SDP}_n^{m,\leq})$. Moreover, it is optimal when $m = n$.*

Proof. The proof has three parts. In the first we focus on n even, and in the second on n odd. In the last part, we exhibit a feasible solution of $(\text{D-SDP}_n^{m,\leq})$ for $m = n$ with the same optimal value F_n^n .

n even, $m \geq n$: we check that \mathbf{F}^n defined in Eq. (130) satisfies all the constraints of $(\text{SDP}_n^{m,\leq})$.

- For all $k \in \mathbb{N}$, $F_k^n \geq 0$.
- We have

$$\begin{aligned} \sum_{k=0}^{\infty} F_k^n &= \sum_{\substack{k=0 \\ \text{even}}}^n \frac{1}{2^n} \binom{k}{\frac{k}{2}} \binom{n-k}{\frac{n-k}{2}} \\ &= \sum_{k=0}^{\frac{n}{2}} \frac{1}{2^n} \binom{2k}{k} \binom{n-2k}{\frac{n}{2}-k} \\ &= 1, \end{aligned} \tag{153}$$

where the last equality follows from [69, (3.90)].

- We have to show that $x \mapsto \sum_{k=0}^n (-1)^k F_k^n L_k(x^2)$ is a positive function on \mathbb{R} . Due to Lemma 3, we aim to find a sum of squares decomposition for this polynomial. Guided by numerical results, we look for a polynomial $P(x) := \sum_{i=0}^n p_i x^i$ such that:

$$\begin{aligned} \sum_{k=0}^n (-1)^k F_k^n L_k(x^2) &= P^2(x^2) \\ &= \left(\sum_{i=0}^n p_i x^{2i} \right)^2 \\ &= \sum_{l=0}^n \left(\sum_{\substack{i+j=l \\ 0 \leq i,j \leq n}} p_i p_j \right) x^{2l}, \end{aligned} \tag{154}$$

and the sought coefficients are given by Lemma 12, which concludes the first part of the proof.

n odd, $m \geq n$: similarly, we check that \mathbf{F}^n defined in Eq. (131) satisfies all the constraints of $(\text{SDP}_n^{m,\leq})$.

- For all $k \in \mathbb{N}$, $F_k^n \geq 0$.
- Writing $n = 2t + 1$, from Eq. (131) we have

$$\begin{aligned} \forall s \leq t, F_{2s}^{2t+1} &= \frac{1}{2^{2t+2}} \left(1 - \frac{s}{t+1} \right) \binom{2s}{s} \binom{2t+2-2s}{t+1-s} \\ \forall s \in \llbracket 1, t+1 \rrbracket, F_{2s-1}^{2t+1} &= \frac{1}{2^{2t+2}} \frac{s}{t+1} \binom{2s}{s} \binom{2t+2-2s}{t+1-s}. \end{aligned} \tag{155}$$

In particular, for all $s \leq t$ we have

$$F_{2s-1}^{2t+1} + F_{2s}^{2t+1} = F_{2s}^{2t+2}, \tag{156}$$

where F_{2s}^{2t+2} is defined in Eq. (130). Since $F_{2s+1}^{2t+2} = 0$ for all $s \leq t$, and $F_k^n = 0$ for all $k > n$, we have:

$$\begin{aligned} \sum_{k=0}^{\infty} F_k^n &= \sum_{k=0}^{\infty} F_k^{n+1} \\ &= 1, \end{aligned} \quad (157)$$

where we used Eq. (153).

- We have to show that $x \mapsto \sum_{k=0}^n (-1)^k F_k^n L_k(x^2)$ is a positive function on \mathbb{R} . Due to Lemma 3, we aim to find a sum of squares decomposition for this polynomial. Guided by numerical results, we look for a polynomial $Q(x) := \sum_{i=0}^n q_i x^i$ such that:

$$\begin{aligned} \sum_{k=0}^n (-1)^k F_k^n L_k(x^2) &= Q^2(x^2) \\ &= \left(\sum_{i=0}^n q_i x^{2i} \right)^2 \\ &= \sum_{l=0}^n \left(\sum_{\substack{i+j=l \\ 0 \leq i, j \leq n}} q_i q_j \right) x^{2l}, \end{aligned} \quad (158)$$

and the sought coefficients are given by Lemma 13, which concludes the second part of the proof.

We thus obtained a feasible solution of (LP_n) for all $n \in \mathbb{N}^*$, which is feasible for $(SDP_n^{m, \leq})$ for all $m \geq n$.

Optimality for $m = n$: we now find an analytical solution of the dual $(D-SDP_n^{n, \leq})$ with the same optimal value as the primal $(SDP_n^{n, \leq})$, by finding a Cholesky decomposition for the matrix appearing in the dual program $(D-SDP_n^{m, \leq})$ for $m = n$. The coefficients of the Cholesky decomposition are given by the triangular matrix L with coefficients:

$$\begin{aligned} \forall j \leq i, \quad l_{2i, 2j} &= 2^i i! \binom{i}{j} \\ \forall j \leq i, \quad l_{2i+1, 2j+1} &= 2^{i+1/2} \frac{(i+1)!}{\sqrt{j+1}} \binom{i}{j} \\ l_{nn} &= 0 \\ l_{ij} &= 0 \quad \text{otherwise.} \end{aligned} \quad (159)$$

Once again, these analytical expressions were extrapolated from numerical values. Then, $A = LL^T$ is a positive semidefinite matrix given by:

$$A_{ij} = \sum_{k=0}^{\min(i, j)} l_{ik} l_{jk}. \quad (160)$$

Now l_{ij} is non-zero only when i and j have the same parity, so for all $k \in \llbracket 0, \min(i, j) \rrbracket$, i and j must have the same parity than k for $l_{ik} l_{jk}$ to be non-zero.

- Suppose $i = 2i'$, $j = 2j'$, $i' \leq j'$ and $(i', j') \neq (n, n)$. Furthermore let $l = \frac{i+j}{2}$.

$$\begin{aligned}
A_{ij} &= \sum_{\substack{k=0 \\ k \text{ even}}}^i l_{2i',k} l_{2j',k'} \\
&= \sum_{k=0}^{i'} 2^{i'} i'! \binom{i'}{k} 2^{j'} j'! \binom{j'}{k} \\
&= 2^l i'! (l - i')! \sum_{k=0}^{i'} \binom{i'}{k} \binom{l - i'}{k} \\
&= 2^l l!.
\end{aligned} \tag{161}$$

- Suppose $i = 2i' + 1$, $j = 2j' + 1$, $i' \leq j'$ and $(i', j') \neq (n, n)$. Furthermore let $l = \frac{i+j}{2}$.

$$\begin{aligned}
A_{ij} &= \sum_{\substack{k=0 \\ k \text{ odd}}}^i l_{2i'+1,k} l_{2j'+1,k} \\
&= \sum_{k=0}^{i'} 2^{i'+1/2} \frac{(i'+1)!}{\sqrt{k+1}} \binom{i'}{k} 2^{j'+1/2} \frac{(j'+1)!}{\sqrt{k+1}} \binom{j'}{k} \\
&= 2^l i'! (l - i')! \sum_{k=0}^{i'} \binom{i'+1}{k+1} \binom{l - i' - 1}{k} \\
&= 2^l l!.
\end{aligned} \tag{162}$$

- Suppose $n = 2t$:

$$\begin{aligned}
A_{nn} &= \sum_{k=0}^t l_{2t,2k}^2 \\
&= 2^{2t} (t!)^2 \sum_{k=0}^{t-1} \binom{t}{k}^2 \\
&= 2^n (t!)^2 \left(\binom{2t}{t} - 1 \right) \\
&= 2^n n! \left(1 - \binom{n}{\lfloor \frac{n}{2} \rfloor}^{-1} \right).
\end{aligned} \tag{163}$$

- Suppose $n = 2t + 1$:

$$\begin{aligned}
A_{nn} &= \sum_{\substack{k=0 \\ k \text{ odd}}}^t l_{2t+1,k}^2 \\
&= 2^{2t+1} (t+1)!^2 \sum_{k=0}^{t-1} \frac{1}{k+1} \binom{t}{k}^2 \\
&= 2^n t! (t+1)! \sum_{k=0}^{t-1} \binom{t+1}{k+1} \binom{t}{k} \\
&= 2^n t! (t+1)! \left(\binom{2t+1}{t} - 1 \right) \\
&= 2^n n! \left(1 - \binom{n}{\lfloor \frac{n}{2} \rfloor}^{-1} \right).
\end{aligned} \tag{164}$$

In both cases, A is indeed constructed as:

$$(A\boldsymbol{\mu})_{i,j} = \begin{cases} \sum_{k=0}^l \mu_k \binom{l}{k} l! & \text{when } i+j = 2l, \\ 0 & \text{otherwise,} \end{cases} \quad (165)$$

for $\boldsymbol{\mu} = (F_n^n, F_n^n, \dots, F_n^n, 1 - F_n^n)$ with $F_n^n = \frac{1}{2^n} \binom{n}{\lfloor \frac{n}{2} \rfloor}$, and this provides a feasible solution of (D-SDP $_n^{m,\leq}$) for $m = n$, with value F_n^n . This shows the optimality of \mathbf{F}^n for (SDP $_n^{n,\leq}$) (and the fact that strong duality holds between the programs (SDP $_n^{n,\leq}$) and (D-SDP $_n^{n,\leq}$), which we already knew from Theorem 2). \square

For (D-SDP $_n^{m,\leq}$), we see numerically that the optimal solution is the same as for (D-SDP $_n^{n,\leq}$), for a few values of m greater than n . However, this is no longer the case for higher values, for example when $n = 3$ and $m = 10$.

G Strong duality between (LP $_n$) and (D-LP $_n$)

Theorem 6 shows that the optimal values of (LP $_n$) and (D-LP $_n$) are equal, i.e., that we have strong duality between those programs both when they are expressed in the search space $L^2(\mathbb{R}_+)$ and $\mathcal{S}(\mathbb{R}_+)$.

In this section we give another proof of strong duality between these linear programs as usually done in the literature of infinite-dimensional optimisation [81, 82], when they are expressed in the search space $L^2(\mathbb{R}_+)$.

Note that this result itself is not enough to provide the convergence of our hierarchy of semidefinite programs, hence the need for another proof technique: strong duality directly results from the fact the feasible set of (LP $_n$) is closed when it is expressed over the search space $L^2(\mathbb{R}_+)$.

We use notations from Appendix D. From [61, IV-(7.2)], there is no duality gap between (LP $_n$) and (D-LP $_n$) if there is a primal feasible plan and the cone:

$$\begin{aligned} \mathcal{K} &= \left\{ \left(A(e_1), \langle e_1, c \rangle_1 \right) : e_1 \in K_1 \right\} \\ &= \left\{ \left(\sum_k u_k, x \in \mathbb{R}_+ \mapsto f(x) - \sum_k u_k \mathcal{L}_k(x), F_n \right) : ((u_k), f) \in K_1 \right\} \end{aligned} \quad (166)$$

is closed in $E_2 \oplus \mathbb{R}$ (for the weak topology).

Proof. The null sequence provides a feasible plan for the primal problem.

