

Benchmarking single-photon sources from an auto-correlation measurement

Pavel Sekatski¹, Enky Oudot², Patrik Caspar¹, Rob Thew¹, and Nicolas Sangouard³

¹Department of Applied Physics, University of Geneva, Geneva, Switzerland

²ICFO - Institut de Ciències Fòniques, The Barcelona Institute of Science and Technology, 08860 Castelldefels (Barcelona), Spain

³Université Paris-Saclay, CEA, CNRS, Institut de physique théorique, 91191, Gif-sur-Yvette, France

Here we argue that the probability that a given source produces exactly a single photon is a natural quantity to benchmark single-photon sources as it certifies the absence of multi-photon components and quantifies the efficiency simultaneously. Moreover, this probability can be bounded simply from an auto-correlation measurement – a balanced beam splitter and two photon detectors. Such a bound gives access to various non-classicality witnesses that can be used to certify and quantify Wigner-negativity, in addition to non-Gaussianity and P-negativity of the state produced by the source. We provide tools that can be used in practice to account for an imperfect beam splitter, non-identical and non-unit detection efficiencies, dark counts and other imperfections, to take finite statistical effects into account without assuming that identical states are produced in all rounds, and optionally allow one to remove the detector inefficiencies from the analysis. We demonstrate the use of the proposed benchmark, non-classicality witness and measure using a heralded single-photon source based on spontaneous parametric down-conversion. We report on an average probability that a single photon is produced $\geq 55\%$ and an average measure of the Wigner negativity ≥ 0.004 with a confidence level of $1 - 10^{-10}$.

1 Introduction

Single-photon sources [1, 2] are key resources for quantum communication [3], photonic quantum computation [4] or radiometry [5, 6]. Not all single-photon sources are alike and to be scaled up, most applications require efficient sources of true single photons (single-photon Fock/number states). The quality of single-photon sources is usually quantified from an auto-correlation measurement [7], that is, by sending the photons to a balanced beam splitter and checking that the ratio between the detected twofold coinci-

Pavel Sekatski: pavel.sekatski@unige.ch

Rob Thew: robert.thew@unige.ch

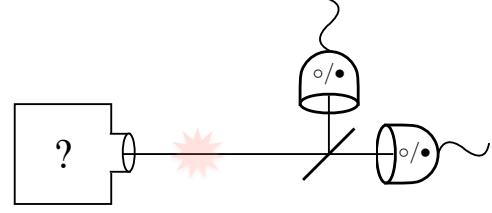


Figure 1: Schematic representation of the measurement that is considered to characterize an unknown photon source which supposedly produces single photons. It is realized with a beam splitter and two non-photon-number-resolving detectors, as in a standard auto-correlation measurement. At each round, each detector either clicks $\bullet$ or not $\circ$. By analyzing the frequency of these events, the probability that the source actually produces exactly a single photon can be lower bounded. Furthermore, if a single radiation mode is detected, various forms of non-classicality can be witnessed and quantified.

dences and the product of singles vanishes, see Fig. 1 and the discussion in Appendix A. This ensures that the source produces no more than one photon. The result is, however, insensitive to loss as the efficiency cancels out of the ratio. These two aspects – the capacity of a source to produce no more than one photon and its efficiency – are thus considered separately. Both aspects are, however, important and are quantified jointly by the probability that the source actually produces exactly a single photon. Characterizing this probability is a direct and more complete way to benchmark single-photon sources.

Interestingly, this probability can be bounded by reconsidering the statistics of detector counts in an auto-correlation measurement. This suggests a systematic way to benchmark the quality and quantify the efficiency of single-photon sources, and to witness and quantify their quantum nature. To motivate benchmarking single-photon sources by the probability that the source actually produces exactly a single photon, we provide a detailed analysis which includes a simple statistical tool to account for finite-size effects without assuming that identical states were produced in all rounds of the experiment. We show how to include imperfections in the measurements apparatus and how to remove the detector ef-

iciencies from the analysis to facilitate the use of the proposed benchmark. An experimental demonstration is presented, illustrating the quality of heralded single-photon sources based on spontaneous parametric down-conversion.

The auto-correlation measurement is also known to be valuable for witnessing various forms of non-classicality, including the non-positivity of the P-distribution and quantum non-Gaussianity [8–10]. Here we show that for single-mode states or measurements a bound on the single-photon probability can be readily used to quantify the negativity of the Wigner representation [7], the strongest form of non-classicality, with respect to a measure proposed in [11]. We apply this method to verify Wigner-negativity in the reported experiments.

2 Measurement apparatus

The measurement apparatus we consider is similar to the one used for the second order auto-correlation measurement. It is a simple measurement consisting of sending the photonic state to be measured (labeled ρ) on to a beam splitter and recording the photon-count correlations between the two outputs, see Fig. 1. We consider that photon detections are made with typical non-photon-number-resolving detectors. In order to draw conclusions from such photon-counts, we introduce a simple quantum model for such a measurement setup. A non-photon-number-resolving detector of efficiency η can be modeled with a two element positive operator-valued measure (POVM) $\{E_\bullet, E_\circ\}$ corresponding to click ($\bullet$) and no-click ($\circ$) outcomes. When the measurement acts on a single mode characterized by bosonic operators a and $a^\dagger$, the POVM elements take the following form

$$E_\circ = (1 - \eta)^{a^\dagger a}, \quad E_\bullet = \mathbb{1} - (1 - \eta)^{a^\dagger a}. \quad (1)$$

This model describes a non-photon-number-resolving detector, for which every incident photon can trigger the detection event with probability η . The no-click outcome then corresponds to the event where none of the incident photons triggered a click, and occurs with probability $p_\circ = \text{tr} \rho (1 - \eta)^{a^\dagger a}$, see e.g. [12]. From now on we assume that the detectors are accurately described by Eq. (1) for *some value* of η . In section 3.1.4 we will discuss how to account for small deviation from this model.

When two such detectors are placed after a beam splitter with reflectance r (transmittance t), it is straightforward to see that the four possible outcomes are given by the POVM elements

$$\begin{aligned} E_{\circ\circ} &= (1 - \eta)^{a^\dagger a} \\ E_{\bullet\circ} &= (1 - \eta t)^{a^\dagger a} - (1 - \eta)^{a^\dagger a} \\ E_{\circ\bullet} &= (1 - \eta r)^{a^\dagger a} - (1 - \eta)^{a^\dagger a} \\ E_{\bullet\bullet} &= \mathbb{1} - E_{\circ\circ} - E_{\bullet\circ} - E_{\circ\bullet}, \end{aligned} \quad (2)$$

where the first (second) label $\circ/\bullet$ refers to the detector after the reflected (transmitted) output of the beam splitter, see Appendix B for a formal derivation. The events where a fixed detector does not click are modeled by the two POVM elements $E_{\circ-} = E_{\circ\bullet} + E_{\circ\circ} = (1 - \eta r)^{a^\dagger a}$ and $E_{- \circ} = (1 - \eta t)^{a^\dagger a}$. The corresponding probabilities are labeled $p_{\circ-}$ and $p_{- \circ}$. Note that the case where the two detectors do not have the same efficiency $\eta_R \neq \eta_T$ can be accounted for by replacing t with $t' = \frac{t\eta_T}{t\eta_T + r\eta_R}$, r with $r' = \frac{r\eta_R}{t\eta_T + r\eta_R}$, and setting $\eta = t\eta_T + r\eta_R$ in Eq. (2).

3 Benchmarking a single-photon source

For simplicity, through the main part of this section we will be considering the case of single-mode sources. In section 3.3, we show that all the presented tools also apply to multi-mode sources.

With the measurement we just described, any single-mode state ρ incident on the beam splitter can be associated with a probability vector

$$\mathbf{p} = (p_{\circ\circ}, p_{\bullet\circ}, p_{\circ\bullet}, p_{\bullet\bullet}) \quad (3)$$

governing the occurrence of clicks. Our goal is to construct an estimator $\hat{P}_1(\mathbf{p})$ that relates this vector $\mathbf{p}$ to the photon number statistics of the state and in particular to the weight of the single-photon component $P_1 = \langle 1 | \rho | 1 \rangle$. Directly lower bounding P_1 is a natural way to benchmark a single-photon source. In particular, it sets a bound on the trace distance between the state ρ prepared by the source and an ideal single photon $|1\rangle$

$$\frac{1}{2} \|\rho - |1\rangle\langle 1|\|_1 \leq 1 - P_1, \quad (4)$$

which can be readily used to bound the errors for various applications of single-photon sources.

3.1 Measurement calibration independent benchmark

We proceed step by step, considering first an ideal measurement apparatus, then considering two identical non-unit detector efficiencies and finally focusing on the most imperfect measurement, where we consider an unbalanced beam splitter and two detectors having different efficiencies. At this point we only assume that the measurement is described by the POVM in Eq. (2) for *some* parameters η, t and r , that do not need to be known. The benchmark that we will derive in this section is thus independent of the calibration of the measurement apparatus. Finally, we show that by adding a small correction term our benchmark can be applied in situations where the POVM in Eq. (2) is only an approximate description of the measurement apparatus.

3.1.1 Ideal measurement apparatus

For clarity, we first assume that the source produces an identical state ρ at each round, the detectors have unit detection efficiency $\eta = 1$ and the beam splitter is balanced $t = r = 1/2$. In this ideal case, the probabilities $p_{\bullet\bullet}$, $p_{\bullet\circ}$ are equal, and $\mathbf{p}$ is described by two independent real parameters. For convenience, we introduce $p_{\bullet} = p_{\bullet\bullet} + p_{\bullet\circ}$ the probability to get exactly one click. The probabilities that a given detector does not click $p_{\circ-} = p_{\circ\circ} + p_{\bullet\circ}$ and $p_{\circ\circ} = p_{\circ\circ} + p_{\bullet\circ}$ are equal in the ideal case and in particular $p_{\circ-} = p_{\circ\circ} = p_{\circ\circ} + \frac{1}{2}p_{\bullet}$. This means that the probabilities of outcomes of the measurement of interest can be fully captured by $(p_{\circ-}, p_{\circ\circ})$. Let

$$P_n = \langle n | \rho | n \rangle \quad (5)$$

be the weight of the n -photon Fock state component of the measured state. For $\eta = 1$, the no-click events can only come from the vacuum state $E_{\circ\circ} = |0\rangle\langle 0|$, hence the probabilities $(p_{\circ-}, p_{\circ\circ})$ can be linked to the photon number distribution P_n . From Eqs. (2), one gets $p_{\circ\circ} = P_0$ and $p_{\circ-} = \sum_n P_n \frac{1}{2^n}$.

The question we ask now is very simple – what are the values $(p_{\circ-}, p_{\circ\circ})$ that are obtainable for states ρ satisfying $P_1 \leq P$, for some parameter $P \in [0, 1]$?

First, we note that $p_{\circ-} \geq p_{\circ\circ}$ holds by definition. Furthermore, the points $(1, 1)$ and $(0, 0)$ are attained by the vacuum and the state with infinitely many photons, respectively. Thus, the line $p_{\circ-} = p_{\circ\circ}$ is also attainable by mixtures of aforementioned states. Then, we look for the maximum value of $p_{\circ-} = \sum_n P_n \frac{1}{2^n}$ for a fixed $p_{\circ\circ}$. We have to solve

$$\begin{aligned} p_{\circ-}^{\uparrow}(p_{\circ\circ}, P) &= \max_{\rho} \sum_n P_n \frac{1}{2^n} \\ \text{such that } P_1 &\leq P \\ P_0 &= p_{\circ\circ}. \end{aligned} \quad (6)$$

As $(1/2)^n$ is decreasing with n , the maximum is attained by saturating the values of P_n starting with P_0 . Hence, it equals

$$p_{\circ-} \leq p_{\circ-}^{\uparrow}(p_{\circ\circ}, P) = \begin{cases} \frac{1+p_{\circ\circ}}{2} & 1 - p_{\circ\circ} \leq P \\ \frac{1+P+3p_{\circ\circ}}{4} & 1 - p_{\circ\circ} > P \end{cases} \quad (7)$$

The set of possible values $(p_{\circ-}, p_{\circ\circ})$ is thus included in a convex polytope with four vertices $\mathbf{Q}_P = \text{Polytope}\{(0, 0), (\frac{1+P}{4}, 0), (\frac{2-P}{2}, 1-P), (1, 1)\}$, sketched in Fig. 2. The only nontrivial facet of this polytope is the edge connecting $(\frac{1+P}{4}, 0)$ and $(\frac{2-P}{2}, 1-P)$ which is associated to the inequality $4p_{\circ-} - 3p_{\circ\circ} - 1 \leq P$, and is given by the colored lines in Fig. 2. Thus, without loss of generality, the condition $\langle 1 | \rho | 1 \rangle \leq P$ implies that the elements of $\mathbf{p}$ satisfy the linear constraint

$$\hat{P}_1^T(\mathbf{p}) = 4p_{\circ-} - 3p_{\circ\circ} - 1 \leq P. \quad (8)$$

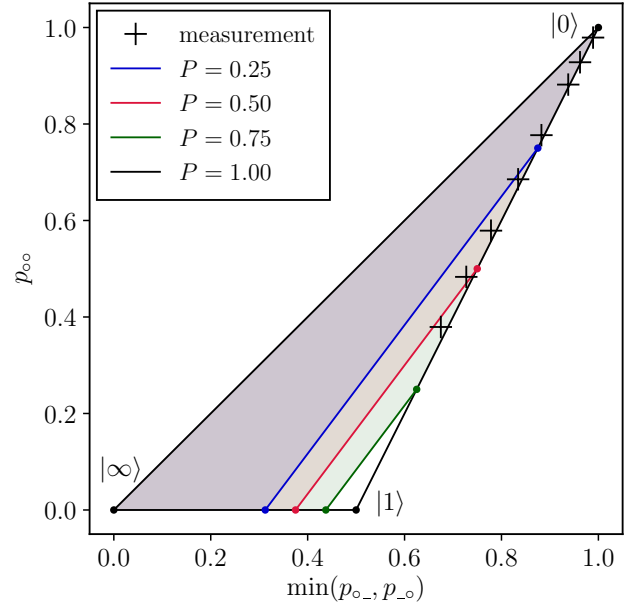


Figure 2: Representation of the polytopes $\mathbf{Q}_P$ (defined after Eq. (7)) containing all possible values $(\min(p_{\circ-}, p_{\circ\circ}), p_{\circ\circ})$ associated to states ρ with $\langle 1 | \rho | 1 \rangle \leq P$. There are four polytopes $\mathbf{Q}_P$ for the values $P \in \{0.25, 0.5, 0.75, 1\}$. The physically possible region $\mathbf{Q}_{P=1}$ is given by the black triangle. The regions for $P = 0.25, 0.5, 0.75$ are given by the part of the black triangle above the corresponding colored line. The black crosses are measurements of $(\min(p_{\circ-}, p_{\circ\circ}), p_{\circ\circ})$ for a heralded single photon undergoing different added attenuation corresponding to transmission efficiencies of $\eta_{\text{att}} \in \{1.0, 0.83, 0.68, 0.51, 0.36, 0.19, 0.12, 0.034\}$.

Conversely, by measuring the pair $(p_{\circ-}, p_{\circ\circ})$ and by computing the resulting value of $\hat{P}_1^T$, we can guarantee that for any value of P such that $\hat{P}_1^T > P$, $\langle 1 | \rho | 1 \rangle > P$ holds, that is we get a lower bound on the probability that the source to be benchmarked produces exactly a single photon.

