PAPER • OPEN ACCESS

Harnessing data augmentation to quantify uncertainty in the early estimation of single-photon source quality

To cite this article: David Jacob Kedziora *et al* 2023 *Mach. Learn.: Sci. Technol.* **4** 045042

View the article online for updates and enhancements.

MACHINE
LEARNING
Science and Technology

PAPER

Harnessing data augmentation to quantify uncertainty in the early estimation of single-photon source quality

David Jacob Kedziora[1,*] , Anna Musiał[2] , Wojciech Rudno-Rudziński[2] and Bogdan Gabrys[1]

[1] Complex Adaptive Systems Lab, University of Technology Sydney, Sydney, New South Wales 2007, Australia
[2] Department of Experimental Physics, Wrocław University of Science and Technology, Wybrzeże Wyspiańskiego 27, Wrocław 50-370, Poland
[*] Author to whom any correspondence should be addressed.

E-mail: david.kedziora@uts.edu.au

Abstract

Novel methods for rapidly estimating single-photon source (SPS) quality have been promoted in recent literature to address the expensive and time-consuming nature of experimental validation via intensity interferometry. However, the frequent lack of uncertainty discussions and reproducible details raises concerns about their reliability. This study investigates the use of data augmentation, a machine learning technique, to supplement experimental data with bootstrapped samples and quantify the uncertainty of such estimates. Eight datasets obtained from measurements involving a single InGaAs/GaAs epitaxial quantum dot serve as a proof-of-principle example. Analysis of one of the SPS quality metrics derived from efficient histogram fitting of the synthetic samples, i.e. the probability of multi-photon emission events, reveals significant uncertainty contributed by stochastic variability in the Poisson processes that describe detection rates. Ignoring this source of error risks severe overconfidence in both early quality estimates and claims for state-of-the-art SPS devices. Additionally, this study finds that standard least-squares fitting is comparable to using a Poisson likelihood, and expanding averages show some promise for early estimation. Also, reducing background counts improves fitting accuracy but does not address the Poisson-process variability. Ultimately, data augmentation demonstrates its value in supplementing physical experiments; its benefit here is to emphasise the need for a cautious assessment of SPS quality.

1. Introduction

The world is presently witnessing the advent of the 'Quantum 2.0' technological revolution, an era distinguished by the development of devices that, in their operation, involve manipulating quantum states of light and matter [1]. Indeed, these devices seek to exploit purely quantum phenomena, such as superposition or entanglement. Unsurprisingly, many research fields are linked to this revolution, e.g. quantum communication networks and cryptography [2–4], quantum computing and simulation [5–7], quantum imaging [8–10], quantum sensing within magnetic [11, 12] and gravitational [13, 14] fields, and so on. Likewise, considerably many engineering designs, both proposed and proven, take advantage of quantum optical phenomena, e.g. by using photonic degrees of freedom as an information carrier, subject to transmission/manipulation, or to probe the state of a system. However, to really exploit the quantum nature of light, non-classical light sources, e.g. a single-photon source (SPS), are required.

Certainly, an SPS constitutes a valuable resource for both fundamental studies in quantum optics and industrial applications of quantum technology. However, engineering an SPS is complicated as it must fulfil the exacting requirements of practicality, e.g. a feasibly attainable operating temperature, as well as a desired emission wavelength, emission rate and collection efficiency. More advanced application schemes may even require polarisation control or indistinguishability between consecutive photons. Nonetheless, the most

crucial characteristic of an intended SPS is the purity of single-photon emission. While other secondary attributes are important, they become irrelevant if the essential criterion for an SPS is not fulfilled. This purity is typically measured by a photon-intensity autocorrelation experiment, which determines the second-order auto-correlation function $g^{(2)}(\tau)$ [15–18].

Notably, many material systems have been proposed for the engineering of an SPS, starting with systems based on atoms and ions. These are clean and reproducible, but require bulky and cumbersome setups, limiting their practical applications [19, 20]. In contrast, solid-state systems appear to have better long-term prospects, and several classes of related structures show promise as single-photon emitters: quantum dots grown by colloidal [21, 22] or epitaxial techniques [4, 23, 24], as well as various defect centres, e.g. in diamonds [25, 26], silicon carbide [27], hexagonal boron nitride [28], and 2D materials [29, 30].

Whatever the practical realisation of an SPS is, the quality of the final device must be precisely assessed. This paper will focus solely on the purity of single-photon emission. Ideally, an SPS works as follows: a quantum emitter with a discrete energy spectrum is repeatedly excited optically or electrically, each time with relaxation following excitation, and thus emits a train of photons. Every excitation pulse it encounters should result in the emission of just a single photon, i.e. it must act as an on-demand source of single photons. Proving that an object is classifiable as an SPS typically requires using an intensity-based 'Hanbury Brown and Twiss' interferometer [31] to experimentally determine $g^{(2)}(\tau)$. Specifically, the emission is split equally between two detection paths. A detector with single-photon sensitivity is placed at the end of each path. Dedicated electronics then build up a histogram of two-photon events as a function of the temporal separation between the output signals from the two detectors. These events occur when both detectors are separately triggered. Any events at an effective time delay of zero imply the SPS underwent multi-photon emission (MPE), i.e. $g^{(2)}(0) > 0$. Indeed, the quality of an SPS depends on how often an MPE event triggers a two-photon detection. Excessive MPE is undesirable as it limits the utility of an intended SPS, e.g. causing errors in quantum communication and cryptography protocols [32].

The problem with determining SPS quality is that the measurement is statistical in nature and, therefore, a sufficiently large amount of data needs to be collected to derive reliable conclusions. This accumulation can take substantial time, even on the order of days, if the probability of observing two-photon events is low. Likewise, measurements can be resource intensive, given that single-photon emission is typically observed at expensive-to-maintain cryogenic temperatures. Specifically, the most efficient single-photon detectors need to be kept below the critical temperature of a superconducting transition, while the highest-quality SPS devices, i.e. III–V semiconductor epitaxial quantum dots, require cryogenic temperatures for optimal performance. Then there is the aim of the measurement. For proof-of-principle experiments, obtaining a value of $g^{(2)}(0) < 0.5$ will be sufficient, while, for state-of-the-art engineering, ensuring a minimal $g^{(2)}(0)$ is critical [33–35]. Obtaining statistical significance in the latter case can take a substantially longer measurement. The situation worsens when working with more complex characteristics, e.g. photon indistinguishability or the degree of quantum entanglement, for which quantum state tomography requires many cross-correlation measurements to be performed in different bases [36, 37]. In short, physicists in the field keenly await any improvements in data acquisition or analysis techniques that might give accurate SPS quality estimates for far fewer two-photon event detections, i.e. significantly reduced experiment times.

