

Resourcefulness of non-classical continuous-variable quantum gates

Massimo Frigerio, Antoine Debray, Nicolas Treps, and Mattia Walschaers

Laboratoire Kastler Brossel, Sorbonne Université, CNRS, ENS-Université PSL, Collège de France, 4 place Jussieu, F-75252 Paris, France
June 12, 2026

In continuous-variable quantum computation, identifying key elements that enable a quantum computational advantage is a long-standing issue. Starting from the standard results on the necessity of Wigner negativity, we develop a comprehensive and versatile approach in which the techniques of (s) -ordered quasiprobabilities are exploited to provide rigorous statements on the simulability of photonic quantum circuits consisting of previously characterized gates and thereby identifying the contribution of each quantum gate to the potential achievement of quantum computational advantage. This is achieved by means of an analysis of the so-called transfer function, allowing us to highlight the resourcefulness of a gate set. As such this technique can be straightforwardly applied to current continuous-variables quantum circuits, while also constraining the tolerable amount of losses above which any potential quantum advantage can be ruled out. We use (s) -ordered quasiprobability distributions on phase-space to capture the non-classical features in the protocol, and focus our technique entirely on the ordering parameter s . This allows us to highlight the resourcefulness and robustness to loss of a universal set of unitary gates comprising three distinct Gaussian gates and any non-Gaussian unitary gate, providing important insight on the role of non-Gaussianity.

1 Introduction

Ever since Shor’s algorithm [1] promised an exponential speed-up in the factorisation of prime numbers, quantum computing has captivated the scientific community. The laws and logic behind quantum physics make it possible to solve certain highly specific problems much more efficiently than what could be achieved with classical devices. While it is not unreasonable to expect that quantum computers hold huge potential as a new computational framework, so far actual algorithmic design has only scratched the surface [2].

Mattia Walschaers: mattia.walschaers@lkb.upmc.fr

A bit over a decade ago, a new approach to study the question of the computational advantage of quantum devices over classical ones was introduced. Specific protocols such as boson sampling [3], IQP [4], and random circuit sampling [5], were proposed as candidates for reaching a provable quantum computational advantage. The foundations of these works within computational complexity theory [6, 7] separates them from earlier works on quantum speed-ups in the sense that there are strong arguments to suggest that these protocols cannot be simulated efficiently by any classical algorithm. These works have ultimately led to a series of experiments claiming to have reached such a quantum computational advantage [8–11].

In the field of photonic quantum computation, the boson sampling protocol and its many variations [12–16] have drawn much attention. The race for a quantum computational advantage quickly imposed the question: when do noise and experimental imperfections make a setup efficiently simulable on a classical computer? This has led to applied results exploiting the fine details of boson sampling [17–20] and Gaussian boson sampling [21–25] setups. Furthermore, this sparked research activities on the simulation of photonic sampling problems in a more general context.

This more general line of research can be seen as a series of attempts to translate and generalise the Gottesman-Knill theorem [26] for bosonic systems. The first such generalisation dates back to the early days of continuous-variable quantum information processing, and it showed simply that a setup involving exclusively Gaussian states, operations, and measurements can be efficiently simulated [27]. For a highly ideal setup with pure states, unitary gates, and projective measurements, Hudson’s theorem [28, 29] teaches us that these Gaussian elements are the only ones that are described by a fully positive Wigner function [30]. When more general sampling setups are considered, allowing states to be mixed, measurements can be general positive operator-valued measure (POVMs), and gates to be described by general channels, it turns out that this positivity of the Wigner function is the element that makes a system efficiently simulable [31, 32]. In other words, the existence of negative regions in the Wigner functions that describe the system is a key resource for reaching a quantum computational advantage.

While such Wigner negativity might be a necessary resource, it clearly is not a sufficient one [33]. A series of recent works has used tools from bosonic codes to completely challenge the paradigm of Wigner negativity as a crucial resource [34, 35], relying instead on discrete-variable resources that come from the qubits that bosonic codes are encoding [36]. However, these sufficient resources may turn out to be more demanding than necessary. In this regard, other works have searched for new necessary resources for reaching a quantum computational advantage, leading to the identification of stellar rank and some non-Gaussian type of entanglement as key ingredients [37]. Other approaches have aimed to generalise the results of Wigner negativity to more general phase-space representations [20, 38].

Most of these recent results have focused primarily on the resources that are present in the initial states and/or the measurements. The resourcefulness of the operations that are implemented in between (typically quantum gates) have not received much attention. In works such as [20, 34] these operations are limited to a very specific class of circuits, or one directly considers measurements on some generic state as in [37]. These works therefore assume, either explicitly or implicitly, that a straightforward approach can be taken when addressing big multimode channels connecting input states to measurements. This is in strong contrast to [31], where a lot of attention is devoted to the development of a Markov-chain model to simulate the action of a sequence of few-mode quantum channels. This can effectively be understood as a hidden-variable model for quantum computing which allows, at least in principle, to study a very wide set of operations.

In the present work, we provide insight and results on the simulatability of quantum gates by arguing, in contrast to [20], that the ordering parameters can be chosen in an optimal way that also allows to treat quantum gates in a decoupled, independent manner. Our gate-by-gate model allows for the precise identification of the sources of quantum computational advantage within a circuit and its tolerance to local losses. As such, we introduce an algorithmic approach to derive the best decomposition of a quantum circuit in terms of (s) -ordered quasiprobability distributions, we demonstrate the significant insights that this can yield. It is worth highlighting that our approach effectively looks for an optimal hidden-variable model for the computation, by optimizing the choice of s -ordering parameters in the representation of subsequent layers of operations. Moreover, this algorithm allows us to effectively deal with losses or noise at precise points in the circuit. Even if we find that a certain protocol is impossible to simulate efficiently, because no appropriate (s) -parameters can be found, our approach allows us to identify how much loss or noise can be tolerated before the setup can be effi-

ciently simulated.

The key results of our work lie in the second part, Sec. 4 and Sec. 5, where we devote attention to specific types of quantum gates and channels that are highly relevant in photonic quantum computing. Through an examination of their respective transfer functions, and the manner in which each function fits within the decoupled approach outlined in section 3, it becomes possible to pinpoint the precise sources of quantum computational advantage within a circuit, or, conversely, where such advantage is crucially lost. This notably allows to understand the role of squeezing operations, photon addition and subtraction, and any deterministic non-Gaussian gate, such as the cubic phase gate, in reaching a quantum computational advantage.

We stress that the primary aim of this work is not to introduce a new classical simulation algorithm, but rather to establish rigorous, gate-level conditions under which a continuous-variable quantum computation can be efficiently simulated classically. Algorithm 1 serves as a concise formalism to state and organise these conditions: by expressing the simulability of the full circuit as the conjunction of local positivity constraints on (s) -ordered transfer functions, it reduces a global problem to a tractable sequence of single-gate analyses. It is in this analytical reduction, and in the concrete characterisation of the resulting conditions for a universal gate set, that the contribution of this work lies.

2 Phase-space Representations

2.1 (s) -ordered representations of states and observables

The main theoretical tool on which our results rely consists of (s) -ordered quasiprobability distributions (quasi-PDF) on phase-space. Let us consider a Fock space of M bosonic modes, each associated with its pair of creation and annihilation operators $\hat{a}_j^\dagger, \hat{a}_j$ such that $[\hat{a}_j, \hat{a}_k^\dagger] = \delta_{jk}$. We denote by $\hat{\mathbf{a}} = (\hat{a}_1, \dots, \hat{a}_M)$ the corresponding vector of operators. For a bounded operator $\hat{O}$ on such space, and for any real number $s \in [-1, 1]$, we then define the (s) -ordered phase-space representation [30, 39]

$$W_{\hat{O}}^{(s)}(\boldsymbol{\alpha}) = \int_{\boldsymbol{\xi} \in \mathbb{C}^M} \frac{d^{2M} \boldsymbol{\xi}}{\pi^{2M}} \chi_{\hat{O}}^{(s)}(\boldsymbol{\xi}) e^{\boldsymbol{\alpha} \boldsymbol{\xi}^\dagger - \boldsymbol{\xi} \boldsymbol{\alpha}^\dagger}, \quad (1)$$

where $\boldsymbol{\xi}, \boldsymbol{\alpha} \in \mathbb{C}^M$ are complex vectors. The (s) -ordered characteristic function $\chi_{\hat{O}}^{(s)}(\boldsymbol{\xi})$ is defined by

$$\chi_{\hat{O}}^{(s)}(\boldsymbol{\xi}) = \text{Tr} [\hat{O} \hat{D}_s(\boldsymbol{\xi})], \quad (2)$$

with $\hat{D}_s(\boldsymbol{\xi})$ the (s) -ordered displacement operator $\hat{D}_s(\boldsymbol{\xi}) = \hat{D}(\boldsymbol{\xi}) e^{\boldsymbol{\xi} s \boldsymbol{\xi}^\dagger / 2}$, which is a generalisation of

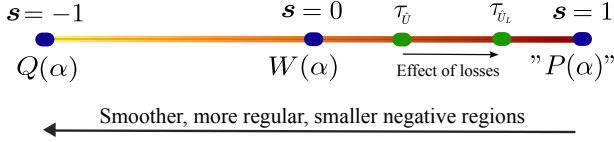


Figure 1: The range of values of the (s) -parameter. The values corresponding to the Q -, P -, and Wigner function are highlighted. We also visually represent the effect of Eq. (18), which describes how losses drive the nonclassical parameter closer to one.

the standard M -modes displacement operator $\hat{D}(\xi) = \exp\{\xi\hat{a}^\dagger - \xi^\dagger\hat{a}\}$. These (s) -ordered representations are useful tools to study CV quantum systems due to the identity [30, 39]

$$\text{Tr}[\hat{O}_1\hat{O}_2] = \int_{\alpha \in \mathbb{C}^M} \frac{d^{2M}\alpha}{\pi^M} W_{\hat{O}_1}^{(-s)}(\alpha) W_{\hat{O}_2}^{(s)}(\alpha). \quad (3)$$

We emphasize the requirement of pairing a quasiprobability distribution with parameter s with one with parameter $-s$. As indicated in Fig. 1, the higher the value of s , the more singular the distributions. Vice-versa, lower values of s indicate more regular phase-space representations.

Phase-space representations are often used to characterise quantum states, in which case we choose $\hat{O}$ in Eq. (1) to be the density matrix $\hat{\rho}$. The (s) -ordered phase-space representation then becomes a quasiprobability distribution, meaning that it is normalised, but not necessarily positive. For the case $s = 0$, we recover the Wigner function. For $s = -1$, instead, we have the Husimi Q -function, which is always non-negative and normalised to 1, being a proper probability density function (PDF) associated with non-orthogonal projectors onto coherent states. On the opposite hand, for $s = 1$, the integral won't always converge to some regular function and the result is instead defined as a distribution, i.e. a linear functional on a space of test functions, and it bears the name of Glauber's P -representation.