3.1.2 Identical non-unit efficiency detectors

To move away from the ideal case, we still consider a perfectly balanced beam splitter and focus on a situation where non-unit efficiency detectors are used. We consider the case where the detector efficiency η is unknown. In the measurement setup shown in Fig. 1, non-unit efficiency detectors can be modeled by taking ideal detectors and placing a beam splitter with transmission η before the balanced beam splitter. As a consequence, observing a violation of Ineq. (8) proves that the state produced by the single-photon source and undergoing losses satisfies $\langle 1 | \rho | 1 \rangle \geq \hat{P}_1^T(\mathbf{p})$. This provides a valid benchmark even though the intrinsic quality of the source is estimated with a lossy measurement apparatus. It is interesting to note that for any state ρ with $P_1 \geq 2/3$, the probability of the single-photon weight P_1 can only decrease with loss, see Appendix C. Therefore, showing that $P_1 \geq 2/3$ with lossy detectors implies that the original state also

satisfies $P_1 \geq \frac{2}{3}$. This is not the case when $P_1 < 2/3$, i.e. for specific states the single-photon weight P_1 can be increased by loss (an intuitive example is the two-photon Fock state).

3.1.3 Unbalanced beam splitter and different non-unit efficiency detectors

We now relax the assumptions that the beam splitter is balanced and the detector efficiencies are the same, i.e. we consider the case with a measurement performed with a beam splitter having an unknown transmission t and reflection $r = 1 - t$ and two detectors having different efficiencies labeled η_R and η_T . In this case, the observed statistics would be equivalently obtained with a beam splitter having a transmission coefficient $t' = t\eta_T/(t\eta_T + r\eta_R)$ and two detectors with the same efficiency $\eta = t\eta_T + r\eta_R$, as already mentioned below Eq. (2). This means that the measurement can be modeled with a first unbalanced beam splitter with transmission coefficient η corresponding to loss on the state to be characterized, an unbalanced beam splitter with transmission t' and two detectors with unit detection efficiency. In this case, the relation between $p_{\circ\circ}$ and P_0 is unchanged, i.e. $p_{\circ\circ} = P_0$. The probabilities $p_{\circ-} = p_{\circ\bullet} + p_{\circ\circ}$ and $p_{-\circ} = p_{\bullet\circ} + p_{\circ\circ}$ are however no longer the same. They are now given by $p_{\circ-} = \sum_n P_n (1 - r')^n$ and $p_{-\circ} = \sum_n P_n (1 - t')^n$. Hence, the quantity $\sum_n P_n \frac{1}{2^n}$ is no longer directly related to the probability $\mathbf{p}$. Nevertheless, it can be bounded from observable quantities, as $\sum_n P_n \frac{1}{2^n} \geq \min(p_{\circ-}, p_{-\circ})$.

We introduce $\hat{P}_1^R(\mathbf{p})$ which is defined analogously to $\hat{P}_1^T(\mathbf{p})$ by $\hat{P}_1^R(\mathbf{p}) = 4p_{-\circ} - 3p_{\circ\circ} - 1$. Using the definition of $p_{\circ-}$ and $p_{-\circ}$, we rewrite them in terms of probabilities of disjoint events as

$$\begin{aligned}\hat{P}_1^T(\mathbf{p}) &= 4p_{\circ\bullet} + p_{\circ\circ} - 1 \\ \hat{P}_1^R(\mathbf{p}) &= 4p_{\bullet\circ} + p_{\circ\circ} - 1.\end{aligned}\quad (9)$$

With this notation in hand, we conclude that the quantity

$$\hat{P}_1(\mathbf{p}) = \min\{\hat{P}_1^T(\mathbf{p}), \hat{P}_1^R(\mathbf{p})\} \quad (10)$$

is a benchmark for single-photon sources, without assumptions on the detector efficiencies and on the fact that the beam splitter is balanced. This means that from the outcome probabilities $\mathbf{p}$ of a usual auto-correlation measurement, we can compute $\hat{P}_1^T(\mathbf{p})$ and $\hat{P}_1^R(\mathbf{p})$, deduce their minimum $\hat{P}_1(\mathbf{p})$ and guarantee the tested source produces states with the weight of the single-photon component satisfying $P_1 = \langle 1|\rho|1\rangle \geq \hat{P}_1(\mathbf{p})$.

3.1.4 Dark counts and other imperfections

Finally let us briefly consider general passive detectors described by a POVM

$$\tilde{E}_\circ = \sum_{n \geq 0} e_\circ(n) |n\rangle\langle n| \quad \text{and} \quad \tilde{E}_\bullet = \sum_{n \geq 0} e_\bullet(n) |n\rangle\langle n|, \quad (11)$$

where $e_\circ(n) + e_\bullet(n) = 1$. Consider combining two such detectors in the $g^{(2)}$ setup, with the POVM $\{\tilde{E}_{\circ/\bullet}\}$ performed on the transmitted mode and $\{\tilde{E}'_{\circ/\bullet}\}$ on the reflected one. The resulting POVM elements for $a, b = \circ, \bullet$ read

$$\begin{aligned}\tilde{E}_{ab} &= \sum_{n \geq 0} \sum_{k=0}^n P_{BS}(k|n) \\ &\quad e_a(k) e'_b(n-k) |k, n-k\rangle\langle k, n-k|.\end{aligned}\quad (12)$$

with $P_{BS}(k|n) = \binom{n}{k} t^k r^{n-k}$.

Now let us assume these detectors are not too different from the textbook single-photon detector model of Eq. (1). Concretely we assume that for some η, t and r

$$|\text{tr}(\tilde{E}_{ab} - E_{ab})\rho| \leq \Delta, \quad (13)$$

where ρ can be any state susceptible to be prepared in the experiment. In particular, consider the case where both detectors satisfy

$$\begin{aligned}|e_\circ(n) - (1 - \eta)^n| &\leq \delta \\ |e'_\circ(n) - (1 - \eta')^n| &\leq \delta'\end{aligned}\quad (14)$$

for some η, η' , and all n on which the input state is supported. For $a = b = \circ$ we then find $|e_\circ(k) e'_\circ(n-k) - (1 - \eta)^k (1 - \eta')^{n-k}| \leq \delta + \delta' + \delta\delta'$ in Eq. (12). Similar inequalities hold for the other possible outcomes $\circ\bullet, \bullet\circ$ and $\bullet\bullet$, and lead to the bound of Eq. (13) with

$$\Delta = \delta + \delta' + \delta\delta'. \quad (15)$$

It follows that the probabilities $\tilde{\mathbf{p}} = (\tilde{p}_{\circ\circ}, \tilde{p}_{\circ\bullet}, \tilde{p}_{\bullet\circ}, \tilde{p}_{\bullet\bullet})$ observed with these detectors described by $\{\tilde{E}_{\circ/\bullet}\}$ and $\{\tilde{E}'_{\circ/\bullet}\}$, are close to the probabilities $\mathbf{p} = (p_{\circ\circ}, p_{\circ\bullet}, p_{\bullet\circ}, p_{\bullet\bullet})$ that would have been observed with textbook detectors, concretely

$$|\tilde{p}_{ab} - p_{ab}| \leq \Delta. \quad (16)$$

With the help of Eq. (9) it is then straightforward to lower bound the value of the benchmark

$$\hat{P}_1(\mathbf{p}) \geq \min\{\hat{P}_1^T(\tilde{\mathbf{p}}), \hat{P}_1^R(\tilde{\mathbf{p}})\} - 5\Delta. \quad (17)$$

given the coincidence probabilities $\tilde{\mathbf{p}}$ observed with any detectors abiding to Eq. (13).

Here, it is worth noting that the POVM model of Eq. (1) does not account for dark counts. These can be modeled by setting a nonzero click probability for the vacuum $e_\bullet(0) = p_{dc}$, implying $|e_\circ(n) - (1 - \eta)^n| \leq$

$\delta = p_{\text{dc}}$. Then accordingly to Eqs. (15,17), to account for the effect of dark counts one can use the bound

$$\hat{P}_1(\mathbf{p}) \geq \min\{\hat{P}_1^T(\tilde{\mathbf{p}}), \hat{P}_1^R(\tilde{\mathbf{p}})\} - 10 p_{\text{dc}}, \quad (18)$$

where we only took the leading order correction in p_{dc} , which is typically fairly small.

3.2 Measurement calibration dependent benchmark

The benchmark we proposed relies on no assumption on the characteristics of the beam splitter or the two detectors used in the auto-correlation measurement. This prevents a miscalibration of the measurement apparatus that could result in an overestimation of the quality of the single-photon source. Nevertheless, as one would expect, the bound on the quality of the tested source (see Eq. (10)) is reduced when using an unbalanced beam splitter or inefficient detectors. We now discuss a way to characterize the intrinsic quality of the source by making additional assumptions on the measurement apparatus. The basic idea is to exploit estimations of different losses in photonic experiments. In particular, we assume that the detector efficiencies $\eta_{R(T)}$ and the beam splitter reflectivity r are bounded, that is $\eta_{R(T)} \leq \hat{\eta}_{R(T)}$ and $r \in [1 - \hat{r}, \hat{r}]$. We show in Appendix D that the condition $\langle 1 | \rho | 1 \rangle \leq P$ implies that the two following inequalities hold

$$\begin{aligned} \hat{P}_1^{T*}(\mathbf{p}) &= C_1(\hat{t}, \hat{\eta}_T) p_{\bullet\bullet} - C_2(\hat{t}, \hat{\eta}_T, \hat{\eta}_R) p_{\bullet\bullet} \leq P, \\ \hat{P}_1^{R*}(\mathbf{p}) &= C_1(\hat{r}, \hat{\eta}_R) p_{\bullet\bullet} - C_2(\hat{r}, \hat{\eta}_R, \hat{\eta}_T) p_{\bullet\bullet} \leq P, \end{aligned} \quad (19)$$

where the coefficients C_1 and C_2 , defined as

$$\begin{aligned} C_1(x, \eta) &= \frac{1}{x\eta} \\ C_2(x, \eta_1, \eta_2) &= \frac{1}{x\eta_1} \left(\frac{2 - x\eta_1}{2(1 - x\eta_2)} - 1 \right) \end{aligned} \quad (20)$$

have been optimized such that $\hat{P}_1^{T*}(\mathbf{p})$ and $\hat{P}_1^{R*}(\mathbf{p})$ give the tightest bound on P . Finally, one can choose the best among the two bounds, giving rise to the benchmark

$$\hat{P}_1^*(\mathbf{p}) = \max\{\hat{P}_1^{T*}(\mathbf{p}), \hat{P}_1^{R*}(\mathbf{p})\} \leq P_1 \quad (21)$$

for the single-photon probability, which takes advantage of the additional experimental knowledge.

Again we can account for deviations of the POVMs $\{\hat{E}_{\circ/\bullet}\}$ and $\{\hat{E}'_{\circ/\bullet}\}$ describing the detectors from the textbook model of Eq. (1), as quantified by Eq. (15). Repeating the analysis of Sec. 3.1.4, we obtain the bounds

$$\begin{aligned} \hat{P}_1^{T*}(\mathbf{p}) &\geq \hat{P}_1^{T*}(\tilde{\mathbf{p}}) - (C_1(\hat{t}, \hat{\eta}_T) + |C_2(\hat{t}, \hat{\eta}_T, \hat{\eta}_R)|) \Delta \\ \hat{P}_1^{R*}(\mathbf{p}) &\geq \hat{P}_1^{R*}(\tilde{\mathbf{p}}) - (C_1(\hat{r}, \hat{\eta}_R) + |C_2(\hat{r}, \hat{\eta}_R, \hat{\eta}_T)|) \Delta. \end{aligned} \quad (22)$$

To account for dark counts one may use $\Delta = 2p_{\text{dc}}$. Recall that here $\tilde{\mathbf{p}} = (\tilde{p}_{\circ\circ}, \tilde{p}_{\circ\bullet}, \tilde{p}_{\bullet\circ}, \tilde{p}_{\bullet\bullet})$ is the vector of probabilities observed with the real detectors, while $\mathbf{p}$ gives the probabilities that would have been observed with the textbook model and that are used in the benchmark.

3.3 Multi-mode sources

We have so far considered sources emitting light in a single mode, or equivalently that the emitted light is filtered in all the auxiliary degrees of freedom so that a single mode of light is detected. We will now briefly consider the situation where the detected state is multi-mode, each mode being associated to an annihilation operator a_k satisfying $[a_k, a_\ell^\dagger] = \delta_{k\ell}$. The no-click and click events for a multi-mode input are associated to the POVM elements $E_\circ = \bigotimes_k (1 - \eta) a_k^\dagger a_k$ and $E_\bullet = \mathbb{1} - E_\circ$ – a detector does not click only if none of the modes triggers a click. To a multi-mode state ρ one associates the distribution P_n of the *total photon number* operator $\hat{n} = \sum_k a_k^\dagger a_k$. Assuming that the beam splitter acts identically on all modes, the measurement apparatus is only sensitive to the total number of photons $\hat{n}$. That is, the POVM elements $\{E_{\circ\circ}, E_{\circ\bullet}, E_{\bullet\circ}, E_{\bullet\bullet}\}$ are given by Eq. (2) albeit with $a^\dagger a$ replaced by $\hat{n}$, see Appendix B. We thus conclude that the quantities $\hat{P}_1(\mathbf{p})$ and $\hat{P}_1^*(\mathbf{p})$ that have been derived, can be readily used to benchmark the probability that a multi-mode source emits a single photon

$$P_1 = \text{tr} \left(\varrho \sum_k a_k^\dagger |0\rangle\langle 0| a_k \right). \quad (23)$$

It is worth noting that a high P_1 for a multi-mode source does not guarantee that the single-photon probability is high in any of the individual modes. As a practical example, consider the multi-mode single-photon state

$$\varrho = \frac{1}{N} \sum_{j=1}^N a_j^\dagger |0\rangle\langle 0| a_j. \quad (24)$$

This state has exactly one photon in total $\hat{n} = 1$. Nevertheless, it is not a good *single-mode* single-photon state. The probability to find exactly one photon in any single mode is only $1/N$, therefore two such states would only exhibit a limited two-photon interference (bunching). In other words ϱ is not a pure single-photon state $|1\rangle\langle 1|_1$, the degree of freedom that distinguishes between the different modes is in a highly mixed states. Depending on the application, the mode-purity (or single-mode character) of the source may be either irrelevant (e.g. for some quantum random number generators) if the interference between different sources plays no role, or crucial (e.g. quantum repeaters, boson sampling or photonic quantum computation) if it is at the heart of the task.

In Appendix G, we show that one can guarantee that the source produces a single-mode state with high P_1 if it is reasonable to assume that the multi-mode state is a product $\varrho = \bigotimes_k \rho_k$. In general, however, the auto-correlation measurement is intrinsically insensitive to the multi-mode characteristics of the source, as illustrated by the example of Eq. (24). In principle, it is possible to extract a single-mode state from a multi-mode source by filtering the auxiliary degrees of freedom, e.g. using a single-mode fiber for spatial degrees of freedom and a spectral filter for frequency domain. In any case, the mode-purity of the detected state has to be verified with a different set of measurements, e.g. via Hong-Ou-Mandel interference [13] or based on a physical model of the source. For example, in the case of a heralded single-photon source, the spectral purity can be determined by measuring the signal-idler joint spectral intensity [14], as will be explained in Sec. 6.3.