Given this context, recent literature, such as an article proposing a maximum *a posteriori* (MAP) scheme [38], has drawn much attention, suggesting that small data samples are sufficient to provide reasonable estimates of SPS quality in terms of MPE probability, i.e. $g^{(2)}(0)$. This work investigates how true such claims are. Exemplary experimental datasets are readily available for such an analysis, i.e. measurements for epitaxial InGaAs/GaAs quantum dots within deterministic photonic nanostructures emitting at 1.3 μm [39], where a 10 s collection interval for two-photon events provides insight into how the correlation histogram can build up over time. It is then not difficult to examine various statistical fitting methods, assessing how estimates of SPS quality and their fitting-based accuracy develop with respect to the growing number of detected two-photon events. However, herein lies a major weakness. Who is to say that one sequence of observations, already very costly, is a typical reflection of how the histogram and associated quality estimates evolve? How can one even be sure that the final long-time estimate, used as a baseline of comparison for a short-time estimate, is itself accurate?

In data science and machine learning, when information is sparse, generative procedures are sometimes used to simulate new data from existing datasets in a process often called 'data augmentation' [40, 41]. This work continues the trend of applying data-science principles to experimental physics [42, 43] by leaning on a technique that is effectively 'bootstrapping'. Each experimental dataset is used to generate arbitrarily many alternative samples that then determine the sampling-based variance of an SPS-quality estimand. After highlighting the non-negligible error bars on even late-stage estimates of $g^{(2)}(0)$, this article reinforces a

cautionary message when evaluating new approaches for fast estimation: *do not underestimate stochastic variability*.

This work is organised and presented in the following manner. Section 2 provides, with a novel reformulation suited for efficient computation, the theoretical equations that describe the expected structure of a two-photon event histogram. Section 3 describes the eight datasets analysed in this work (3.1), the two investigated methods of fitting data for SPS quality estimation (3.2), and the procedure for generating synthetic data in support of calculating error bounds (3.3). Section 4 then presents the results of numerically fitting theoretical equations to the data, both observed and synthetic, assessing (1) the expected sample-based errors of $g^{(2)}(0)$, (2) whether a MAP scheme [38] is superior to least-squares histogram fitting, (3) the possibility of using expanding averages for early $g^{(2)}(0)$ estimation, and (4) the impact of background counts on accuracy. Section 5 subsequently takes the lessons learned here to critically reexamine and discuss the article that motivated this work [38]. Finally, section 6 summarises the conclusions of this research.

2. Theory

SPS quality estimation typically leverages the technique of intensity interferometry with single-photon detectors [31], albeit ones that cannot resolve the number of photons. Because of this limitation, the photon flux must be kept sufficiently low for the measured histogram to reflect actual photon statistics. Thus, in many cases, only a handful of two-photon events is detected, per second, within a temporal-separation window of interest. Consequently, in ideal operating conditions, these independent events can be modelled as elements of Poisson processes. Specifically, given any particular temporal separation τ between the two photons incident on the two detectors, the probability of encountering n such events after t seconds is

$$Pr(n; R, t) = \frac{(Rt)^n}{n!} e^{-Rt}, \tag{1}$$

where R is a mean detection rate in Hz. This rate technically depends on how τ values are binned, but matters of discretisation are ignored in this section; they are trivial to apply during numerical implementation.

In theory, because the mean detection rate R is fixed for any given τ, an accumulated histogram of detected events should closely approximate $R(\tau) \times t$ in the long-time limit of t. Moreover, the expected form of $R(\tau)$ is well-known in many cases. However, before defining this function, it is essential to note that many experimental datasets report two-photon event delays in raw uncalibrated fashion, τ_r, so an offset that aligns MPE events with a time separation of zero must often be determined, i.e. $\tau = \tau_r - \tau_0$. Part of the offset is due to the difference in length between optical paths from the beam splitter to the detectors, as well as the length of the electrical connections. It is constant for a given experimental configuration. Additionally, as the triggering order of detectors determines the sign of a time delay, an electronic offset applied during measurements will also typically be present, thus capturing a portion of the negative time delays. Now, in the case of SPS emission driven by a pulsed laser with period Λ, detected two-photon events should peak in number for temporal separations of $k\Lambda$, where $k \neq 0$. This outcome represents the detector identifying pairs of photons released by temporally neighbouring laser pulses. In scenarios where bunching on long time scales is observed, the envelope of these peaks can be considered to decay over extended ranges of time delays, i.e. by a factor of γ_e. As for the individual peaks, their spread across τ values can be modelled as two-sided exponential functions with decay factor γ_p.

Putting this all together, the mean rate of observing two-photon events with time separation τ is

$$R(\tau; \theta) = R_b + R_p e^{-\gamma_e |\tau|} \left(g e^{-\gamma_p |\tau|} + \sum_{k \neq 0} e^{-\gamma_p |\tau - k\Lambda|} \right), \tag{2}$$

where R_b represents background detections predominantly caused by detector dark counts, R_p denotes the peak rate of pulse-driven events, and g defines the ratio of MPE events to non-MPE events. Overall, the function has seven parameters when compared against experimental data defined over τ_r, i.e. $\theta = \{\tau_0, R_b, R_p, g, \gamma_e, \gamma_p, \Lambda\}$. In particular, g is one of the indicators of SPS quality, and it is often written as $g^{(2)}(0)$ to denote second-order coherence/correlation at zero time delay; the short form will be used in the rest of the text for convenience. A value of $g = 0$ indicates a perfect SPS.

Now, much of the work in this article involves fitting equation (2) to experimental observations of two-photon events and their time delays. However, infinite sums, even finitely truncated, are not

computationally efficient to calculate. So, using the floor function to apply a modulo operation, i.e. $\tau_m = \mathrm{mod}(\tau, \Lambda) = \tau - \Lambda \lfloor \tau/\Lambda \rfloor$, a useful insight is as follows:

$$
\sum_{k \in \mathbb{Z}} e^{-\gamma_{\mathrm{p}}|\tau - k\Lambda|} = \sum_{k \in \mathbb{Z}} e^{-\gamma_{\mathrm{p}}|\tau_m - k\Lambda|},
$$

$$
= e^{-\gamma_{\mathrm{p}}\tau_m} \sum_{k \in \mathbb{Z}_0^+} e^{-k\gamma_{\mathrm{p}}\Lambda} + e^{\gamma_{\mathrm{p}}(\tau_m - \Lambda)} \sum_{k \in \mathbb{Z}_0^+} e^{-k\gamma_{\mathrm{p}}\Lambda}. \tag{3}
$$

Effectively, for any value of τ, this term has the same answer as it does for some τ between 0 and Λ, i.e. an infinite sum of concave-right exponentials to the left and concave-left exponentials to the right. Here, a particular equality proves useful: $1 + e^{-x} + e^{-2x} + \ldots = e^x/(e^x - 1)$. Hence, defining $T = \gamma_{\mathrm{p}}\tau_m$ and $L = \gamma_{\mathrm{p}}\Lambda$ for convenience, equation (3) can be further manipulated as follows:

$$
\left(e^{-T} + e^{(T-L)} \right) \left(\frac{e^L}{e^L - 1} \right) = \frac{e^{\left(-T + \frac{L}{2}\right)} + e^{\left(T - \frac{L}{2}\right)}}{e^{\frac{L}{2}} - e^{-\frac{L}{2}}}
$$

$$
= \frac{\cosh\left(T - \frac{L}{2}\right)}{\sinh\left(\frac{L}{2}\right)}. \tag{4}
$$

Thus, by trading an infinite series for reliance on the modulo function, equation (2) can be written as a much more computationally amenable expression:

$$
R(\tau; \theta) = R_{\mathrm{b}} + R_{\mathrm{p}} e^{-\gamma_e |\tau|}
$$

$$
\times \left((g-1) e^{-\gamma_{\mathrm{p}}|\tau|} + \frac{\cosh\left(\gamma_{\mathrm{p}}\left(\tau_m - \frac{\Lambda}{2}\right)\right)}{\sinh\left(\frac{\gamma_{\mathrm{p}}\Lambda}{2}\right)} \right). \tag{5}
$$

This form of the equation makes it explicit that the mean rate for detected two-photon events appears as a repeating cosh function over the domain of τ, albeit with the peak around $\tau = 0$ diminished according to SPS quality; this is the peak that represents MPE events.