Importantly, it is possible to transform a (s) -ordered representation into a (s') -ordered representation with $s' < s$ by convoluting it with a Gaussian function:

$$W_{\hat{O}}^{(s')}(\alpha) = \frac{2}{s-s'} \int_{\beta \in \mathbb{C}} \frac{d^2\beta}{\pi} e^{-\frac{2|\alpha-\beta|^2}{s-s'}} W_{\hat{O}}^{(s)}(\beta). \quad (4)$$

Clearly then, $W_{\hat{O}}^{(s')}(\alpha)$ will be more regular than $W_{\hat{O}}^{(s)}(\alpha)$ for $s' < s$ [40]. If we are dealing with a quantum state, $\hat{O} = \hat{\rho}$, with a negative Wigner function ($s = 0$), looking for the highest value of $s' < 0$ such that $W_{\hat{\rho}}^{(s')}$ is a PDF amounts to finding the narrowest Gaussian kernel that will erase all the negative regions in the original Wigner function. As such, it is a means of quantifying Wigner negativity, referred to as the non-classical depth, which will be addressed in the

following paragraph. Unlike most common measures of Wigner negativity, such as the negativity volume, though, it also takes into account how the negative regions are shaped in phase-space: a very elongated negativity region would require a much broader Gaussian to be fully cancelled, when compared to another region with the same negative volume but concentrated in a more symmetric area. To quantify this idea, we can use the negativity volume:

$$\mathcal{N}^{(s)}(\hat{\rho}) = \int_{\alpha \in \mathbb{C}^M} d^{2M}\alpha |W_{\hat{\rho}}^{(s)}(\alpha)| - 1. \quad (5)$$

Combining with Eq. (4) and using the fact that the modulus of an integral is always majorized by the integral of the modulus, we directly find that

$$\mathcal{N}^{(s')}(\hat{\rho}) \leq \mathcal{N}^{(s)}(\hat{\rho}) \quad \text{for all } s' \leq s. \quad (6)$$

More generally, also for Wigner-positive states, we can define τ as the maximal value for which $\mathcal{N}^{(\tau)}(\hat{\rho}) = 0$, and thus $W_{\hat{\rho}}^{(\tau)}(\alpha)$ is a PDF. Given that the Husimi Q -function is obtained by projecting $\hat{\rho}$ onto coherent states, it is clear that such a $\tau \geq -1$ always exists. This value τ defines the *nonclassical depth* [41] of the state $\hat{\rho}$:

$$\mathfrak{t}[\hat{\rho}] := \frac{1-\tau}{2} \quad (7)$$

All non-Gaussian pure states have nonclassical depth $\mathfrak{t} = 1$ (meaning that only their Q -function is a PDF). For Gaussian states, since they have positive Wigner functions, $\mathfrak{t} \leq 1/2$, and for the particular case of coherent states we obtain $\mathfrak{t} = 0$. The latter are the only pure states whose P -function is a PDF, thus they are commonly referred to as the only classical pure state. The nonclassical depth can be employed as a measure of nonclassicality of quantum states and is particularly informative for mixed states and noisy quantum processes.

Up to now, we have only considered the case with a single ordering parameter. However, dealing with multimode states, it is much fruitful to define the $(\mathbf{s})$ -ordered phase-space representation such that $\mathbf{s} = \text{diag}(s_1, s_2, \dots, s_M)$ is a diagonal matrix, in such a way that each mode has its own independent ordering parameter, instead of choosing a single ordering parameter for all modes. This simply boils down to replacing $\hat{D}_{\mathbf{s}}(\xi)$ with $\hat{D}_{\mathbf{s}}(\xi) = \hat{D}(\xi)e^{\xi\mathbf{s}\xi^\dagger/2}$ in Eq. (2). Consequently, the $(\mathbf{s})$ -ordered representation of any operator $\hat{O}$ is designated by $W_{\hat{O}}^{(\mathbf{s})}$. For $\mathbf{s} = -\mathbb{1}_M$, $\mathbf{s} = 0_M$ and $\mathbf{s} = \mathbb{1}_M$ one finds respectively the Husimi Q -function, the Wigner function and the Glauber-Sudarshan P -function.

2.2 (s) -ordered representations of quantum operations

We first consider a generic quantum optical scheme: M input bosonic modes, described by a density op-

erator $\hat{\rho}_{in}$, are processed by a trace-preserving completely positive (CPTP) quantum channel $\mathcal{E}$. The overall output state $\hat{\rho}_{out} = \mathcal{E}(\hat{\rho}_{in})$ is then measured by a *factorized* POVM $\hat{\Pi}_{\mathbf{x}}$, $\mathbf{x}$ being the generic label for each possible measurement result. It was shown in [20] that the probability to obtain the outcome $\mathbf{x}$ is given by

$$p(\mathbf{x}) = \iint_{\alpha, \beta \in \mathbb{C}^M} \pi^M d^{2M} \alpha d^{2M} \beta W_{\hat{\rho}_{in}}^{(s')}(\alpha) \times \Lambda_{\mathcal{E}}^{(-s', s)}(\alpha, \beta) W_{\hat{\Pi}_{\mathbf{x}}}^{(-s)}(\beta), \quad (8)$$

where $\Lambda_{\mathcal{E}}^{(s', s)}(\alpha, \beta)$ is the (s', s) -ordered transfer function of $\mathcal{E}$, defined as:

$$\Lambda_{\mathcal{E}}^{(s', s)}(\alpha, \beta) = \iint_{\zeta, \xi \in \mathbb{C}^M} \frac{d^{2M} \zeta}{\pi^M} \frac{d^{2M} \xi}{\pi^{2M}} e^{\beta \zeta^\dagger - \zeta \beta^\dagger + \xi \alpha^\dagger - \alpha \xi^\dagger} \times e^{\zeta s \zeta^\dagger / 2 + \xi s' \xi^\dagger / 2} \text{Tr}[\mathcal{E}(\hat{D}^\dagger(\xi)) \hat{D}(\zeta)]. \quad (9)$$

As such, the transfer function $\Lambda_{\mathcal{E}}^{(s', s)}(\alpha, \beta)$ is a phase-space representation of the trace-preserving quantum operation $\mathcal{E}$. Moreover, the transfer function in Eq. (9) comes with an interesting additional property that derives from the fact that it represents a CPTP map:

$$\int_{\beta \in \mathbb{C}^M} d^{2M} \beta \Lambda_{\mathcal{E}}^{(s', s)}(\alpha, \beta) = 1. \quad (10)$$

In Section 3, we will use that (10) implies, whenever $\Lambda_{\mathcal{E}_i}^{(s', s)}$ is a regular and positive function, that we can interpret $\Lambda_{\mathcal{E}_i}^{(s', s)}(\alpha, \beta)$ as a probability distribution for β , for every possible value of α .

For a general multimode system Eq. (9) is usually very hard to work out explicitly. This should not come as a surprise, given that, for setups with factorized multi-mode input states and single-mode measurements, the map $\mathcal{E}$ should provide the necessary entanglement to make the setup hard to simulate on a classical computer. It is therefore common to make some assumptions on $\mathcal{E}$.

In [20], the authors focus on boson sampling and, thus, concentrate primarily on the case where $\mathcal{E}_L$ is a passive linear-optics channel. In this case, it is possible to find an analytical expression for $\Lambda_{\mathcal{E}_L}^{(s', s)}(\alpha, \beta)$ by using that $\mathcal{E}_L(\hat{D}^\dagger(\xi)) = \hat{D}^\dagger(L^\dagger \xi)$. Here L is an $M \times M$ matrix with $L^\dagger L \leq \mathbb{1}_M$ that captures, both, the mixing of the input modes and losses. This led to the key result that $\Lambda_{\mathcal{E}_L}^{(s', s)}(\alpha, \beta)$ is positive if and only if s and s' satisfy the following condition

$$\mathbb{1}_M - L^\dagger L - s + L^\dagger s' L \geq 0. \quad (11)$$

When it is possible to find s and s' such that condition (11) is satisfied, and such that $W_{\hat{\rho}_{in}}^{(s')}(\alpha)$ and $W_{\hat{\Pi}_{\mathbf{x}}}^{(-s)}(\beta)$ are both positive and can therefore be treated as PDFs, one can efficiently simulate the sampling process.

Despite the interest of this result, the class of operations that can be described by such an $\mathcal{E}_L$ is rather limited by the fact that it is a linear-optical operation; in contrast, when we aim at studying quantum advantage in CV systems, it is crucial to include a wider range of operations that can induce squeezing and non-Gaussianity.

In this work, we will also make an assumption on $\mathcal{E}$, namely that the total channel is decomposed into many operations: $\mathcal{E} = \mathcal{E}_1 \circ \mathcal{E}_2 \circ \dots \circ \mathcal{E}_k$. In the context of quantum computing, we typically understand this as a decomposition into quantum gates, i.e. other CPTP quantum channels involving at most $n \ll M$ modes each. In Section 4 we restrict to the common cases of single- and two-mode operations, but the results of Section 3 can be applied more generally. To study this sequence of operations, we rely on the composition rule

$$\Lambda_{\mathcal{E}_1 \circ \mathcal{E}_2}^{(s', s)}(\alpha, \beta) = \int_{\gamma \in \mathbb{C}^M} d^{2M} \gamma \Lambda_{\mathcal{E}_1}^{(s', s'')}(\alpha, \gamma) \Lambda_{\mathcal{E}_2}^{(-s'', s)}(\gamma, \beta). \quad (12)$$

The fact that an operation only acts on n modes is reflected by the fact that these functions involve at most n components of α, β , and γ each, and behave like delta distributions on the remaining $M - n$ modes. Notice that s'' drops out from the left-hand side of the previous equation, according to the cancellation of the terms $e^{\pm s'' |\gamma|^2 / 2}$ in the $(\pm s'')$ -ordered displacement operators appearing in Eq. (9).

3 Simulation of Sampling Algorithms

In Fig. 2, we show a schematic representation of the full sampling protocol that is considered in our work: a factorised input state $\hat{\rho}_{in} = \hat{\rho}_1 \otimes \dots \otimes \hat{\rho}_M$ is sent through a CPTP channel $\mathcal{E} = \mathcal{E}_1 \circ \mathcal{E}_2 \circ \dots \circ \mathcal{E}_k$, and is ultimately measured by a set of single-mode detectors with POVM elements $\hat{\Pi}_{\mathbf{x}} = \hat{\Pi}_{x_1} \otimes \dots \otimes \hat{\Pi}_{x_M}$. In phase-space, we can generalise Eq. (8) to describe the probability to obtain a string of measurement outcomes $\mathbf{x} = (x_1, \dots, x_M)$, associated with POVM element $\hat{\Pi}_{\mathbf{x}}$, as

$$p(\mathbf{x}) = \pi^M \int \dots \int W_{\hat{\rho}_{in}}^{(s_1)}(\alpha_1) \Lambda_{\mathcal{E}_1}^{(-s_1, s_2)}(\alpha_1, \alpha_2) \Lambda_{\mathcal{E}_2}^{(-s_2, s_3)}(\alpha_2, \alpha_3) \dots \Lambda_{\mathcal{E}_k}^{(-s_k, s_{k+1})}(\alpha_k, \alpha_{k+1}) W_{\hat{\Pi}_x}^{(-s_{k+1})}(\alpha_{k+1}) d\alpha_1 \dots d\alpha_{k+1}. \quad (13)$$

Explicit calculation for Eq. (13) can be found in Appendix (A). Here, the factorized input state $\hat{\rho}_{in}$ of M modes can be represented by a factorized (s_1) -ordered quasi-PDF $W_{\hat{\rho}_{in}}^{(s_1)}(\alpha_1)$, and the measurement process is described by a quasi-PDF $W_{\hat{\Pi}_x}^{(-s_{k+1})}(\alpha_{k+1})$.

If there exists a set $\{s_i\}_{i \in [1, k+1]}$ so that each phase-space representation appearing in the decomposition (13) is regular and positive, the whole scheme can be classically simulated in an efficient way, using the same strategy as the one followed in [31], based on the fact that we can perform each step by sampling a probability distribution (see the second part of Algorithm 1). We thus seek to optimise each s_i , which will eventually lead us to Algorithm 1 to check whether a scheme that follows the general structure of Fig. 2 can be efficiently simulated by classical means.

We note that the gate-by-gate decomposition provides sufficient conditions for classical simulability: if positivity is achieved at every step, the full circuit is simulable. Conversely, the failure of Algorithm 1 does not imply that the circuit is non-simulable, since a different decomposition or a global optimization over all ordering parameters might still succeed. The trade-off is between tractability — our local conditions are analytically or efficiently computable for each characterised gate — and optimality, which would require a global search over an exponentially large parameter space. As discussed in the outlook, regrouping gates whose individual transfer functions are negative into larger blocks is a natural strategy to tighten the conditions, at the cost of increasing the dimensionality of the sampling space. One should keep into account that simulating the full unitary is believed to be hard task under the assumption that quantum computation is hard to simulate.

To explain this optimisation of the s_i parameters, we start with the input state. By virtue of Eq. ((6)), for any single-mode state $\hat{\rho}$ there exists a maximal value $\tau_{\hat{\rho}} \in [-1, +1]$ such that $W_{\hat{\rho}}^{(s)}$ is a probability-density function for any $s \leq \tau_{\hat{\rho}}$. The positivity of the Husimi Q-representation on top guarantees that such a parameter $\tau_{\hat{\rho}}$ can always be found.