A systematic analysis of the characterization of the mode-purity of a source is beyond the scope of this paper. Nevertheless, it is worth mentioning that with the additional information on the multi-mode characteristics of the source and the knowledge of the total single-photon probability P_1 it is possible to bound the probability to find a single photon in the mode of interest $P_1^{[1]}$ (denoted mode 1). In particular, given a bound on the probability that all the modes but one are empty $\text{tr} \varrho(1_1 \otimes |0\rangle\langle 0|_2 \otimes \cdots \otimes |0\rangle\langle 0|_N) \geq 1 - \varepsilon$, one can conclude that

$$P_1^{[1]} = \text{tr} |1\rangle\langle 1|_1 \rho_k \geq P_1 - \varepsilon, \quad (25)$$

since the maximal contribution to P_1 from the other modes is ε . Here, $\rho_1 = \text{tr}_{2,\dots,N} \varrho$ is the marginal state of the first mode, which can in principle be filtered from the source. A bound on $P_1^{[1]}$ can then be used to, e.g., quantify the Wigner-negativity of the state ρ_1 , see Sec. 5.

4 Finite statistics

In this section, we analyze finite size effects for the benchmark.

We sketch an analysis to account for finite statistics in any experiment aiming to evaluate $\hat{P}_1(\mathbf{p})$, the benchmark for single-photon sources derived in Sec. 3. For a measurement round described by $\mathbf{p}$ we associate a random variable X_T that takes different real values depending on the measurement result

$$X_T = \begin{cases} 3 & (\circ\bullet) \\ 0 & (\circ\circ) \\ -1 & (\bullet\circ) \text{ or } (\bullet\bullet) \end{cases} \quad (26)$$

This random variable satisfies $\mathbb{E}(X_T) = \hat{P}_1^T(\mathbf{p})$ in Eq. (9) (here and further $\mathbb{E}$ denotes the expected value of a random variable). Analogously, we define X_R by

exchanging the role of the two detectors, such that it satisfies $\mathbb{E}(X_R) = \hat{P}_1^R(\mathbf{p})$.

In general, the source may prepare a different state $\rho^{(i)}$ at each round, corresponding to different probabilities $\mathbf{p}^{(i)}$ of the measurement outcomes. This means that in each round, we sample different random variables $X_T^{(i)}$ and $X_R^{(i)}$, which are independent between rounds given the sequence of states $\rho^{(1)}, \dots, \rho^{(n)}$ produced in the experiment. In this case, a reasonable figure of merit is the average quality of the state prepared by the source $\bar{P}_1 = \frac{1}{n} \sum_{i=1}^n P_1^{(i)}$ where $P_1^{(i)}$ is the probability of the single-photon component of the state $\rho^{(i)}$. Because the transmission and reflection coefficients of the beam splitter can be considered to be constant, either $\mathbb{E}(X_T^{(i)}) \geq \mathbb{E}(X_R^{(i)})$ or $\mathbb{E}(X_T^{(i)}) \leq \mathbb{E}(X_R^{(i)})$ holds for all i . This means that the average single-photon weight fulfills

$$\bar{P}_1 \geq \min\{\mathbb{E}(\bar{X}_T), \mathbb{E}(\bar{X}_R)\}, \quad (27)$$

where $\bar{X}_{T(R)} = \frac{1}{n} \sum_{i=1}^n X_{T(R)}^{(i)}$. Finally, we use the Hoeffding 1963 theorem [15] to show that

$$\hat{q}_\alpha = \min\{\bar{X}_T, \bar{X}_R\} - \sqrt{\frac{16 \log(1/\alpha)}{2n}} \quad (28)$$

is a one-sided confidence interval on $\bar{P}_1$ with confidence α (see Appendix F). Precisely, with probability $1 - \alpha$ the observed value of $\hat{q}_\alpha$ lower bounds $\bar{P}_1$. It might be convenient to note that the quantity $\min\{\bar{X}_T, \bar{X}_R\}$ of this confidence interval can be computed using

$$\min\{\bar{X}_T, \bar{X}_R\} = \frac{4 \min\{n_{\circ\bullet}, n_{\bullet\circ}\} - n_{\circ\bullet} - n_{\bullet\circ} - n_{\bullet\bullet}}{n} \quad (29)$$

with e.g. $n_{\bullet\bullet}$ counting the number of outcomes $\bullet\bullet$.

Analogously, one can derive a one-sided confidence intervals associated to the calibration-dependent benchmark derived in Sec. 3.2. With very similar arguments one can show that both

$$\begin{aligned} \hat{q}_\alpha^{(T)*} &= \frac{C_1(\hat{t}, \hat{\eta}_T) n_{\circ\bullet} - C_2(\hat{t}, \hat{\eta}_T, \hat{\eta}_R) n_{\bullet\bullet}}{n} \\ &\quad - (C_1(\hat{t}, \hat{\eta}_T) + C_2(\hat{t}, \hat{\eta}_T, \hat{\eta}_R)) \sqrt{\frac{\log(1/\alpha)}{2n}}, \\ \hat{q}_\alpha^{(R)*} &= \frac{C_1(\hat{r}, \hat{\eta}_R) n_{\bullet\circ} - C_2(\hat{r}, \hat{\eta}_R, \hat{\eta}_T) n_{\bullet\bullet}}{n} \\ &\quad - (C_1(\hat{r}, \hat{\eta}_R) + C_2(\hat{r}, \hat{\eta}_R, \hat{\eta}_T)) \sqrt{\frac{\log(1/\alpha)}{2n}}, \end{aligned} \quad (30)$$

where the functions C_1 and C_2 defined in Eq. (20) are confidence intervals for $\bar{P}_1$. Since they are derived from the same data, for a statistically meaningful statement one has to chose parameter α before computing the confidence interval. On the other hand, it is straightforward to see that

$$\hat{q}_{2\alpha}^* = \max\{\hat{q}_\alpha^{(T)*}, \hat{q}_\alpha^{(R)*}\} \quad (31)$$

is also a confidence interval at confidence level $1 - 2\alpha$. See App. F for a detailed derivation.

5 Relation to the non-classicality of the source

The data obtained from auto-correlation type measurements are known to be valuable for witnessing and quantifying various forms of non-classicality, including the non-positivity of the P-function and quantum non-Gaussianity [8–10]. We now show that the knowledge of P_1 in a single bosonic mode (as e.g. provided by our benchmarks $\hat{P}_1(\mathbf{p})$ and $\hat{P}_1^*(\mathbf{p})$) can reveal Wigner-negativity [16], arguably the strongest form of non-classicality for a bosonic mode. In particular, Wigner-negativity implies the non-positivity of the P-function [17]. Similarly, it implies that the corresponding state is non-Gaussian, as a Gaussian state has a Gaussian (and thus positive) Wigner function¹. Thus, demonstrating Wigner-negativity for a light source brings evidence of its strong quantum nature. Note that it has been shown recently that witnesses of Wigner-negativity can be derived systematically using a hierarchy of semidefinite programs [19]. Our contribution is more specific and aims at witnessing Wigner-negativity simply and directly from $\hat{P}_1(\mathbf{p})$ or $\hat{P}_1^*(\mathbf{p})$ observed on a single mode.

5.1 Wigner-negativity witness

The Wigner function is a representation of a single-mode state ρ in terms of the following quasi-probability distribution [20]

$$W_\rho(\beta) = \frac{2}{\pi} \text{Tr}(\mathbf{D}_\beta(-1)^{a^\dagger a} \mathbf{D}_\beta^\dagger \rho), \quad (32)$$

with $\int d\beta^2 W_\rho(\beta) = 1$. Here, $\mathbf{D}_\beta = e^{a^\dagger \beta - a \beta^*}$ is the displacement operator with a complex amplitude β . Applying Eq. (32) to a Fock state gives [17]

$$W_{|n\rangle\langle n|}(\beta) = \frac{2(-1)^n}{\pi} e^{-2|\beta|^2} L_n(4|\beta|^2) \quad (33)$$

where L_n is the Laguerre polynomial. Note that the following bound on the Laguerre polynomials $e^{-x/2}|L_n(x)| \leq 1$, see e.g. Eq. (18.14.8) in [21], leads to a bound on the Wigner function of Fock states $|W_{|n\rangle\langle n|}(\beta)| \leq \frac{2}{\pi}$. Note also that $L_1(x) = 1 - x$.

With the help of Eq. (33), the upper bound on the Wigner function of Fock states and the definition of the Laguerre polynomial $L_1(x)$, it is easy to see that the Wigner function of any mixture of Fock states

¹In addition, Hudson's theorem [18] tells us that any pure state with a positive Wigner function is Gaussian.

$\rho = \sum p_n |n\rangle\langle n|$ satisfies ²

$$\begin{aligned} W_\rho(\beta) &= P_1 W_{|1\rangle\langle 1|}(\beta) + \sum_{n \neq 1} P_n W_{|n\rangle\langle n|}(\beta) \\ &\leq \frac{2}{\pi} \left(-P_1(1 - 4|\beta|^2)e^{-2|\beta|^2} + (1 - P_1) \right). \end{aligned} \quad (34)$$

Focusing on the origin $\beta = 0$, we get $W_\rho(0) \leq 2\frac{1-2P_1}{\pi}$ which is negative if P_1 is larger than $\frac{1}{2}$. Hence, if one concludes from the measurement of $\mathbf{p}$ that $\hat{P}_1(\mathbf{p}) > \frac{1}{2}$, one can conclude that the measured state is Wigner-negative (recall that we assumed that the state ρ is single-mode).

5.2 Wigner-negativity measure

A natural way to quantify the negativity of the Wigner representation of a given state ρ is to measure the total quasi-probability for which the function $W_\rho(\beta)$ takes negative values [11], i.e.

$$N_W(\rho) = \int d\beta^2 \frac{|W_\rho(\beta)| - W_\rho(\beta)}{2}, \quad (35)$$

which is manifestly zero for states with a positive Wigner function. In the Appendix E we show that $N_W(\rho)$ is non-increasing under Gaussian operations, which justifies its use as a measure of Wigner-negativity. Note that with the help of Ineq. (34), we show that $N_W(\rho)$ satisfies

$$\begin{aligned} N_W(\rho) &\geq F(P_1) = \begin{cases} \frac{3(1-P_1)(4w^2+3)}{8w} + P_1 - 2 & P_1 > \frac{1}{2} \\ 0 & P_1 \leq \frac{1}{2} \end{cases} \\ \text{with } w &= w_0 \left(\frac{\sqrt{e}}{2} \frac{1 - P_1}{P_1} \right), \end{aligned} \quad (36)$$

where w_0 is the principal branch of the Lambert W function. The function $F(P_1)$ is non-decreasing. Hence, from the measurement of $\mathbf{p}$, we get a lower bound $\hat{P}_1(\mathbf{p})$ on P_1 that can be used to lower bound $N_W(\rho)$ using $F(\hat{P}_1(\mathbf{p}))$. The bound (36) is tight by construction in the ideal case $N_W(|1\rangle) = F(1) = \frac{9}{4\sqrt{e}} - 1 \approx 0.36$.

By computing $F''(P_1) \geq 0$ we show that the function $F(P_1)$ in Eq. (36) is convex. This property will be used in the following section, where we discuss the finite statistics effects.

²For a general state $\varrho = \sum_{nm} c_{nm} |n\rangle\langle m|$ with Wigner function $W_\varrho(\beta)$, one can always define the corresponding Fock state mixture $\rho = \sum_n P_n |n\rangle\langle n|$ with $p_n = c_{nn}$. Its Wigner function reads $W_\rho(\beta) = W_\varrho(|\beta|) = \int d\varphi W_\varrho(|\beta|e^{i\varphi}) = \langle W_\varrho(|\beta|e^{i\varphi}) \rangle_\varphi$. The two functions coincide at the origin $W_\rho(0) = W_\varrho(0)$. Furthermore, W_ρ can only be non-positive if W_ϱ is non-positive, and $N_{W_\rho} \leq N_{W_\varrho}$ (introduced at the end of the section) follows from $|\langle W(|\beta|e^{i\varphi}) \rangle_\varphi| \leq \langle |W(|\beta|e^{i\varphi})| \rangle_\varphi$.

5.3 p-value to witness Wigner-negativity

We now present the extension of finite statistics analysis presented in Sec. 4 to the task of Wigner-negativity detection and quantification.

First, let us now consider the witness of Wigner-negativity discussed in Sec. 5.1, that is $W_\rho(0) \geq 0 \implies \hat{P}_1(\mathbf{p}) \leq 1/2$, and quantify the statistical significance of its contrapositive given the measurement data. This can be done by computing the p-value associated to the hypothesis that the Wigner function of the state is positive. As before, we consider the general case where a different state $\rho^{(i)}$ may be prepared at each run. Nevertheless, at each round the bound $W_{\rho^{(i)}}(0) \leq \frac{2}{\pi}(1 - 2\bar{P}_1)$ holds. Therefore, for the sequence of states prepared in the experiment, the average Wigner function at the origin satisfies

$$\bar{W}(0) = \frac{1}{n} \sum_{i=1}^n W_{\rho^{(i)}}(0) \leq \frac{2}{\pi}(1 - 2\bar{P}_1), \quad (37)$$

and is negative if $\bar{P}_1 > \frac{1}{2}$. Given some value of $\min\{\bar{X}_T, \bar{X}_R\}$ recorded after n measurement rounds, we show in Appendix F that for any collection of n states with $\bar{W}(0) \geq 0$, the probability that the results are equal or exceed the observed value of $\min\{\bar{X}_T, \bar{X}_R\}$ is given by

$$\text{p-value} \leq \exp\left(-\frac{2n(\min\{\bar{X}_T, \bar{X}_R\} - \frac{1}{2})^2}{16}\right), \quad (38)$$

for $\min\{\bar{X}_T, \bar{X}_R\} > \frac{1}{2}$. In other words, given the observed value of $\min\{\bar{X}_T, \bar{X}_R\}$, the probability that it is coming from states that are Wigner-positive on average is bounded by the right-hand side of Ineq. (38). In App. F one finds a bound on the p-value for the calibration-dependent setting.

5.4 Confidence interval on the measure of Wigner-negativity

Finally, the convexity of the function $F(P_1)$ in Eq. (36) implies that the average Wigner-negativity satisfies $\bar{N}_W = \frac{1}{n} \sum_i N_W(\rho^{(i)}) \geq \frac{1}{n} \sum_i F(P_1^{(i)}) \geq F(\bar{P}_1)$. Therefore, a confidence interval q_α for $\bar{P}_1$

$$nw_\alpha = F(q_\alpha) \quad (39)$$

is a one-sided confidence interval on $\bar{N}_W$, that is, with probability $1 - \alpha$, the average Wigner-negativity as quantified by $\bar{N}_W$ is lower bounded by $nw_\alpha = F(q_\alpha)$. This can be used both in the calibration-independent $\hat{n}w_\alpha = F(\hat{q}_\alpha)$ and calibration-dependent $\hat{n}w_\alpha^* = F(\hat{q}_\alpha^*)$ settings.

5.5 The crucial role of the single-mode hypothesis

It is important to emphasize that the single-mode hypothesis is crucial in order to relate P_1 (or its esti-

mated value $\hat{P}_1(\mathbf{p}), \hat{P}_1^*(\mathbf{p})$) to Wigner-negativity. In particular, the multi-mode single-photon state ϱ of Eq. (24) becomes Wigner-positive for $N > 2$. In section 3.3 we have discussed how the single-mode character of the emitted radiation can be verified in practice. Here we merely recall that a bound of the form $P_1^{[1]} \geq P_1 - \varepsilon$, where $P_1^{[1]}$ is the single photon probability for a given mode, can be readily used to verify Wigner-negativity. One simply has

$$nw_\alpha^{[1]} \geq F(q_\alpha - \varepsilon)$$

$$\text{p-value} \leq \exp\left(-\frac{2n(\min\{\bar{X}_T, \bar{X}_R\} - \varepsilon - \frac{1}{2})^2}{16}\right), \quad (40)$$

if $\min\{\bar{X}_T, \bar{X}_R\} > \frac{1}{2} + \varepsilon$, for Wigner-negativity of the said mode.