Next, we consider a sequence $(e_{1j})_j = (((u_k^j)_k)_j, (f_j)_j) \in K_1^{\mathbb{N}} = \ell^2(\mathbb{N}) \times L_+^2(\mathbb{R}_+)^{\mathbb{N}}$ and we want to show that the accumulation point $(b, g, a) = \lim_{j \rightarrow \infty} (A(e_{1j}), \langle e_{1j}, c \rangle_1)$ belongs to \mathcal{K} where $a, b \in \mathbb{R}$ and $g \in L^2(\mathbb{R}_+)$.

For all $j \in \mathbb{N}$, $(u_k^j)_k \in \ell^2$ and for all $k \in \mathbb{N}$, u_k^j is bounded. Thus, for all $k \in \mathbb{N}$, the sequence $(u_k^j)_j$ is bounded and via diagonal extraction there exists $\phi : \mathbb{N} \rightarrow \mathbb{N}$ strictly increasing such that $(u_k^j)_{\phi(j)}$ converges. We denote \tilde{u}_k its limit. Since ℓ^2 is closed, the sequence $(\tilde{u}_k)_k$ belongs to ℓ^2 and we have $b = \sum_k \tilde{u}_k$ and $a = \tilde{u}_n$.

Now $f_j - \sum_k u_k^j \mathcal{L}_k \rightarrow g$ so that $f_j \rightarrow g + \sum_k \tilde{u}_k \mathcal{L}_k \in L_+^2(\mathbb{R}_+)$ since $L_+^2(\mathbb{R}_+)$ is closed. Thus, for $\tilde{e}_1 = ((\tilde{u}_k)_k, g + \sum_k \tilde{u}_k \mathcal{L}_k) \in K_1$, $(b, g, a) = (A(\tilde{e}_1), \langle \tilde{e}_1, c \rangle_1)$ and $(b, g, a) \in \mathcal{K}$. \square

H Multimode case

In this section, we provide the technical background for proving the convergence of the hierarchies of semidefinite programs in the multimode setting. The results obtained are summarised in Fig. 6.

⁸Because we are dealing with a sequence of sequences, we use the upper index to refer to the embracing sequence.

H.1 Multimode notations and definitions

We use bold math for multi-index notations. The main advantage of these notations is that the proofs of most technical results are easily extended to the multimode setting by replacing standard notations by multi-index notations.

Let M denote the number of modes. We consider M copies $\mathcal{H}^{\otimes M}$ of a separable Hilbert space \mathcal{H} . We denote the corresponding multimode orthonormal Fock basis by $\{|\mathbf{n}\rangle\}_{\mathbf{n} \in \mathbb{N}^M}$. For all $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_M) \in \mathbb{C}^M$, $\mathbf{k} = (k_1, \dots, k_M) \in \mathbb{N}^M$, $\mathbf{n} = (n_1, \dots, n_M) \in \mathbb{N}^M$ and $m \in \mathbb{N}$, we introduce the notations:

$$\begin{aligned}
\mathbf{0} &= (0, \dots, 0) \in \mathbb{N}^M \\
\mathbf{1} &= (1, \dots, 1) \in \mathbb{N}^M \\
m\mathbf{1} &= (m, \dots, m) \\
m\mathbf{k} &= (mk_1, \dots, mk_M) \\
\pi_{\mathbf{k}} &= \prod_{i=1}^M (k_i + 1) \\
\hat{D}(\boldsymbol{\alpha}) &= \hat{D}(\alpha_1) \otimes \dots \otimes \hat{D}(\alpha_M) \\
|\mathbf{k}\rangle &= |k_1\rangle \otimes \dots \otimes |k_M\rangle \\
\langle \mathbf{k}| &= \langle k_1| \otimes \dots \otimes \langle k_M| \\
\mathbf{k} \leq \mathbf{n} &\Leftrightarrow k_i \leq n_i \quad \forall i = 1, \dots, M \\
L_{\mathbf{k}}(\boldsymbol{\alpha}) &= L_{k_1}(\alpha_1) \dots L_{k_M}(\alpha_M) \\
\mathcal{L}_{\mathbf{k}}(\boldsymbol{\alpha}) &= \mathcal{L}_{k_1}(\alpha_1) \dots \mathcal{L}_{k_M}(\alpha_M) \\
|\mathbf{k}| &= k_1 + \dots + k_M \\
\boldsymbol{\alpha}^{\mathbf{k}} &= \alpha_1^{k_1} \dots \alpha_M^{k_M} \\
\mathbf{k}! &= k_1! \dots k_M! \\
\binom{\mathbf{n}}{\mathbf{k}} &= \binom{n_1}{k_1} \dots \binom{n_M}{k_M} \\
\mathbf{k} + \mathbf{n} &= (k_1 + n_1, \dots, k_M + n_M) \\
e^{\boldsymbol{\alpha}} &= e^{\alpha_1} \dots e^{\alpha_M}.
\end{aligned} \tag{167}$$

A multivariate polynomial of degree $p \in \mathbb{N}$ may then be written in the compact form $P(\mathbf{x}) = \sum_{|\mathbf{l}| \leq p} p_{\mathbf{l}} \mathbf{x}^{\mathbf{l}}$, where the sum is over all the tuples $\mathbf{l} \in \mathbb{N}^M$ such that $|\mathbf{l}| \leq m$, also known as the weak compositions of the integer m . There are $\binom{M+m}{m}$ such tuples. In what follows, we will also consider sums over all the tuples $\mathbf{l} \in \mathbb{N}^M$ such that $\mathbf{l} \leq \mathbf{k}$, for $\mathbf{k} = (k_1, \dots, k_M) \in \mathbb{N}^M$. There are $\pi_{\mathbf{k}}$ such tuples. In particular, for all $\mathbf{x} \in \mathbb{R}_+^M$ and all $\mathbf{k} \in \mathbb{N}^M$,

$$L_{\mathbf{k}}(\mathbf{x}) = \sum_{\mathbf{l} \leq \mathbf{k}} \frac{(-1)^{|\mathbf{l}|}}{\mathbf{l}!} \binom{\mathbf{k}}{\mathbf{l}} \mathbf{x}^{\mathbf{l}} \quad \text{and} \quad \mathbf{x}^{\mathbf{k}} = \sum_{\mathbf{l} \leq \mathbf{k}} (-1)^{|\mathbf{l}|} \binom{\mathbf{k}}{\mathbf{l}} \mathbf{k}! L_{\mathbf{l}}(\mathbf{x}). \tag{168}$$

We extend a few definitions from the single-mode case.

For $\mathbf{s} = (s_{\mathbf{k}})_{\mathbf{k} \in \mathbb{N}^M} \in \mathbb{R}^{\mathbb{N}^M}$, we define the associated formal series of multivariate Laguerre functions:

$$f_{\mathbf{s}} := \sum_{\mathbf{k}} s_{\mathbf{k}} \mathcal{L}_{\mathbf{k}}, \tag{169}$$

where \mathbf{s} is the so-called sequence of Laguerre moments of $f_{\mathbf{s}}$. For $m \in \mathbb{N}$, we also define the associated $\binom{M+m}{m} \times \binom{M+m}{m}$ matrix $A_{\mathbf{s}}$ by

$$(A_{\mathbf{s}})_{\mathbf{i}, \mathbf{j}} = \begin{cases} \sum_{\mathbf{k} \leq \mathbf{l}} s_{\mathbf{k}} \binom{\mathbf{l}}{\mathbf{k}} \mathbf{l}! & \text{when } \mathbf{i} + \mathbf{j} = 2\mathbf{l}, \\ 0 & \text{otherwise,} \end{cases} \tag{170}$$

where $\mathbf{i}, \mathbf{j} \in \mathbb{N}^M$ with $|\mathbf{i}| \leq m$ and $|\mathbf{j}| \leq m$.

The multimode Laguerre functions $(\mathcal{L}_{\mathbf{k}})_{\mathbf{k} \in \mathbb{N}^M}$ form an orthonormal basis of the space $L^2(\mathbb{R}_+^M)$ of real square-integrable functions over \mathbb{R}_+^M equipped with the usual scalar product:

$$\langle f, g \rangle = \int_{\mathbb{R}_+^M} f(\mathbf{x})g(\mathbf{x})d\mathbf{x}, \quad (171)$$

for $f, g \in L^2(\mathbb{R}_+^M)$. We denote by $L_+^2(\mathbb{R}_+^M)$ its subset of non-negative elements. The space $L^2(\mathbb{R}_+^M)$ is isomorphic to its dual space $L^{2'}(\mathbb{R}_+^M)$: elements of $L^{2'}(\mathbb{R}_+^M)$ can be identified by the Lebesgue measure on \mathbb{R}_+^M times the corresponding function in $L^2(\mathbb{R}_+^M)$.

Moreover, the elements of the space $\mathcal{S}(\mathbb{R}_+^M)$ of Schwartz functions over \mathbb{R}_+^M , i.e., the space of C^∞ functions that go to 0 at infinity faster than any inverse polynomial, as do their derivatives, can be written as series of Laguerre functions with a sequence indexed by \mathbb{N}^M of rapidly decreasing coefficients (which go to 0 at infinity faster than any inverse M -variate polynomial). Its dual space $\mathcal{S}'(\mathbb{R}_+^M)$ of tempered distributions over \mathbb{R}_+^M is characterised as the space of formal series of Laguerre functions with a slowly increasing sequence indexed by \mathbb{N}^M of coefficients (sequences that are upper bounded by an M -variate polynomial) [64]. We also extend the definition of the duality $\langle -, - \rangle$ in Eq. (171) to these spaces.

For all $m \in \mathbb{N}$, the set of series of Laguerre functions over \mathbb{R}_+^M truncated at m is denoted $\mathcal{R}_m(\mathbb{R}_+^M)$. This is the set of M -variate polynomials $P(\mathbf{x}) = \sum_{|\mathbf{k}| \leq m} p_{\mathbf{k}} \mathbf{x}^{\mathbf{k}}$ of degree at most m multiplied by the function $\mathbf{x} \mapsto e^{-\frac{\mathbf{x}}{2}}$. Let $\mathcal{R}_{m,+}(\mathbb{R}_+^M)$ denotes its subset of non-negative elements where the polynomial P is such that $\mathbf{x} \mapsto P(\mathbf{x}^2)$ has a sum-of-squares decomposition.

Similarly, for $\mathbf{m} \in \mathbb{N}^M$, the set of truncated series of Laguerre functions over \mathbb{R}_+^M with monomials smaller than \mathbf{m} is denoted $\mathcal{R}_{\mathbf{m}}(\mathbb{R}_+^M)$. This is the set of M -variate polynomials $P(\mathbf{x}) = \sum_{\mathbf{k} \leq \mathbf{m}} p_{\mathbf{k}} \mathbf{x}^{\mathbf{k}}$, multiplied by the function $\mathbf{x} \mapsto e^{-\frac{\mathbf{x}}{2}}$. Let $\mathcal{R}_{\mathbf{m},+}(\mathbb{R}_+^M)$ denotes its subset of non-negative elements where the polynomial P is such that $\mathbf{x} \mapsto P(\mathbf{x}^2)$ has a sum-of-squares decomposition.