The study of the transfer functions $\Lambda_{\mathcal{E}_i}^{(-s_i, s_{i+1})}(\alpha_i, \alpha_{i+1})$ is somewhat more subtle. We note, first of all, that we can easily derive a result similar to Eq. (4) for $s'_i \succ s_i$ and $s'_{i+1} \prec s_{i+1}$ (as

matrix inequalities):

$$\Lambda_{\mathcal{E}_i}^{(-s'_i, s'_{i+1})}(\alpha_i, \alpha_{i+1}) = \frac{2}{\det(s_{i+1} - s'_{i+1}) \det(s'_i - s_i)} \times \int \int_{\beta, \gamma \in \mathbb{C}^M} \frac{d^{2M} \beta d^{2M} \gamma}{\pi^{4M}} e^{-2(\alpha_i - \beta)^\dagger (s'_i - s_i)^{-1} (\alpha_i - \beta)} \times e^{-2(\alpha_{i+1} - \gamma)^\dagger (s_{i+1} - s'_{i+1})^{-1} (\alpha_{i+1} - \gamma)} \times \Lambda_{\mathcal{E}_i}^{(-s_i, s_{i+1})}(\beta, \gamma). \quad (14)$$

Since we are convoluting with a Gaussian, the function gets smoother and negative regions become shallower (see Eq. (4)). Due to the sign difference, the same happens for $s'_{i+1} \prec s_{i+1}$. In other words, the phase-space representation $\Lambda_{\mathcal{E}_i}^{(-s_i, s_{i+1})}$ behaves more and more like a positive kernel as we increase s_i and decrease s_{i+1} .

At the i th step in the for-loop of Algorithm 1, the input ordering parameters τ_i are fixed in the previous step, $i - 1$, to be as high as possible. As implied by Eq. (14), this also makes $\Lambda_{\mathcal{E}_i}^{(-\tau_i, s_{i+1})}$ as close to being regular and positive as possible, therefore we shall now focus just on the effect of the output parameters s_{i+1} to show the consistency of our algorithm.

This maximal value will then become τ_{i+1} . By virtue of Eq. (14), the procedure of maximising s_{i+1} in step i of Algorithm 1 guarantees the best possible choice for the first ordering parameters in the transfer function of step $i + 1$. Nevertheless, at any stage of the for-loop, it is possible that no appropriate ordering parameter can be found to guarantee that the transfer function is a positive kernel. In that case, the algorithm fails, and we cannot simulate (13) through sampling in this manner, one would need to find another way to compile the gates, as it is discussed in the outlook.

Let us note that the concept of 'maximizing a vector' is ambiguous. Therefore, in the context of multi-mode gates, where one is confronted with ambiguity, we sought to maximize the minimal value of the vector. One will find an illustration, and a discussion on this issue, in Section 5.1.2, where the beam-splitter is studied. The order in which the algorithm is applied can be impacted by this ambiguity. Initiating the process in reverse, beginning with the measurement, has the potential to yield divergent results when such a gate is encountered. Consequently, this can result in different constraints being imposed on the (s) -parameter at the last step.

If we do make it all the way to the final step, we find a fixed parameter τ_{k+1} for which we have to evaluate the POVM elements $W_{\hat{\Pi}_x}^{(-\tau_{k+1})}(\alpha_{k+1})$. The func-

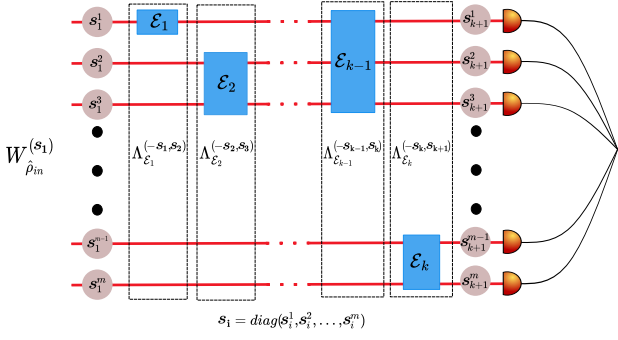


Figure 2: Generic quantum-optical scheme depicted by M input modes, described by a density operator $\hat{\rho}_{in}$ processed through a trace-preserving quantum channel $\mathcal{E}$, that can be decomposed into a sequence of trace-preserving quantum channels $\mathcal{E} = \mathcal{E}_1 \circ \mathcal{E}_2 \circ \dots \circ \mathcal{E}_k$. This produces the output state $\hat{\rho}_{out} = \mathcal{E}(\hat{\rho}_{in})$ and an output probability distribution $p(x) = \text{Tr}[\hat{\rho}_{out} \hat{\Pi}_x]$ sampled by measuring the POVM $\hat{\Pi}_x$.

tion $W_{\hat{\Pi}_x}^{(-\tau_{k+1})}$ is a product of single-mode functions $W_{\hat{\Pi}_x}^{(-\tau_{k+1}^{(j)})}$, with $\int_{x \in \mathcal{X}_j} dx \hat{\Pi}_x = \mathbb{1}$. The set $\mathcal{X}_j$ denotes the set of all possible measurement outcomes for the detector that measures the j th mode. This implies that every local detector is described by a set of phase-space representations $\{W_{\hat{\Pi}_x}^{(-s)}(\alpha) \mid x \in \mathcal{X}_j\}$ that satisfy the condition

$$\pi \int_{x \in \mathcal{X}_j} dx W_{\hat{\Pi}_x}^{(-\tau_{k+1}^{(j)})}(\alpha) = 1, \quad (15)$$

which must hold for all $\alpha \in \mathbb{C}$. For the simulation algorithm to work, we must be able to sample a final measurement outcome x that describes the click in all the detectors. When the resulting parameter τ_{k+1} is such that the functions $W_{\hat{\Pi}_x}^{(-\tau_{k+1}^{(j)})}(\alpha)$ are positive functions for all possible measurement $x \in \mathcal{X}_j$, for all the modes j , we can treat the functions $\pi W_{\hat{\Pi}_x}^{(-\tau_{k+1}^{(j)})}(\alpha)$ as probability distributions on the space $x \in \mathcal{X}_j$. If this is the case, we can sample the measurement outcomes for all the individual detectors and finally obtain x in a way that is consistent with probability density (13).

The previously described optimisation scheme for the ordering parameters explains the foundations of Algorithm 1. This algorithm effectively reduces the problem of the positivity of the quasiprobability distributions to a study of the ordering parameters. We are only concerned with those parameters and how they change during the computation. Indeed, Algorithm 1 fixes ordering parameters layer by layer, and we can understand this as an evolution of the initial ordering parameters (and thus nonclassical depth) of the input state throughout the different stages of the algorithm. Once the optimization part of the algorithm has identified the optimal ordering parameters, the second part of the algorithm, doing the

Algorithm 1: Classical simulation algorithm for a quantum computational scheme following the general structure of Figure (2). Notice that in the last step β_{k+1} is fixed and $\pi^M W_{\hat{\Pi}_x}^{(-\tau_{k+1})}(\beta_{k+1})$ is a PDF in x .

input : set of functions $W_{\hat{\rho}_{in}}^{(s_1)}, \{\Lambda_{\mathcal{E}_k}^{(-s_i, s_{i+1})}\}_{i \in [1, k]}, W_{\hat{\Pi}_x}^{(-s_{k+1})}$

output: a sample of measurement outcomes according to $p(\hat{x})$ (result of classical simulation), *failed* (classical simulation not possible)

begin

▷ First the optimization

find the max s_1 s.t. $W_{\hat{\rho}_{in}}^{(s_1)}(\alpha_1)$ is non-negative ;

let $\tau_1 := \max(s_1)$;

for $i \in [1, k]$ **do**

find max s_{i+1} s.t. $\Lambda_{\mathcal{E}_i}^{(-\tau_i, s_{i+1})}(\alpha_i, \alpha_{i+1})$ is non-negative ;

if no such s_{i+1} can be found **then**

return: *failed* ;

end

else

let $\tau_{i+1} := \max(s_{i+1})$;

end

end

if $W_{\hat{\Pi}_x}^{(-\tau_{k+1})}(\alpha_{k+1})$ is non-negative for τ_{k+1} **then**

proceed to sampling ;

end

else

return: *failed* ;

end

▷ Then the sampling

sample a point α_1 according to $W_{\hat{\rho}_{in}}^{(\tau_1)}(\alpha_1)$;

$\beta_1 := \alpha_1$;

for $i \in [1, k]$ **do**

sample a point α_{i+1} according to $\Lambda_{\mathcal{E}_i}^{(-\tau_i, \tau_{i+1})}(\beta_i, \alpha_{i+1})$;

$\beta_{i+1} := \alpha_{i+1}$;

end

sample a string of outcomes x according to $\pi^M W_{\hat{\Pi}_x}^{(-\tau_{k+1})}(\beta_{k+1})$;

return: x ;

end

sampling, can proceed in the same way as the one in Ref. [31]. Considerations on computational complexity of the sampling algorithm were already provided there, whereas the focus of our work is on the characterization of the gates transfer functions in terms of the ordering parameters and on their optimal choice.

We emphasise that when the optimisation phase succeeds, the second phase of Algorithm 1 generates full measurement outcome strings $\mathbf{x}$ by sequentially sampling from the certified positive probability density functions, following a Markov-chain strategy. This constitutes a bona fide classical simulation of the sampling problem, producing outcome strings distributed according to $p(\mathbf{x})$, and not merely an estimation of individual outcome probabilities.

The pseudo-algorithm presented here is intended to serve as a conceptual framework, to which mathematical considerations will be applied in the following sections. Nevertheless, the interested reader may refer to the explicit code provided [42] to check simulatability for arbitrary photonic quantum circuits involving a universal set of gates, specifically the one characterized in Section 5. Given such a circuit and the non-classical depths of the (multimode, factorized) input states, the code provides two possible outcomes: either the quest for positivity conditions fails at some point in the computation, or succeeds until the end. In the first instance, the code outputs the gate whose transfer function could not be made positive. In the latter case, the user is asked to input the nonclassical depths associated with the measurements POVM and the code checks the classical simulatability of the full scheme by comparing the latter with the final optimized s parameters. We emphasise that, for the universal gate set characterised in Sec. 5, the maximisation of the output ordering parameter at each step reduces to evaluating closed-form inequalities (Eqs. (25), (27)–(29), and Theorem 2), rather than requiring a generic numerical optimisation. The code [42] implements these analytic conditions and chains them sequentially through a user-specified circuit.

We should emphasise some key assumptions from the computational point of view. First of all, we assume that the circuit decomposition $\mathcal{E} = \mathcal{E}_1 \circ \mathcal{E}_2 \circ \dots \circ \mathcal{E}_k$ is known. Furthermore, we assume that for all the elements in the circuit (input states, gates, and measurements) the (s) -parameterised phase space representations in (13) are known (or can be calculated at a negligible overhead). Finally, we assume that it is computationally efficient to sample from these PDFs once they are positive. This final assumption may break down once the functions oscillate too rapidly, an effect which is in principle captured by the stellar rank [37].

To encapsulate, in broad terms, the fundamental principle underlying the present method we propose the following simplified thought experiment. Let us assume that a single-mode highly non-classical state

goes through a specific channel whose transfer function, $\Lambda_{\mathcal{E}}^{(-\tau_{in}, \tau_{out})}$, is positive only for τ_{out} arbitrarily close to 1. The inherent characteristic of the input state would force the first parameter to assume a value τ_{in} close to -1 . Indeed, if one wishes to erase all negativity from the Wigner function of such a state, one would have to proceed all the way to the smooth, completely positive Husimi Q -function by convolving said Wigner function with a broad Gaussian. As such, the channel has effectively transformed an input state that is highly non-classical into an output state that is highly classical; a greater range of quasi-PDFs behave as PDFs at the output than at the input. Such a channel would represent a point in the computation where all potential quantum advantage that might have been provided by the nonclassical input is effectively lost. This channel is ineffective in terms of providing a quantum advantages, as it effectively negates the potential benefits of imputing a Wigner-negative state. This highlights that it is crucial to understand the relation between the parameters τ_i and τ_{i+1} that are selected by Algorithm 1. Therefore, in the next section, we will examine different elementary gates and observe how the ordering parameter evolves through them.