It is worth mentioning that a possibility to ensure that the detected radiation is single-mode is to perform homodyne measurements. In such a measurement the incoming beam is mixed with a strong local oscillator on a beam splitter, the intensity of the output beams are then measured with linear detectors and subtracted. Under the assumption that the local oscillator is single-mode³ the obtained signal is only sensitive to the single input mode identified by the local oscillator. Photon number statistics, and P_1 in particular, can be reconstructed from the statistics of a phase-averaged quadrature measurement [22], i.e. a homodyne measurement with a phase-randomized local oscillator. This offers the possibility to use our bound on the Wigner-negativity $N_W(\rho)$ in Eq. (36) with measurements that are guaranteed to pick up a single mode.

6 Experiment

To demonstrate the feasibility of our tools, we experimentally benchmark, witness and quantify the non-classical nature of a heralded single-photon source [23] that is optimized for high efficiency of the heralded photon [24]. A periodically poled potassium titanyl phosphate (PPKTP) crystal is pumped by a Ti:Sapphire laser at $\lambda_p = 771.8$ nm in the picosecond pulsed regime with a repetition rate of 76 MHz to create nondegenerate photon pairs at $\lambda_s = 1541.3$ nm (signal) and $\lambda_i = 1546.1$ nm (idler) via type-II spontaneous parametric down-conversion (SPDC). The pair creation probability per pump pulse is set to $P_{\text{pair}} \approx 1.0 \times 10^{-3}$ and high-purity heralded signal photons are ensured by spectrally filtering the heralding idler photons using a dense wavelength division multiplexer

³Note that the situation where n modes $a_1, \dots, a_n$ are prepared in coherent states $|\alpha_1\rangle |\alpha_2\rangle \dots |\alpha_n\rangle$ with fixed phase relations, can be viewed as a coherent state of the mode $\bar{a} = \frac{\sum_k a_k \alpha_k}{\sqrt{\sum_k |\alpha_k|^2}}$, plus $n - 1$ modes in the vacuum states.

at ITU channel 39. From a joint spectral intensity measurement [14], we estimate the spectral purity of the heralded photon to be $98.59\% \pm 0.04\%$. In this way, we herald signal photons at a rate of 19.1 kcps.

For the heralded auto-correlation measurement, the signal photon is sent to a 50/50 fiber coupler (AFW FOBC). All photons are detected by MoSi superconducting nanowire single-photon detectors [25] and time-correlated single-photon counting in a programmable time-to-digital converter (ID Quantique ID900) is used to register the detection events. Data are acquired for 200s in order to evaluate $\mathbf{p} = (p_{\circ\circ}, p_{\bullet\circ}, p_{\circ\bullet}, p_{\bullet\bullet})$ for the signal photons after the 50/50 beam splitter.

The dark count probabilities of the detectors we used for the auto-correlation measurement are fairly small, $p_{dc} \leq 4 \times 10^{-7}$, and can be completely neglected. That is, the dark count corrections ($\approx 4 \times 10^{-6}$) to the estimated values of the benchmark in Eqs. (18,22) are more than two orders of magnitude lower than the statistical noise, see Tables 1 and 3.

6.1 Calibration-independent benchmark

In a first step we apply our benchmark to the experimental results without taking the splitting ratio of the beam splitter and the detector efficiencies into account. The overall efficiency is 25 % for the heralding idler photons and 62 % for the heralded signal photons. In order to simulate a less efficient single-photon source, we introduce loss by inserting a fiber coupled variable attenuator (JDS Uniphase MV47W) into the heralded photon path before the 50/50 beam splitter and repeat the auto-correlation measurement for eight different transmission efficiencies η_{att} . Each transmission efficiency leads to a value for the pair $(\min(p_{\circ\circ}, p_{\bullet\circ}), p_{\circ\circ})$ that is represented by a black cross in Fig. 2. In the same figure, we represent the polytope $\mathbf{Q}_P$ (defined after Eq. (7)) containing all possible values $(\min(p_{\circ\circ}, p_{\bullet\circ}), p_{\circ\circ})$ associated to states ϱ with $\langle 1|\varrho|1\rangle \leq P$. Four polytopes are represented corresponding to the values $P \in \{0.25, 0.5, 0.75, 1\}$. A measurement result associated to a black cross lying outside a polytope $\mathbf{Q}_P$ is guaranteed to come from a state with a single-photon component satisfying $P_1 > P$.

$\eta_{s,tot}$	$\hat{P}_1^T$	$\hat{P}_1^R$	$\hat{q}_{\alpha=10^{-10}}$
62 %	0.561(1)	0.678(1)	0.554
52 %	0.460(1)	0.573(1)	0.453
42 %	0.376(1)	0.465(1)	0.369

Table 1: Results of the measurement for the three highest transmission efficiencies $\eta_{s,tot}$ of the heralded single-photon state. The values for $\hat{P}_1^T$ and $\hat{P}_1^R$ are calculated according to Eq. (9). For the finite statistics analysis we calculate the confidence interval $\hat{q}_\alpha \leq \bar{P}_1$ from Eq. (28) for the confidence level $1 - \alpha = 1 - 10^{-10}$.

Mode	η_{tot}	η_c	η_f	η_t	η_d
Idler	25 %	80 %	50 %	83 %	75 %
Signal	62 %	80 %	-	R	43 %
				T	44 %

Table 2: Characterization of the loss for idler (heralding) and signal (heralded) modes. η_{tot} , total efficiency; η_c , fiber coupling efficiency; η_f , spectral filter transmission; η_t , fiber transmission including the insertion loss of the 50/50 fiber coupler, connectors and telecom fiber isolators for further pump rejection; η_d , detector efficiency.

For the measurements with the three highest transmission efficiencies, we give the results of our benchmark in Tab. 1. We conclude for the highest transmission for example, that the measured states have on average a single-photon component with a weight $\bar{P}_1 \geq 0.554$ with a confidence level of $1 - 10^{-10}$.

6.2 Calibration-dependent benchmark

To compute the value of the calibration-dependent benchmark one needs to estimate the detector efficiencies and the reflection/transmission coefficient of the beam splitter. In order to characterize the detectors, we use the standard method, see e.g. [25] for a detailed description. For our setup we find that the beam splitter coefficients are bounded by $\hat{r} \in [0.49, 0.50]$ and the detector efficiencies are upper bounded by $(\hat{\eta}_R, \hat{\eta}_T) = (0.95, 0.88)$. The upper bounds for the detector efficiencies are obtained from results of the measured detection efficiencies given in Tab. 2 by adding three times the measurement uncertainty of around 0.01, see Supplementary Material of [25].

Under the assumptions that the $(\hat{r}, \hat{\eta}_R, \hat{\eta}_T)$ belong to these intervals, the values of $\hat{P}_1^{T*}(\mathbf{p})$ and $\hat{P}_1^{R*}(\mathbf{p})$ as measured in our experiment are given in Tab. 3 for the three highest transmission efficiencies. The confidence interval $\hat{q}_\alpha^* \leq \bar{P}_1$ is also reported for a confidence level of $1 - \alpha = 1 - 10^{-10}$.

$\eta_{s,tot}$	$\hat{P}_1^{T*}$	$\hat{P}_1^{R*}$	$\hat{q}_{\alpha=10^{-10}}^*$
62 %	0.658(1)	0.683(1)	0.677
52 %	0.544(1)	0.573(1)	0.566
42 %	0.444(1)	0.466(1)	0.459

Table 3: Results of the measurement including the imperfect beam splitter ratio with $(1 - \hat{r}, \hat{r}) = (0.49, 0.50)$ and the non-unit detection efficiencies by using the upper bounds $(\hat{\eta}_R, \hat{\eta}_T) = (0.95, 0.88)$. The values for $\hat{P}_1^{T*}$ and $\hat{P}_1^{R*}$ are calculated with Eq. (19). The confidence intervals $\hat{q}_\alpha^* \leq \bar{P}_1$ in the finite statistics analysis are calculated for a confidence level of $1 - \alpha = 1 - 10^{-10}$.

6.3 Non-classicality of the source

As already mentioned, there is no guarantee that the single-mode assumption is exactly satisfied in our experiment, hence we have to estimate the mode purity of the source. We assume that the spatial mode purity is guaranteed by coupling the photons into a single-mode fiber. The polarization-purity is ensured by the fact that the signal and idler photons are separated with a polarizing beam splitter. To estimate the spectral purity, we apply the standard approach that relies on a physical model of the source, which we believe to properly describe the experiment. Precisely, we assume that the SPDC process responsible for the generation of the photon pairs is of the form

$$H_{\text{SPDC}} \propto \int dw_s dw_i f(w_s, w_i) a^\dagger(w_s) b^\dagger(w_i) + \text{h.c.}, \quad (41)$$

where $a(w_s)$ and $b(w_i)$ are the frequency field-modes of the signal/idler photons with $[a(w), a^\dagger(w')] = [b(w), b^\dagger(w')] = \delta(w - w')$. At low pumping power, we reconstruct $f(w_s, w_i)$ with a signal-idler joint spectral intensity measurement. Via a 2D-Gaussian fit we perform a singular value decomposition of $f(w_s, w_i)$ to rewrite the interaction in the form

$$H_{\text{SPDC}} \propto \sum_k \sqrt{\lambda_k} a_k^\dagger b_k^\dagger + \text{h.c.}, \quad (42)$$

where $[a_k, a_\ell^\dagger] = [b_k, b_\ell^\dagger] = \delta_{k\ell}$ now describe discrete spectral modes. With this procedure the largest Schmidt coefficient λ_1 is computed to be $\lambda_1 = 0.99292(18)$, where the standard deviation $\sigma_{\lambda_1} = 1.8 \times 10^{-4}$ is obtained from a Monte Carlo method assuming Poissonian count statistics in the joint spectral intensity measurement. Therefore, assuming that the efficiency of the trigger detector is the same for all idler modes, we obtain the leading order estimate

$$1 - \varepsilon \approx \lambda_1 - 3\sigma_{\lambda_1} = 99.24\%, \quad (43)$$

of the probability that the modes $a_{k \geq 2}$ are empty conditional to the detection of an idler photon. This corresponds to the mode-purity of $\sum_k \lambda_k^2 \approx 98.59\%$.

With the help of Eq. (25) we can take this into account for the quantification of Wigner-negativity, resulting in $n\hat{w}_{\alpha=10^{-10}} = 0.0046$ for the case of no added loss on the heralded single-photon state and no assumptions on the calibration of the measurement apparatus. The corresponding p-value and the results for the measurement-apparatus-dependent case are given in Tab. 4.

7 Conclusion

Auto-correlation measurements are commonly used to assess the quality and the quantum nature of single-photon sources, that is, they are used to check that a given source does not emit more than one photon

$\eta_{s,\text{tot}}$	$n\hat{w}_{\alpha=10^{-10}}$	p-value	$n\hat{w}_{\alpha=10^{-10}}^*$	p-value*
62 %	0.0046	10^{-603}	0.053	10^{-6420}
52 %	0	$\times$	0.0072	10^{-894}
42 %	0	$\times$	0	$\times$

Table 4: Wigner-negativity in our experiment for the three highest transmission efficiencies $\eta_{s,\text{tot}}$ of the heralded single-photon state. The confidence interval on the measure of Wigner-negativity $n\hat{w}_\alpha \leq \bar{N}_W$ is obtained from Eq. (39) for the confidence level $\alpha = 10^{-10}$, assuming the reduced single-mode $P_1^{[1]}$ as given in Eq. (25). Further, we give the p-value according to Eq. (38) associated with the hypothesis that the measured states are on average Wigner-positive. The quantities with a * are taking the detector efficiencies into account and are obtained accordingly from Eqs. (19) and (25).

and its emission is non-classical in the sense that its P-distribution is non-positive or that its state is non-Gaussian. We have shown that the statistics obtained from these measurements is actually richer. They can be used to lower bound the probability that a given source actually produces a single photon. We argued that this probability is a good benchmark for single-photon sources as it captures both its quality and its efficiency. Moreover, we showed that if the mode purity of the source can be assessed the lower bound on the single-photon emission probability can be used to witness and quantify the negativity of the Wigner function, a stronger form of non-classicality than the negativity of the P-distribution and the non-Gaussianity. We have proposed practical tools to benchmark single-photon sources and characterize its Wigner-negativity this way. With this material in hand, we hope that the community which is developing single-photon sources could exploit the statistics of their auto-correlation measurements in a more enlightening way.

Acknowledgments

We thank R.J. Warburton for fruitful discussions at an early stage of the project. This work was supported by the Swiss National Science Foundation (SNSF) under Grant No. 200020-182664. E.O acknowledges support from the Government of Spain (FIS2020-TRANQI and Severo Ochoa CEX2019-000910-S), Fundació Cellex, Fundació Mir-Puig, Generalitat de Catalunya (CERCA, AGAUR SGR 1381) and from the ERC AdGCERQUT.

Data availability

The data supporting the experimental results within this paper are available on the Zenodo data repository [26].

A Auto-correlation function and auto-correlation measurement

Here we briefly recall the definitions of the auto-correlation function $g^{(2)}$. We will use the notation introduced in Section 2, 3 and 3.1.1 of the main text, and assume that the beam splitter is balanced and the two detectors have equal efficiency, i.e. $t = r = \frac{1}{2}$ in Eq. (2).

Historically [27], the auto-correlation function was defined as the ratio

$$g^{(2)} = \frac{\langle a^{\dagger 2} a^2 \rangle}{\langle a^{\dagger} a \rangle^2} = \frac{\langle (a^{\dagger} a)^2 \rangle - \langle a^{\dagger} a \rangle}{\langle a^{\dagger} a \rangle^2} \quad (44)$$

and the efficiency of the source can be characterized by the average number of photons it emits $I = \langle a^{\dagger} a \rangle$. Then it is not difficult to see that the bound

$$P_1 \geq 2\langle a^{\dagger} a \rangle - \langle (a^{\dagger} a)^2 \rangle = I - I^2 g^{(2)} \quad (45)$$

is a tight benchmark. To see this note that in the $(\langle a^{\dagger} a \rangle, \langle (a^{\dagger} a)^2 \rangle)$ plane the quantity $w = 2\langle a^{\dagger} a \rangle - \langle (a^{\dagger} a)^2 \rangle$ measures the distance from the line connecting the points $(0, 0)$ and $(2, 4)$ corresponding to Fock states $|0\rangle$ and $|2\rangle$. With $w = 1$ for the point $(1, 1)$ corresponding to the single-photon state $|1\rangle$. In practice, one can not directly measure $\langle (a^{\dagger} a)^2 \rangle$ and $\langle a^{\dagger} a \rangle$, but can approximate these values by increasing the loss artificially, as $E_{\bullet-} = \frac{a^{\dagger} a}{2} \eta + O(\eta^2)$ and $E_{\bullet\bullet} = \frac{a^{\dagger 2} a^2}{4} \eta^2 + O(\eta^3)$ for $\frac{1}{\eta} \gg a^{\dagger} a$ (recall that $E_{\bullet-} = E_{\bullet\bullet} + E_{\bullet\circ}$). Such an approach thus requires a precise control of the efficiency η and is statistically inefficient, since additional losses are introduced.