We recall here the expressions of the linear program (LP $_{\mathbf{n}}$) and its dual (D-LP $_{\mathbf{n}}$):

$$(\text{LP}_{\mathbf{n}}) \quad \left\{ \begin{array}{l} \sup_{(F_{\mathbf{k}})_{\mathbf{k} \in \mathbb{N}^M} \in \mathbb{R}^{\mathbb{N}^M}} F_{\mathbf{n}} \\ \text{subject to} \quad \sum_{\mathbf{k}} F_{\mathbf{k}} = 1 \\ \text{and} \quad \forall \mathbf{k} \in \mathbb{N}^M, \quad F_{\mathbf{k}} \geq 0 \\ \text{and} \quad \forall \mathbf{x} \in \mathbb{R}_+^M, \quad \sum_{\mathbf{k}} F_{\mathbf{k}} \mathcal{L}_{\mathbf{k}}(\mathbf{x}) \geq 0, \end{array} \right.$$

where the optimisation is over real sequences indexed by elements of \mathbb{N}^M . Its dual linear program reads

$$(\text{D-LP}_{\mathbf{n}}) \quad \left\{ \begin{array}{l} \inf_{\substack{y \in \mathbb{R} \\ \mu \in L^{2'}(\mathbb{R}_+^M)}} y \\ \text{subject to} \quad \forall \mathbf{k} \neq \mathbf{n} \in \mathbb{N}^M, \quad y \geq \int_{\mathbb{R}_+^M} \mathcal{L}_{\mathbf{k}} d\mu \\ \text{and} \quad y \geq 1 + \int_{\mathbb{R}_+^M} \mathcal{L}_{\mathbf{n}} d\mu \\ \text{and} \quad \forall f \in L_+^2(\mathbb{R}_+^M), \quad \langle \mu, f \rangle \geq 0. \end{array} \right.$$

H.2 Multimode semidefinite programs

Hereafter, we state without proofs the technical results used to derive the semidefinite programs in section 6 and their dual programs. It is a straightforward exercise to obtain the proofs from their single-mode version by using multi-index notations.

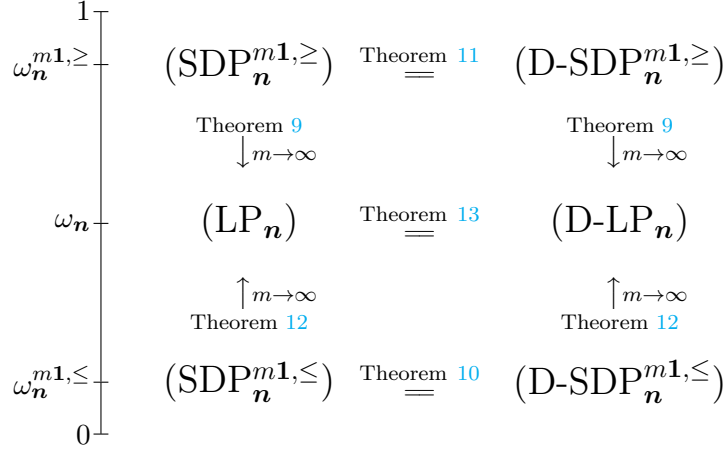


Figure 6: Multimode hierarchies of semidefinite relaxations and restrictions converging to the linear program (LP_n) , together with their dual programs. The upper index m denotes the level of the relaxation or restriction. On the left are the associated optimal values. The equal sign denotes strong duality, i.e., equality of optimal values, and the arrows denote convergence of the corresponding sequences of optimal values. The hierarchies $(\text{SDP}_n^{m1, \geq})$ and $(\text{SDP}_n^{m1, \leq})$ in the main text are different from the ones appearing in the figure, but equivalent by Lemma 17.

There are two natural ways to obtain relaxations and restrictions by replacing constraints on non-negative functions by constraints on non-negative polynomials: either by considering polynomials $P(\mathbf{x}) = \sum_{|\mathbf{k}| \leq m} p_{\mathbf{k}} \mathbf{x}^{\mathbf{k}}$ of degree at most m for $m \in \mathbb{N}$, or by considering polynomials $P(\mathbf{x}) = \sum_{\mathbf{k} \leq \mathbf{m}} p_{\mathbf{k}} \mathbf{x}^{\mathbf{k}}$ with monomials smaller than \mathbf{m} for $m \in \mathbb{N}^M$. Note that when $M = 1$ these two are the equivalent.

All the results below containing conditions of the form $|\mathbf{k}| \leq m$ for $m \in \mathbb{N}$ are also valid when replaced by conditions of the form $\mathbf{k} \leq \mathbf{m}$ for $\mathbf{m} = (m_1, \dots, m_M) \in \mathbb{N}^M$, with the same proofs, by replacing $\binom{M+m}{m}$ by $\pi_{\mathbf{m}} = \prod_{i=1}^M (m_i + 1)$.

Lemma 14 (Equivalent of Lemma 3). *Let $p \in \mathbb{N}$ and let P be a multivariate polynomial of degree $2p$. Let $\mathbf{X} = (\mathbf{x}^{\mathbf{k}})_{|\mathbf{k}| \leq p}$ be the vector of monomials. Then, P has a sum-of-squares decomposition if and only if there exists a real $\binom{M+p}{p} \times \binom{M+p}{p}$ positive semidefinite matrix Q such that for all $x \in \mathbb{R}^M$,*

$$P(\mathbf{x}) = \mathbf{X}^T Q \mathbf{X}. \quad (172)$$

Lemma 15 (Generalisation of Lemma 4). *Let P be a non-negative polynomial over \mathbb{R}_+^M such that $\mathbf{x} \mapsto P(\mathbf{x}^2)$ has a sum-of-squares decomposition. Then, P can be written as a sum of polynomials of the form $\sum_{|\mathbf{l}| \leq p} \mathbf{x}^{\mathbf{l}} \sum_{\mathbf{i}+\mathbf{j}=2\mathbf{l}} y_{\mathbf{i}} y_{\mathbf{j}}$ for $p \in \mathbb{N}$ and $y_{\mathbf{i}} \in \mathbb{R}$ for all $\mathbf{i} \in \mathbb{N}^M$ such that $|\mathbf{i}| \leq p$.*

Lemma 16 (Generalisation of Lemma 5). *Let $m \in \mathbb{N}$ and let $\mathbf{s} = (s_{\mathbf{k}})_{\mathbf{k} \in \mathbb{N}^M} \in \mathbb{R}^{\mathbb{N}^M}$. The following propositions are equivalent:*

(i) $\forall g \in \mathcal{R}_{m,+}(\mathbb{R}_+^M), \langle f_{\mathbf{s}}, g \rangle \geq 0,$

(ii) $A_{\mathbf{s}} \succeq 0.$

Using Lemma 16 we obtain the semidefinite *relaxations*:

$$(\text{SDP}_{\mathbf{n}}^{m,\geq}) \left\{ \begin{array}{l} \sup_{\substack{A \in \text{Sym}^{\binom{M+m}{m}} \\ \mathbf{F} \in \mathbb{R}^{\binom{M+m}{m}}}} F_{\mathbf{n}} \\ \text{subject to } \sum_{|\mathbf{k}| \leq m} F_{\mathbf{k}} = 1 \\ \text{and } \forall |\mathbf{k}| \leq m, \quad F_{\mathbf{k}} \geq 0 \\ \text{and } \forall |\mathbf{l}| \leq m, \forall \mathbf{i} + \mathbf{j} = 2\mathbf{l}, \quad A_{\mathbf{i}\mathbf{j}} = \sum_{\mathbf{k} \leq \mathbf{l}} F_{\mathbf{k}} \binom{\mathbf{l}}{\mathbf{k}} \mathbf{l}! \\ \text{and } \forall |\mathbf{r}| \leq 2m, \mathbf{r} \neq 2\mathbf{l}, \forall |\mathbf{l}| \leq m, \forall \mathbf{i} + \mathbf{j} = \mathbf{r}, \quad A_{\mathbf{i}\mathbf{j}} = 0 \\ \text{and } A \succeq 0. \end{array} \right.$$

for all $m \geq |\mathbf{n}|$. We denote its optimal value by $\omega_{\mathbf{n}}^{m,\geq}$. The corresponding dual programs are given by:

$$(\text{D-SDP}_{\mathbf{n}}^{m,\geq}) \left\{ \begin{array}{l} \inf_{\substack{Q \in \text{Sym}^{\binom{M+m}{m}} \\ y \in \mathbb{R}, \mu \in \mathbb{R}^{\binom{M+m}{m}}}} y \\ \text{subject to } y \geq 1 + \mu_{\mathbf{n}} \\ \text{and } \forall |\mathbf{k}| \leq m, \mathbf{k} \neq \mathbf{n}, \quad y \geq \mu_{\mathbf{k}} \\ \text{and } \forall |\mathbf{l}| \leq m, \quad \sum_{\mathbf{i}+\mathbf{j}=2\mathbf{l}} Q_{\mathbf{i}\mathbf{j}} = \sum_{\mathbf{k} \geq \mathbf{l}} \frac{(-1)^{|\mathbf{k}|+|\mathbf{l}|}}{\mathbf{l}!} \binom{\mathbf{k}}{\mathbf{l}} \mu_{\mathbf{k}} \\ \text{and } Q \succeq 0, \end{array} \right.$$

for all $m \geq |\mathbf{n}|$. Similarly, using Lemma 14 we obtain the semidefinite *restrictions*:

$$(\text{SDP}_{\mathbf{n}}^{m,\leq}) \left\{ \begin{array}{l} \sup_{\substack{Q \in \text{Sym}^{\binom{M+m}{m}} \\ \mathbf{F} \in \mathbb{R}^{\binom{M+m}{m}}}} F_{\mathbf{n}} \\ \text{subject to } \sum_{|\mathbf{k}| \leq m} F_{\mathbf{k}} = 1 \\ \text{and } \forall |\mathbf{k}| \leq m, \quad F_{\mathbf{k}} \geq 0 \\ \text{and } \forall |\mathbf{l}| \leq m, \quad \sum_{\mathbf{i}+\mathbf{j}=2\mathbf{l}} Q_{\mathbf{i}\mathbf{j}} = \sum_{\mathbf{k} \geq \mathbf{l}} \frac{(-1)^{|\mathbf{k}|+|\mathbf{l}|}}{\mathbf{l}!} \binom{\mathbf{k}}{\mathbf{l}} F_{\mathbf{k}} \\ \text{and } \forall |\mathbf{r}| \leq 2m, \mathbf{r} \neq 2\mathbf{l}, \forall |\mathbf{l}| \leq m, \quad \sum_{\mathbf{i}+\mathbf{j}=\mathbf{r}} Q_{\mathbf{i}\mathbf{j}} = 0 \\ \text{and } Q \succeq 0, \end{array} \right.$$

for all $m \geq |\mathbf{n}|$. We denote its optimal value by $\omega_{\mathbf{n}}^{m,\leq}$. The corresponding dual programs are given by:

$$(\text{D-SDP}_{\mathbf{n}}^{m,\leq}) \left\{ \begin{array}{l} \inf_{\substack{A \in \text{Sym}^{\binom{M+m}{m}} \\ y, \mu \in \mathbb{R} \times \mathbb{R}^{\binom{M+m}{m}}}} y \\ \text{subject to } y \geq 1 + \mu_{\mathbf{n}} \\ \text{and } \forall |\mathbf{k}| \leq m, \mathbf{k} \neq \mathbf{n}, \quad y \geq \mu_{\mathbf{k}} \\ \text{and } \forall |\mathbf{l}| \leq m, \forall \mathbf{i} + \mathbf{j} = 2\mathbf{l}, \quad A_{\mathbf{i}\mathbf{j}} = \sum_{\mathbf{k} \leq \mathbf{l}} \mu_{\mathbf{k}} \binom{\mathbf{l}}{\mathbf{k}} \mathbf{l}! \\ \text{and } A \succeq 0, \end{array} \right.$$

for all $m \geq |\mathbf{n}|$. Like in the single-mode case, note that without loss of generality the condition $y \leq 1$, and thus $\mu_{\mathbf{k}} \leq 1$ for all \mathbf{k} , can be added to the optimisation, since setting $A = 0$, $y = 1$ and $\boldsymbol{\mu} = 0$ gives a feasible solution with objective value 1.