4 Resourcefulness of Gates

4.1 Conditions for a general unitary

Before delving into less trivial examples, important conclusions on the ordering parameters can be drawn from the study of the identity channel. Since it acts independently on each mode, we will detail it for a single mode. Moreover, given that this channel leaves any input state unchanged, it cannot modify the non-classical depths. Consider then an input state $\hat{\rho}$ such that $W_{\hat{\rho}}^{(s_i)}$ is a regular probability density function and s_i is the largest possible value, i.e. corresponding to the nonclassical depth. If the transition kernel $\Lambda_{\mathbf{1}}^{(-s_i, s_{i+1})}$ were to be positive and regular for $s_{i+1} > s_i$, we would obtain as an output a function $W_{\hat{\rho}}^{(s_{i+1})}$ which is still regular and positive, contradicting the hypothesis that s_i was the largest value. This implies that a necessary condition for $\Lambda_{\mathbf{1}}^{(-s_i, s_{i+1})}$ to be regular and positive is that $s_{i+1} \leq s_i$, which is also a sufficient condition by Eq. (11) with $L = \mathbf{1}$ being the identity matrix. When this condition is met $\Lambda_{\mathbf{1}}^{(-s_i, s_{i+1})}$ is a Gaussian function in its input and output variables, while for $s_{i+1} > s_i$ it is a singular distribution attaining divergent values almost everywhere.

Likewise, it can be foreseen that for a single-mode unitary channel $\hat{U}$ the condition $s_{i+1} \leq s_i$ is still necessary in order for $\Lambda_{\hat{U}}^{(-s_i, s_{i+1})}$ to be regular and positive, otherwise $\hat{U}$ could increase the nonclassical depth of *any* input state with initial nonclassical depth $\mathfrak{t}$. Let us prove this formally:

Theorem 1. Let $\hat{U}$ be a single-mode unitary operator. If the associated $(-s_i, s_{i+1})$ -ordered transfer function $\Lambda_{\hat{U}}^{(-s_i, s_{i+1})}$ is a regular function, then $s_{i+1} \leq s_i$.

Proof. Consider the equality $\hat{U}\hat{U}^\dagger = \mathbb{1}$ rewritten in terms of s -ordered quasi-PDF:

$$\int_{\alpha_2 \in \mathbb{C}} d^2\alpha_2 \Lambda_{\hat{U}}^{(-s_i, s_{i+1})}(\alpha_1, \alpha_2) \Lambda_{\hat{U}^\dagger}^{(-s_{i+1}, s_{i+1})}(\alpha_2, \alpha_3) = \Lambda_{\mathbb{1}}^{(-s_i, s_{i+1})}(\alpha_1, \alpha_3). \quad (16)$$

Formally, we can write the right-hand side as:

$$\Lambda_{\mathbb{1}}^{(-s_i, s_{i+1})}(\alpha_1, \alpha_3) \propto \int_{\xi \in \mathbb{C}} d^2\xi e^{(\alpha_1 - \alpha_3)\xi^\dagger - (\alpha_1 - \alpha_3)^\dagger \xi} \times e^{\xi(s_{i+1} - s_i)\xi^\dagger/2}. \quad (17)$$

If $s_{i+1} > s_i$, the former expression diverges almost everywhere in α_1, α_3 and it is more singular than a delta distribution, while for $s_{i+1} = s_i$ it reduces to the delta distribution, imposing $\alpha_1 = \alpha_3$. Now if $\Lambda_{\hat{U}}^{(-s_i, s_{i+1})}$ were to be a regular, positive semidefinite function¹, the result of the left-hand side in Eq. (16) cannot be divergent, even if $\Lambda_{\hat{U}^\dagger}^{(-s_{i+1}, s_{i+1})}$ is a very singular distribution, which implies that indeed $\Lambda_{\hat{U}}^{(-s_i, s_{i+1})}$ can never be a regular probability density function if $s_{i+1} > s_i$. Notice that when $s_{i+1} = s_i$ we could have, in the best case, that $\Lambda_{\hat{U}}^{(-s_i, s_{i+1})}$ is as singular as a delta distribution and then Eq. (16) still holds even if the right-hand side diverges for $\alpha_1 = \alpha_3$. $\square$

4.2 Losses

Before providing the characterization of a universal set of unitary quantum gates, it is instructive to contrast the general result that we got for unitary channels with the prototypical example of dissipative quantum channels, i.e. losses which are crucial in modelling optical setups and deserves an additional discussion in its own right. The optical loss channel is implicitly already covered in Ref. [20] where Eq. (11) is derived to describe a linear optics network with losses. To consider optical losses in the spirit of our gate-based protocol of Fig. 2, we can just define a single-mode loss channel $\Lambda_{\mathcal{E}_\eta}^{(-s_i, s_{i+1})}(\alpha, \beta)$ as a single-mode version of Eq. (11) setting $L = \sqrt{\eta}$, where $\eta \in [0, 1]$ describes the efficiency of the channel ($\eta = 1$ reflects the absence of losses, while $\eta = 0$ means that the state is completely lost). We can from this find a decreasing linear relation with respect to s_i :

$$s_{i+1} \geq 1 - \eta(1 - s_i). \quad (18)$$

¹We shall assume that whenever an s -ordered quasiprobability is regular and non-negative, it also belongs to the Schwartz space of rapidly decreasing function, so that no issues of convergence can arise with respect to tempered distributions. This is typically the case, since this probability density functions decay as Gaussians for large absolute values of their arguments.

Since $s_i = (1 - \eta)s_i + \eta s_i < 1 - \eta(1 - s_i)$, this implies that we can always pick $s_{i+1} > s_i$ for $\eta < 1$ and, the smaller the value of η , the closer the maximum value of s_{i+1} approaches 1, as shown in Fig. 1. It can thus be concluded that losses always result in a more classical output, in stark contrast with the opposite result that we derived for unitary gates 16, thereby emphasizing the peculiar nature of losses for nonclassicality. This relatively straightforward result, moreover, allows us to understand how losses can hinder a possible quantum computational advantage. In a real circuit, loss channels of the type $\Lambda_{\mathcal{E}_\eta}^{(s_i, s_{i+1})}$ can occur everywhere, effectively increasing the chances of success of Algorithm 1 at any step. As is commonly done in quantum optics, we can now describe realistic operations and measurements by taking ideal unitary ones and adding a loss channel at the input and output. When we add them in between the gates and run Algorithm 1, we can now deal with mode-dependent losses, which are often overlooked in studies of sampling problems.

Finally, it is worth mentioning that losses allow us to consider in principle values of the s parameters outside the range $[-1, 1]$: indeed a coherent state going through a loss channel will become a displaced thermal state, whose s -ordered quasi-PDF will be a regular and positive function also for some values of $s > 1$, defined again through Eq. (1). In general, problems might arise because the s -ordered quasi-PDFs of channels and states for values of $s > 1$ are not guaranteed to exist, even in the sense of distributions, and they are not associated with an ordering of the creation and annihilation operators. However, if we have input states that are mixed and noisy, such that they can be parametrized with values of s larger than 1, the matching rule in Eq. (3) ensures that at the next step we will have an even more regular function (since s matches with $-s$) and we can possibly have more chances to find a positive decomposition of the whole circuit by going beyond the limitation of $s \in [-1, 1]$.

5 Characterization of a Universal Set of Unitary Gates

Algorithm 1 is based on the idea that we have at hand a decomposition of the full photonic quantum circuit into a set of previously characterized gates, such that the computational effort needed to compute the (s) -ordered quasi-PDFs for such elements constitutes just a fixed preliminary overhead with respect to the actual simulation cost of the computation through sampling. It is thus consequential to characterize a universal set of unitary gates, with which all CV unitary operations can be approximated efficiently [43]; one such set is provided by *two-mode Gaussian unitary gates* and one single-mode non-Gaussian gate, such as *the cubic phase gate* [44], whose characterizations

are provided in the following.

5.1 Gaussian unitaries

Among Gaussian unitary gates, displacements and single-mode phase shifters clearly do not affect the nonclassicality (hence they leave the value of the s -parameters unchanged), since they act as Euclidean transformations on quasi-PDFs. Thus, we only need to consider squeezers and beam splitters. We will show that the action of a squeezing gate can directly alter the s parameter of the output, whereas beam splitters can “mix” the s parameters of their input modes.

5.1.1 Squeezing gate

Starting with the squeezing gate, denoted by $\mathcal{S}$, the trace-preserving quantum gate $\mathcal{S}$ acting on the displacement operator is:

$$\mathcal{S}(\hat{D}^\dagger(\xi)) = \hat{S}^\dagger(r) \hat{D}^\dagger(\xi) \hat{S}(r) = \hat{D}^\dagger(\chi), \quad (19)$$

where $\hat{S}(r)$ is the squeezing operator for a squeezing parameter r and $\chi = \xi \cosh r + \xi^\dagger \sinh r$. Thus implementing a simple re-scaling of the displacement operator. For $\mathcal{S}$ the transfer function take the form

$$\begin{aligned} \Lambda_S^{(s_i, s_{i+1})}(\alpha, \beta) &= \iint \frac{d^2\zeta}{\pi^2} \frac{d^2\xi}{\pi} e^{\beta\zeta^\dagger - \zeta\beta^\dagger} e^{\zeta s_{i+1}\zeta^\dagger/2} e^{\xi\alpha^\dagger - \alpha\xi^\dagger} e^{\xi s_i \xi^\dagger/2} \\ &\quad \times \text{Tr}[\mathcal{S}(\hat{D}^\dagger(\xi)) \hat{D}(\zeta)]. \end{aligned} \quad (20)$$

We can write

$$\mathcal{S}(\hat{D}^\dagger(\xi)) = e^{\chi\chi^\dagger} \int \frac{d^2\gamma}{\pi} e^{\gamma\xi^\dagger - \xi\gamma^\dagger} |\gamma\rangle \langle\gamma|. \quad (21)$$

And we can use the normally ordered form of the displacement operator, $\hat{D}(\zeta) = e^{-\zeta\zeta^\dagger/2} e^{\zeta\hat{a}^\dagger} e^{-\hat{a}\zeta^\dagger}$ to obtain

$$\text{Tr}[\mathcal{S}(\hat{D}^\dagger(\xi)) \hat{D}(\zeta)] = \pi e^{\chi\chi^\dagger/2} e^{\zeta\zeta^\dagger/2} \delta(\zeta - \chi) \quad (22)$$

Thus the transfer function (20) becomes

$$\Lambda_S^{(-s_i, s_{i+1})}(\alpha, \beta) = \int d\xi e^{\chi^\dagger\beta - \chi\beta^\dagger + \frac{\chi s_{i+1}\chi^\dagger}{2}} e^{\xi^\dagger\alpha - \xi\alpha^\dagger + \frac{\xi s_i \xi^\dagger}{2}} \quad (23)$$

As one can notice from this form the transfer function is well behave and non-negative if and only if

$$\sigma^{-1} = \begin{aligned} &[(\xi \cosh r + \xi^\dagger \sinh r) s_{i+1} (\xi^\dagger \cosh r + \xi \sinh r) \\ &- s_i |\xi|^2]^{-1} \geq 0 \end{aligned} \quad (24)$$

Finally, we find that one can sample from the transfer function, and thus classically simulate this block if and only if

$$\begin{cases} s_{i+1} \leq \frac{s_i}{e^{2|r|}} & \text{if } s_i \geq 0 \\ s_{i+1} \leq \frac{s_i}{e^{-2|r|}} & \text{if } s_i \leq 0 \end{cases} \quad (25)$$

Therefore, in general, the squeezing channel requires us to parametrize the output expecting it to have a greater nonclassical depth (lower maximum value allowed for s) in order for the transfer function to be a positive transfer probability kernel. It is interesting to note that the conditions linearly bound one ordering parameter with respect to the other and they are fully symmetric under the exchange $-s_i \leftrightarrow s_{i+1}$, as can be seen more clearly from Fig. (3(a)), where we plot in gray the region in the (s_i, s_{i+1}) plane where $\Lambda_S^{(-s_i, s_{i+1})}$ is non-negative, for $r = 0.3$.

Notice that, for $s_i < 0$, the upper bound $s_i/e^{-2|r|}$ can be smaller than -1 , namely if $|r| > -\frac{1}{2} \ln |s_i|$, therefore for sufficiently nonclassical, Wigner negative input states the squeezing channel cannot be simulated through Algorithm 1, no matter how the output is parametrized. This observation provides a first hint of the fact that non-Gaussian resources are pivotal for possible quantum advantages.

It is crucial to stress that one could include the squeezing channel into the previous channel that was applied to the multi-mode system and try to simulate it as a single block with a positive transfer probability kernel, possibly finding less strict conditions for the output parameter s_{i+1} ; however, when the previous channel is multi-mode and potentially entangling, this forces us to apply the squeezing to a multi-mode state, converting the challenge of simulating the possibly negative squeezing transfer function into that of simulating a positive transfer function on a larger space.