Alternatively [28], the auto-correlation function can be directly defined as

$$\tilde{g}^{(2)} = \frac{\langle E_{\bullet\bullet} \rangle}{\langle E_{\bullet-} \rangle \langle E_{\bullet\circ} \rangle}. \quad (46)$$

The efficiency can also be characterized by probability that a source produces a click $\tilde{I} = \langle E_{\bullet\bullet} + E_{\bullet\circ} + E_{\circ\bullet} \rangle$. Given the two values $\tilde{g}^{(2)}$ and $\tilde{I}$ it is then possible to reconstruct the full probability distribution $\mathbf{p} = (\langle E_{\circ\circ} \rangle, \langle E_{\bullet\circ} \rangle, \langle E_{\circ\bullet} \rangle, \langle E_{\bullet\bullet} \rangle)$, since we assumed $\langle E_{\bullet\circ} \rangle = \langle E_{\circ\bullet} \rangle$ so that $\mathbf{p}$ is defined by two parameters. As argued above the two functions coincide $g^{(2)} = \tilde{g}^{(2)}$ in the limit $\eta \rightarrow 0$.

In both cases the auto-correlation measurements relies on the setup of the Fig. 1. The measurement data can thus be readily used to estimate $\mathbf{p}$ and compute the benchmarks proposed in this paper.

To finish the discussion of the auto-correlation functions we recall that both $g^{(2)}$ and $\tilde{g}^{(2)}$ are witnesses of the non-classicality of the state ρ [28], i.e. $g^{(2)}, \tilde{g}^{(2)} < 1$ is only possible for states whose P-function admits negative values. In fact, this is true in a more general context, as given by the following observation.

Observation. For any two binary POVMs $\{M_{\bullet}, M_{\circ}\}$ and $\{M'_{\bullet}, M'_{\circ}\}$ measured at the two outputs of a beam splitter (Fig. 1), the inequality

$$G^{(2)} = \frac{\langle M_{\bullet} \otimes M'_{\bullet} \rangle}{\langle M_{\bullet} \otimes \mathbb{1} \rangle \langle \mathbb{1} \otimes M'_{\bullet} \rangle} < 1 \quad (47)$$

is a witness of non-classicality (P-function taking negative values), as long as the POVM element of individual detectors are only functions of the number of photons $\langle n | M_{\bullet} | m \rangle = \delta_{n,m} p_{\bullet}(|n\rangle)$ and the click probabilities are increasing functions of the photon number $p_{\bullet}(|n\rangle) \geq p_{\bullet}(|m\rangle)$ for $n \geq m$ (and the same for $M'_{\bullet}$).

For the sake of completeness we prove the above statement here. A coherent state splits into two coherent states on a beam splitter $|\alpha\rangle \mapsto_{BS} |\sqrt{r}\alpha\rangle |\sqrt{t}\alpha\rangle$. Hence for a coherent state

$$\begin{aligned} \langle M_{\bullet} \otimes M'_{\bullet} \rangle &= \langle \sqrt{r}\alpha | M_{\bullet} | \sqrt{r}\alpha \rangle \langle \sqrt{t}\alpha | M'_{\bullet} | \sqrt{t}\alpha \rangle \\ &= p_{\bullet}(\sqrt{r}|\alpha|^2) p'_{\bullet}(\sqrt{t}|\alpha|^2) \\ &= \langle M_{\bullet} \otimes \mathbb{1} \rangle \langle \mathbb{1} \otimes M'_{\bullet} \rangle \end{aligned} \quad (48)$$

and $G^{(2)} = 1$. Then for any mixture of coherent states $\rho_{cl} = \int d^2\alpha P(\alpha) |\alpha\rangle\langle\alpha|$ one finds

$$\langle M_{\bullet} \otimes M'_{\bullet} \rangle = \int d^2\alpha P(\alpha) p_{\bullet}(\sqrt{r}|\alpha|^2) p'_{\bullet}(\sqrt{t}|\alpha|^2). \quad (49)$$

To shorten the equations let us denote $z = |\alpha|$, $f(z) = p_{\bullet}(\sqrt{r}|\alpha|^2)$, $g(z) = p'_{\bullet}(\sqrt{t}|\alpha|^2)$, $\mu(z) = 2z \int d\varphi P(z e^{i\varphi})$ and $d\mu(z) = dz\mu(z)$ with $\int d\mu(z) = 1$, such that

$$\begin{aligned} \langle M_{\bullet} \otimes M'_{\bullet} \rangle &= \int d\mu(z) f(z) g(z) \\ &= \int d\mu(z) d\mu(z') \frac{1}{2} (f(z)g(z) + f(z')g(z')), \\ \langle M_{\bullet} \otimes \mathbb{1} \rangle \langle \mathbb{1} \otimes M'_{\bullet} \rangle &= \int d\mu(z) f(z) \int d\mu(z') g(z') \\ &= \int d\mu(z) d\mu(z') \frac{1}{2} (f(z)g(z') + f(z')g(z)). \end{aligned} \quad (50)$$

Without loss of generality consider $z' \geq z$, by $p_{\bullet}(|n\rangle) \geq p_{\bullet}(|m\rangle)$ for $n \geq m$ it follows that $f(z') \geq f(z)$ for $z' \geq z$ (using the fact that a coherent state has a Poissonian photon number distribution). Hence, one can write

$$\begin{aligned} f(z') &= f(z) + \Delta_f & \Delta_f &\geq 0 \\ g(z') &= g(z) + \Delta_g & \Delta_g &\geq 0 \end{aligned} \quad (51)$$

and

$$\begin{aligned} &(f(z)g(z) + f(z')g(z')) - (f(z)g(z') + f(z')g(z)) \\ &= \Delta_f \Delta_g \geq 0 \end{aligned} \quad (52)$$

for each z and z' . Therefore the inequality is also true for the integrals

$$\langle M_\bullet \otimes M'_\bullet \rangle \geq \langle M_\bullet \otimes \mathbb{1} \rangle \langle \mathbb{1} \otimes M'_\bullet \rangle. \quad (53)$$

Showing that $G^{(2)} \geq 1$ for any state with a positive P-function. Notably, the same proof works if both $\langle n | M_\bullet | n \rangle$ and $\langle n | M'_\bullet | n \rangle$ are decreasing functions of n , e.g. if they are replaced with $M_\circ$ and $M'_\circ$. The above observation has been used to propose an experiment where non-classicality of light would be demonstrated by directly using human eyes as detectors [29].

B POVM corresponding to photon detection preceded by a beam splitter

Since two single-photon detectors after a beam splitter cannot detect any coherence between Fock states of an incoming state, we can consider a state $\rho = \sum_n \frac{c_n}{n!} (a^\dagger)^n |0\rangle \langle 0| a^n$ without loss of generality. The state ρ arrives at a beam splitter with transmittance t , the resulting state is (ignoring again the coherence between the Fock states)

$$\rho_r = \sum_{n,k} c_n \binom{n}{k} (t^k) (1-t)^{n-k} |k, n-k\rangle \langle k, n-k|. \quad (54)$$

One can then compute the probability of the different events for the state ρ_r , for example

$$P_{\circ\bullet} = \text{Tr}(E_\circ \otimes E_\bullet \rho_r) = \text{Tr}(E_\circ \rho_r) - \text{Tr}(E_\circ \otimes E_\circ \rho_r) \quad (55)$$

where

$$\begin{aligned} \text{Tr}(E_\circ \otimes E_\circ \rho_r) &= \sum_n c_n^2 (1-\eta)^n \sum_{k=0}^n \binom{n}{k} t^k (1-t)^{n-k} \\ &= \sum_n c_n^2 (1-\eta)^n \\ &= \text{Tr}((1-\eta)^{a^\dagger a} \rho) \end{aligned}$$

and

$$\begin{aligned} \text{Tr}(E_\circ \rho_r) &= \sum_n c_n^2 \sum_{k=0}^n \binom{n}{k} t^k (1-t)^{n-k} (1-\eta)^k \\ &= \sum_n c_n^2 (1-t\eta)^n \\ &= \text{Tr}((1-t\eta)^{a^\dagger a} \rho). \end{aligned}$$

The resulting POVM is thus $E_{\circ\bullet} = (1-t\eta)^{a^\dagger a} - (1-\eta)^{a^\dagger a}$.

Let us consider a two-mode state (generalization to n modes is straightforward) $\rho = \sum_{n_1, n_2} \frac{c_{n_1, n_2}}{n_1! n_2!} (a_1^\dagger)^{n_1} (a_2^\dagger)^{n_2} |0\rangle \langle 0| a_1^{n_1} a_2^{n_2}$. The POVM element corresponding to no-clicks after the beam

splitter on each mode now reads $E_\circ^{(2)} = E_\circ \otimes E_\circ$. The state after the beam splitter is now

$$\begin{aligned} \rho_r &= \sum_{n_1, k_1, n_2, k_2} C_{n_1, k_1, n_2, k_2} \\ &|k_1, n_1 - k_1, k_2, n_2 - k_2\rangle \langle k_1, n_1 - k_1, k_2, n_2 - k_2| \end{aligned} \quad (56)$$

where $C_{n_1, k_1, n_2, k_2} = \binom{n_1}{k_1} \binom{n_2}{k_2} t^{k_1+k_2} (1-t)^{n_1+n_2-k_1-k_2}$. We focus on the event click on the first detector and no click on the second

$$\begin{aligned} \text{Tr}(E_\circ^2 \otimes E_\circ^2 \rho_r) &= \sum_{n_1, n_2} c_{n_1, n_2}^2 (1-\eta)^{n_1+n_2} \\ &\sum_{k_1, k_2=0}^{n_1, n_2} \binom{n_2}{k_2} \binom{n_1}{k_1} t^{k_1+k_2} (1-t)^{n_1+n_2-k_1-k_2} \\ &= \sum_n c_n^2 (1-\eta)^{n_1+n_2} \\ &= \text{Tr}((1-\eta)^{\hat{n}} \rho) \end{aligned}$$

similarly

$$\text{Tr}(E_\circ^2 \rho_r) = \text{Tr}((1-t\eta)^{\hat{n}} \rho). \quad (57)$$

We retrieve that the POVM elements in the multi-mode case are given by

$$\begin{aligned} E_{\circ\circ} &= (1-\eta)^{\hat{n}} \\ E_{\bullet\circ} &= (1-\eta t)^{\hat{n}} - (1-\eta)^{\hat{n}} \\ E_{\circ\bullet} &= (1-\eta r)^{\hat{n}} - (1-\eta)^{\hat{n}} \\ E_{\bullet\bullet} &= \mathbb{1} - E_{\circ\circ} - E_{\bullet\circ} - E_{\circ\bullet}. \end{aligned} \quad (58)$$

C The effect of loss on P_1

We show here that the set of states ρ with $P_1 \geq \frac{2}{3}$ is closed under losses. Consider a state ρ associated with a single-photon component $P_1 = P$. Let us apply infinitesimal transmission losses $\eta = 1 - d\epsilon$. After the loss, the photon number distribution $P_n = \langle n | \rho | n \rangle$ reads

$$P_n(\epsilon) = (1 - nd\epsilon)P_n + d\epsilon(n+1)P_{n+1}. \quad (59)$$

In particular,

$$\begin{aligned} \frac{d}{d\epsilon} P_1 &= -P_1 + 2P_2 \\ &\leq -P + 2(1-P) \\ &= 2-3P \end{aligned} \quad (60)$$

which is negative for $P \geq 2/3$.

On the other hand, there are states with $P_1 < 2/3$ for which the single-photon probability can be increased substantially by losses. Consider a channel with transmission efficiency η and apply it to the state

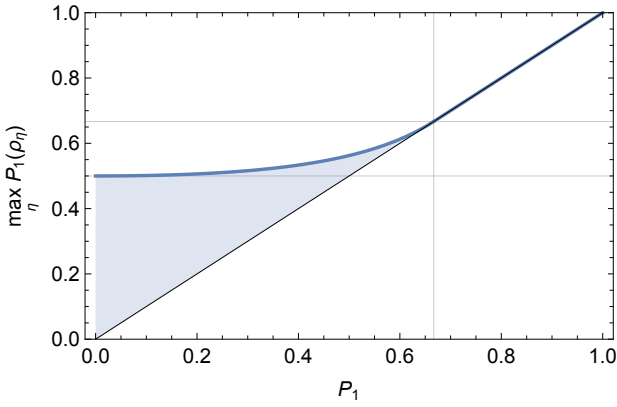


Figure 3: Consider the state $\rho = P_1 |1\rangle\langle 1| + (1 - P_1) |2\rangle\langle 2|$, which is transformed into ρ_η by a loss channel of transmission η . For any P_1 there is a value of η which maximizes the single-photon probability $P_1(\rho_\eta)$ after losses, see Eq. (62). The blue line depicts $\max_\eta P_1(\rho_\eta)$ as the function P_1 , and the black line is simply P_1 . One sees that for this example the single-photon component can be increased by losses if $P_1 < 2/3$, as illustrated by the shaded area. Conversely, we show that for any single-mode state ρ' with $P_1 \geq 2/3$, the single-photon component can not be increased by losses.

$\rho = P_1 |1\rangle\langle 1| + (1 - P_1) |2\rangle\langle 2|$. This leads to a state ρ_η having a single-photon component

$$P_1(\rho_\eta) = \eta P_1 + 2\eta(1 - \eta)(1 - P_1). \quad (61)$$

This quantity is maximized at

$$\max_\eta P_1(\rho_\eta) = \begin{cases} \frac{(2-P_1)^2}{8(1-P_1)} & P_1 \leq \frac{2}{3} \\ P_1 & P_1 > \frac{2}{3} \end{cases} \quad (62)$$

the maximum being depicted in Fig. 3 as a function of P_1 . To give a concrete example, for the initial $\rho = \frac{1}{2}(|1\rangle\langle 1| + |2\rangle\langle 2|)$, the weight of the single-photon component can be increased to $P_1(\rho_{\eta=3/4}) = 1/2 + 1/16 = 0.562$ while for the Fock state $\rho = |2\rangle\langle 2|$, it is possible to reach $P_1(\rho_{\eta=1/2}) = 1/2$.