These are the relaxations and restrictions of $(LP_{\mathbf{n}})$ obtained by considering polynomials of degree less or equal to m , where the optimisation is over matrices and vectors indexed by elements of \mathbb{N}^m with sum of coefficients lower than m . Alternatively, we may also consider the relaxations and restrictions obtained by considering polynomials with monomials smaller than $\mathbf{m} \in \mathbb{N}^M$, where the optimisation is over matrices and vectors indexed by elements of \mathbb{N}^m lower than \mathbf{m} . Recalling the notation $\pi_{\mathbf{m}} = \prod_{i=1}^M (m_i + 1)$, the corresponding semidefinite *relaxations* are given by

$$(SDP_{\mathbf{n}}^{m, \geq}) \quad \left\{ \begin{array}{l} \sup_{\substack{A \in \text{Sym}_{\pi_{\mathbf{m}}} \\ \mathbf{F} \in \mathbb{R}^{\pi_{\mathbf{m}}} \\ \mathbf{F} \in \mathbb{R}^{\pi_{\mathbf{m}}}}} F_{\mathbf{n}} \\ \text{subject to} \quad \sum_{\mathbf{k} \leq \mathbf{m}} F_{\mathbf{k}} = 1 \\ \text{and} \quad \forall \mathbf{k} \leq \mathbf{m}, \quad F_{\mathbf{k}} \geq 0 \\ \text{and} \quad \forall \mathbf{l} \leq \mathbf{m}, \forall \mathbf{i} + \mathbf{j} = 2\mathbf{l}, \quad A_{\mathbf{i}\mathbf{j}} = \sum_{\mathbf{k} \leq \mathbf{l}} F_{\mathbf{k}} \binom{\mathbf{l}}{\mathbf{k}} \mathbf{l}! \\ \text{and} \quad \forall \mathbf{r} \leq 2\mathbf{m}, \mathbf{r} \neq 2\mathbf{l}, \forall \mathbf{l} \leq \mathbf{m}, \forall \mathbf{i} + \mathbf{j} = \mathbf{r}, \quad A_{\mathbf{i}\mathbf{j}} = 0 \\ \text{and} \quad A \succeq 0. \end{array} \right.$$

for all $\mathbf{m} \geq \mathbf{n}$. We denote its optimal value by $\omega_{\mathbf{n}}^{m, \geq}$. The corresponding dual programs are given by:

$$(D\text{-}SDP_{\mathbf{n}}^{m, \geq}) \quad \left\{ \begin{array}{l} \inf_{\substack{Q \in \text{Sym}_{\pi_{\mathbf{m}}} \\ y, \boldsymbol{\mu} \in \mathbb{R} \times \mathbb{R}^{\pi_{\mathbf{m}}}}} y \\ \text{subject to} \quad y \geq 1 + \mu_{\mathbf{n}} \\ \text{and} \quad \forall \mathbf{k} \leq \mathbf{m}, \mathbf{k} \neq \mathbf{n}, \quad y \geq \mu_{\mathbf{k}} \\ \text{and} \quad \forall \mathbf{l} \leq \mathbf{m}, \quad \sum_{\mathbf{i} + \mathbf{j} = 2\mathbf{l}} Q_{\mathbf{i}\mathbf{j}} = \sum_{\mathbf{k} \geq \mathbf{l}} \frac{(-1)^{|\mathbf{k}| + |\mathbf{l}|}}{\mathbf{l}!} \binom{\mathbf{k}}{\mathbf{l}} \mu_{\mathbf{k}} \\ \text{and} \quad Q \succeq 0, \end{array} \right.$$

for all $\mathbf{m} \geq \mathbf{n}$. Similarly, the semidefinite *restrictions* are given by:

$$(SDP_{\mathbf{n}}^{m, \leq}) \quad \left\{ \begin{array}{l} \sup_{\substack{Q \in \text{Sym}_{\pi_{\mathbf{m}}} \\ \mathbf{F} \in \mathbb{R}^{\pi_{\mathbf{m}}} \\ \mathbf{F} \in \mathbb{R}^{\pi_{\mathbf{m}}} \\ \mathbf{F} \in \mathbb{R}^{\pi_{\mathbf{m}}}}} F_{\mathbf{n}} \\ \text{subject to} \quad \sum_{\mathbf{k} \leq \mathbf{m}} F_{\mathbf{k}} = 1 \\ \text{and} \quad \forall \mathbf{k} \leq \mathbf{m}, \quad F_{\mathbf{k}} \geq 0 \\ \text{and} \quad \forall \mathbf{l} \leq \mathbf{m}, \quad \sum_{\mathbf{i} + \mathbf{j} = 2\mathbf{l}} Q_{\mathbf{i}\mathbf{j}} = \sum_{\mathbf{k} \geq \mathbf{l}} \frac{(-1)^{|\mathbf{k}| + |\mathbf{l}|}}{\mathbf{l}!} \binom{\mathbf{k}}{\mathbf{l}} F_{\mathbf{k}} \\ \text{and} \quad \forall \mathbf{r} \leq 2\mathbf{m}, \mathbf{r} \neq 2\mathbf{l}, \forall \mathbf{l} \leq \mathbf{m}, \quad \sum_{\mathbf{i} + \mathbf{j} = \mathbf{r}} Q_{\mathbf{i}\mathbf{j}} = 0 \\ \text{and} \quad Q \succeq 0, \end{array} \right.$$

for all $\mathbf{m} \geq \mathbf{n}$. We denote its optimal value by $\omega_{\mathbf{n}}^{m, \leq}$. The corresponding dual programs are given by:

$$(\text{D-SDP}_{\mathbf{n}}^{m, \leq}) \left\{ \begin{array}{l} \inf_{\substack{A \in \text{Sym}_{\pi_{\mathbf{n}}}^m \\ y, \mu \in \mathbb{R} \times \mathbb{R}^{\pi_{\mathbf{n}}}^m}} y \\ \text{subject to } y \geq 1 + \mu_{\mathbf{n}} \\ \text{and } \forall \mathbf{k} \leq \mathbf{m}, \mathbf{k} \neq \mathbf{n}, \quad y \geq \mu_{\mathbf{k}} \\ \text{and } \forall \mathbf{l} \leq \mathbf{m}, \forall \mathbf{i} + \mathbf{j} = 2\mathbf{l}, \quad A_{\mathbf{i}\mathbf{j}} = \sum_{\mathbf{k} \leq \mathbf{l}} \mu_{\mathbf{k}} \binom{\mathbf{l}}{\mathbf{k}} \mathbf{l}! \\ \text{and } A \succeq 0, \end{array} \right.$$

for all $\mathbf{m} \geq \mathbf{n}$.

The programs $(\text{SDP}_{\mathbf{n}}^{m, \geq})$ and $(\text{SDP}_{\mathbf{n}}^{m, \leq})$ respectively provide hierarchies of relaxations and restrictions of $(\text{LP}_{\mathbf{n}})$, since the set of M -variate polynomials of degree m is included in the set of M -variate polynomials of degree $m + 1$. On the other hand, there is no natural ordering in \mathbb{N}^M of the relaxations $(\text{SDP}_{\mathbf{n}}^{m, \geq})$ or the restrictions $(\text{SDP}_{\mathbf{n}}^{m, \leq})$ (consider for instance $\mathbf{m} = (2, 1)$ and $\mathbf{m}' = (1, 2)$). In order to obtain proper hierarchies of semidefinite programs, we thus consider the subset of these programs where the tuple \mathbf{m} is of the form $m\mathbf{1} = (m, \dots, m) \in \mathbb{N}^M$, for $m \in \mathbb{N}$. We have $\pi_{m\mathbf{1}} = (m+1)^M$, and the *relaxations* are then given by

$$(\text{SDP}_{\mathbf{n}}^{m\mathbf{1}, \geq}) \left\{ \begin{array}{l} \sup_{\substack{A \in \text{Sym}_{(m+1)^M} \\ \mathbf{F} \in \mathbb{R}^{(m+1)^M}}} F_{\mathbf{n}} \\ \text{subject to } \sum_{\mathbf{k} \leq m\mathbf{1}} F_{\mathbf{k}} = 1 \\ \text{and } \forall \mathbf{k} \leq m\mathbf{1}, \quad F_{\mathbf{k}} \geq 0 \\ \text{and } \forall \mathbf{l} \leq m\mathbf{1}, \forall \mathbf{i} + \mathbf{j} = 2\mathbf{l}, \quad A_{\mathbf{i}\mathbf{j}} = \sum_{\mathbf{k} \leq \mathbf{l}} F_{\mathbf{k}} \binom{\mathbf{l}}{\mathbf{k}} \mathbf{l}! \\ \text{and } \forall \mathbf{r} \leq 2m\mathbf{1}, \mathbf{r} \neq 2\mathbf{l}, \forall \mathbf{l} \leq m\mathbf{1}, \forall \mathbf{i} + \mathbf{j} = \mathbf{r}, \quad A_{\mathbf{i}\mathbf{j}} = 0 \\ \text{and } A \succeq 0. \end{array} \right.$$

for $m \geq \max_i n_i$. We denote its optimal value by $\omega_{\mathbf{n}}^{m\mathbf{1}, \geq}$. The corresponding dual programs are given by:

$$(\text{D-SDP}_{\mathbf{n}}^{m\mathbf{1}, \geq}) \left\{ \begin{array}{l} \inf_{\substack{Q \in \text{Sym}_{(m+1)^M} \\ y, \mu \in \mathbb{R} \times \mathbb{R}^{(m+1)^M}}} y \\ \text{subject to } y \geq 1 + \mu_{\mathbf{n}} \\ \text{and } \forall \mathbf{k} \leq m\mathbf{1}, \mathbf{k} \neq \mathbf{n}, \quad y \geq \mu_{\mathbf{k}} \\ \text{and } \forall \mathbf{l} \leq m\mathbf{1}, \quad \sum_{\mathbf{i} + \mathbf{j} = 2\mathbf{l}} Q_{\mathbf{i}\mathbf{j}} = \sum_{\mathbf{k} \geq \mathbf{l}} \frac{(-1)^{|\mathbf{k}| + |\mathbf{l}|}}{\mathbf{l}!} \binom{\mathbf{k}}{\mathbf{l}} \mu_{\mathbf{k}} \\ \text{and } Q \succeq 0, \end{array} \right.$$