A special point deserves now our attention, the output state of the squeezing channel may not be more non-classical than the input. It is evident that the conditions set out in Eq. (25) pertain to the positivity and well-behavedness of the transfer function. In the event that the aforementioned conditions are not met, the transfer function will exhibit negative region(s) and/or will be highly singular, thereby implying that one cannot sample from it efficiently [45]. Nevertheless, it is possible for the output state to be, in all respects and according to all characterisations, classical. This can be understood by imagining a highly squeezed input state going through a squeezing channel with the opposite squeezing parameter and resulting in a more classical output. Through the action of $\mathcal{S}$, we then squeeze in the orthogonal direction, by the same factor of r . In this case, the output state would be very classical. However, if the conditions set out in (25) are not met, it is not possible to sample from the transfer function describing the process.

5.1.2 Beam-splitter

We shall now proceed to identify the conditions under which the beam-splitter (BS) can be considered a resourceful gate. Assume that we encounter a BS

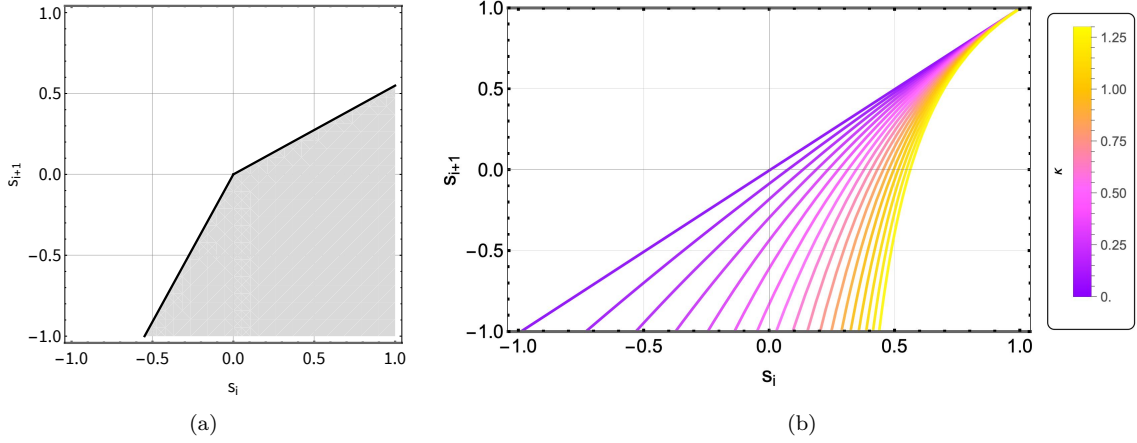


Figure 3: (a) On the squeezing gate: the region in the (s_i, s_{i+1}) plane where $\Lambda_S^{(-s_i, s_{i+1})}$ is non-negative according to Eq. (25) is displayed in gray, for $r = 0.3$. (b) On the single-photon subtraction: limit lines according to Eq. (32) for various values of $\alpha \in [0, 1]$, represented by different shadings. For each value of α , the allowed values of s_{i+1} to have a non-negative transfer functions of the BS lie under the respective curve.

gate on block i , where $1 < i \leq k$. In a formal sense, our objective is to construct a channel $\mathcal{E}^{BS}$ that acts as a beam-splitter transformation between two modes and trivially on the remaining $M - 2$ modes. Denoting by a and b the two modes on which the BS acts non-trivially, according to Eq. (11) the limiting condition relating output to input ordering parameters is obtained by taking L is a unitary matrix describing the beam-splitter transformation,

$$L = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix}, \quad (26)$$

and $\theta \in \mathbb{R}$. Potential phase factors $e^{i\phi}$, representing the phase shift between the input modes of the beam splitter, has not been taken into account since the phase shift channel has no impact on the ordering parameter, merely rotating phase-space distributions without altering their values. By virtue of unitarity, one needs to solve the following equation maximizing the output parameters:

$$\min \text{eig}(-\text{diag}(s_{i+1}^a, s_{i+1}^b) + L^\dagger \text{diag}(s_i^a, s_i^b) L) = 0. \quad (27)$$

While the relation between the input and output parameters is not very insightful at first glance, we do see that s_{i+1}^a will generally depend on both s_i^a and s_i^b (and analogously for s_{i+1}^b). Accordingly, the process of maximizing the output (s)-parameters for a given input is not well defined and a continuous range of solutions may exists, depending on the relative balance in the combined maximization. To keep consistency with the proposed algorithm we impose the symmetric condition $s_{i+1}^a = s_{i+1}^b$, which is equivalent to maximizing the minimum of the two. It is useful to notice that from Eq. (27) one can derive a necessary condition on the sum of the output s -parameters:

$$s_{i+1}^a + s_{i+1}^b \leq s_i^a + s_i^b - |s_i^a - s_i^b| \sin 2\theta \quad (28)$$

assuming $0 \leq \theta \leq \frac{\pi}{2}$. In particular, if one imposes $s_{i+1}^a = s_{i+1}^b$ for a balanced BS ($\theta = \pi/4$), one gets the necessary and sufficient condition for the positivity of the transfer function:

$$s_{i+1}^a = s_{i+1}^b \leq \min\{s_i^a, s_i^b\} \quad (29)$$

Thus the most nonclassical of the two inputs fully dictates the maximum value for the output ordering parameters.

5.2 Realistic protocol for single-photon subtraction

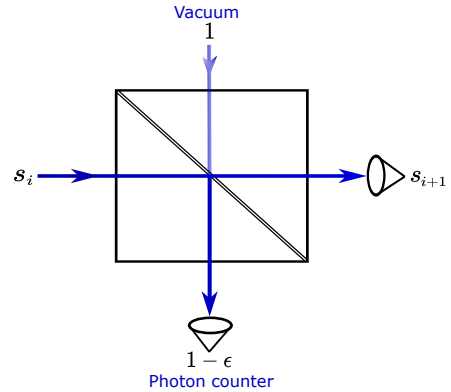


Figure 4: Realistic single-photon subtraction protocol seen from the perspective of our framework, where we considers solely the order parameter of the states on each branch.

In order to showcase the usefulness of the proposed protocol, we will apply it to the single-photon subtraction operation which is commonly considered as one of the most accessible de-Gaussifying procedures. We model realistic single-photon subtraction with a BS whose input modes are respectively the input state

on which we wish to subtract a photon on one branch, and the vacuum state on the other. The transmittance $\eta = \cos^2 \theta$ of the BS with respect to the first mode is close to 1, meaning that only a small fraction of the signal will be reflected. On this weak output signal, a photon counter with some finite efficiency will be implemented, corresponding to the ideal POVM ($\epsilon = 0$) with elements $\{\hat{\Pi}_1 = |1\rangle\langle 1|, \hat{\Pi}_0 = \mathbb{1} - |1\rangle\langle 1|\}$. At the last step, only the outcomes corresponding to $\hat{\Pi}_1$ are post-selected. This measurement will be modelled choosing an ordering parameter close to -1 , since the only value such that the s -ordered quasi-PDF of $|1\rangle\langle 1|$ is positive is -1 , being a projector on a highly non-classical state, but we shall assume that its value is $-1 + \epsilon$ to account for imperfect measurements. This proposed model is illustrated in Fig. 4.

This results in the following set of input-output or-

dering parameters:

$$\mathbf{s}_{in} = \begin{pmatrix} s_i & 0 \\ 0 & 1 \end{pmatrix}; \quad \mathbf{s}_{out} = \begin{pmatrix} s_{i+1} & 0 \\ 0 & 1 - \epsilon \end{pmatrix} \quad (30)$$

recalling that the output parameter for the mode that is measured has the opposite sign with respect to the ordering parameter that we had to choose for the highly nonclassical measurement. The resulting output parameter, s_{i+1} , with which we have to parameterize the output to have a positive transfer function for the beam-splitter will depend on the only free parameter : the ordering parameter of the input state. Clearly, if we take the limit $\theta \rightarrow 0$ for finite ϵ the result will be trivial, having $s_{i+1} \leq s_i$ as the only condition. Going back to Eq. (27), inserting the diagonal matrices above and imposing that the smallest eigenvalue is non-negative, for an ideal detection ($\epsilon = 0$) with generic mixing angle θ we find instead:

$$\lambda_{min} = \frac{1}{2} \left[s_i - s_{i+1} - \sqrt{(1 - s_i)^2 + (1 - s_{i+1})^2 - 2(1 - s_i)(1 - s_{i+1}) \cos 2\theta} \right] \geq 0. \quad (31)$$

For $-1 \leq s_{i+1} \leq s_i \leq 1$, this quantity is zero if θ is a multiple of π , including $\theta = 0$, and it is negative otherwise. Therefore we conclude that there exists no set of input-output parameters such that the transfer function is well-behaved in the case of perfect photodetection. This result is of great importance as it demonstrates that ideal single-photon detectors are sufficient to render this building-block impossible to simulate classically.

There exists however a nontrivial limit, tuning in on the scaling $\theta \propto \sqrt{\epsilon}$. In this case, taking the limit $\epsilon \rightarrow 0$ with $\theta = \kappa\sqrt{\epsilon}$ for a free tuning parameter $\kappa \in \mathbb{R}$, we can rewrite the condition $\lambda_{min} \geq 0$ as:

$$s_{i+1} \leq \frac{s_i - (1 - s_i)\kappa^2}{1 - (1 - s_i)\kappa^2}. \quad (32)$$

In Fig. 3(b) these curves for various values of κ are displayed, with smaller values corresponding to shades towards the violet. We notice that, for each value of $\kappa > 0$, there is a minimum s_i under which no solution exist (each curve reaches the value $s_{i+1} = -1$ for some s_i strictly greater than -1). Furthermore, for $\kappa > \frac{1}{2}$ this limiting value for s_i is strictly positive, meaning that the single-photon subtraction modelled by this limit cannot be described with a non-negative transfer function whenever we input a Wigner-negative state. Significantly though, this implies that if we input noisy squeezed states, at fixed detection efficiency and transmittance, there is a maximum ratio of squeezing to noise beyond which the transfer function will never be positive, rendering this block non-simulatable even if the input is Wigner-positive.

Finally, let us note that the above arguments are purely based on the order parameters and do not require any further details of the phase-space representations. This means that this reasoning can be extended to virtually all other measurement-based operations.

5.3 Non-Gaussian unitary gates

Having seen that only Wigner-negative inputs can force the squeezing channel to be represented only by a negative transfer function, we will now tackle non-Gaussian single-mode unitary channels, which are needed to enable universal quantum computation. The key result that we obtained is the following:

Theorem 2. *Let $\hat{U}$ be a single-mode unitary non-Gaussian gate. Then $\Lambda_{\hat{U}}^{(s_1, s_2)}(\alpha, \beta)$ cannot be an everywhere non-negative function for any $s_1, s_2 > -1$, while it is always non-negative for $s_1 = s_2 = -1$.*

Proof. Consider the application of $\hat{U}$ on coherent states. If each coherent state were mapped by $\hat{U}$ to some Gaussian state, it follows that every Gaussian state would be mapped to a Gaussian state via the Glauber representation theorem and the channel would be Gaussian. Since we supposed $\hat{U}$ to be non-Gaussian, there must then exist a coherent state $|\alpha_0\rangle$ such that $|\psi\rangle = \hat{U}|\alpha_0\rangle$ is non-Gaussian. Since it is pure, non-Gaussianity implies that it has maximal nonclassical depth and only its Husimi Q-function is a proper probability distributions, while all the other s -ordered quasiprobability representations of $|\psi\rangle$ are non-positive for any $s > -1$. On the other hand,

coherent states are P -classical and their Glauber P -distribution is a delta distribution, therefore we can choose $s_{in} = 1$ and $s_1 = -s_{in} = -1$. Then we have:

$$\begin{aligned} W_{|\psi\rangle}^{(s_2)}(\beta) &\propto \int \delta(\alpha - \alpha_0) \Lambda_{\hat{U}}^{(-1, s_2)}(\alpha, \beta) d^2\alpha = \\ &= \Lambda_{\hat{U}}^{(-1, s_2)}(\alpha_0, \beta) \end{aligned} \quad (33)$$

And since we showed that the left-hand side must attain negative values for some β whenever $s_2 > -1$, the first part of the thesis follows. Finally, notice that for $s_1 = s_2 = -1$ we have the very general result:

$$\Lambda_{\hat{U}}^{(-1, -1)}(\alpha, \beta) \propto |\langle \beta | \hat{U} | \alpha \rangle|^2 \quad (34)$$

therefore the $(-1, -1)$ -ordered transfer function is always non-negative. $\square$

5.3.1 The cubic phase gate

Having seen that only Wigner negative inputs can force the squeezing channel to be represented only by a negative transfer function, we can tackle an instance of a non-Gaussian single-mode unitary channel of utmost relevance in the theory of CV quantum computation: the cubic phase gate. It can be described by the unitary operator $\hat{U}_\gamma = e^{-i\frac{\gamma}{3!}\hat{q}^3}$, where $\gamma \in \mathbb{R}$ is known as the *cubicity*. It is well known [43, 46, 47] that, together with Gaussian gates, the cubic gate is sufficient to attain universality with CV quantum computation.