3. Methodology

The research described in this article revolves around the computational analysis of experimental interferometry measurements. To support reproducibility, the data and Python scripts are available at https://github.com/UTS-CASLab/sps-quality, commit f1782ff. Here, the experimental datasets are detailed in section 3.1, two fitting methods for estimating g are described in section 3.2, and the procedure for augmenting the experimental data, i.e. generating synthetic samples of observations, is elaborated in section 3.3.
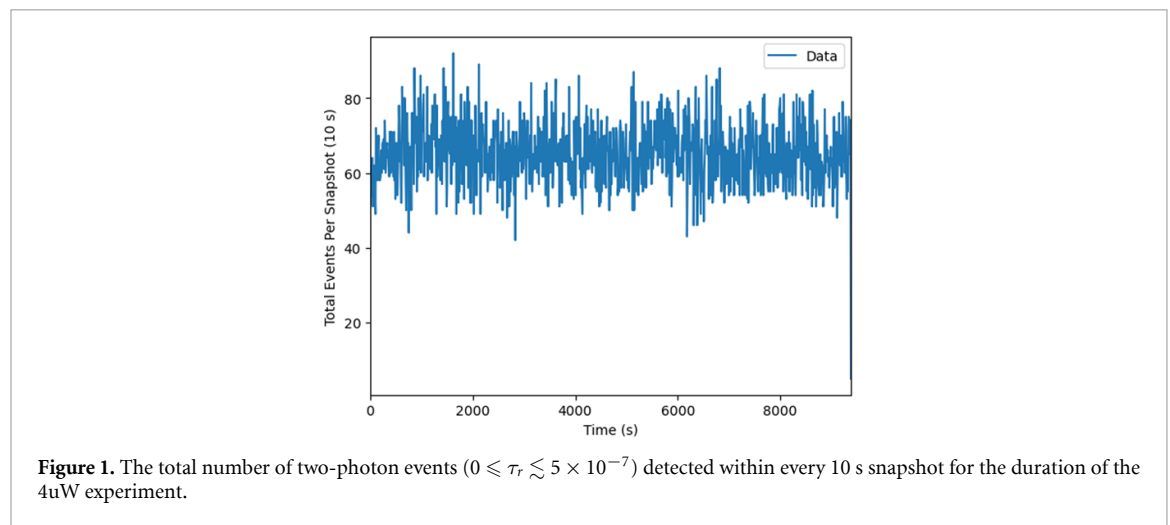
3.1. The datasets

This work selects eight experimental datasets for investigation, all sourced from prior attempts to engineer/assess a 'fibre-coupled semiconductor SPS for secure quantum-communication in the 1.3 μm range' (FI-SEQUR). Specifically, they all involve the same transition for the same single InGaAs/GaAs epitaxial quantum dot, which is positioned deterministically with respect to the centre of a photonic mesa structure. The growth and deterministic fabrication methods [44, 45], as well as the optical setup for assessing SPS quality [39], are detailed elsewhere.

Each dataset, summarised in table 1, results from a separate experiment where an 80 MHz laser of a fixed output power excites the SPS into emitting a train of photons. Here, the excitation is above-band, with the semiconductor laser operating at a wavelength of 805 nm and a pulse length of 50 ps. In contrast, many SPS devices reaching state-of-the-art $g^{(2)}(0)$ values [33–35] rely on coherent excitation schemes. Their trade-off is that the source laser tends to be more complex, both in physical implementation and size, and therefore less practical, e.g. requiring precise wavelength tuning for each quantum dot and a heavy spectral filtering of laser photons that reduces source brightness. However, excitation details are irrelevant to the data augmentation/analysis technique in this paper; only the obtained histogram matters.

After emission, the actual number of photons received by each detector is then measured in counts per second (CPS). This value depends on the efficiency and associated losses inherent within the experimental setup, e.g. the quantum efficiency of the single-photon detectors, as well as the source itself. On the source side, three factors are important: (1) the probability of occupation for the emitting state that is driven by the excitation power, (2) internal quantum efficiency, i.e. the probability that the excited quantum dot emits a

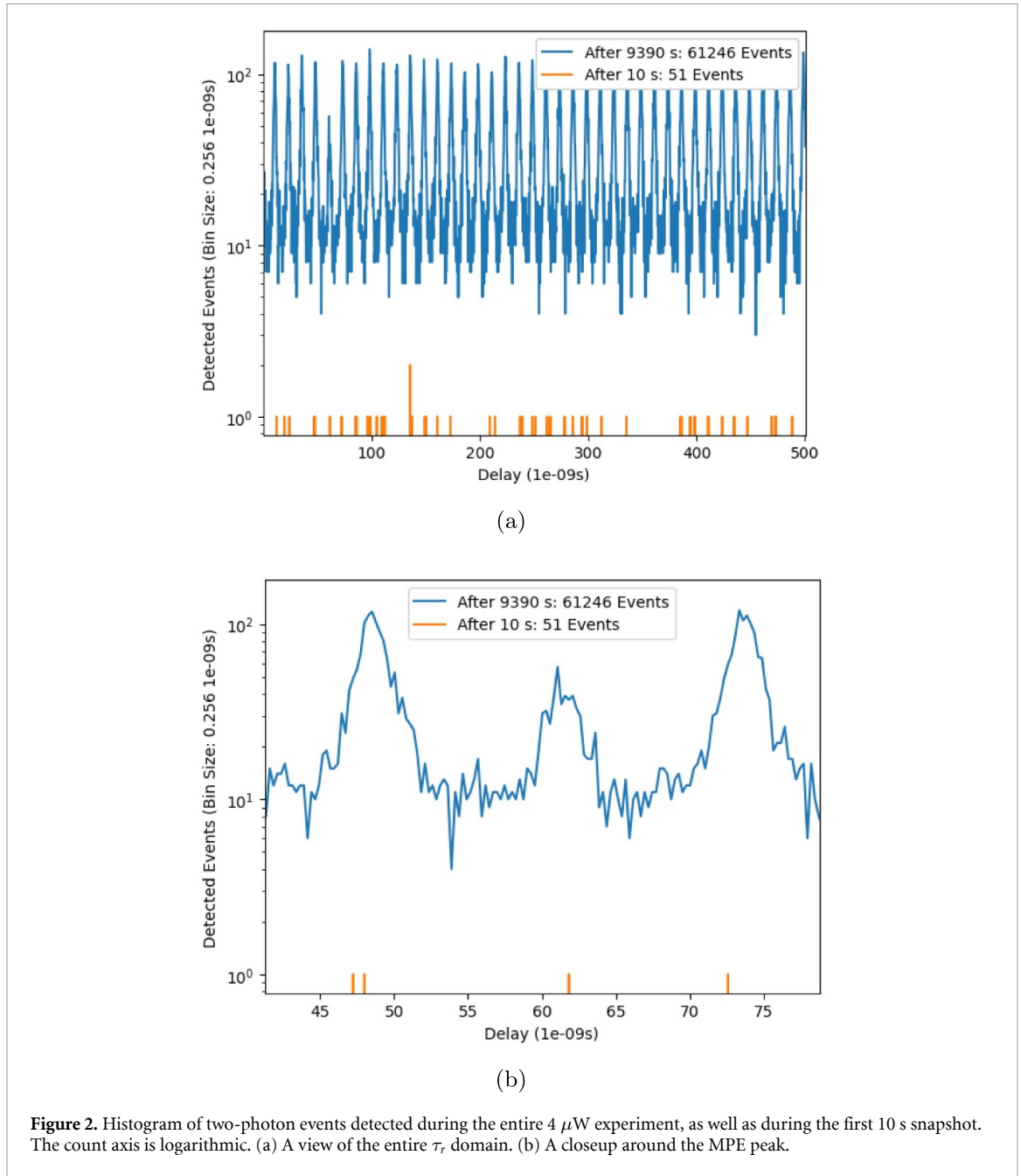**Table 1.** Descriptive summary of the eight FI-SEQUR datasets.