for $m \geq \max_i n_i$. Similarly, the *restrictions* are given by:

$$(\text{SDP}_{\mathbf{n}}^{m\mathbf{1}, \leq}) \left\{ \begin{array}{l} \sup_{\substack{Q \in \text{Sym}_{(m+1)M} \\ \mathbf{F} \in \mathbb{R}^{(m+1)M}}} F_{\mathbf{n}} \\ \text{subject to } \sum_{\mathbf{k} \leq m\mathbf{1}} F_{\mathbf{k}} = 1 \\ \text{and } \forall \mathbf{k} \leq m\mathbf{1}, \quad F_{\mathbf{k}} \geq 0 \\ \text{and } \forall \mathbf{l} \leq m\mathbf{1}, \quad \sum_{i+j=2\mathbf{l}} Q_{ij} = \sum_{\mathbf{k} \geq \mathbf{l}} \frac{(-1)^{|\mathbf{k}|+|\mathbf{l}|}}{\mathbf{l}!} \binom{\mathbf{k}}{\mathbf{l}} F_{\mathbf{k}} \\ \text{and } \forall \mathbf{r} \leq 2m\mathbf{1}, \mathbf{r} \neq 2\mathbf{l}, \forall \mathbf{l} \leq m\mathbf{1}, \quad \sum_{i+j=\mathbf{r}} Q_{ij} = 0 \\ \text{and } Q \succeq 0, \end{array} \right.$$

for $m \geq \max_i n_i$. We denote its optimal value by $\omega_{\mathbf{n}}^{m\mathbf{1}, \leq}$. The corresponding dual programs are given by:

$$(\text{D-SDP}_{\mathbf{n}}^{m\mathbf{1}, \leq}) \left\{ \begin{array}{l} \inf_{\substack{A \in \text{Sym}_{(m+1)M} \\ y, \boldsymbol{\mu} \in \mathbb{R} \times \mathbb{R}^{(m+1)M}}} y \\ \text{subject to } y \geq 1 + \boldsymbol{\mu}_{\mathbf{n}} \\ \text{and } \forall \mathbf{k} \leq m\mathbf{1}, \mathbf{k} \neq \mathbf{n}, \quad y \geq \boldsymbol{\mu}_{\mathbf{k}} \\ \text{and } \forall \mathbf{l} \leq m\mathbf{1}, \forall i+j=2\mathbf{l}, \quad A_{ij} = \sum_{\mathbf{k} \leq \mathbf{l}} \boldsymbol{\mu}_{\mathbf{k}} \binom{\mathbf{l}}{\mathbf{k}} \mathbf{l}! \\ \text{and } A \succeq 0, \end{array} \right.$$

for $m \geq \max_i n_i$.

The programs $(\text{SDP}_{\mathbf{n}}^{m\mathbf{1}, \geq})$ and $(\text{SDP}_{\mathbf{n}}^{m\mathbf{1}, \leq})$ are the relaxations and restrictions of $(\text{LP}_{\mathbf{n}})$ obtained by considering polynomials of individual degree in each variable less or equal to m . These programs respectively provide hierarchies of relaxations and restrictions of $(\text{LP}_{\mathbf{n}})$, since the set of M -variate polynomials with monomials lower than $m\mathbf{1}$ is included in the set of M -variate polynomials with monomials lower than $(m+1)\mathbf{1}$.

Note that these hierarchies of programs obtained by setting \mathbf{m} of the form $m\mathbf{1}$ capture the behaviour of all bounds that can be obtained from the more general family of programs indexed by $\mathbf{m} = (m_1, \dots, m_M)$, since $\mathbf{m} \leq (\max_i m_i)\mathbf{1}$, i.e., the bound obtained by considering the program indexed by $(\max_i m_i)\mathbf{1}$ supersedes the bound obtained by considering the program indexed by \mathbf{m} . Formally, for all $\mathbf{m} = (m_1, \dots, m_M) \in \mathbb{N}^M$,

$$\omega_{\mathbf{n}}^{\mathbf{m}, \geq} \geq \omega_{\mathbf{n}}^{(\max_i m_i)\mathbf{1}, \geq} \quad \text{and} \quad \omega_{\mathbf{n}}^{\mathbf{m}, \leq} \leq \omega_{\mathbf{n}}^{(\max_i m_i)\mathbf{1}, \leq}. \quad (173)$$

Finally, we show that both ways of defining the hierarchies are equivalent:

Lemma 17. *For all $m \in \mathbb{N}$,*

$$\omega_{\mathbf{n}}^{m, \geq} \geq \omega_{\mathbf{n}}^{m\mathbf{1}, \geq} \geq \omega_{\mathbf{n}}^{Mm, \geq}, \quad (174)$$

and

$$\omega_{\mathbf{n}}^{m, \leq} \leq \omega_{\mathbf{n}}^{m\mathbf{1}, \leq} \leq \omega_{\mathbf{n}}^{Mm, \leq}. \quad (175)$$

Proof. For all $\mathbf{k} \in \mathbb{N}^M$ and all $m \in \mathbb{N}$ we have

$$|\mathbf{k}| \leq m \quad \Rightarrow \quad \mathbf{k} \leq m\mathbf{1} = (m, \dots, m) \quad \Rightarrow \quad |\mathbf{k}| \leq Mm. \quad (176)$$

We thus obtain the corresponding inclusions between sets of M -variate polynomials: (i) M -variate polynomials of degree less or equal to m have all their monomials lower than $m\mathbf{1}$, and (ii) all M -variate polynomials with monomials lower than $m\mathbf{1}$ have degree less or equal to Mm . Hence,

$$\mathcal{R}_{m,+}(\mathbb{R}_+^M) \stackrel{(i)}{\subset} \mathcal{R}_{m\mathbf{1},+}(\mathbb{R}_+^M) \stackrel{(ii)}{\subset} \mathcal{R}_{Mm,+}(\mathbb{R}_+^M). \quad (177)$$

In particular, $(\text{SDP}_{\mathbf{n}}^{m,\geq})$ is a relaxation of $(\text{SDP}_{\mathbf{n}}^{m\mathbf{1},\geq})$ which is itself a relaxation of $(\text{SDP}_{\mathbf{n}}^{Mm,\geq})$, and $(\text{SDP}_{\mathbf{n}}^{m,\leq})$ is a restriction of $(\text{SDP}_{\mathbf{n}}^{m\mathbf{1},\leq})$ which is itself a restriction of $(\text{SDP}_{\mathbf{n}}^{Mm,\leq})$. \square

This result implies that the two versions of the hierarchies of relaxations are interleaved (Eq. (174)), and that the two versions of the hierarchies of restrictions are also interleaved (Eq. (175)). As such, for any bound obtained with one version of the hierarchy at some fixed level, a better bound can be obtained with the other version at some other level. While this means that the hierarchies are equivalent, note that in practice it may be simpler to solve numerically the version where the parameter space is smaller.

H.3 Convergence of the multimode hierarchies

For $\mathbf{n} = (n_1, \dots, n_M) \in \mathbb{N}^M$, the sequences $(\omega_{\mathbf{n}}^{m,\geq})_{m \geq |\mathbf{n}|}$ and $(\omega_{\mathbf{n}}^{m\mathbf{1},\geq})_{m \geq \max_i n_i}$ (resp. $(\omega_{\mathbf{n}}^{m,\leq})_{m \geq |\mathbf{n}|}$ and $(\omega_{\mathbf{n}}^{m\mathbf{1},\leq})_{m \geq \max_i n_i}$) are decreasing (resp. increasing) sequences, lower bounded (resp. upper bounded) by $\omega_{\mathbf{n}}$. Hence, these sequences are converging.

We show in what follows that $(\omega_{\mathbf{n}}^{m\mathbf{1},\geq})_{m \geq \max_i n_i}$ (resp. $(\omega_{\mathbf{n}}^{m\mathbf{1},\leq})_{m \geq \max_i n_i}$) converges to $\omega_{\mathbf{n}}$. By Lemma 17, this implies that the sequence $(\omega_{\mathbf{n}}^{Mm,\geq})_{m \geq \max_i n_i}$ (resp. $(\omega_{\mathbf{n}}^{Mm,\leq})_{m \geq \max_i n_i}$) also converges to $\omega_{\mathbf{n}}$. Since this is a subsequence of the converging sequence $(\omega_{\mathbf{n}}^{m,\geq})_{m \geq |\mathbf{n}|}$ (resp. $(\omega_{\mathbf{n}}^{m,\leq})_{m \geq |\mathbf{n}|}$), it implies that the sequence $(\omega_{\mathbf{n}}^{m,\geq})_{m \geq |\mathbf{n}|}$ (resp. $(\omega_{\mathbf{n}}^{m,\leq})_{m \geq |\mathbf{n}|}$) also converges to $\omega_{\mathbf{n}}$.

With similar proofs to the single-mode case using multi-index notations, we obtain the following result:

Theorem 8 (Generalisation of Theorem 1). *Let $\boldsymbol{\mu} = (\mu_{\mathbf{k}})_{\mathbf{k} \in \mathbb{N}^M} \in \mathbb{R}^{\mathbb{N}^M}$. Then, $\boldsymbol{\mu}$ is the sequence of Laguerre moments $\int_{\mathbb{R}_+^M} \mathcal{L}_{\mathbf{k}}(\mathbf{x}) d\boldsymbol{\mu}(\mathbf{x})$ of a non-negative distribution μ supported on \mathbb{R}_+^M if and only if $\forall m \in \mathbb{N}, \forall g \in \mathcal{R}_{m,+}(\mathbb{R}_+^M), \langle f_{\boldsymbol{\mu}}, g \rangle \geq 0$.*

The proof of this theorem is identical to the univariate case, with the use of Riesz–Haviland theorem over \mathbb{R}_+^M [67] rather than \mathbb{R}_+ .

With Eq. (177), the proof of convergence of the multimode hierarchy of upper bounds is then obtained directly from its single-mode counterpart using multi-index notations:

Theorem 9 (Generalisation of Theorem 4). *The decreasing sequence of optimal values $\omega_{\mathbf{n}}^{m\mathbf{1},\geq}$ of $(\text{SDP}_{\mathbf{n}}^{m\mathbf{1},\geq})$ converges to the optimal value $\omega_{\mathbf{n}}$ of $(LP_{\mathbf{n}})$:*

$$\lim_{m \rightarrow +\infty} \omega_{\mathbf{n}}^{m\mathbf{1},\geq} = \omega_{\mathbf{n}}. \quad (178)$$

With Lemma 17, we also obtain

$$\lim_{m \rightarrow +\infty} \omega_{\mathbf{n}}^{m,\geq} = \omega_{\mathbf{n}}. \quad (179)$$

On the other hand, the proof of convergence of the single-mode hierarchy of lower bounds crucially exploits analytical feasible solutions of the programs $(\text{SDP}_{\mathbf{n}}^{m,\leq})$ in order to obtain two results:

- Strong duality between programs $(\text{SDP}_{\mathbf{n}}^{m,\leq})$ and $(\text{D-SDP}_{\mathbf{n}}^{m,\leq})$ (Theorem 2).
- The fact that the feasible set of $(\text{D-SDP}_{\mathbf{n}}^{m,\leq})$ is compact with coefficients bounded independently of m (Eq. (63)).