The explicit calculations are shown in Appendix (B), where the (s_1, s_2) -ordered transfer function is given in two versions, one involving a single integral with a cubic phase term and another involving a double convolution of the Airy function with Gaussian kernels. This example also showcases how our method can be applied to the characterization of more complex unitary gates.

6 Conclusions

6.1 Simulability of photonic circuits

Using the language of (s) -ordered quasiprobability distributions on optical phase-space, we detailed a procedure to systematically assess the possibility of framing a computational photonic quantum process in terms of probability distributions updated through positive-semidefinite transition kernels. Whenever such a decomposition is possible, the process can be simulated efficiently by a classical algorithm, namely by sampling the starting probability distribution and applying the steps of the algorithm in order to compute the output. In doing so, we go beyond the already established necessity of Wigner negativity by enlarging the class of functions, allowing us to unveil bounds in which new quantum resources may be

identified. In particular, our method is able to determine the exact amount of losses that are sufficient, at each step in the quantum computation, to make the corresponding process efficiently simulatable. We emphasize that the strength of our method heavily relies on the fact that the conditions for positivity of each transfer function are determined just by the ordering parameters, and therefore only by the non-classical depth of the input and not on the specific quantum state. This property allows us to explicitly derive these positivity conditions for all Gaussian unitary channels, as well as for non-Gaussian unitary channels for ordering parameters in the range $[-1, 1]$. In particular, a single-mode squeezing channel can always be implemented by a positive transfer function by some choice of the second ordering parameter, unless the input is Wigner negative. Studying how a beam splitter mixes the nonclassicalities of the input and fixing a maximally nonclassical measurement, we deduce that (realistic) single-photon subtraction is not efficiently simulatable if the squeezing of the input is sufficiently high. Finally, we tackled the harder computation of the (s_1, s_2) -ordered function of the cubic gate to showcase the power of our method to provide quantitative answers also for nontrivial non-Gaussian gates. We emphasize that, in the most general setting, we always assume that definite sets of gates, input states and measurements have been fixed and their (s) -ordered QPDs have been characterized. Even when those steps imply a non-negligible numerical cost, it is crucial that this is needed only once for each type of element (state, gate or measurement) and it can thus be considered as a constant overhead to the simulation cost.

6.2 Loss tolerance of photonic quantum gates and their resourcefulness

The presented techniques and results can not only directly be applied to existing photonic quantum circuits to test their simulability beyond the requirement of Wigner negativity, but they also probe the robustness of different circuit designs to losses. They allow us to pinpoint the most sensitive gates in the circuit and provide constraints for the quality of technological components that implement them. For example, taking Figs. 3(a) and 3(b) for the squeezing and beam-splitter channels, respectively, we can interpret the minimal value of s_i attained by each curve (for $s_{i+1} = -1$) as quantifying the maximum amount of losses above which said gate can always be represented by a regular PDF, no matter the input state.

6.3 Outlook

Our results provide a significant step forward towards the goal of singling out the salient quantum features needed to hinder any attempt of simulating

a photonic quantum process. Many directions remain to be explored, including but not limited to: the role of entanglement at preventing larger combinations of quantum gates, acting on fewer modes each, to be simulated in a single shot; the implications of the conditions on (s) -ordered quasiprobability distributions on the existence of local hidden-variable models; the relation between common measures of Wigner negativity and the nonclassical depth of Wigner-negative states, to determine which one best captures the quantitative aspects of this resource.

In particular, concerning the role of entanglement, one may try to regroup gates whose transfer functions cannot be made individually positive into larger ones, but if these gates act on more than one mode and can thus create entanglement, this requires sampling from a probability distribution over a larger set of variables, increasing the sampling costs accordingly.

It is also interesting to explore what happens when we drop some of the computational assumptions. A first key step is to connect our approach to the problem of compiling. In the present work, we have assumed that the decomposition of the system into a sequence of fundamental gates is known. In general, there may be many different decompositions that lead to the same global operation, a problem which concerns more the *compiling* aspect of the algorithmic design. Nevertheless, to each of these instances one can readily apply our procedure to determine the resourcefulness of each gate independently of the specific instance itself.

One may also argue that the computational cost of sampling from a probability distribution is not really taken into account. While from a probability-theory point of view this is more a classical than a quantum problem, it is clear that quantum properties of states, gates, and measurements, may lead to highly oscillating PDFs on phase space. These features may render our simulation protocol extremely costly, even when Algorithm 1 would technically succeed. Moreover, the sampling cost has been explored through the lens of (s) -ordered quasiprobabilities in [48]. Given sufficient classicality, the output parameter can be made arbitrarily close to one for a single step. This suggests that the step in question may be simulatable in polynomial time. Furthermore, allowing for additive errors in the sampling allows for more space in the ordering parameter’s optimisation algorithm.

Finally, one may want to challenge altogether the idea that one cannot construct a sampling procedure if the phase space representation manifests negativity. Indeed, it has been shown that we can use sampling techniques to deal with non-positive quasiprobabilities by paying an additional sampling cost [45] that is proportional to the negativity volume. Some recent work on the role of phase space negativity on quantum kernel methods [49] may help us

get a more nuanced idea of the role of the negativity volume in cases where our Algorithm 1 fails.

Acknowledgments

We would like to thank our fellow theoreticians Carlos Lopetegui and Mathieu Isoard for the numerous informative and supportive discussions. We thank Ulysse Chabaud for valuable comments on the manuscript. We acknowledge financial support from the ANR JCJC project NoRdiC (ANR-21-CE47-0005), the Plan France 2030 through the project OQuLus (ANR-22-PETQ-0013), and the HORIZON-EIC-2022- PATHFINDERCHALLENGES-01 programme under Grant Agreement Number 101114899 (Veriqub).

References

- [1] Peter W. Shor. “Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer”. *SIAM Review* **41**, 303–332 (1999).
- [2] John M. Martyn, Zane M. Rossi, Andrew K. Tan, and Isaac L. Chuang. “Grand unification of quantum algorithms”. *PRX Quantum* **2**, 040203 (2021).
- [3] Scott Aaronson and Alex Arkhipov. “The computational complexity of linear optics”. In *Proceedings of the Forty-Third Annual ACM Symposium on Theory of Computing*. Page 333–342. STOC ’11New York, NY, USA (2011). Association for Computing Machinery.
- [4] Michael J. Bremner, Richard Jozsa, and Dan J. Shepherd. “Classical simulation of commuting quantum computations implies collapse of the polynomial hierarchy”. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* **467**, 459–472 (2011).
- [5] Sergio Boixo, Sergei V. Isakov, Vadim N. Smelyanskiy, Ryan Babbush, Nan Ding, Zhang Jiang, Michael J. Bremner, John M. Martinis, and Hartmut Neven. “Characterizing quantum supremacy in near-term devices”. *Nature Physics* **14**, 595–600 (2018).
- [6] Aram W Harrow and Ashley Montanaro. “Quantum computational supremacy”. *Nature* **549**, 203–209 (2017).
- [7] Dominik Hangleiter and Jens Eisert. “Computational advantage of quantum random sampling”. *Rev. Mod. Phys.* **95**, 035001 (2023).
- [8] Frank Arute, Kunal Arya, Ryan Babbush, Dave Bacon, Joseph C. Bardin, Rami Barends, et al. “Quantum supremacy using a programmable superconducting processor”. *Nature* **574**, 505–510 (2019).

- [9] Han-Sen Zhong, Hui Wang, Yu-Hao Deng, Ming-Cheng Chen, Li-Chao Peng, Yi-Han Luo, Jian Qin, et al. “Quantum computational advantage using photons”. *Science* **370**, 1460–1463 (2020).
- [10] Lars S. Madsen, Fabian Laudenbach, Mohsen Falamarzi, Askarani, Fabien Rortais, Trevor Vincent, Jacob F. F. Bulmer, Filippo M. Miatto, Leonhard Neuhaus, Helt, et al. “Quantum computational advantage with a programmable photonic processor”. *Nature* **606**, 75–81 (2022).
- [11] M. DeCross, R. Haghshenas, M. Liu, E. Rinaldi, J. Gray, Y. Alexeev, C. H. Baldwin, J. P. Bartolotta, M. Bohn, E. Chertkov, J. Cline, J. Colina, D. DelVento, J. M. Dreiling, C. Foltz, J. P. Gaebler, T. M. Gatterman, C. N. Gilbreth, J. Giles, D. Gresh, A. Hall, A. Hankin, A. Hansen, N. Hewitt, I. Hoffman, C. Holliman, R. B. Hutson, T. Jacobs, J. Johansen, P. J. Lee, E. Lehman, D. Lucchetti, D. Lykov, I. S. Madjarov, B. Mathewson, K. Mayer, M. Mills, P. Niroula, J. M. Pino, C. Roman, M. Schecter, P. E. Siegfried, B. G. Tiemann, C. Volin, J. Walker, R. Shaydulin, M. Pistoia, S. A. Moses, D. Hayes, B. Neyenhuis, R. P. Stutz, and M. Foss-Feig. “Computational power of random quantum circuits in arbitrary geometries”. *Phys. Rev. X* **15**, 021052 (2025).
- [12] Craig S. Hamilton, Regina Kruse, Linda Sansoni, Sonja Barkhofen, Christine Silberhorn, and Igor Jex. “Gaussian boson sampling”. *Phys. Rev. Lett.* **119**, 170501 (2017).
- [13] A. P. Lund, A. Laing, S. Rahimi-Keshari, T. Rudolph, J. L. O’Brien, and T. C. Ralph. “Boson sampling from a gaussian state”. *Phys. Rev. Lett.* **113**, 100502 (2014).
- [14] U. Chabaud, T. Douce, D. Markham, P. van Loock, E. Kashefi, and G. Ferrini. “Continuous-variable sampling from photon-added or photon-subtracted squeezed states”. *Phys. Rev. A* **96**, 062307 (2017).
- [15] L. Chakhmakhchyan and N. J. Cerf. “Boson sampling with gaussian measurements”. *Phys. Rev. A* **96**, 032326 (2017).
- [16] Nicolás Quesada, Juan Miguel Arrazola, and Nathan Killoran. “Gaussian boson sampling using threshold detectors”. *Physical Review A* **98**, 062322 (2018).
- [17] Raúl García-Patrón, Jelmer J Renema, and Valery Shchesnovich. “Simulating boson sampling in lossy architectures”. *Quantum* **3**, 169 (2019).
- [18] Jelmer J Renema. “Simulability of partially distinguishable superposition and gaussian boson sampling”. *Physical Review A* **101**, 063840 (2020).
- [19] J. J. Renema, A. Menssen, W. R. Clements, G. Triginer, W. S. Kolthammer, and I. A. Walmsley. “Efficient classical algorithm for boson sampling with partially distinguishable photons”. *Phys. Rev. Lett.* **120**, 220502 (2018).
- [20] Saleh Rahimi-Keshari, Timothy C Ralph, and Carlton M Caves. “Sufficient conditions for efficient classical simulation of quantum optics”. *Physical Review X* **6**, 021039 (2016).
- [21] Jacob F. F. Bulmer, Bryn A. Bell, Rachel S. Chadwick, Alex E. Jones, Diana Moise, Alessandro Rigazzi, Jan Thorbecke, Utz-Uwe Haus, Thomas Van Vaerenbergh, Raj B. Patel, Ian A. Walmsley, and Anthony Laing. “The boundary for quantum advantage in gaussian boson sampling”. *Science Advances* **8**, eabl9236 (2022).
- [22] Haoyu Qi, Daniel J. Brod, Nicolás Quesada, and Raúl García-Patrón. “Regimes of classical simulability for noisy gaussian boson sampling”. *Phys. Rev. Lett.* **124**, 100502 (2020).
- [23] Nicolás Quesada, Rachel S. Chadwick, Bryn A. Bell, Juan Miguel Arrazola, Trevor Vincent, Haoyu Qi, and Raúl García-Patrón. “Quadratic speed-up for simulating gaussian boson sampling”. *PRX Quantum* **3**, 010306 (2022).
- [24] Peter D. Drummond, Bogdan Opanchuk, A. Dellios, and M. D. Reid. “Simulating complex networks in phase space: Gaussian boson sampling”. *Phys. Rev. A* **105**, 012427 (2022).
- [25] Changhun Oh, Minzhao Liu, Yuri Alexeev, Bill Fefferman, and Liang Jiang. “Classical algorithm for simulating experimental gaussian boson sampling”. *Nature Physics* (2024).
- [26] Daniel Gottesman. “The heisenberg representation of quantum computers” (1998). [arXiv:quant-ph/9807006](https://arxiv.org/abs/quant-ph/9807006).
- [27] Stephen D. Bartlett, Barry C. Sanders, Samuel L. Braunstein, and Kae Nemoto. “Efficient classical simulation of continuous variable quantum information processes”. *Physical Review Letters* **88**, 097904 (2002).
- [28] Robin L Hudson. “When is the Wigner quasiprobability density non-negative?”. *Reports on Mathematical Physics* **6**, 249–252 (1974).
- [29] Francisco Soto and Pierre Claverie. “When is the Wigner function of multidimensional systems nonnegative?”. *Journal of Mathematical Physics* **24**, 97–100 (1983).
- [30] Kevin E Cahill and Roy J Glauber. “Density operators and quasiprobability distributions”. *Physical Review* **177**, 1882 (1969).
- [31] Andrea Mari and Jens Eisert. “Positive wigner functions render classical simulation of quantum computation efficient”. *Physical review letters* **109**, 230503 (2012).
- [32] Victor Veitch, Christopher Ferrie, David Gross, and Joseph Emerson. “Negative quasi-probability as a resource for quantum computation”. *New Journal of Physics* **14**, 113011 (2012).