D Parameter dependent witness

We consider the case where the two detectors have efficiencies η_T and η_R and the beam splitter has reflectance r and transmittance t (with $t + r = 1$). For an incoming Fock state $|n\rangle$ the probabilities of clicks are given by

$$\begin{aligned} f_n &= p_{\bullet\bullet}^{(n)} = (1 - \eta_R r)^n - (1 - \eta_T t - \eta_R r)^n \\ h_n &= p_{\bullet\circ}^{(n)} = (1 - \eta_T t)^n - (1 - \eta_R r - \eta_T t)^n \\ g_n &= p_{\circ\bullet}^{(n)} = 1 + (1 - \eta_R r - \eta_T t)^n - (1 - \eta_R r)^n - (1 - \eta_T t)^n. \end{aligned} \quad (63)$$

For a mixture of Fock states $\rho = \sum_n P_n |n\rangle\langle n|$ one has

$$p_{\bullet\bullet} = \sum_{n=1}^{\infty} P_n f_n, \quad (64)$$

from which we get

$$P_1 = \frac{1}{f_1} \left(p_{\circ\bullet} - \sum_{n \geq 2} P_n f_n \right). \quad (65)$$

In order to set a lower bound on P_1 we thus need to upper bound the term $\sum_{n \geq 2} P_n f_n$. We also have

$$p_{\bullet\bullet} = \sum_{n \geq 2} P_n g_n. \quad (66)$$

Therefore, to derive a benchmark for P_1 we are looking for a function $f_*(p_{\bullet\bullet})$ such that

$$\begin{aligned} f_*(p_{\bullet\bullet}) &= \max_{\mathbf{p} > \mathbf{0}} \sum_{n \geq 2} P_n f_n \\ \text{s.t. } \sum_{n \geq 2} P_n g_n &= p_{\bullet\bullet} \end{aligned} \quad (67)$$

To find $f_*(p_{\bullet\bullet})$ we define $q_n = P_n g_n$, so that the maximization can be rewritten as

$$\begin{aligned} f_*(p_{\bullet\bullet}) &= \max \sum_{n \geq 2} q_n \frac{f_n}{g_n} \\ \text{s.t. } \sum_{n \geq 2} q_n &= p_{\bullet\bullet}. \end{aligned} \quad (68)$$

Using $\sum_n q_n \frac{f_n}{g_n} \leq (\sum_n q_n) \left(\max_n \frac{f_n}{g_n} \right)$ we see that the solution of Eq. (67) satisfies

$$f_*(p_{\bullet\bullet}) \leq p_{\bullet\bullet} \left(\max_{n \geq 2} \frac{f_n}{g_n} \right). \quad (69)$$

We prove right after that the maximum is achieved for $n = 2$. By plugging $\sum_{n \geq 2} P_n f_n \leq f_*(p_{\bullet\bullet}) \leq p_{\bullet\bullet} \frac{f_2}{g_2}$ in Eq. (65), we get the desired inequality

$$P_1 \geq \frac{1}{f_1} \left(p_{\circ\bullet} - p_{\bullet\bullet} \frac{f_2}{g_2} \right). \quad (70)$$

The same bound holds with $p_{\circ\circ}$ instead of $p_{\bullet\bullet}$ and h_n instead of f_n . The right-hand side of these two inequalities are a function of the observed probabilities $\mathbf{p}$ and a lower bound on P_1 , hence defining two benchmarks

$$\begin{aligned} \hat{P}_1^{T*}(\mathbf{p}) &= \frac{1}{f_1} p_{\circ\bullet} - \frac{f_2}{f_1 g_2} p_{\bullet\bullet}, \\ \hat{P}_1^{R*}(\mathbf{p}) &= \frac{1}{h_1} p_{\circ\circ} - \frac{h_2}{h_1 g_2} p_{\bullet\bullet}. \end{aligned} \quad (71)$$

The best option is to consider the larger value of $p_{\circ\circ}$ and $p_{\bullet\bullet}$.

The proof of $\max_n \frac{f_n}{g_n} = \frac{f_2}{g_2}$. To maximize the ratio $\frac{f_n}{g_n}$ express it as

$$\begin{aligned} \frac{f_n}{g_n} &= \frac{(1 - \eta_R r)^n - (1 - \eta_T t - \eta_R r)^n}{1 - (1 - \eta_T t)^n - (1 - \eta_R r)^n + (1 - \eta_R r - \eta_T t)^n} \\ &= \frac{1}{\frac{1 - (1 - \eta_T t)^n}{(1 - \eta_R r)^n - (1 - \eta_R r - \eta_T t)^n} - 1}. \end{aligned} \quad (72)$$

Manifestly, maximizing $\frac{f_n}{g_n}$ is equivalent to minimizing $\frac{1-(1-\eta_T t)^n}{(1-\eta_{Rr})^n - (1-\eta_{Rr}-\eta_T t)^n}$. In other words we want to show that for $n \geq 2$ the fraction

$$\frac{1^n - (1-x)^n}{y^n - (y-x)^n}, \quad (73)$$

with $x = \eta_T t$ and $y = 1 - \eta_{Rr}$ satisfying $0 < x < y < 1$, is minimized at $n = 2$. However, it is enough to show that the expression $\frac{1^n - (1-x)^n}{y^n - (y-x)^n}$ is increasing with n . To do so, let us derive this quantity with respect to n . We have

$$\begin{aligned} \frac{d}{dn} \frac{1 - (1-x)^n}{y^n - (y-x)^n} &= \frac{1}{(y^n - (y-x)^n)^2} \\ &\times \left((1-x)^n - 1 \right) (y^n \log(y) - (y-x)^n \log(y-x)) \\ &- (1-x)^n \log(1-x) (y^n - (y-x)^n). \end{aligned} \quad (74)$$

To show that it is positive we can omit the denominator $(y^n - (y-x)^n)^2$. Labeling $a = (1-x)^n$, $b = y^n$, $c = (y-x)^n$ and noting that $\log(x^{\frac{1}{n}}) = \frac{1}{n} \log(x)$ we get

$$\frac{d}{dn} \frac{1 - (1-x)^n}{y^n - (y-x)^n} \geq 0 \iff f(a, b, c) \geq 0,$$

with

$$f(a, b, c) = (a-1)(b \log(b) - c \log(c)) - a \log(a)(b-c). \quad (75)$$

It remains to show that the function $f(a, b, c)$ is positive for $a, b > c$. Note that it is a decreasing function of c , as

$$\begin{aligned} \frac{d}{dc} f(a, b, c) &= (1-a)(\log(c) + 1) + a \log(a) \\ &\leq (1-a)(\log(a) + 1) + a \log(a) \\ &= 1 - a + \log(a) \\ &\leq 0 \end{aligned} \quad (76)$$

using a standard inequality for the logarithm $\log(a) \leq 1 - a$. We can thus only verify the positivity of the function for the maximal possible value of c . There are, however, two possibilities $a \geq b$ and $b > a$. For $a \geq b$ we set $c = b$ and obtain

$$f(a, b, c) \geq f(a, b, b) = 0. \quad (77)$$

For $b > a$ we set $c = a$ and get

$$\begin{aligned} f(a, b, c) &\geq f(a, b, a) \\ &= (1-b)a \log(a) - (1-a)b \log(b). \end{aligned} \quad (78)$$

To show that the last expression is positive, we divide it by $(1-a)(1-b)$ to get

$$\frac{a}{1-a} \log(a) - \frac{b}{1-b} \log(b), \quad (79)$$

and note that the function $\frac{x}{1-x} \log(x)$ is decreasing ($\frac{d}{dx} \frac{x}{1-x} \log(x) = \frac{1-x+\log(x)}{(1-x)^2} \leq 0$ by Eq. (76)). Therefore, $b \geq a$ implies

$$\frac{a}{1-a} \log(a) - \frac{b}{1-b} \log(b) \geq 0 \implies f(a, b, c) \geq 0. \quad (80)$$

Hence, the fraction $\frac{1-(1-x)^n}{y^n - (y-x)^n}$ is increasing with n and attains its minimum at the boundary $n = 2$ of the interval $[2, \infty)$. Therefore, $\frac{f_n}{g_n}$ is maximized at $n = 2$, which concludes the proof.

E Wigner-negativity measure

For a single-mode state ρ , the Wigner function $W_\rho(\beta)$ is a quasi-probability distribution satisfying $\int d\beta^2 W_\rho(\beta) = 1$. The negativity of the Wigner function ($W(\beta) < 0$ for some $\beta \in \mathbb{C}$) is an important non-classical feature of the state, as argued in the main text. A natural way to quantify this negativity is to measure the total quasi-probability where the function W_ρ takes negative values, that is to compute

$$N_W(\rho) = \int d\beta^2 \frac{|W_\rho(\beta)| - W_\rho(\beta)}{2}. \quad (81)$$

This intuitive quantity was introduced in [11]. We now show that $N_W(\rho)$ is a good "measure" of Wigner negativity in the sense that it can not be increased by Gaussian operations.

Pure Gaussian operations are displacements $D_\gamma = e^{\gamma a^\dagger - \gamma^* a}$, single-mode squeezing $SMS_g = e^{\frac{g}{2}(a^{\dagger 2} - a^2)}$, phase rotations $e^{i\varphi a^\dagger a}$, or combination thereof. Consider a single-mode state ρ with its Wigner function $W_\rho(\beta)$ and its Wigner negativity measure $N_W(\rho)$. The effect of a displacement $\varrho = D_\gamma \rho D_\gamma^\dagger$ on the Wigner function is a mere translation in phase space $W_\varrho(\beta) = W_\rho(\beta - \gamma)$, which does not affect the Wigner negativity measure $N_W(\rho) = N_W(\varrho)$. The same goes for a phase rotation, which merely transform $W_\varrho(\beta) = W_\rho(\beta e^{i\varphi})$. For a squeezing operation $\varrho = SMS_g \rho SMS_g^\dagger$, the Wigner function is transformed as

$$W_\varrho(\beta) = W_\rho(\tilde{\beta}) \quad (82)$$

where $\beta = \beta' + i\beta''$ and $\tilde{\beta} = e^g \beta' + e^{-g} i\beta''$. This implies for the Wigner negativity measure that

$$\begin{aligned} N_W(\varrho) &= \frac{1}{2} \int d\beta^2 (|W_\varrho(\beta)| - W_\varrho(\beta)) \\ &= \frac{1}{2} \int d\beta' d\beta'' (|W_\rho(\tilde{\beta})| - W_\rho(\tilde{\beta})) \\ &= \frac{1}{2} \int e^{-g} d\tilde{\beta}' e^g d\tilde{\beta}'' (|W_\rho(\tilde{\beta})| - W_\rho(\tilde{\beta})) \\ &= \frac{1}{2} \int d\tilde{\beta}^2 (|W_\rho(\tilde{\beta})| - W_\rho(\tilde{\beta})) \\ &= N_W(\rho). \end{aligned} \quad (83)$$

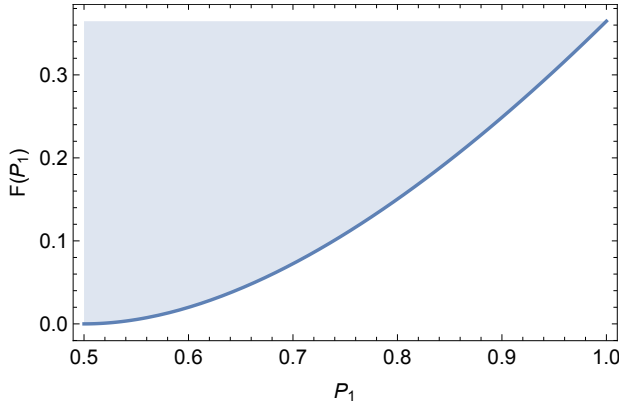


Figure 4: A sketch of the function $F(P_1)$ in Eq. (89) that lower bounds the Wigner-negativity of a state $N_W(\rho)$, for $P_1 = \langle 1|\rho|1\rangle \in [0.5, 1]$. Any single-mode state ρ with single-photon probability P_1 (or higher) has Wigner-negativity $N_W(\rho)$ above the blue line, as illustrated by the shaded area.

Hence, $N_W(\rho)$ is also unchanged by squeezing. Moreover, the quantity

$$N_W(p_1\rho_1 + p_2\rho_2) \leq p_1 N_W(\rho_1) + p_2 N_W(\rho_2) \quad (84)$$

is manifestly convex as $|p_1 W_1(\beta) + p_2 W_2(\beta)| \leq p_1 |W_1(\beta)| + p_2 |W_2(\beta)|$. Hence, $N_W(\rho)$ is non-increasing under mixtures of pure Gaussian operations. We conclude that $N_W(\rho)$ is a reasonable measure of Wigner negativity.

Let us now show how the Wigner negativity measure of a given state can be related to the weight of its single-photon component. The Wigner function for an arbitrary Fock state $|n\rangle$ reads [17]

$$W_n(\beta) = \frac{2(-1)^n}{\pi} e^{-2|\beta|^2} L_n(4|\beta|^2), \quad (85)$$

with $|W_n(\beta)| \leq \frac{2}{\pi}$ since $e^{-x/2} |L_n(x)| \leq 1$ (L_n are Laguerre polynomials). Hence, for any mixture of Fock states $\rho = \sum P_n |n\rangle\langle n|$, we have

$$\begin{aligned} W_\rho(\beta) &= P_1 W_1(\beta) + \sum_{n \neq 1} P_n W_n(\beta) \\ &\leq -\frac{2}{\pi} \left(P_1 (1 - 4|\beta|^2) e^{-2|\beta|^2} - (1 - P_1) \right). \end{aligned} \quad (86)$$

For $P_1 \geq 1/2$, the Wigner function is negative in the phase-space region with

$$\begin{aligned} (1 - 4|\beta|^2) e^{-2|\beta|^2} &> \frac{1 - P_1}{P_1} \iff \\ |\beta|^2 &< \underbrace{\frac{1}{4} \left(1 - 2w_0 \left(\frac{\sqrt{e}}{2} \frac{1 - P_1}{P_1} \right) \right)}_{\equiv \ell(P_1)}, \end{aligned} \quad (87)$$

where $w_0(x)$ is the principal branch of the Lambert W function. Eq. (87) defines a disk

$$\text{disk}(P_1) = \left\{ \beta \in \mathbb{C} \mid |\beta|^2 \leq \ell(P_1) \right\} \quad (88)$$

in phase-space centered at the origin, where the Wigner function is negative. We can now compute the integral over this region

$$\begin{aligned} N_W(\rho) &\geq \int_{\text{disk}(P_1)} d\beta^2 |W_\rho(\beta)| \\ &= \int_0^{\sqrt{\ell(P_1)}} 2\pi r dr |W_\rho(r)| \\ &\geq 4 \int_0^{\sqrt{\ell(P_1)}} r dr \left(P_1 (1 - 4r^4) e^{-2r^2} - (1 - P_1) \right) \\ &= F(P_1) = \frac{3(1 - P_1)(4w^2 + 3)}{8w} + P_1 - 2 \\ &\text{with } w = w_0 \left(\frac{\sqrt{e}}{2} \frac{1 - P_1}{P_1} \right) \quad \text{for} \end{aligned} \quad (89)$$

that is plotted in Fig. 4. Notably, for $P_1 = 1$ the bound becomes tight $N_W(|1\rangle) = F(1) = \frac{9}{4\sqrt{e}} - 1 \approx 0.36$. To show that the function $F(P_1)$ is convex, one computes

$$F''(P_1) = \frac{3w(4w(w+2)+5)}{8(1-P_1)P_1^2(w+1)^3}, \quad (90)$$

which is positive since $w \geq 0$. In the main text, we defined the function $F(P_1)$ continued on the whole interval $P_1 \in [0, 1]$ by simply setting $F(P_1) = 0$ for $P_1 \leq \frac{1}{2}$. Obviously, the continued function remains convex. Furthermore, at $P_1 = \frac{1}{2}$ the derivative of the function F is zero $F'(1/2) = 0$, and since $F''(P_1) \geq 0$ we can conclude that $F(P_1)$ is non-decreasing on the whole interval.

F Finite statistics

Consider n independent random variables $X^{(i)} \in [a, b]$ with its mean $\mathbb{E}(\bar{X}) = \mathbb{E}(\frac{1}{n} \sum X^{(i)})$. The Hoeffding theorem [15] gives a simple bound on the deviation of the observed average $\bar{X}$ after n trials from the expected value $\mathbb{E}(\bar{X})$

$$\mathbb{P}(\bar{X} - t \geq \mathbb{E}(\bar{X})) \leq \exp\left(-\frac{2nt^2}{(b-a)^2}\right). \quad (91)$$

In our case, the observables $X^{(i)}$ takes values in the interval $[-1, 3]$ so that $(b-a)^2 = 16$.