In what follows, we generalise these two results to the multimode setting by obtaining multimode analytical feasible solutions from products of single-mode ones.

Lemma 18. For $m, n \in \mathbb{N}$ with $m \geq n$, let $Q(m, n) \in \text{Sym}_{m+1}$ and $\mathbf{F}(m, n) = (F_{\mathbf{k}}(m, n))_{\mathbf{k}} \in \mathbb{R}^{m+1}$ be feasible solutions of $(\text{SDP}_{\mathbf{n}}^{m, \leq})$. Let $\mathbf{m} = (m_1, \dots, m_M) \in \mathbb{N}^M$ and $\mathbf{n} = (n_1, \dots, n_M) \in \mathbb{N}^M$ with $\mathbf{m} \geq \mathbf{n}$. Let $Q := Q(m_1, n_1) \otimes \dots \otimes Q(m_M, n_M) \in \text{Sym}_{\pi_{\mathbf{m}}}$ and $\mathbf{F} = (F_{\mathbf{k}})_{\mathbf{k} \leq \mathbf{m}} \in \mathbb{R}^{\pi_{\mathbf{m}}}$, where for all $\mathbf{k} = (k_1, \dots, k_M) \leq \mathbf{m}$, $F_{\mathbf{k}} := \prod_{i=1}^M F_{k_i}(m_i, n_i)$. Then, (Q, \mathbf{F}) is a feasible solution of $(\text{SDP}_{\mathbf{n}}^{m, \leq})$. Moreover, if $(Q(m_i, n_i), \mathbf{F}(m_i, n_i))$ is strictly feasible for all $i = 1, \dots, M$ then (Q, \mathbf{F}) is a strictly feasible solution of $(\text{SDP}_{\mathbf{n}}^{m, \leq})$.

Proof. With the notations of the Lemma, we show the feasibility of (Q, \mathbf{F}) (resp. strict feasibility). We have $Q \succeq 0$, $F_{\mathbf{k}} \geq 0$ (resp. $Q \succ 0$, $F_{\mathbf{k}} > 0$) for all $\mathbf{k} \leq \mathbf{m}$, and $Q_{i\mathbf{j}} = \prod_{p=1}^M Q_{i_p j_p}(m_p, n_p)$ for all $i = (i_1, \dots, i_M) \leq \mathbf{m}$ and $\mathbf{j} = (j_1, \dots, j_M) \leq \mathbf{m}$. Hence, for all $\mathbf{r} = (r_1, \dots, r_M) \leq 2\mathbf{m}$,

$$\begin{aligned} \sum_{i+\mathbf{j}=\mathbf{r}} Q_{i\mathbf{j}} &= \sum_{i_1+j_1=r_1, \dots, i_M+j_M=r_M} \prod_{p=1}^M Q_{i_p j_p}(m_p, n_p) \\ &= \prod_{p=1}^M \sum_{i_p+j_p=r_p} Q_{i_p j_p}(m_p, n_p). \end{aligned} \quad (180)$$

In particular, if $\mathbf{r} \neq 2\mathbf{l}$ for all $\mathbf{l} \leq \mathbf{m}$, then at least one coefficient r_p is odd, and the corresponding sum gives 0 since $(Q(m_p, n_p), \mathbf{F}(m_p, n_p))$ is feasible for $(\text{SDP}_{n_p}^{m_p, \leq})$. In that case, $\sum_{i+\mathbf{j}=\mathbf{r}} Q_{i\mathbf{j}} = 0$. Otherwise, for all $\mathbf{l} = (l_1, \dots, l_M) \leq \mathbf{m}$,

$$\begin{aligned} \sum_{i+\mathbf{j}=2\mathbf{l}} Q_{i\mathbf{j}} &= \prod_{p=1}^M \sum_{i_p+j_p=2l_p} Q_{i_p j_p}(m_p, n_p) \\ &= \prod_{p=1}^M \sum_{k_p \geq l_p} \frac{(-1)^{k_p+l_p}}{l_p!} \binom{k_p}{l_p} F_{k_p}(m_p, n_p) \\ &= \sum_{l_1 \leq k_1 \leq m_1, \dots, l_M \leq k_M \leq m_M} \prod_{p=1}^M \frac{(-1)^{k_p+l_p}}{l_p!} \binom{k_p}{l_p} F_{k_p}(m_p, n_p) \\ &= \sum_{\mathbf{k} \geq \mathbf{l}} \frac{(-1)^{|\mathbf{k}|+|\mathbf{l}|}}{\mathbf{l}!} \binom{\mathbf{k}}{\mathbf{l}} F_{\mathbf{k}}, \end{aligned} \quad (181)$$

where we used the feasibility of $(Q(m_p, n_p), \mathbf{F}(m_p, n_p))$ in the second line. Finally,

$$\begin{aligned} \sum_{\mathbf{k} \leq \mathbf{m}} F_{\mathbf{k}} &= \sum_{k_1 \leq m_1, \dots, k_M \leq m_M} \prod_{i=1}^M F_{k_i}(m_i, n_i) \\ &= \prod_{i=1}^M \sum_{k_i=0}^{m_i} F_{k_i}(m_i, n_i) \\ &= 1, \end{aligned} \quad (182)$$

since $\sum_{k=0}^m F_k(m, n) = 1$ for all $m, n \in \mathbb{N}$ with $m \geq n$. This shows that (Q, \mathbf{F}) is a feasible solution of $(\text{SDP}_{\mathbf{n}}^{m, \leq})$ (resp. strictly feasible). \square

A direct consequence of this construction is the following result:

Theorem 10 (Generalisation of Theorem 2). *Strong duality holds between the programs $(\text{SDP}_{\mathbf{n}}^{m, \leq})$ and $(D\text{-SDP}_{\mathbf{n}}^{m, \leq})$.*

Proof. The proof of Theorem 2 gives a strictly feasible solution $(Q(m, n), \mathbf{F}(m, n))$ of $(\text{SDP}_{\mathbf{n}}^{m, \leq})$ for all $m \geq n$. By Lemma 18, the program $(\text{SDP}_{\mathbf{n}}^{m, \leq})$ thus has a strictly feasible solution, for all $\mathbf{m} \geq \mathbf{n}$. By Slater condition, this implies that strong duality holds between the programs $(\text{SDP}_{\mathbf{n}}^{m, \leq})$ and $(D\text{-SDP}_{\mathbf{n}}^{m, \leq})$. \square

In particular, strong duality holds between the programs $(\text{SDP}_{\mathbf{n}}^{m1,\leq})$ and $(\text{D-SDP}_{\mathbf{n}}^{m1,\leq})$. Note that the multimode generalisation of Theorem 3 is a direct consequence of Theorem 10:

Theorem 11 (Generalisation of Theorem 3). *Strong duality holds between the programs $(\text{SDP}_{\mathbf{n}}^{m,\geq})$ and $(\text{D-SDP}_{\mathbf{n}}^{m,\geq})$.*

Proof. the strictly feasible solution of $(\text{SDP}_{\mathbf{n}}^{m,\leq})$ derived in the proof of Theorem 10 yields a strictly feasible solution for $(\text{SDP}_{\mathbf{n}}^{m,\geq})$. With Slater condition, this shows again that strong duality holds between the programs $(\text{SDP}_{\mathbf{n}}^{m,\geq})$ and $(\text{D-SDP}_{\mathbf{n}}^{m,\geq})$. \square

In particular, strong duality holds between the programs $(\text{SDP}_{\mathbf{n}}^{m1,\geq})$ and $(\text{D-SDP}_{\mathbf{n}}^{m1,\geq})$.

We recall the following definition from the main text: for all $n \in \mathbb{N}$, $\mathbf{F}^n = (F_k^n)_{k \in \mathbb{N}} \in \mathbb{R}^{\mathbb{N}}$ where

• if n is even:

$$F_k^n := \begin{cases} \frac{1}{2^n} \binom{k}{\frac{k}{2}} \binom{n-k}{\frac{n-k}{2}} & \text{when } k \leq n, k \text{ even,} \\ 0 & \text{otherwise,} \end{cases} \quad (183)$$

• if n is odd:

$$F_k^n := \begin{cases} \frac{1}{2^n} \frac{\binom{n}{\lfloor \frac{n}{2} \rfloor} \binom{\lfloor \frac{n}{2} \rfloor}{\lfloor \frac{k}{2} \rfloor}}{\binom{n}{k}}, & \text{when } k \leq n, \\ 0 & \text{otherwise.} \end{cases} \quad (184)$$

Let us define, for all $\mathbf{n} = (n_1, \dots, n_M) \in \mathbb{N}^m$, $\mathbf{F}^{\mathbf{n}} = (F_{\mathbf{k}}^{\mathbf{n}})_{\mathbf{k} \in \mathbb{N}^M} \in \mathbb{R}^{\mathbb{N}^M}$ where

$$F_{\mathbf{k}}^{\mathbf{n}} := \begin{cases} \prod_{i=1}^M F_{k_i}^{n_i} & \text{when } \mathbf{k} \leq \mathbf{n}, \\ 0 & \text{otherwise.} \end{cases} \quad (185)$$

By (60), for all $n \in \mathbb{N}$, $F_n^n \geq \frac{1}{n+1}$, so for all $\mathbf{n} = (n_1, \dots, n_M) \in \mathbb{N}^M$,

$$F_{\mathbf{n}}^{\mathbf{n}} \geq \frac{1}{\pi_{\mathbf{n}}}. \quad (186)$$

Like in the single-mode case, the program $(\text{SDP}_{\mathbf{n}}^{m,\leq})$ is equivalent to

$$(\text{SDP}_{\mathbf{n}}^{m,\leq}) \quad \begin{cases} \sup_{\mathbf{F} \in \mathbb{R}^{\tau_m}} F_{\mathbf{n}} \\ \text{subject to} & \sum_{\mathbf{k} \leq \mathbf{m}} F_{\mathbf{k}} = 1 \\ \text{and} & \forall \mathbf{k} \leq \mathbf{m}, \quad F_{\mathbf{k}} \geq 0 \\ \text{and} & \sum_{\mathbf{k} \leq \mathbf{m}} F_{\mathbf{k}} \mathcal{L}_{\mathbf{k}} \in \mathcal{R}_{m,+}(\mathbb{R}_+^M), \end{cases}$$

with the dual program given by

$$(\text{D-SDP}_{\mathbf{n}}^{m,\leq}) \quad \begin{cases} \inf_{y, \mu \in \mathbb{R} \times \mathbb{R}^{\tau_m}} y \\ \text{subject to} & y \geq 1 + \mu_{\mathbf{n}} \\ \text{and} & \forall \mathbf{k} \leq \mathbf{m}, \mathbf{k} \neq \mathbf{n}, \quad y \geq \mu_{\mathbf{k}} \\ \text{and} & \forall g \in \mathcal{R}_{m,+}(\mathbb{R}_+^M), \left\langle \sum_{\mathbf{k} \leq \mathbf{m}} \mu_{\mathbf{k}} \mathcal{L}_{\mathbf{k}}, g \right\rangle \geq 0, \end{cases}$$

for all $\mathbf{m} \geq \mathbf{n}$. Moreover, adding the condition $\mu_{\mathbf{k}} \leq 1$ for all $\mathbf{k} \leq \mathbf{m}$ does not change the optimal value of the program. We enforce this condition in what follows. With Lemma 7 and Lemma 18, we thus obtain the following result:

Lemma 19 (Generalisation of Lemma 7). For all $m \geq n$, \mathbf{F}^m (defined in Eq. (185)) is a feasible solution of $(SDP_n^{m,\leq})$.