- [33] Laura García-Álvarez, Cameron Calcluth, Alessandro Ferraro, and Giulia Ferrini. “Efficient simulatability of continuous-variable circuits with large Wigner negativity”. *Physical Review Research* **2**, 043322 (2020).
- [34] Cameron Calcluth, Alessandro Ferraro, and Giulia Ferrini. “Efficient simulation of Gottesman-Kitaev-Preskill states with Gaussian circuits”. *Quantum* **6**, 867 (2022).
- [35] Cameron Calcluth, Alessandro Ferraro, and Giulia Ferrini. “Vacuum provides quantum advantage to otherwise simulatable architectures”. *Phys. Rev. A* **107**, 062414 (2023).
- [36] Cameron Calcluth, Nicolas Reichel, Alessandro Ferraro, and Giulia Ferrini. “Sufficient condition for universal quantum computation using bosonic circuits”. *PRX Quantum* **5**, 020337 (2024).
- [37] Ulysse Chabaud and Mattia Walschaers. “Resources for bosonic quantum computational advantage”. *Phys. Rev. Lett.* **130**, 090602 (2023).
- [38] Ulysse Chabaud, Roohollah Ghobadi, Salman Beigi, and Saleh Rahimi-Keshari. “Phase-space negativity as a computational resource for quantum kernel methods” (2024). [arXiv:2405.12378](https://arxiv.org/abs/2405.12378).
- [39] Kevin E Cahill and Roy J Glauber. “Ordered expansions in boson amplitude operators”. *Physical Review* **177**, 1857 (1969).
- [40] N. Lütkenhaus and Stephen M. Barnett. “Non-classical effects in phase space”. *Phys. Rev. A* **51**, 3340–3342 (1995).
- [41] Ching Tsung Lee. “Measure of the nonclassicality of nonclassical states”. *Physical Review A* **44**, R2775 (1991).
- [42] Frigerio, Massimo and Debray, Antoine. “(s)-QPDs optimization tool”. <https://github.com/EQ15T/Resourcefulness-s-QPDs>.
- [43] Seth Lloyd and Samuel L. Braunstein. “Quantum computation over continuous variables”. *Phys. Rev. Lett.* **82**, 1784–1787 (1999).
- [44] Timo Hillmann, Fernando Quijandría, Göran Johansson, Alessandro Ferraro, Simone Gasparinetti, and Giulia Ferrini. “Universal gate set for continuous-variable quantum computation with microwave circuits”. *Phys. Rev. Lett.* **125**, 160501 (2020).
- [45] Hakop Pashayan, Joel J Wallman, and Stephen D Bartlett. “Estimating outcome probabilities of quantum circuits using quasiprobabilities”. *Physical Review Letters* **115**, 070501 (2015).
- [46] Daniel Gottesman, Alexei Kitaev, and John Preskill. “Encoding a qubit in an oscillator”. *Physical Review A* **64**, 012310 (2001).
- [47] Mile Gu, Christian Weedbrook, Nicolas C. Menicucci, Timothy C. Ralph, and Peter van Loock. “Quantum computing with continuous-variable clusters”. *Phys. Rev. A* **79**, 062318 (2009).
- [48] Youngrong Lim and Changhun Oh. “Approximating outcome probabilities of linear optical circuits”. *npj Quantum Information* **9**, 124 (2023).
- [49] Ulysse Chabaud, Roohollah Ghobadi, Salman Beigi, and Saleh Rahimi-Keshari. “Phase-space negativity as a computational resource for quantum kernel methods”. *Quantum* **8**, 1519 (2024).

A Explicit calculations for Eq. (13)

Similarly to its construction for operators, it is possible to construct ordered quasiprobability distributions for quantum channels (see e.g. [20]). Figure 2 illustrates a generic quantum process; the simulation task would amount to sampling from $p(\mathbf{n}) = \text{Tr}[\hat{\rho}_{\text{out}}\Pi_{\mathbf{n}}]$. It is possible to rewrite this probability distribution in terms of $(\mathbf{s})$ -ordered quasi-probability distributions using a generalization of the trace rules for Wigner functions [39]: for any two bounded operators $\hat{O}_1, \hat{O}_2$ of trace class we can write

$$\text{Tr}[\hat{O}_1\hat{O}_2] = \int_{\xi \in \mathbb{C}} \frac{d^{2M}\xi}{\pi^M} W_{\hat{O}_1}^{(-s)} W_{\hat{O}_2}^{(s)}. \quad (35)$$

Thus leading to

$$p(\mathbf{x}) = \text{Tr}[\hat{\rho}_{\text{out}}\hat{\Pi}_{\mathbf{x}}] = \pi^{2M} \int_{\beta \in \mathbb{C}} d^{2M}\beta W_{\hat{\Pi}_{\mathbf{x}}}^{(-s)}(\beta) W_{\hat{\rho}_{\text{out}}}^{(s)}(\beta), \quad (36)$$

where

$$W_{\hat{\Pi}_{\mathbf{x}}}^{(-s)}(\beta) = \int_{\xi \in \mathbb{C}} \frac{d^{2M}\xi}{\pi^{2M}} \text{Tr}[\hat{\Pi}_{\mathbf{n}} \hat{D}_{-s}(\xi)] e^{\beta\xi^\dagger - \xi\beta^\dagger} \quad (37)$$

and

$$W_{\hat{\rho}_{\text{out}}}^{(s)}(\beta) = \int_{\xi \in \mathbb{C}} \frac{d^{2M}\xi}{\pi^{2M}} \chi_{\hat{\rho}_{\text{out}}}^{(s)}(\xi) e^{\beta\xi^\dagger - \xi\beta^\dagger}. \quad (38)$$

One can write the M modes input state in terms of displacement operators as

$$\hat{\rho}_{\text{in}} = \int_{\mathbf{x} \in \mathbb{C}} \chi_{\hat{\rho}_{\text{in}}}^{(s')}(\xi) e^{-\xi s' \xi^\dagger / 2} \hat{D}^\dagger(\xi) \frac{d^{2M}\xi}{\pi^M}. \quad (39)$$

This implies, by linearity of the quantum process, that the output state is

$$\hat{\rho}_{\text{out}} = \int_{\xi \in \mathbb{C}} \chi_{\hat{\rho}_{\text{in}}}^{(s')}(\xi) e^{-\xi s' \xi^\dagger / 2} \mathcal{E}(\hat{D}^\dagger(\xi)) \frac{d^{2M}\xi}{\pi^M}. \quad (40)$$

Moreover, by virtue of the inverse Fourier transform, the $(\mathbf{s})$ -ordered characteristic function of the output state is

$$\chi_{\hat{\rho}_{\text{out}}}^{(s)}(\zeta) = \text{Tr}[\hat{\rho}_{\text{out}} \hat{D}_s(\zeta)] \quad (41)$$

By plugging Eq. (40) into the last we find

$$\chi_{\hat{\rho}_{\text{out}}}^{(s)}(\zeta) = \int_{\xi \in \mathbb{C}} \chi_{\hat{\rho}_{\text{in}}}^{(s')}(\xi) \text{Tr}[\mathcal{E}(\hat{D}_{s'}^\dagger(\xi)) \hat{D}_s(\zeta)] \frac{d^{2M}\xi}{\pi^M} \quad (42)$$

Which can be re-written in terms of the $(\mathbf{s}')$ -ordered Wigner function of the input state

$$\chi_{\hat{\rho}_{\text{out}}}^{(s)}(\zeta) = \int_{\xi \in \mathbb{C}} \frac{d^{2M}\xi}{\pi^{2M}} \int_{\alpha \in \mathbb{C}} d^{2M}\beta W_{\hat{\rho}_{\text{in}}}^{(s')}(\beta) \text{Tr}[\mathcal{E}(\hat{D}_{s'}^\dagger(\xi)) \hat{D}_s(\zeta)]. \quad (43)$$

Using the Fourier transform we can find a relation between the input and output ordered Wigner functions,

$$W_{\hat{\rho}_{\text{out}}}^{(s)}(\beta) = \int_{\alpha \in \mathbb{C}} d^{2M}\alpha W_{\hat{\rho}_{\text{in}}}^{(s')}(\alpha) \Lambda_{\mathcal{E}}^{(-s',s)}(\alpha, \beta) \quad (44)$$

where we define the $(\mathbf{s}', \mathbf{s})$ -ordered transfer function $\Lambda_{\mathcal{E}}^{(s',s)}(\alpha, \beta)$ as

$$\Lambda_{\mathcal{E}}^{(s',s)}(\alpha, \beta) = \int_{\zeta \in \mathbb{C}} \frac{d^{2M}\zeta}{\pi^M} \int_{\xi \in \mathbb{C}} \frac{d^{2M}\xi}{\pi^{2M}} e^{\beta\xi^\dagger - \zeta\beta^\dagger + \xi\alpha^\dagger - \alpha\xi^\dagger} e^{\zeta s \zeta^\dagger / 2 + \xi s' \xi^\dagger / 2} \text{Tr}[\mathcal{E}(\hat{D}^\dagger(\xi)) \hat{D}(\zeta)]. \quad (45)$$

Finally, by plugging last equation into Eq. (36) we find the highlighted result of [20]:

$$p(\mathbf{x}) = \iint_{\alpha, \beta \in \mathbb{C}} \pi^M d^{2M}\alpha d^{2M}\beta W_{\hat{\rho}_{\text{in}}}^{(s')}(\alpha) \Lambda_{\mathcal{E}}^{(-s',s)}(\alpha, \beta) W_{\hat{\Pi}_{\mathbf{x}}}^{(-s)}(\beta) \quad (46)$$

The last step needed to find Eq. (13) is to replace $\mathcal{E}$ by a sequence of trace-preserving quantum channels $\mathcal{E}_1 \circ \mathcal{E}_2 \circ \dots \circ \mathcal{E}_k$ and notice that

$$\Lambda_{\mathcal{E}_1 \circ \mathcal{E}_2}^{(-s', s)}(\alpha, \beta) = \int_{\gamma \in \mathbb{C}} \Lambda_{\mathcal{E}_1}^{(-s', s'')}(\alpha, \gamma) \Lambda_{\mathcal{E}_2}^{(-s'', s)}(\gamma, \beta) \quad (47)$$

Leading naturally to Eq. (13).