Shorthand label	Laser intensity (μW)	Single detector counts per second (Hz)	Total events	Total duration (s)	Average event rate (Hz)
1p2uW	1.2	3 000	65 037	23 950	2.716
2p5uW	2.5	4 000	51 534	10 510	4.903
4uW	4	4 100	61 246	9 390	6.522
8uW	8	5 100	21 611	2 920	7.401
10uW−	10	6 000	28 817	3 010	9.574
10uW+	10	12 000	45 198	1 210	37.354
20uW	20	7 000	55 469	4 780	11.604
30uW	30	7 000	52 088	4 780	10.897

Figure 1. The total number of two-photon events ($0 \leqslant \tau_r \lesssim 5 \times 10^{-7}$) detected within every 10 s snapshot for the duration of the 4uW experiment.

photon rather than non-radiatively relaxing, and (3) collection efficiency, i.e. the percentage of emitted photons that can be collected by the detection optics. Only the first factor should vary between the investigated datasets. Specifically, increased excitation power means a higher number of carriers are available to be captured by a quantum dot, resulting in shorter periods of time when the quantum dot is unoccupied. Photons cannot be generated without this occupation. However, two of the datasets have the same excitation power and different CPS values. This variation is possibly due to de-adjustment within the experimental setup, e.g. from temperature fluctuations in the lab.

Regarding contents, each dataset is a 2D matrix of two-photon event counts, binned in 10-second detection 'snapshots' along one axis and intervals of time separation along the other. Specifically, the delay domain of interest ranges from 0 ns to approximately 500 ns, with each bin typically covering a $\Delta\tau$ of 0.256 ns; it is 0.128 ns for the 1.2 μW experiment. Importantly, these are raw τ_r values, and the actual zero delay, i.e. MPE, occurs at about 60 ns due to the electronic offset discussed previously.

Notably, the dataset matrix can be summed along one axis to display how many two-photon events within the delay domain of interest are detected over time. Such a summation also enables calculating an average event rate within this domain. It is routine and unremarkable for each event rate to be far lower than its corresponding CPS value; encountering a *pair* of photons across detectors with sub-microsecond temporal separation is relatively rare. An example signal constructed from the 4 μW experiment is displayed in figure 1, where the visually discernible mean of approximately 65 events per 10 s snapshot corresponds with the average event rate of 6.522 Hz listed in table 1. However, the histogram compiled along the other axis of the dataset, i.e. the τ_r domain, is of greater interest. As figure 2(a) shows, one snapshot of 10 s provides a limited amount of information. In contrast, if the events from all snapshots are combined, a 40-peak 'comb' structure appears in the histogram. One of these peaks is smaller than all others, clarified by the figure 2(b) closeup, and, after accounting for the background, the relative amplitude of this MPE peak characterises the quality of an SPS.

(a)

(b)

Figure 2. Histogram of two-photon events detected during the entire 4 μW experiment, as well as during the first 10 s snapshot. The count axis is logarithmic. (a) A view of the entire τ_r domain. (b) A closeup around the MPE peak.

3.2. The fitting procedure

Core to this research is the notion of fitting equation (5) to histograms of data. Specifically, this work uses the Python LMFIT package [46] to, by default, minimise a sum of squared errors, thus obtaining the following optimised 'least-squares' (LS) parameters:

$$\theta_{\text{LS}} = \arg\min_{\theta} \sum_i \left(R\left(\tau_i; \theta\right) t - d_i \right)^2, \tag{6}$$

where i is the index of a bin centred at τ_i—this implicitly redefines R as a discrete function—and d_i is the number of events detected in that bin after t seconds. However, it has also been argued that there may be a better way than least-squares fitting. In short, an MAP proposal [38] suggests using the following 'Poisson-likelihood' (P) objective function instead:

$$\theta_{\text{P}} = \arg\min_{\theta} \sum_i \left(R\left(\tau_i; \theta\right) t - d_i \ln(R\left(\tau_i; \theta\right) t) \right). \tag{7}$$

Table 2. Initial values and parameter constraints used when fitting the mean-rate function in equation (5), times duration t, to the data d_i accumulated over that period.

Parameter	Vary	Min	Initial Value	Max
τ_0 (s)	✔	55×10^{-9}	60×10^{-9}	65×10^{-9}
R_b (Hz)	✔	0	$\left(\sum_i d_i/t\right) / \left(5 \times 10^{-7}/\Delta\tau\right)$	∞
R_p (Hz)	✔	0	$\left(\sum_i d_i/t\right) /40$	∞
g	✔	0	0.5	1
γ_e (s^{-1})	✘	—	0	—
$1/\gamma_p$ (s)	✔	$\Delta\tau/10$	$\Delta\tau$	∞
Λ (s)	✘	—	1.25×10^{-8}	—

Conveniently, by way of defining a `reduce_fcn` argument, the LMFIT package allows fitting procedures to minimise either of the two objectives described by equations (6) and (7) without affecting other processes, e.g. the automatic calculation of standard errors.

To maintain consistency with the MAP proposal [38], this work employs the Powell method for the fitting procedure, although paired with a subsequent Trust Region Reflective optimisation (`method = "least_squares"`). For completeness, the optimisation bounds and initial values for the parameters are listed in table 2, noting that pulse period Λ is experimentally fixed and γ_e, the envelope decay, is assumed to be negligible. Given these settings, a single-threaded completion of the five-parameter fitting procedure generally takes a second or less on a 3.50 GHz Intel Core i9-9900X CPU with 32 GB of RAM. Such speed allows histogram fitting to be part of any online method for estimating g, easily running alongside an actual event-detection experiment.