Let us now defined $X^{(i)}$ as the minimum of two variables $X^{(i)} = \min\{X_T^{(i)}, X_R^{(i)}\}$ such that $\bar{X} = \min\{\bar{X}_T, \bar{X}_R\}$. We have for the probability

$$\begin{aligned} \mathbb{P}(\bar{X} \geq x) &= \mathbb{P}(\bar{X}_T \geq x \text{ and } \bar{X}_R \geq x) \\ &\leq \mathbb{P}(\bar{X}_T \geq x), \mathbb{P}(\bar{X}_R \geq x). \end{aligned} \quad (92)$$

We now use $\bar{P}_1 + t \geq x = \min\{\mathbb{E}(\bar{X}_T), \mathbb{E}(\bar{X}_R)\} + t$ such that

$$P(\bar{X} \geq \bar{P}_1 + t) \leq P(\bar{X} \geq \min\{\mathbb{E}(\bar{X}_T), \mathbb{E}(\bar{X}_R)\} + t) \quad (93)$$

and consider two cases.

If $\mathbb{E}(\bar{X}_T) \leq \mathbb{E}(\bar{X}_R)$ we use

$$\begin{aligned} P(\bar{X} \geq \min\{\mathbb{E}(\bar{X}_T), \mathbb{E}(\bar{X}_R)\} + t) \\ &= P(\bar{X} \geq \mathbb{E}(\bar{X}_T) + t) \\ &\leq P(\bar{X}_T \geq \mathbb{E}(\bar{X}_T) + t) \\ &\leq \exp\left(-\frac{2nt^2}{16}\right). \end{aligned} \quad (94)$$

Otherwise, we do the same with $\bar{X}_R$. For both cases we find that

$$P(\min\{\bar{X}_T, \bar{X}_R\} - t \geq \bar{P}_1) \leq \exp\left(-\frac{2nt^2}{16}\right) \quad (95)$$

or equivalently

$$P(\min\{\bar{X}_T, \bar{X}_R\} - t < \bar{P}_1) \geq 1 - \exp\left(-\frac{2nt^2}{16}\right). \quad (96)$$

Writing the last expression in the form

$$P(\hat{q}_\alpha(\bar{X}_T, \bar{X}_R) < \bar{P}_1) \geq 1 - \alpha \quad (97)$$

we find

$$\hat{q}_\alpha(\bar{X}_T, \bar{X}_R) = \min\{\bar{X}_T, \bar{X}_R\} - \sqrt{\frac{16 \log(1/\alpha)}{2n}}, \quad (98)$$

the latter being a confidence interval for $\bar{P}_1$.

For the calibration-dependent benchmark, one naturally defines the random variable

$$Z_T = \begin{cases} C_1(\hat{t}, \hat{\eta}_T) & (\circ\bullet) \\ 0 & (\circ\circ) \text{ or } (\bullet\bullet) \\ -C_2(\hat{t}, \hat{\eta}_T, \hat{\eta}_R) & (\bullet\bullet) \end{cases} \quad (99)$$

for the quantity $\hat{P}_1^{T*}(\mathbf{p})$ of Eq. (19), with positive constants $C_1, C_2 \geq 0$ giving rise to the confidence interval

$$\hat{q}_\alpha^{T*} = \bar{Z}_T - (C_1(\hat{t}, \hat{\eta}_T) + C_2(\hat{t}, \hat{\eta}_T, \hat{\eta}_R)) \sqrt{\frac{\log(1/\alpha)}{2n}} \quad (100)$$

on the average single-photon weight $\bar{P}_1$. Defining Z_R similarly (with detector's roles exchanged) gives rise to the confidence interval

$$\hat{q}_\alpha^{R*} = \bar{Z}_R - (C_1(\hat{r}, \hat{\eta}_R) + C_2(\hat{r}, \hat{\eta}_R, \hat{\eta}_T)) \sqrt{\frac{\log(1/\alpha)}{2n}}, \quad (101)$$

by exchanging the roles of the detectors. Both are confidence intervals on $\bar{P}_1$, that is

$$P(\hat{q}_\alpha^{T(R)*} < \bar{P}_1) \geq 1 - \alpha. \quad (102)$$

It follows that

$$\begin{aligned} P(\max\{\hat{q}_\alpha^{T*}, \hat{q}_\alpha^{R*}\} < \bar{P}_1) \\ &= P(\hat{q}_\alpha^{T*} < \bar{P}_1 \text{ \& } \hat{q}_\alpha^{R*} < \bar{P}_1) \\ &= P(\hat{q}_\alpha^{T*} < \bar{P}_1) \\ &\quad - P(\hat{q}_\alpha^{T*} < \bar{P}_1 \text{ \& } \hat{q}_\alpha^{R*} \geq \bar{P}_1) \\ &\geq P(\hat{q}_\alpha^{T*} < \bar{P}_1) - P(\hat{q}_\alpha^{R*} \geq \bar{P}_1) \\ &\geq 1 - 2\alpha, \end{aligned} \quad (103)$$

hence

$$\hat{q}_\alpha^* = \max\{\hat{q}_\alpha^{T*/2}, \hat{q}_\alpha^{R*/2}\} \quad (104)$$

is also a confidence interval on $\bar{P}_1$ with confidence level $1 - \alpha$.

Finally, let us discuss the witness of Wigner negativity. First, we label by Q the measured value of $\min\{\bar{X}_T, \bar{X}_R\}$ after n measurement rounds and consider the case $Q > 1/2$. Given that $W_{\rho^{(i)}}(0) \leq \frac{1}{\pi}(1 - 2P_1^{(i)})$ for each state, for any collection of states that have a positive average Wigner function at the origin $\bar{W}(0) = \frac{1}{n} \sum_{i=1}^n W_{\rho^{(i)}}(0) \geq 0$, the average single-photon weight is $\bar{P}_1 \leq \frac{1}{2}$. For such a collection we thus have

$$\begin{aligned} P(\min\{\bar{X}_T, \bar{X}_R\} \geq Q) \\ &\leq P(\min\{\bar{X}_T, \bar{X}_R\} - Q \geq \bar{P}_1 - \frac{1}{2}) \\ &\leq P(\min\{\bar{X}_T, \bar{X}_R\} - \left(Q - \frac{1}{2}\right) \geq \bar{P}_1) \\ &\leq \exp\left(-\frac{2n(Q - \frac{1}{2})^2}{16}\right) \end{aligned} \quad (105)$$

by virtue of Eq. (95). In other words, for any collection of states with $\bar{W}(0) \geq 0$ the probability to get a benchmark value exceeding the observation $Q > 1/2$ is upper bounded by

$$\text{p-value} \leq \exp\left(-\frac{2n(\min\{\bar{X}_T, \bar{X}_R\} - \frac{1}{2})^2}{16}\right). \quad (106)$$

In the calibration-dependent setting with the same argument, the p-value can be obtained analogously. For any value Q and an ensemble of random variables $Z_T^{(1)}, \dots, Z_T^{(n)}$ such that $\bar{P}_1 \leq \frac{1}{2}$ one has

$$\begin{aligned} P(\bar{Z}_T \geq Q) \\ &\leq P(\bar{Z}_T - Q \geq \bar{P}_1 - \frac{1}{2}) \\ &\leq P(\bar{Z}_T - \left(Q - \frac{1}{2}\right) \geq \bar{P}_1) \\ &\leq \exp\left(-\frac{2n(Q - \frac{1}{2})^2}{(C_1 + C_2)^2}\right). \end{aligned} \quad (107)$$

Hence, the hypothesis $\bar{P}_1 \leq 1/2$ is associated to the p-value

$$\text{p-value} \leq \exp \left(-\frac{2n(\bar{Z}_T - \frac{1}{2})^2}{(C_1(\hat{t}, \hat{\eta}_T) + C_2(\hat{t}, \hat{\eta}_T, \hat{\eta}_R))^2} \right), \quad (108)$$

where $\bar{Z}_T$ is now the value of $\bar{Z}_T = \frac{1}{n} \sum_{i=1}^n Z_T^{(i)}$ observed in the experiment. Exchanging the roles of the detectors we also obtain

$$\text{p-value} \leq \exp \left(-\frac{2n(\bar{Z}_R - \frac{1}{2})^2}{(C_1(\hat{r}, \hat{\eta}_R) + C_2(\hat{r}, \hat{\eta}_R, \hat{\eta}_T))^2} \right). \quad (109)$$

Given the observed data $(n_{\circ\circ}, n_{\bullet\circ}, n_{\circ\bullet}, n_{\bullet\bullet})$ both bounds are valid statements about all collections of states with $\bar{P}_1 \leq \frac{1}{2}$. One can then simply choose the most favorable bound, that is

$$\begin{aligned} \text{p-value} &\leq \min \left\{ \exp \left(-\frac{2n(\bar{Z}_T - \frac{1}{2})^2}{(C_1(\hat{t}, \hat{\eta}_T) + C_2(\hat{t}, \hat{\eta}_T, \hat{\eta}_R))^2} \right), \right. \\ &\quad \left. \exp \left(-\frac{2n(\bar{Z}_R - \frac{1}{2})^2}{(C_1(\hat{r}, \hat{\eta}_R) + C_2(\hat{r}, \hat{\eta}_R, \hat{\eta}_T))^2} \right) \right\} \end{aligned} \quad (110)$$

where $\bar{Z}_T = \frac{1}{n}(C_1(\hat{t}, \hat{\eta}_T)n_{\circ\bullet} - C_2(\hat{t}, \hat{\eta}_T, \hat{\eta}_R)n_{\bullet\bullet})$ and $\bar{Z}_R = \frac{1}{n}(C_1(\hat{r}, \hat{\eta}_R)n_{\bullet\circ} - C_2(\hat{r}, \hat{\eta}_R, \hat{\eta}_T)n_{\bullet\bullet})$.

G Multi-mode product states

We now consider multi-mode product states of the form

$$\varrho = \bigotimes_k \rho_k, \quad (111)$$

where the state of each mode $\rho^{[k]}$ is associated to a probability vector $\mathbf{p}^{[k]}$, as defined by the expected values of the operators in Eq. (2). For multi-mode states, we do not have access to individual values of $\mathbf{p}^{[k]}$. Instead, a detector does not click if none of the modes triggers a click. Hence, for the state ϱ , the observed probabilities satisfy $p_{\circ\circ} = \prod_k p_{\circ\circ}^{[k]}$, $p_{\circ\cdot} = \prod_k p_{\circ\cdot}^{[k]}$, and $p_{\cdot\circ} = \prod_k p_{\cdot\circ}^{[k]}$. Denote $P_1^{[k]} = \langle 1 | \rho_k | 1 \rangle$ the single-photon probability for the mode k . Under the assumption that the beam splitter is balanced $p_{\cdot\circ}^{[k]} = p_{\circ\cdot}^{[k]} = \sum_n P_n^{[k]} \frac{1}{2^n}$ we will show in the next section that

$$\tilde{P}_1^T(\mathbf{p}) = \frac{1}{2}(12p_{\circ\cdot} - 9p_{\circ\circ} - 4) \leq \max_k P_1^{[k]}. \quad (112)$$

The case of unbalanced beam splitter is analogous with the single-mode case, we introduce $\tilde{P}_1^R(\mathbf{p})$ which is obtained from the definition of $\tilde{P}_1^T(\mathbf{p})$ (given in Eq. (112)) by replacing $p_{\circ\cdot}$ by $p_{\cdot\circ}$. Since either $p_{\circ\cdot}^{[k]} \geq p_{\cdot\circ}^{[k]}$ or $p_{\circ\cdot}^{[k]} \leq p_{\cdot\circ}^{[k]}$ holds for all modes k , the minimum

of $p_{\circ\cdot}^{[k]}$ and $p_{\cdot\circ}^{[k]}$ is a lower bound on $\sum_n P_n^{[k]}(\lambda) \frac{1}{2^n}$. We deduce that

$$\max_k P_1^{[k]} \geq \tilde{P}_1(\mathbf{p}) = \min\{\tilde{P}_1^T(\mathbf{p}), \tilde{P}_1^R(\mathbf{p})\}. \quad (113)$$

Under the assumption that the source produces a multi-mode product state ϱ of the form given in Eq. (111), we thus deduce that there is a mode k_* , that can in principle be filtered out, such that the corresponding state $\rho^{[k_*]}$ satisfies $\langle 1 | \rho^{[k_*]} | 1 \rangle \geq \tilde{P}_1(\mathbf{p})$.

G.1 The proof of Eq. (112)

Consider a multi-mode product state $\varrho = \bigotimes_k \rho_k$ with $P_1^{[k]} \leq P$ in each mode. We label $(p_{\circ\cdot}^{[k]}, p_{\circ\circ}^{[k]})$ the statistics associated to ρ_k and $(p_{\circ\cdot}, p_{\circ\circ})$ the statistics associated to ϱ . For a balanced beam splitter, we have $(p_{\circ\cdot}^{[k]}, p_{\circ\circ}^{[k]}) \in \mathcal{Q}_P$, with

$$\mathcal{Q}_P = \text{Polytope} \left\{ (0, 0), \left(\frac{1+P}{4}, 0 \right), \left(\frac{2-P}{2}, 1-P \right), (1, 1) \right\}.$$

The probabilities $p_{\circ\cdot}$ and $p_{\circ\circ}$ satisfy

$$\begin{aligned} p_{\circ\cdot} &= \prod_{k=1}^n p_{\circ\cdot}^{[k]}, \\ p_{\circ\circ} &= \prod_{k=1}^n p_{\circ\circ}^{[k]}. \end{aligned} \quad (114)$$

Our first aim is to analyze the possible set of values $\mathcal{Q}_P^\infty = \{(p_{\circ\cdot}, p_{\circ\circ})\}$ for all n from 1 to ∞ and in particular, to show that $\mathcal{Q}_P^\infty \subset \mathcal{Q}_P$ for $P \geq \frac{1}{2}$. Naturally, we are interested in the extreme points of this set. Eq. (114) is linear in all points $(p_{\circ\cdot}^{[k]}, p_{\circ\circ}^{[k]})$, hence the extreme points of $\mathcal{Q}_P^\infty$ are obtained by combining the vertices of $\mathcal{Q}_P$.

Whenever a single vertex $(p_{\circ\cdot}^{[k]}, p_{\circ\circ}^{[k]}) = (0, 0)$ appears in the product of Eq. (114), it results in $(p_{\circ\cdot}, p_{\circ\circ}) = (0, 0)$. Similarly, if the vertex $(p_{\circ\cdot}^{[k]}, p_{\circ\circ}^{[k]}) = (\frac{1+P}{4}, 0)$ is chosen for at least one mode k , $p_{\circ\circ} = 0$ and $p_{\circ\cdot} = \frac{1+P}{4} \prod_{j \neq k} p_{\circ\cdot}^{[j]} \leq \frac{1+P}{4}$. This means that the point $(p_{\circ\cdot}, p_{\circ\circ} = 0)$ remains inside the original polytope $\mathcal{Q}_P$. We can thus remember that $(0, 0)$ and $(\frac{1+P}{4}, 0)$ are points of $\mathcal{Q}_P^\infty$, but ignore these vertices in the further construction. Analogously, all modes with $(p_{\circ\cdot}^{[k]}, p_{\circ\circ}^{[k]}) = (1, 1)$ do not change the value of the product, and we can also ignore this vertex. Hence, the only products in Eq. (114) that are potentially not in $\mathcal{Q}_P$ are of the form

$$(p_{\circ\cdot}, p_{\circ\circ})_n = \left(\left(\frac{2-P}{2} \right)^n, (1-P)^n \right) \quad (115)$$

for $n \geq 2$. Let us first consider the point $(p_{\circ\cdot}, p_{\circ\circ})_2$.