In particular, for all $\mathbf{m} \in \mathbb{N}^M$, $\sum_{k \leq m} F_k^m \mathcal{L}_k \in \mathcal{R}_{m,+}(\mathbb{R}_+^M)$. For $\mathbf{m}, \mathbf{n} \in \mathbb{N}^M$ with $\mathbf{m} \geq \mathbf{n}$, let $\boldsymbol{\mu} \in \mathbb{R}^{\pi m}$ be a feasible solution of $(D\text{-}SDP_n^{m,\leq})$. Then, for all $\mathbf{l} \leq \mathbf{m}$

$$\left\langle \sum_{k \leq m} \mu_k \mathcal{L}_k, \sum_{k \leq l} F_k^l \mathcal{L}_k \right\rangle \geq 0, \quad (187)$$

so that

$$\sum_{k \leq l} \mu_k F_k^l \geq 0. \quad (188)$$

Hence, for all $\mathbf{l} \geq \mathbf{m}$,

$$\begin{aligned} \mu_l &\geq -\frac{1}{F_l^l} \sum_{\substack{k \leq l \\ k \neq l}} \mu_k F_k^l \\ &\geq -\frac{1}{F_l^l} \sum_{\substack{k \leq l \\ k \neq l}} F_k^l \\ &= 1 - \frac{1}{F_l^l} \\ &\geq 1 - \pi_l, \end{aligned} \quad (189)$$

where we used $F_l^l > 0$ in the first line, $\mu_k \leq 1$ and $F_k^l \geq 0$ in the second line, $\sum_{k \leq l} F_k^l = 1$ in the third line, and Eq. (186) in the last line.

With these additional results, the proof of convergence of the multimode hierarchy of lower bounds $(SDP_n^{m1,\leq})$ is then obtained directly from its single-mode counterpart using multi-index notations:

Theorem 12 (Generalisation of Theorem 5). The increasing sequence of optimal values $\omega_n^{m1,\leq}$ of $(SDP_n^{m1,\leq})$ converges to the optimal value ω_n of (LP_n) :

$$\lim_{m \rightarrow +\infty} \omega_n^{m1,\leq} = \omega_n. \quad (190)$$

With Lemma 17, we also obtain

$$\lim_{m \rightarrow +\infty} \omega_n^{m,\leq} = \omega_n. \quad (191)$$

Like in the single-mode case, this result implies strong duality between the linear programs:

Theorem 13 (Generalisation of Theorem 6). Strong duality holds between the programs (LP_n) and $(D\text{-}LP_n)$.

I Bounds on threshold values of several witness for $n = 3$

Below, we provide tables of numerical upper bounds and lower bounds obtained on the threshold values for witnesses of the form:

$$\hat{\Omega}_{(a_1, a_2, a_3)} = a_1 |1\rangle\langle 1| + a_2 |2\rangle\langle 2| + a_3 |3\rangle\langle 3| \quad (192)$$

where $\forall i \in \{1, 2, 3\}$, $0 \leq a_i \leq 1$ and $\max_i a_i = 1$. We focused on these particular witnesses for experimental considerations as it is challenging to obtain fidelities with higher Fock states. We vary each a_i from 0 to 1 with a step of 0.1.

a_1	a_2	a_3	$\omega_{\mathbf{a}}^<$	$\omega_{\mathbf{a}}^>$	a_1	a_2	a_3	$\omega_{\mathbf{a}}^<$	$\omega_{\mathbf{a}}^>$	a_1	a_2	a_3	$\omega_{\mathbf{a}}^<$	$\omega_{\mathbf{a}}^>$
1.0	0.0	0.0	0.500	0.529	1.0	0.3	0.8	0.589	0.590	1.0	0.7	0.5	0.705	0.715
1.0	0.0	0.1	0.500	0.529	1.0	0.3	0.9	0.606	0.610	1.0	0.7	0.6	0.718	0.720
1.0	0.0	0.2	0.500	0.529	1.0	0.3	1.0	0.626	0.633	1.0	0.7	0.7	0.735	0.738
1.0	0.0	0.3	0.500	0.529	1.0	0.4	0.0	0.610	0.615	1.0	0.7	0.8	0.754	0.758
1.0	0.0	0.4	0.500	0.529	1.0	0.4	0.1	0.610	0.615	1.0	0.7	0.9	0.774	0.781
1.0	0.0	0.5	0.500	0.528	1.0	0.4	0.2	0.610	0.615	1.0	0.7	1.0	0.795	0.805
1.0	0.0	0.6	0.500	0.529	1.0	0.4	0.3	0.610	0.615	1.0	0.8	0.0	0.739	0.751
1.0	0.0	0.7	0.500	0.529	1.0	0.4	0.4	0.610	0.615	1.0	0.8	0.1	0.739	0.751
1.0	0.0	0.8	0.500	0.529	1.0	0.4	0.5	0.610	0.615	1.0	0.8	0.2	0.739	0.751
1.0	0.0	0.9	0.500	0.530	1.0	0.4	0.6	0.610	0.615	1.0	0.8	0.3	0.739	0.751
1.0	0.0	1.0	0.500	0.563	1.0	0.4	0.7	0.614	0.615	1.0	0.8	0.4	0.739	0.751
1.0	0.1	0.0	0.526	0.529	1.0	0.4	0.8	0.629	0.631	1.0	0.8	0.5	0.742	0.751
1.0	0.1	0.1	0.526	0.529	1.0	0.4	0.9	0.648	0.653	1.0	0.8	0.6	0.759	0.761
1.0	0.1	0.2	0.526	0.529	1.0	0.4	1.0	0.669	0.677	1.0	0.8	0.7	0.777	0.780
1.0	0.1	0.3	0.526	0.529	1.0	0.5	0.0	0.640	0.649	1.0	0.8	0.8	0.796	0.801
1.0	0.1	0.4	0.526	0.529	1.0	0.5	0.1	0.640	0.649	1.0	0.8	0.9	0.816	0.824
1.0	0.1	0.5	0.526	0.529	1.0	0.5	0.2	0.640	0.649	1.0	0.8	1.0	0.837	0.848
1.0	0.1	0.6	0.526	0.529	1.0	0.5	0.3	0.640	0.649	1.0	0.9	0.0	0.773	0.790
1.0	0.1	0.7	0.526	0.529	1.0	0.5	0.4	0.640	0.649	1.0	0.9	0.1	0.773	0.790
1.0	0.1	0.8	0.526	0.529	1.0	0.5	0.5	0.640	0.649	1.0	0.9	0.2	0.773	0.790
1.0	0.1	0.9	0.526	0.528	1.0	0.5	0.6	0.640	0.649	1.0	0.9	0.3	0.773	0.790
1.0	0.1	1.0	0.542	0.563	1.0	0.5	0.7	0.654	0.655	1.0	0.9	0.4	0.773	0.790
1.0	0.2	0.0	0.552	0.555	1.0	0.5	0.8	0.671	0.673	1.0	0.9	0.5	0.783	0.790
1.0	0.2	0.1	0.552	0.555	1.0	0.5	0.9	0.690	0.696	1.0	0.9	0.6	0.800	0.803
1.0	0.2	0.2	0.552	0.555	1.0	0.5	1.0	0.711	0.719	1.0	0.9	0.7	0.818	0.822
1.0	0.2	0.3	0.552	0.555	1.0	0.6	0.0	0.672	0.682	1.0	0.9	0.8	0.838	0.844
1.0	0.2	0.4	0.552	0.555	1.0	0.6	0.1	0.672	0.682	1.0	0.9	0.9	0.858	0.867
1.0	0.2	0.5	0.552	0.555	1.0	0.6	0.2	0.672	0.682	1.0	0.9	1.0	0.879	0.891
1.0	0.2	0.6	0.552	0.555	1.0	0.6	0.3	0.672	0.683	1.0	1.0	0.0	0.809	0.830
1.0	0.2	0.7	0.552	0.555	1.0	0.6	0.4	0.672	0.682	1.0	1.0	0.1	0.809	0.830
1.0	0.2	0.8	0.552	0.555	1.0	0.6	0.5	0.672	0.682	1.0	1.0	0.2	0.809	0.830
1.0	0.2	0.9	0.565	0.567	1.0	0.6	0.6	0.678	0.682	1.0	1.0	0.3	0.809	0.830
1.0	0.2	1.0	0.584	0.594	1.0	0.6	0.7	0.694	0.696	1.0	1.0	0.4	0.809	0.830
1.0	0.3	0.0	0.581	0.585	1.0	0.6	0.8	0.712	0.716	1.0	1.0	0.5	0.824	0.830
1.0	0.3	0.1	0.581	0.585	1.0	0.6	0.9	0.732	0.739	1.0	1.0	0.6	0.841	0.845
1.0	0.3	0.2	0.581	0.584	1.0	0.6	1.0	0.753	0.762	1.0	1.0	0.7	0.860	0.865
1.0	0.3	0.3	0.581	0.584	1.0	0.7	0.0	0.705	0.716	1.0	1.0	0.8	0.879	0.887
1.0	0.3	0.4	0.581	0.584	1.0	0.7	0.1	0.705	0.715	1.0	1.0	0.9	0.900	0.910
1.0	0.3	0.5	0.581	0.584	1.0	0.7	0.2	0.705	0.715	1.0	1.0	1.0	0.922	0.934
1.0	0.3	0.6	0.581	0.584	1.0	0.7	0.3	0.705	0.715					
1.0	0.3	0.7	0.581	0.584	1.0	0.7	0.4	0.705	0.715					

Table 2: Upper and lower bounds on the threshold values of witnesses of the form $\hat{\Omega}_{(1,a_2,a_3)}$.