3.3. Data augmentation

Each of the eight datasets in table 1 represents a resource-expensive run of over 20 minutes, usually several hours, and each experiment observes between 20 000 and 70 000 two-photon events of interest in that time, i.e. with $0 \leqslant \tau_r \lesssim 5 \times 10^{-7}$. In theory, histograms compiled from the maximum number of events observed are likely to provide the best estimates of quality parameter g. A naive approach then would be to engineer a novel fitting algorithm and assess its virtues simply by how quickly its estimate of g—the fewer two-photon events in a histogram, the better—lands within fitting-based error bars of the maximum-time g estimate. However, the problem with this is as obvious as it is easily overlooked. What guarantee is there that the evolution of a second-order auto-correlation function $g^{(2)}(\tau)$, as exhibited by any one experiment, is typical? It would be disingenuous to hype up an estimation approach based on the off-chance that early detections of coincidences occur in just the right pattern. Then there is a more conservative question: are even 45 000 $\pm$ 25 000 observations sufficient to treat derived $g^{(2)}(0)$ values as ground truths?

One way to address these questions is to synthetically generate new data based on the existing datasets, i.e. engage in data augmentation [40, 41]. Early versions of the research in this article considered doing so by simply shuffling the 10 s snapshots available, assuming them to be independent. However, with such a method, the benefit of statistical independence is gradually lost over the summation of snapshots. Instead, the approach selected for this work leans on the accuracy of the assumptions introduced in section 2, namely the Poisson statistics governing two-photon event detections. Expressly, per dataset, assume the optimised fit of equation (5) applied to all observations is the ground-truth mean rate of detections, i.e. fit $R(\tau_i; \theta_{\text{best}}) \times t_{\text{total}}$. It is then possible to generate new data based on these ground truths as follows. For a given duration t_{new} and a temporal-separation bin centred at τ_i, randomly sample from a Poisson distribution with mean-value parameter $R(\tau_i; \theta_{\text{best}}) \times t_{\text{new}}$. Then, repeat this process for every bin index i until a new histogram is generated across the entire delay domain.

The main benefits of this technique, reminiscent of bootstrapping approaches used in the field [47, 48], are that (1) the new samples are independent, supporting statistical rigour, (2) thousands of synthetic datasets can be generated computationally at speed, (3) the method can extrapolate real-world stochastic data to simulate what is out of feasible reach for a typical experiment, e.g. the accumulation of a million observations, and (4) subsequent fits of the synthetic histograms indicate expected sampling-based errors for quality parameter g. Best of all, this procedure is generally applicable, provided that (1) two-photon event detection adheres to Poisson statistics and (2) the long-term shape of a two-photon event histogram can be theoretically described for an emission context. Thus, although eight variations involving a single quantum dot are used in this paper to showcase the technique, hence enabling convenient comparative analysis with minimal conflating factors, it can just as easily be applied to an SPS stimulated by a continuous-wave laser, a thermal light source [38], and so on.