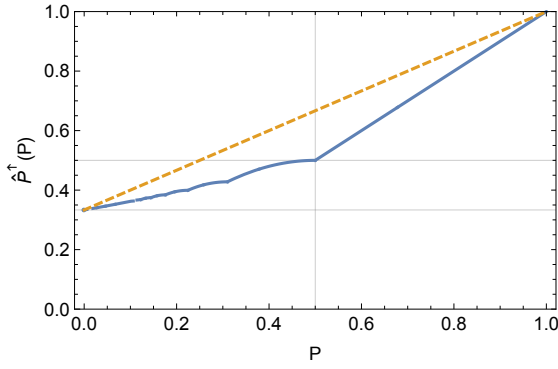


Figure 5: (Full line) Representation of $\hat{P}_1^{T\uparrow}(P)$ as a function of P (full blue line) and an upper-bound (dashed orange line) given by a simple linear function of the form $\frac{1}{3} + \frac{2}{3}P \geq \hat{P}_1^{T\uparrow}(P)$.

It remains inside $\mathcal{Q}_P$ if and only if

$$\begin{aligned} 4p_{o-} - 3p_{o0} - 1 &\leq P \\ 4\left(\frac{2-P}{2}\right)^2 - 3(1-P)^2 - 1 &\leq P \\ P - 2P^2 &\leq 0 \\ P &\geq \frac{1}{2}. \end{aligned} \quad (116)$$

Naturally, if $(p_{o-}, p_{o0})_2 \in \mathcal{Q}_P$ the next points $(p_{o-}, p_{o0})_n$ are also in $\mathcal{Q}_P$. Therefore, $\mathcal{Q}_P^\infty \subset \mathcal{Q}_P$ for $P \geq \frac{1}{2}$. This means that for any state of the form $\varrho = \bigotimes_k \rho_k$, with $P_1^{[k]} \leq 1/2 \leq P$, $4p_{o-} - 3p_{o0} - 1 \leq P$.

In order to extend the analysis to any value of P , we would need to analyze $\mathcal{Q}_P^\infty$ for a arbitrary P , which is cumbersome. Instead, we analyze the maximal value that $\hat{P}_1^T(\mathbf{p})$ takes on $\mathcal{Q}_P^\infty$. We know that it takes its maximum value on one of the vertices $(p_{o-}, p_{o0})_n$, and denote these values

$$\hat{P}_1^T(n) = 4\left(\frac{2-P}{2}\right)^n - 3(1-P)^n - 1, \quad (117)$$

for $n \geq 1$. Its maximal value in $\mathcal{Q}_P^\infty$ is thus given by

$$\begin{aligned} \hat{P}_1^{T\uparrow}(P) &= \sup_{\mathbf{p} \in \mathcal{Q}_P^\infty} \hat{P}_1^T(\mathbf{p}) \\ &= \sup_{n \geq 1} \hat{P}_1^T(n). \end{aligned} \quad (118)$$

Let us now look at $\hat{P}_1^T(n)$ as a function of a continuous parameter $n \in [0, \infty)$, and compute its derivative

$$\frac{d}{dn} \hat{P}_1^T(n) = 4X^n \log(X) - 3Y^n \log(Y) \quad (119)$$

with $X = \frac{2-P}{2}$ and $Y = 1-P$. $\hat{P}_1^T(n)$ admits a unique local extremum $\frac{d}{dn} \hat{P}_1^T(n) = 0$ at

$$n_* = \frac{\log\left(\frac{4\log(X)}{3\log(Y)}\right)}{\log(X) - \log(Y)}. \quad (120)$$

Furthermore, one easily sees that $\hat{P}_1^T(0) = 0$, $\hat{P}_1^T(\infty) = -1$ and $\hat{P}_1^T(1) = P$ and hence $\hat{P}_1^T(n_*)$ is the global maximum of the function.

Next, we recall that n can only take integer values. Thus, the maximal value reads

$$\hat{P}_1^{T\uparrow}(P) = \max\{\hat{P}_1^T(\lfloor n_* \rfloor), \hat{P}_1^T(\lfloor n_* \rfloor + 1)\}. \quad (121)$$

It is quite an irregular function, as can be seen in Fig. 5. The boundary value $\hat{P}_1^{T\uparrow}(0) = \lim_{P \rightarrow 0} \hat{P}_1^{T\uparrow}(P) = \frac{1}{3}$ can be computed analytically. We show numerically that it is upper-bounded by a simple linear function

$$\hat{P}_1^T(\mathbf{p}) \leq \hat{P}_1^{T\uparrow}(P) \leq \frac{1}{3} + \frac{2}{3}P, \quad (122)$$

also can be seen in Fig. 5. By inverting the last inequality we find that

$$\begin{aligned} \tilde{P}_1^T(\mathbf{p}) &\leq P \quad \text{for} \\ \tilde{P}_1^T(\mathbf{p}) &= \frac{3\hat{P}_1^T(\mathbf{p}) - 1}{2} = \frac{1}{2}(12p_{o-} - 9p_{o0} - 4). \end{aligned} \quad (123)$$

Therefore, for any value P such that $\tilde{P}_1^T(\mathbf{p}) > P$, we can conclude that at least one mode satisfies $\max_k \langle 1|\rho^{[k]}|1 \rangle > P$.

G.2 Finite statistics

Here, in order to obtain a confidence interval we define an independent random variable

$$Y_T = \begin{cases} 4 & (\circ\bullet) \\ -1/2 & (\circ\circ) \\ -2 & (\bullet\bullet) \text{ or } (\bullet\circ) \end{cases}, \quad (124)$$

with $\mathbb{E}(Y_T) = \tilde{P}_1^T(\mathbf{p})$. Similarly, we define Y_R by exchanging the role of the two detectors and get $\mathbb{E}(Y_R) = \tilde{P}_1^R(\mathbf{p})$ with $Y_{T(R)} \in [-2, 4]$. The same exact analysis as above yields a one-sided confidence interval

$$\tilde{q}_\alpha(\bar{Y}_T, \bar{Y}_R) = \min\{\bar{Y}_T, \bar{Y}_R\} - \sqrt{\frac{36 \log(1/\alpha)}{2n}} \quad (125)$$

on the quantity $\max_k P_1^{[k]}$ averaged over all states produced by the source. In the general multi-mode case, the same quantity is a one-sided interval on the quantity $\sum_\lambda p(\lambda) \max_k P_1^{[k]}(\lambda)$ averaged over all states produced by the source.

References

- [1] M. D. Eisaman, J. Fan, A. Migdall, and S. V. Polyakov. Invited review article: Single-photon sources and detectors. *Review of Scientific Instruments*, 82(7):071101, 2011. DOI: [10.1063/1.3610677](https://doi.org/10.1063/1.3610677).

- [2] Sarah Thomas and Pascale Senellart. The race for the ideal single-photon source is on. *Nature Nanotechnology*, 16(4):367–368, 2021. DOI: [10.1038/s41565-021-00851-1](https://doi.org/10.1038/s41565-021-00851-1).
- [3] Nicolas Sangouard and Hugo Zbinden. What are single photons good for? *Journal of Modern Optics*, 59(17):1458–1464, 2012. DOI: [10.1080/09500340.2012.687500](https://doi.org/10.1080/09500340.2012.687500).
- [4] Pieter Kok, W. J. Munro, Kae Nemoto, T. C. Ralph, Jonathan P. Dowling, and G. J. Milburn. Linear optical quantum computing with photonic qubits. *Rev. Mod. Phys.*, 79:135–174, Jan 2007. DOI: [10.1103/RevModPhys.79.135](https://doi.org/10.1103/RevModPhys.79.135).
- [5] Christopher J. Chunnillall, Ivo Pietro Degiovanni, Stefan Kück, Ingmar Müller, and Alastair G. Sinclair. Metrology of single-photon sources and detectors: a review. *Optical Engineering*, 53(8):081910–081910, July 2014. DOI: [10.1117/1.oe.53.8.081910](https://doi.org/10.1117/1.oe.53.8.081910).
- [6] Stefan Kück. Single photon sources for absolute radiometry – a review about the current state of the art. *Measurement: Sensors*, 18:100219, 2021. ISSN 2665-9174. DOI: [10.1016/j.measen.2021.100219](https://doi.org/10.1016/j.measen.2021.100219).
- [7] D. F. Walls and Gerard J. Milburn. *Quantum Optics*. Springer, Berlin, Heidelberg, 2 edition, 2008. ISBN 978-3-540-28574-8. DOI: [10.1007/978-3-540-28574-8](https://doi.org/10.1007/978-3-540-28574-8).
- [8] Miroslav Ježek, Ivo Straka, Michal Mičuda, Miloslav Dušek, Jaromír Fiurášek, and Radim Filip. Experimental test of the quantum non-gaussian character of a heralded single-photon state. *Phys. Rev. Lett.*, 107:213602, Nov 2011. DOI: [10.1103/PhysRevLett.107.213602](https://doi.org/10.1103/PhysRevLett.107.213602).
- [9] Ana Predojević, Miroslav Ježek, Tobias Huber, Harishankar Jayakumar, Thomas Kauten, Glenn S. Solomon, Radim Filip, and Gregor Weihs. Efficiency vs. multi-photon contribution test for quantum dots. *Opt. Express*, 2014. DOI: [10.1364/OE.22.004789](https://doi.org/10.1364/OE.22.004789).
- [10] Ivo Straka, Ana Predojević, Tobias Huber, Lukáš Lachman, Lorenz Butschek, Martina Miková, Michal Mičuda, Glenn S. Solomon, Gregor Weihs, Miroslav Ježek, and Radim Filip. Quantum non-gaussian depth of single-photon states. *Phys. Rev. Lett.*, 113:223603, Nov 2014. DOI: [10.1103/PhysRevLett.113.223603](https://doi.org/10.1103/PhysRevLett.113.223603).
- [11] Anatole Kenfack and Karol Życzkowski. Negativity of the wigner function as an indicator of non-classicality. *Journal of Optics B: Quantum and Semiclassical Optics*, 6(10):396, 2004. DOI: [10.1088/1464-4266/6/10/003](https://doi.org/10.1088/1464-4266/6/10/003).
- [12] Alan Migdall, Sergey V Polyakov, Jingyun Fan, and Joshua C Bienfang. *Single-photon generation and detection: physics and applications*. Academic Press, 2013. ISBN 978-0-12-387695-9.
- [13] Katiúscia N Cassemiro, Kaisa Laiho, and Christine Silberhorn. Accessing the purity of a single photon by the width of the hong-ou-mandel interference. *New Journal of Physics*, 12(11):113052, 2010. DOI: [10.1088/1367-2630/12/11/113052](https://doi.org/10.1088/1367-2630/12/11/113052).
- [14] Kevin Zielnicki, Karina Garay-Palmett, Daniel Cruz-Delgado, Hector Cruz-Ramirez, Michael F. O’Boyle, Bin Fang, Virginia O. Lorenz, Alfred B. U’Ren, and Paul G. Kwiat. Joint spectral characterization of photon-pair sources. *Journal of Modern Optics*, 65(10):1141–1160, Jun 2018. DOI: [10.1080/09500340.2018.1437228](https://doi.org/10.1080/09500340.2018.1437228).
- [15] Wassily Hoeffding. Probability Inequalities for Sums of Bounded Random Variables. *Journal of the American Statistical Association*, 58(301):13–30, mar 1963. DOI: [10.1080/01621459.1963.10500830](https://doi.org/10.1080/01621459.1963.10500830).
- [16] E. Wigner. On the quantum correction for thermodynamic equilibrium. *Phys. Rev.*, 40:749–759, Jun 1932. DOI: [10.1103/PhysRev.40.749](https://doi.org/10.1103/PhysRev.40.749).
- [17] Werner Vogel and Dirk-Gunnar Welsch. *Quantum optics*. John Wiley & Sons, 2006. DOI: [10.1002/3527608524](https://doi.org/10.1002/3527608524).
- [18] R. L. Hudson. When is the wigner quasi-probability density non-negative? *Reports on Mathematical Physics*, 6(2):249–252, 1974. ISSN 0034-4877. DOI: [10.1016/0034-4877\(74\)90007-X](https://doi.org/10.1016/0034-4877(74)90007-X).
- [19] U. Chabaud, PE. Emeriau, and F. Grosshans. Witnessing wigner negativity. *Quantum*, 471:230503, Jun 2021. DOI: [10.22331/q-2021-06-08-471](https://doi.org/10.22331/q-2021-06-08-471).
- [20] Antoine Royer. Wigner function as the expectation value of a parity operator. *Phys. Rev. A*, 15:449–450, Feb 1977. DOI: [10.1103/PhysRevA.15.449](https://doi.org/10.1103/PhysRevA.15.449).
- [21] F. W. J. Olver, A. B. Olde Daalhuis, D. W. Lozier, B. I. Schneider, R. F. Boisvert, C. W. Clark, B. R. Miller, B. V. Saunders, H. S. Cohl, M. A. McClain, and eds. Nist digital library of mathematical functions. <http://dlmf.nist.gov/>, Release 1.1.3 of 2021-09-15.
- [22] M. Munroe, D. Boggavarapu, M. E. Anderson, and M. G. Raymer. Photon-number statistics from the phase-averaged quadrature-field distribution: Theory and ultrafast measurement. *Phys. Rev. A*, 52:R924–R927, Aug 1995. DOI: [10.1103/PhysRevA.52.R924](https://doi.org/10.1103/PhysRevA.52.R924).
- [23] N. Bruno, A. Martin, T. Guerreiro, B. Sanguinetti, and R. T. Thew. Pulsed source of spectrally uncorrelated and indistinguishable photons at telecom wavelengths. *Optics Express*, 22(14):17246, jul 2014. DOI: [10.1364/OE.22.017246](https://doi.org/10.1364/OE.22.017246).
- [24] T. Guerreiro, A. Martin, B. Sanguinetti, N. Bruno, H. Zbinden, and R. T. Thew. High efficiency coupling of photon pairs in practice. *Optics Express*, 21(23):27641, nov 2013. DOI: [10.1364/OE.21.027641](https://doi.org/10.1364/OE.21.027641).
- [25] Misael Caloz, Matthieu Perrenoud, Claire Autebert, Boris Korzh, Markus Weiss, Christian

- Schönenberger, Richard J. Warburton, Hugo Zbinden, and Félix Bussi eres. High-detection efficiency and low-timing jitter with amorphous superconducting nanowire single-photon detectors. *Applied Physics Letters*, 112(6), 2018. DOI: [10.1063/1.5010102](https://doi.org/10.1063/1.5010102).
- [26] P. Sekatski, E. Oudot, P. Caspar, R.T. Thew, and N. Sangouard. Benchmarking single-photon sources from an auto-correlation measurement [data set]. 2022. DOI: [10.5281/zenodo.7245446](https://doi.org/10.5281/zenodo.7245446).
- [27] Marlan O Scully and M Suhail Zubairy. *Quantum optics*. American Association of Physics Teachers, 1999. DOI: [10.1017/CBO9780511813993](https://doi.org/10.1017/CBO9780511813993).
- [28] Pavel Sekatski, N Sangouard, F elix Bussieres, Christoph Clausen, Nicolas Gisin, and Hugo Zbinden. Detector imperfections in photon-pair source characterization. *Journal of Physics B: Atomic, Molecular and Optical Physics*, 45(12):124016, 2012. DOI: [10.1088/0953-4075/45/12/124016](https://doi.org/10.1088/0953-4075/45/12/124016).
- [29] A. Dodel, A. Mayinda, E. Oudot, A. Martin, P. Sekatski, J.-D. Bancal, and N. Sangouard. Proposal for witnessing non-classical light with the human eye. *Quantum*, 1:7, April 2017. ISSN 2521-327X. DOI: [10.22331/q-2017-04-25-7](https://doi.org/10.22331/q-2017-04-25-7).