a_1	a_2	a_3	$\omega_{\mathbf{a}}^{\leq}$	$\omega_{\mathbf{a}}^{\geq}$	a_1	a_2	a_3	$\omega_{\mathbf{a}}^{\leq}$	$\omega_{\mathbf{a}}^{\geq}$	a_1	a_2	a_3	$\omega_{\mathbf{a}}^{\leq}$	$\omega_{\mathbf{a}}^{\geq}$
0.0	1.0	0.0	0.500	0.546	0.3	1.0	0.4	0.594	0.603	0.6	1.0	0.8	0.767	0.779
0.0	1.0	0.1	0.500	0.551	0.3	1.0	0.5	0.617	0.629	0.6	1.0	0.9	0.792	0.806
0.0	1.0	0.2	0.500	0.546	0.3	1.0	0.6	0.642	0.655	0.6	1.0	1.0	0.817	0.849
0.0	1.0	0.3	0.500	0.546	0.3	1.0	0.7	0.667	0.683	0.7	1.0	0.0	0.680	0.701
0.0	1.0	0.4	0.517	0.546	0.3	1.0	0.8	0.693	0.712	0.7	1.0	0.1	0.680	0.702
0.0	1.0	0.5	0.543	0.559	0.3	1.0	0.9	0.719	0.741	0.7	1.0	0.2	0.680	0.702
0.0	1.0	0.6	0.569	0.588	0.3	1.0	1.0	0.747	0.789	0.7	1.0	0.3	0.691	0.702
0.0	1.0	0.7	0.597	0.618	0.4	1.0	0.0	0.570	0.603	0.7	1.0	0.4	0.710	0.714
0.0	1.0	0.8	0.625	0.650	0.4	1.0	0.1	0.570	0.600	0.7	1.0	0.5	0.729	0.735
0.0	1.0	0.9	0.654	0.682	0.4	1.0	0.2	0.579	0.603	0.7	1.0	0.6	0.750	0.759
0.0	1.0	1.0	0.683	0.740	0.4	1.0	0.3	0.600	0.607	0.7	1.0	0.7	0.772	0.782
0.1	1.0	0.0	0.505	0.551	0.4	1.0	0.4	0.622	0.631	0.7	1.0	0.8	0.794	0.806
0.1	1.0	0.1	0.505	0.551	0.4	1.0	0.5	0.644	0.653	0.7	1.0	0.9	0.817	0.830
0.1	1.0	0.2	0.505	0.551	0.4	1.0	0.6	0.667	0.679	0.7	1.0	1.0	0.842	0.870
0.1	1.0	0.3	0.517	0.551	0.4	1.0	0.7	0.692	0.705	0.8	1.0	0.0	0.722	0.744
0.1	1.0	0.4	0.542	0.556	0.4	1.0	0.8	0.717	0.733	0.8	1.0	0.1	0.722	0.744
0.1	1.0	0.5	0.567	0.583	0.4	1.0	0.9	0.743	0.763	0.8	1.0	0.2	0.722	0.744
0.1	1.0	0.6	0.593	0.610	0.4	1.0	1.0	0.769	0.813	0.8	1.0	0.3	0.724	0.744
0.1	1.0	0.7	0.619	0.641	0.5	1.0	0.0	0.604	0.632	0.8	1.0	0.4	0.741	0.745
0.1	1.0	0.8	0.647	0.671	0.5	1.0	0.1	0.604	0.638	0.8	1.0	0.5	0.760	0.764
0.1	1.0	0.9	0.675	0.701	0.5	1.0	0.2	0.610	0.632	0.8	1.0	0.6	0.779	0.786
0.1	1.0	1.0	0.704	0.750	0.5	1.0	0.3	0.629	0.635	0.8	1.0	0.7	0.800	0.809
0.2	1.0	0.0	0.519	0.556	0.5	1.0	0.4	0.650	0.658	0.8	1.0	0.8	0.822	0.832
0.2	1.0	0.1	0.519	0.556	0.5	1.0	0.5	0.672	0.681	0.8	1.0	0.9	0.844	0.857
0.2	1.0	0.2	0.522	0.556	0.5	1.0	0.6	0.694	0.705	0.8	1.0	1.0	0.867	0.891
0.2	1.0	0.3	0.544	0.556	0.5	1.0	0.7	0.717	0.729	0.9	1.0	0.0	0.765	0.787
0.2	1.0	0.4	0.567	0.578	0.5	1.0	0.8	0.742	0.756	0.9	1.0	0.1	0.765	0.787
0.2	1.0	0.5	0.592	0.605	0.5	1.0	0.9	0.767	0.784	0.9	1.0	0.2	0.765	0.787
0.2	1.0	0.6	0.617	0.633	0.5	1.0	1.0	0.793	0.831	0.9	1.0	0.3	0.765	0.787
0.2	1.0	0.7	0.643	0.660	0.6	1.0	0.0	0.641	0.665	0.9	1.0	0.4	0.774	0.787
0.2	1.0	0.8	0.669	0.690	0.6	1.0	0.1	0.641	0.665	0.9	1.0	0.5	0.791	0.795
0.2	1.0	0.9	0.697	0.722	0.6	1.0	0.2	0.641	0.665	0.9	1.0	0.6	0.810	0.815
0.2	1.0	1.0	0.725	0.770	0.6	1.0	0.3	0.660	0.665	0.9	1.0	0.7	0.829	0.836
0.3	1.0	0.0	0.542	0.575	0.6	1.0	0.4	0.679	0.685	0.9	1.0	0.8	0.850	0.860
0.3	1.0	0.1	0.542	0.575	0.6	1.0	0.5	0.700	0.708	0.9	1.0	0.9	0.872	0.883
0.3	1.0	0.2	0.550	0.575	0.6	1.0	0.6	0.722	0.732	0.9	1.0	1.0	0.894	0.913
0.3	1.0	0.3	0.572	0.580	0.6	1.0	0.7	0.744	0.756					

Table 3: Upper and lower bounds on the threshold values of witnesses of the form $\hat{\Omega}_{(a_1,1,a_3)}$.

a_1	a_2	a_3	$\omega_{\mathbf{a}}^<$	$\omega_{\mathbf{a}}^>$	a_1	a_2	a_3	$\omega_{\mathbf{a}}^<$	$\omega_{\mathbf{a}}^>$	a_1	a_2	a_3	$\omega_{\mathbf{a}}^<$	$\omega_{\mathbf{a}}^>$
0.0	0.0	1.0	0.377	0.428	0.3	0.4	1.0	0.529	0.547	0.6	0.8	1.0	0.734	0.752
0.0	0.1	1.0	0.401	0.437	0.3	0.5	1.0	0.561	0.583	0.6	0.9	1.0	0.775	0.793
0.0	0.2	1.0	0.429	0.452	0.3	0.6	1.0	0.594	0.621	0.7	0.0	1.0	0.463	0.521
0.0	0.3	1.0	0.458	0.474	0.3	0.7	1.0	0.628	0.657	0.7	0.1	1.0	0.490	0.522
0.0	0.4	1.0	0.488	0.510	0.3	0.8	1.0	0.667	0.694	0.7	0.2	1.0	0.521	0.542
0.0	0.5	1.0	0.519	0.533	0.3	0.9	1.0	0.707	0.733	0.7	0.3	1.0	0.555	0.580
0.0	0.6	1.0	0.550	0.572	0.4	0.0	1.0	0.426	0.480	0.7	0.4	1.0	0.594	0.615
0.0	0.7	1.0	0.583	0.605	0.4	0.1	1.0	0.451	0.485	0.7	0.5	1.0	0.635	0.653
0.0	0.8	1.0	0.616	0.642	0.4	0.2	1.0	0.480	0.501	0.7	0.6	1.0	0.676	0.692
0.0	0.9	1.0	0.649	0.678	0.4	0.3	1.0	0.511	0.528	0.7	0.7	1.0	0.717	0.733
0.1	0.0	1.0	0.390	0.440	0.4	0.4	1.0	0.543	0.567	0.7	0.8	1.0	0.758	0.774
0.1	0.1	1.0	0.414	0.448	0.4	0.5	1.0	0.576	0.602	0.7	0.9	1.0	0.800	0.815
0.1	0.2	1.0	0.441	0.464	0.4	0.6	1.0	0.610	0.637	0.8	0.0	1.0	0.475	0.535
0.1	0.3	1.0	0.471	0.486	0.4	0.7	1.0	0.649	0.674	0.8	0.1	1.0	0.504	0.535
0.1	0.4	1.0	0.501	0.519	0.4	0.8	1.0	0.689	0.713	0.8	0.2	1.0	0.537	0.562
0.1	0.5	1.0	0.532	0.549	0.4	0.9	1.0	0.729	0.752	0.8	0.3	1.0	0.576	0.597
0.1	0.6	1.0	0.565	0.588	0.5	0.0	1.0	0.439	0.494	0.8	0.4	1.0	0.617	0.633
0.1	0.7	1.0	0.597	0.623	0.5	0.1	1.0	0.464	0.497	0.8	0.5	1.0	0.658	0.672
0.1	0.8	1.0	0.631	0.659	0.5	0.2	1.0	0.494	0.513	0.8	0.6	1.0	0.700	0.713
0.1	0.9	1.0	0.665	0.696	0.5	0.3	1.0	0.525	0.544	0.8	0.7	1.0	0.742	0.754
0.2	0.0	1.0	0.402	0.453	0.5	0.4	1.0	0.558	0.583	0.8	0.8	1.0	0.784	0.808
0.2	0.1	1.0	0.426	0.460	0.5	0.5	1.0	0.592	0.617	0.8	0.9	1.0	0.825	0.849
0.2	0.2	1.0	0.454	0.476	0.5	0.6	1.0	0.631	0.655	0.9	0.0	1.0	0.488	0.549
0.2	0.3	1.0	0.484	0.499	0.5	0.7	1.0	0.671	0.693	0.9	0.1	1.0	0.518	0.549
0.2	0.4	1.0	0.515	0.530	0.5	0.8	1.0	0.711	0.732	0.9	0.2	1.0	0.559	0.578
0.2	0.5	1.0	0.547	0.570	0.5	0.9	1.0	0.752	0.773	0.9	0.3	1.0	0.600	0.613
0.2	0.6	1.0	0.579	0.604	0.6	0.0	1.0	0.451	0.506	0.9	0.4	1.0	0.642	0.652
0.2	0.7	1.0	0.613	0.640	0.6	0.1	1.0	0.477	0.509	0.9	0.5	1.0	0.684	0.693
0.2	0.8	1.0	0.646	0.676	0.6	0.2	1.0	0.507	0.527	0.9	0.6	1.0	0.726	0.736
0.2	0.9	1.0	0.685	0.714	0.6	0.3	1.0	0.540	0.564	0.9	0.7	1.0	0.768	0.779
0.3	0.0	1.0	0.414	0.466	0.6	0.4	1.0	0.573	0.598	0.9	0.8	1.0	0.810	0.821
0.3	0.1	1.0	0.439	0.505	0.6	0.5	1.0	0.613	0.635	0.9	0.9	1.0	0.852	0.866
0.3	0.2	1.0	0.467	0.488	0.6	0.6	1.0	0.653	0.673					
0.3	0.3	1.0	0.497	0.513	0.6	0.7	1.0	0.693	0.712					

Table 4: Upper and lower bounds on the threshold values of witnesses of the form $\hat{\Omega}_{(a_1, a_2, 1)}$.