Table 3. Optimised parameters with standard errors for the mean-rate function in equation (5) when fitting $R(\tau;\theta) \times t$ to all the data available in a dataset.
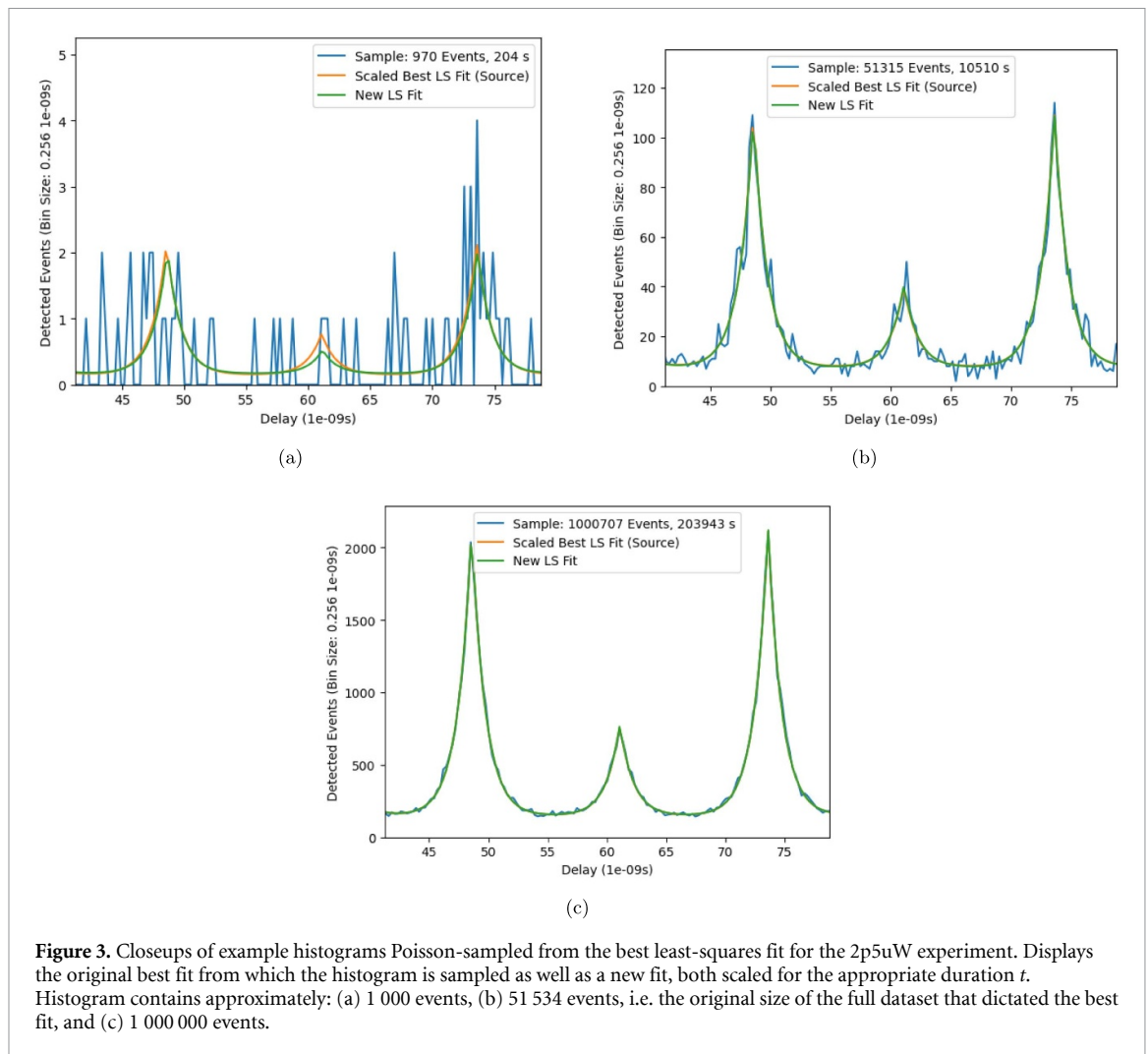
Best Fit	g	τ_0 (s)	R_b (Hz)	R_p (Hz)	$1/\gamma_p$ (s)
1p2uW LS	0.23 ($\pm$11.72%)	6.11×10^{-8} ($\pm$0.05%)	2.54×10^{-4} ($\pm$24.37%)	2.55×10^{-3} ($\pm$1.20%)	1.10×10^{-9} ($\pm$13.20%)
1p2uW P	0.22 ($\pm$12.17%)	6.11×10^{-8} ($\pm$0.05%)	2.66×10^{-4} ($\pm$22.90%)	2.59×10^{-3} ($\pm$1.17%)	1.05×10^{-9} ($\pm$13.40%)
2p5uW LS	0.32 ($\pm$8.79%)	6.11×10^{-8} ($\pm$0.06%)	7.24×10^{-4} ($\pm$39.84%)	9.76×10^{-3} ($\pm$1.37%)	1.16×10^{-9} ($\pm$15.13%)
2p5uW P	0.30 ($\pm$9.23%)	6.11×10^{-8} ($\pm$0.06%)	7.79×10^{-4} ($\pm$36.30%)	1.01×10^{-2} ($\pm$1.32%)	1.10×10^{-9} ($\pm$15.33%)
4uW LS	0.36 ($\pm$7.31%)	6.11×10^{-8} ($\pm$0.05%)	9.13×10^{-4} ($\pm$42.84%)	1.30×10^{-2} ($\pm$1.35%)	1.18×10^{-9} ($\pm$15.08%)
4uW P	0.34 ($\pm$8.01%)	6.11×10^{-8} ($\pm$0.05%)	1.09×10^{-3} ($\pm$34.90%)	1.34×10^{-2} ($\pm$1.31%)	1.07×10^{-9} ($\pm$16.01%)
8uW LS	0.49 ($\pm$9.15%)	6.11×10^{-8} ($\pm$0.10%)	1.14×10^{-3} ($\pm$79.85%)	1.36×10^{-2} ($\pm$2.75%)	1.23×10^{-9} ($\pm$32.09%)
8uW P	0.48 ($\pm$9.27%)	6.11×10^{-8} ($\pm$0.09%)	1.09×10^{-3} ($\pm$81.37%)	1.39×10^{-2} ($\pm$2.66%)	1.23×10^{-9} ($\pm$30.62%)
10uW− LS	0.53 ($\pm$7.20%)	6.11×10^{-8} ($\pm$0.08%)	1.20×10^{-3} ($\pm$91.99%)	1.74×10^{-2} ($\pm$2.54%)	1.34×10^{-9} ($\pm$27.54%)
10uW− P	0.50 ($\pm$7.44%)	6.11×10^{-8} ($\pm$0.08%)	1.19×10^{-3} ($\pm$87.34%)	1.79×10^{-2} ($\pm$2.37%)	1.31×10^{-9} ($\pm$26.15%)
10uW+ LS	0.51 ($\pm$5.92%)	6.11×10^{-8} ($\pm$0.07%)	4.46×10^{-3} ($\pm$78.09%)	7.33×10^{-2} ($\pm$1.93%)	1.26×10^{-9} ($\pm$22.14%)
10uW+ P	0.50 ($\pm$6.15%)	6.11×10^{-8} ($\pm$0.06%)	4.92×10^{-3} ($\pm$68.21%)	7.48×10^{-2} ($\pm$1.85%)	1.20×10^{-9} ($\pm$22.11%)
20uW LS	0.56 ($\pm$5.94%)	6.11×10^{-8} ($\pm$0.07%)	2.18×10^{-3} ($\pm$49.65%)	1.78×10^{-2} ($\pm$2.39%)	1.34×10^{-9} ($\pm$26.62%)
20uW P	0.57 ($\pm$6.05%)	6.11×10^{-8} ($\pm$0.07%)	2.44×10^{-3} ($\pm$44.00%)	1.81×10^{-2} ($\pm$2.34%)	1.22×10^{-9} ($\pm$28.61%)
30uW LS	0.68 ($\pm$5.97%)	6.11×10^{-8} ($\pm$0.09%)	2.62×10^{-3} ($\pm$50.86%)	1.34×10^{-2} ($\pm$3.68%)	1.39×10^{-9} ($\pm$41.80%)
30uW P	0.68 ($\pm$5.96%)	6.11×10^{-8} ($\pm$0.09%)	2.76×10^{-3} ($\pm$48.89%)	1.36×10^{-2} ($\pm$3.67%)	1.31×10^{-9} ($\pm$44.31%)

4. Results

The key question is as follows: at what point will an expensive SPS quality-assurance experiment have collected enough data to warrant shutting down?

If allowed to run for a seemingly long time, detecting $45\,000 \pm 25\,000$ coincidences, the experimental setups described by table 1 enable what will be called least-squares and Poisson-likelihood 'best fits' for the parameters listed in table 2. These apparent ground truths are shown in table 3. Sure enough, as reported earlier [39], increasing the output power of the exciting laser diminishes the SPS quality of the studied InGaAs/GaAs epitaxial quantum dot. Also of note, the fitting errors for g are already not insubstantial. While the location of the MPE peak and the maximum detection rate—recall the comb structure in figure 2(a)—are easily identified, the background and the width of the peaks are harder to characterise with certainty. But what of sampling-based errors?

For the present analysis, the generative method described in section 3.3 works to 'Poisson-sample' histograms representing the following numbers of total events: 1×10^3, 1×10^4, 1×10^5, and 1×10^6. Naturally, every unique experiment is expected to accumulate these observations at different times. For instance, based on the average event rate in table 1, 1p2uW should encounter 10 000 two-photon events at $t \simeq 3\,682$ s, while 10uW+ should observe 100 000 at $t \simeq 2\,677$ s. Nonetheless, given that a unit of information is more directly tied with an individual observation than the passage of a second, this dataset-dependent normalisation is essential for a fair comparison.

Figure 3. Closeups of example histograms Poisson-sampled from the best least-squares fit for the 2p5uW experiment. Displays the original best fit from which the histogram is sampled as well as a new fit, both scaled for the appropriate duration *t*. Histogram contains approximately: (a) 1 000 events, (b) 51 534 events, i.e. the original size of the full dataset that dictated the best fit, and (c) 1 000 000 events.

Accordingly, using the appropriate values of *t*, the next step is to optimise equation (5) for any synthetic histograms generated. Immediate inspection shows, as exemplified by figure 3 for the 2p5uW experiment and least-squares fitting methodology, that typical fits for sampled data become less and less visually discrepant with the appropriately scaled ground truth as more events are 'observed'. Of course, examining a single fit is of limited use; synthetically generated histograms prove most informative in bulk. For instance, figure 4 shows how the least-squares fitted parameter *g* varies across 250 Poisson-samplings for each of the following sizes: ∼1 000, ∼10 000, ∼100 000, and ∼1 000 000. It also covers artificial histograms of the same size as the full experimental dataset, i.e. ∼51 534 detections for 2p5uW.