These functions are all defined on the complex plane, rather than on the real symplectic phase-space that is commonly encountered when dealing with Wigner functions. To convert to quadrature notation, one can use the following identities:

$$\begin{cases} \hat{q} = \hat{a} + \hat{a}^\dagger, & \hat{p} = -i(\hat{a} - \hat{a}^\dagger) \\ \vec{q} = \frac{\boldsymbol{\xi} + \boldsymbol{\xi}^\dagger}{2}, & \vec{p} = \frac{\boldsymbol{\xi} - \boldsymbol{\xi}^\dagger}{2i} \end{cases} \quad (48)$$

B Explicit calculations for the cubic gate

Let $\hat{U}_\gamma = \exp(-i\frac{\gamma}{3!}\hat{q}^3)$. Firstly, we need to compute the characteristic function of this gate:

$$\text{Tr} \left[\hat{U}_\gamma \hat{D}^\dagger(\xi) \hat{U}_\gamma^\dagger \hat{D}(\zeta) \right] \quad (49)$$

It is convenient to start by rewriting $D(\xi)$ in terms of the real, quadratures variables:

$$\hat{D}(\xi) = \exp \{ \xi \hat{a}^\dagger - \xi^* \hat{a} \} = \exp \{ -i(q_1 \hat{p} - p_1 \hat{q}) \} \quad (50)$$

We have that $\hat{U}_\gamma \hat{q} \hat{U}_\gamma^\dagger = \hat{q}$, while $\hat{U}_\gamma \hat{p} \hat{U}_\gamma^\dagger = \hat{p} + \gamma \hat{q}^2$, from which it follows that:

$$\hat{U}_\gamma \hat{D}^\dagger(\xi) \hat{U}_\gamma^\dagger = \exp \{ i [q_1 (\hat{p} + \gamma \hat{q}^2) - p_1 \hat{q}] \} \quad (51)$$

where $q_1 = \text{Re} \xi$ and $p_1 = \text{Im} \xi$. To compute the trace using the eigenbasis of $\hat{q}$, it is useful to rearrange this result as:

$$\exp \{ i [q_1 (\hat{p} + \gamma \hat{q}^2) - p_1 \hat{q}] \} = e^{a\hat{p}} e^{b\hat{q} + c\hat{q}^2 + d} \quad (52)$$

for a proper choice of $a, b, c \in \mathbb{C}$. Notice that there must be a solution, since commutators of quadratic operators in the quadratures form a closed subalgebra and a term in $\hat{p}^2$ cannot appear, as it is not present in the exponent on the left-hand side. We can compute a, b, c by applying the Baker-Campbell-Hausdorff formula to the right-hand side, to get:

$$\begin{aligned} & e^{a\hat{p}} e^{b\hat{q} + c\hat{q}^2} \\ &= \exp \left\{ a\hat{p} + b\hat{q} + c\hat{q}^2 + \frac{1}{2}[a\hat{p}, b\hat{q} + c\hat{q}^2] + \dots \right\} \\ &= \exp \left\{ a\hat{p} + (b - 2iac)\hat{q} + c\hat{q}^2 - iab + \frac{1}{12}[a\hat{p}, -4iac\hat{q}] \right\} \\ &= \exp \left\{ a\hat{p} + (b - 2iac)\hat{q} + c\hat{q}^2 - iab - \frac{2}{3}a^2c \right\} \end{aligned} \quad (53)$$

and we used $[\hat{q}, \hat{p}] = 2i$. All higher order commutators vanish, since the last second-order nested commutator is a scalar. By direct comparison with the left-hand side of Eq. (52) we conclude that:

$$\begin{aligned} a &= iq_1, \quad c = iq_1\gamma \\ b &= -ip_1 - 2iq_1^2\gamma, \quad d = iq_1p_1 + \frac{4}{3}i\gamma q_1^3 \end{aligned} \quad (54)$$

If we similarly rewrite $\hat{D}(\zeta) = e^{iq_2 p_2} e^{-iq_2 \hat{p}} e^{ip_2 \hat{q}}$, with $q_2 = \text{Re } \zeta$ and $p_2 = \text{Im } \zeta$, we finally have:

$$\begin{aligned}
& \text{Tr} \left[\hat{U}_\gamma \hat{D}^\dagger(\xi) \hat{U}_\gamma^\dagger \hat{D}(\zeta) \right] = \\
&= e^\phi \int_{q \in \mathbb{R}} dq \langle q | e^{a\hat{p}} e^{b\hat{q}+c\hat{q}^2} e^{-iq_2 \hat{p}} e^{ip_2 \hat{q}} | q \rangle = \\
&= e^\phi \int_{q \in \mathbb{R}} dq e^{ip_2 q} \langle q + 2q_1 | e^{b\hat{q}+c\hat{q}^2} | q + 2q_2 \rangle = \\
&= \frac{1}{2} e^\phi \delta(q_1 - q_2) \int_{q \in \mathbb{R}} dq e^{ip_2 q + b(q+2q_1) + c(q+2q_1)^2} = \\
&= \frac{1}{2} e^{\tilde{\phi}} \delta(q_1 - q_2) \int_{q \in \mathbb{R}} dq e^{i(p_2 - p_1 - 2\gamma q_1^2)q + i\gamma q_1 q^2}
\end{aligned} \tag{55}$$

with $\phi = i(q_1 p_1 + q_2 p_2) + \frac{4}{3} i \gamma q_1^3$ and $\tilde{\phi} = -i q_1 (p_1 - p_2) + \frac{4}{3} i \gamma q_1^3$. The remaining integration is the Fourier transform of a Gaussian with an imaginary variance, and it can be computed by regularizing the integral, to find:

$$\int_{q \in \mathbb{R}} dq e^{i(p_2 - p_1 - 2\gamma q_1^2)q + i\gamma q_1 q^2} = \sqrt{\frac{i\pi}{\gamma q_1}} e^{-i \frac{(p_2 - p_1 - 2\gamma q_1^2)^2}{4\gamma q_1}} \tag{56}$$

Putting everything together, we arrive at:

$$\text{Tr} \left[\hat{U}_\gamma \hat{D}^\dagger(\xi) \hat{U}_\gamma^\dagger \hat{D}(\zeta) \right] = \frac{1}{2} \delta(q_1 - q_2) \sqrt{\frac{i\pi}{\gamma q_1}} e^{-\frac{i}{4\gamma q_1} (p_1 - p_2)^2 + \frac{4}{3} i \gamma q_1^3} \tag{57}$$

The next steps consist in performing the Fourier transform of this function with respect to q_1, q_2, p_1, p_2 , with additional Gaussian weights that impose the (s_1, s_2) ordering, that is, multiplying the result of Eq. (57) by:

$$e^{\alpha^* \xi - \alpha \xi^* + \beta \zeta^* - \beta^* \zeta} e^{\frac{s_1 |\xi|^2}{2} + \frac{s_2 |\zeta|^2}{2}} = \exp[2i(q_\alpha p_1 - p_\alpha q_1 + p_\beta q_2 - q_\beta p_2)] \exp\left[\frac{s_1}{2}(q_1^2 + p_1^2) + \frac{s_2}{2}(q_2^2 + p_2^2)\right] \tag{58}$$

Because of the delta distribution, the integration with respect to q_2 is trivial. Also, the integrations with respect to p_1 and p_2 amounts to computing the Fourier transform of Gaussian functions with complex variance; these converge if and only if the real part of the variance is positive, such that the integrand decays exponentially in p_1 and p_2 . Also, observing that the Gaussian factor in $\text{Tr} \left[\hat{U}_\gamma \hat{D}^\dagger(\xi) \hat{U}_\gamma^\dagger \hat{D}(\zeta) \right]$ is purely imaginary, the conditions of convergence are just $s_1 < 0$ and $s_2 < 0$. The general formula that should be applied for the integrations in p_1 and p_2 is then:

$$\forall A, B \in \mathbb{C} : \text{Re}(A) > 0 \implies \int_{x \in \mathbb{R}} e^{-Ax^2 + iBx} dx = \sqrt{\frac{\pi}{A}} e^{-\frac{B^2}{4A}} \tag{59}$$

and the end result is:

$$\begin{aligned}
\Lambda_{\hat{U}_\gamma}^{(s_1, s_2)}(q_\alpha, p_\alpha, q_\beta, p_\beta) &= \int_{q_1 \in \mathbb{R}} \frac{dq_1}{\sqrt{2\pi}} \sqrt{\frac{i}{2\gamma q_1 s_1 s_2 - i(s_1 + s_2)}} \\
&\times \exp \left[\frac{4\gamma q_1 (s_2 q_\alpha^2 + s_1 q_\beta^2) - 2i(q_\alpha - q_\beta)^2}{2\gamma q_1 s_1 s_2 - i(s_1 + s_2)} + \frac{s_1 + s_2}{2} q_1^2 + \frac{i}{3} \gamma q_1^3 + 2i(p_\beta - p_\alpha) q_1 \right]
\end{aligned} \tag{60}$$

Again, for $s_1 < 0$ and $s_2 < 0$ the integral must converge. Notice, moreover, that the integrand is Gaussian in q_α, q_β and the real part of the variance is positive as long as $s_1, s_2 < 0$. Finally, it depends only upon the combination $p_\alpha - p_\beta$ and not on the single variables independently. We can then conclude that $\Lambda_{\hat{U}_\gamma}^{(s_1, s_2)}(q_\alpha, p_\alpha, q_\beta, p_\beta)$ is a well-defined function whenever $s_1, s_2 < 0$ and, by Theorem 2, we know that it cannot be positive-semidefinite for any $s_1 > -1$ and $s_2 > -1$. At the same time, it must be positive semidefinite for $s_1 = s_2 = -1$, since it then reduces to:

$$\Lambda_{\hat{U}_\gamma}^{(-1, -1)}(q_\alpha, p_\alpha, q_\beta, p_\beta) \propto |\langle \beta | \hat{U}_\gamma | \alpha \rangle|^2 \tag{61}$$

A somewhat simpler expression for $\Lambda_{\hat{U}_\gamma}^{(s_1, s_2)}$ for $s_1, s_2 < 0$ can be computed also using Eq. (14) for the change of the s -parameters starting with $\Lambda_{\hat{U}_\gamma}^{(0,0)}$, which can be directly computed to be:

$$\Lambda_{\hat{U}_\gamma}^{(0,0)}(q_\alpha, p_\alpha, q_\beta, p_\beta) = \frac{2}{\pi} |\gamma|^{-1/3} \delta(q_\alpha - q_\beta) \text{Ai} \left(\frac{2(p_\beta - p_\alpha)}{\gamma^{1/3}} + 4|\gamma|^{2/3} q_\alpha^2 \right) \tag{62}$$

one can then finally get the following form for $s_1, s_2 < 0$:

$$\begin{aligned}
\Lambda_{\hat{U}_\gamma}^{(s_1, s_2)}(q_\alpha, p_\alpha, q_\beta, p_\beta) &= \\
&= 4 \iint \frac{dq'_\alpha dp'_\alpha dq'_\beta dp'_\beta}{\pi^2 s_1 s_2} e^{2 \frac{(q_\alpha - q'_\alpha)^2 + (p_\alpha - p'_\alpha)^2}{s_1} + 2 \frac{(q_\beta - q'_\beta)^2 + (p_\beta - p'_\beta)^2}{s_2}} \Lambda_{\hat{U}_\gamma}^{(0,0)}(q'_\alpha, p'_\alpha, q'_\beta, p'_\beta) = \\
&= \frac{4}{\pi^{5/2} |\gamma|^{1/3} \Sigma} e^{-\frac{2(q_\alpha - q_\beta)^2}{\Sigma}} \int dx \sqrt{\frac{2}{\sigma}} e^{-\frac{2}{\sigma}(x - \mu)^2} \int du e^{-\frac{2(u - p)^2}{\Sigma}} \text{Ai}\left(\frac{2u}{\gamma^{1/3}} + 4|\gamma|^{2/3} x^2\right) \quad (63)
\end{aligned}$$

where $p = p_\alpha - p_\beta$, $\Sigma = |s_1 + s_2|$, $\sigma = \frac{s_1 s_2}{|s_1 + s_2|}$ and $\mu = \frac{s_2 q_\alpha + s_1 q_\beta}{\Sigma}$. Notice that this is a double Gaussian smoothing of the Airy function. Due to the highly oscillatory behavior of the latter, however, it is still very challenging to find conditions on (s_1, s_2) as functions of γ such that the final result is everywhere non-negative.