This result reveals an unfortunate implication. Even after observing tens of thousands of two-photon events during the characterisation of an SPS under 2.5 μW laser excitation power, the estimate of quality parameter *g* will not just have a fitting-related uncertainty of ∼0.025; it could be off from a true characterisation by over ∼0.1, simply due to variability in Poisson processes. Indeed, this sampling-revealed variance impacts all eight experiments, as figure 5 demonstrates for histograms of size 1×10^5. That said, in fairness, the box plots in the figure indicate that most of the sampled fits are distributed much more tightly than they appear in the scatter plots. Standard deviations are listed in table 4. Accordingly, if there is a threshold for SPS quality, such as *g* = 0.5, most of the listed experiments will be clearly classifiable even after observing only 10 000 events. Nevertheless, in the case of 8–10 μW, the stochastic nature of detection can easily swing determinations one way or the other.

Turning to the fitting methodologies described in section 3.2, this work was unable to identify any significant or systematic benefit to optimising the Poisson-likelihood objective function in equation (7) proposed by previous research [38]. Generally, the fitting and sample-based errors appear reasonably similar regardless of the method used. Worse yet, while minimising the least-squares residual appears robust, almost universally aligning the mean value of fitted parameters for Poisson-sampled histograms with the original
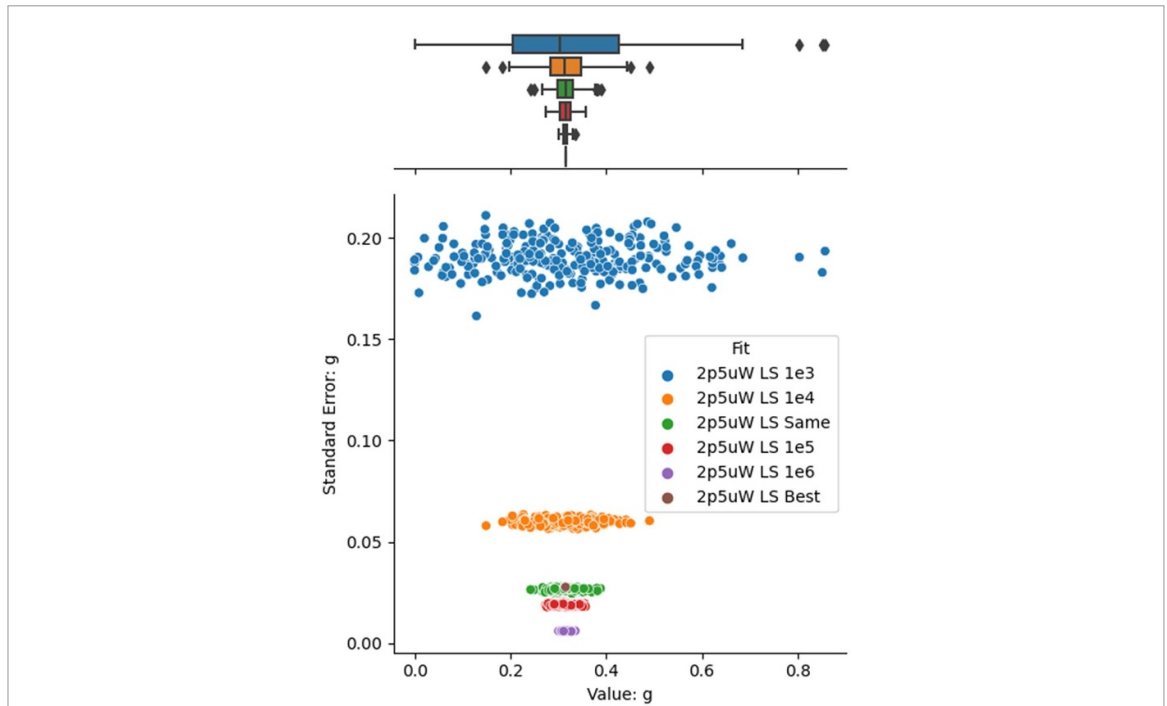
Figure 4. Comparison of least-squares fitted parameter *g* between the 'ground-truth' best fit for the 2p5uW experiment and fits applied to subsequent Poisson-sampled histograms. There are 250 histograms generated for each of the following approximate numbers of total events: 1×10^3, 1×10^4, 1×10^5, and 1×10^6. There are also 250 generated for approximately the same number of events as used for the best fit, i.e. 51 534 for 2p5uW.
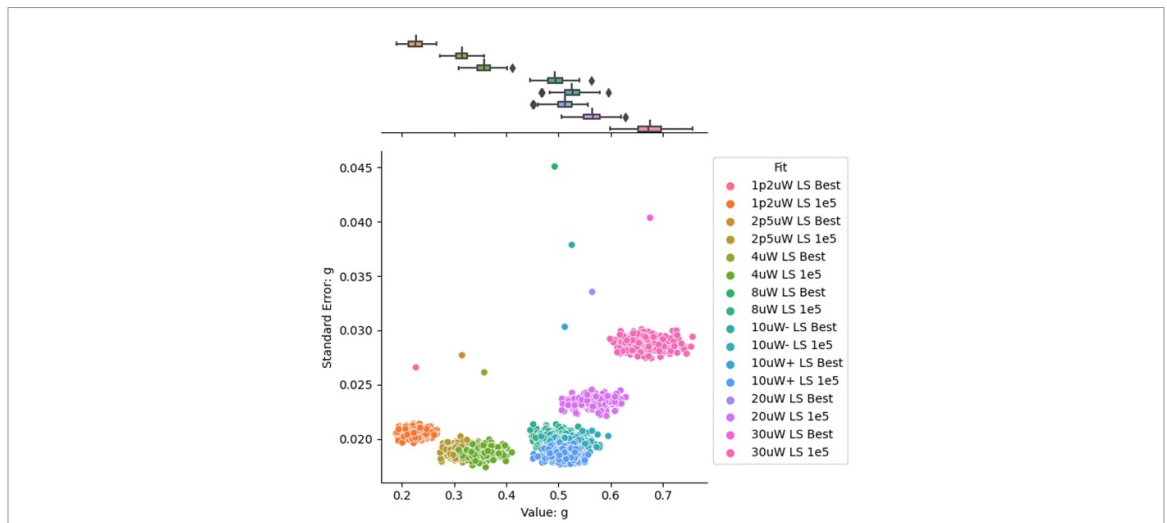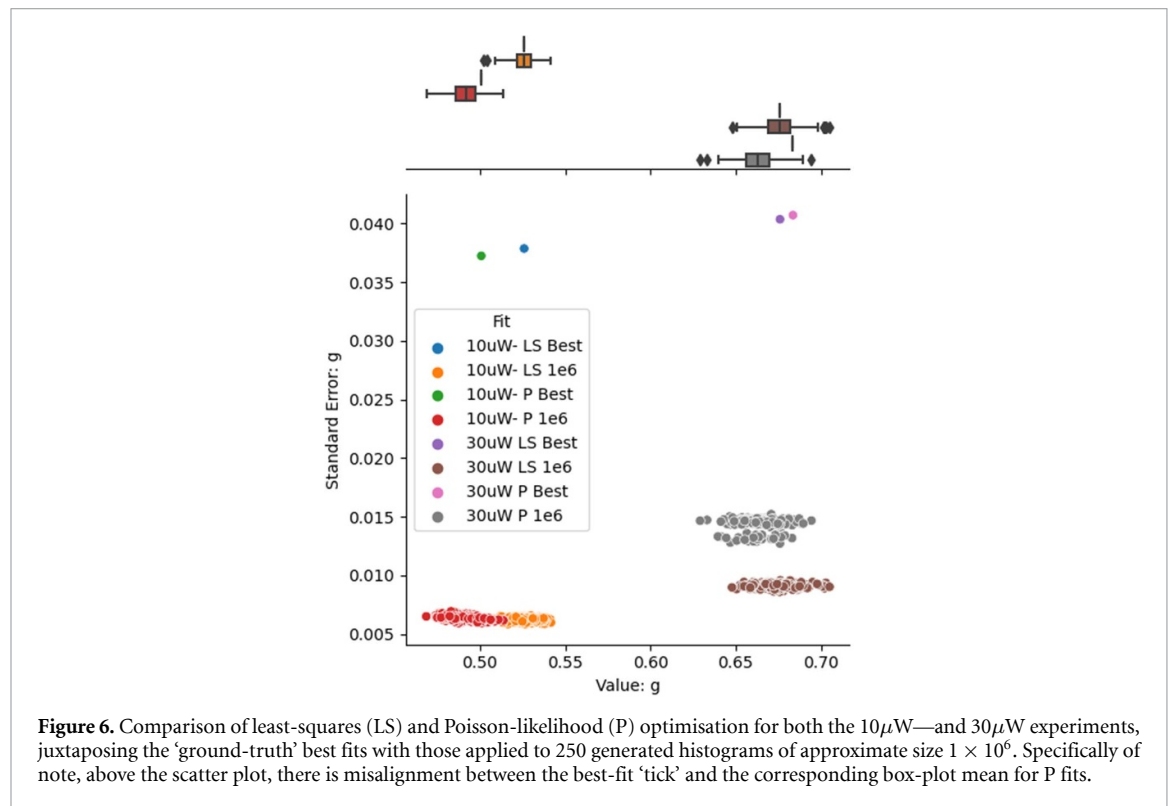

Figure 5. Comparison of least-squares fitted parameter *g*, per dataset, between the 'ground-truth' best fit and fits applied to 250 subsequent Poisson-sampled histograms of approximate size 1×10^5.

'best fit' value—see the boxplots in figure 4 or figure 5—this is not the case for the Poisson-likelihood method. For instance, figure 6 exemplifies cases where generated histograms should closely approximate the shape of the function they were Poisson-sampled from, being of size 1×10^6, yet their P fits do not align with the best-fit parameter values. It is unclear why this is; one hypothesis is that converging to an optimum for equation (7), where logarithms are involved, can face challenges of numerical stability. More investigation is required to confirm this. Nonetheless, as this work cannot validate a substantial comparative advantage of the MAP procedure, the rest of this article focusses solely on least-squares results.

Now, the takeaway thus far is necessarily cautioning. The quality parameter *g* is highly sensitive, crucially dependent on a small fraction of detected events across the τ delay domain and easily distorted by an incorrect background characterisation. Hoping for a rapid and authoritative estimate of *g* may be overly optimistic when even hours of accumulated observations leave a sizeable uncertainty. Nonetheless, this work indicates how much uncertainty to expect at different two-photon event counts, providing, at a minimum, a

Table 4. Mean values and standard deviations for fitted parameter *g*, as derived from Poisson-sampled histograms of size 1×10^4 and 1×10^5 for different datasets. Both least-squares and Poisson-likelihood fits are represented. Also includes values for Poisson-sampled histograms containing the same number of events as each full dataset. Importantly, the sample-based standard deviation here is not the same as a fit-specific standard error.

Best Fit	1×10^4 Event Histograms		Same Size Histograms		1×10^5 Event Histograms	
	μ_g	σ_g	μ_g	σ_g	μ_g	σ_g
1p2uW LS	0.227	0.056	0.230	0.021	0.226	0.018
1p2uW P	0.218	0.059	0.219	0.022	0.211	0.017
2p5uW LS	0.315	0.054	0.317	0.025	0.315	0.017
2p5uW P	0.299	0.055	0.295	0.023	0.294	0.017
4uW LS	0.357	0.064	0.358	0.023	0.358	0.018
4uW P	0.333	0.056	0.332	0.023	0.332	0.019
8uW LS	0.486	0.072	0.492	0.052	0.494	0.021
8uW P	0.477	0.065	0.474	0.044	0.473	0.020
10uW− LS	0.519	0.065	0.520	0.036	0.527	0.021
10uW− P	0.492	0.060	0.494	0.039	0.494	0.020
10uW+ LS	0.511	0.073	0.507	0.034	0.512	0.020
10uW+ P	0.491	0.063	0.492	0.030	0.495	0.019
20uW LS	0.558	0.078	0.563	0.034	0.565	0.025
20uW P	0.567	0.074	0.566	0.035	0.567	0.026
30uW LS	0.664	0.100	0.674	0.045	0.674	0.032
30uW P	0.672	0.092	0.684	0.043	0.685	0.030

Figure 6. Comparison of least-squares (LS) and Poisson-likelihood (P) optimisation for both the 10μW—and 30μW experiments, juxtaposing the 'ground-truth' best fits with those applied to 250 generated histograms of approximate size 1×10^6. Specifically of note, above the scatter plot, there is misalignment between the best-fit 'tick' and the corresponding box-plot mean for P fits.

'bootstrapping' procedure to apply in varying experimental contexts. Since different practical applications involving single photons have varying requirements, one must decide what confidence level to embrace for a particular quantum optical measurement, subsequently accumulating detected events until the fitting and expected sampling errors are small enough to declare an SPS is 'good' or 'bad' for a particular purpose.

Figure 7. The expected distributions for expanding averages of *g*, displayed for select datasets. Each expanding average is applied to a random iterative selection of histograms, where each histogram contains ∼1 000 events; the dotted lines exemplify one such expanding average per dataset. Each shaded distribution is defined by the mean of 1 500 expanding averages, plus/minus one standard deviation. The black vertical 'whiskers' mark the mean of *g*, plus/minus one standard deviation, derived from fitting histograms that contain as many events as each original dataset.

Granted, while small samples of two-photon events are highly unreliable and unlikely to carry discernible information about SPS quality, it is notable that averages over collections of small samples align well with individual fits over larger sample sizes. The boxplots in figure 4 demonstrate this fact. Admittedly, there is a lower limit for which this insight holds—histograms of size 100 are so sparse that there are systematic errors in fitting—but sampling 1 000 events seems to meet this threshold decently. One might then lean on an idea known well in machine learning and other fields, i.e. ensembling weak estimators to create a singular powerful one.
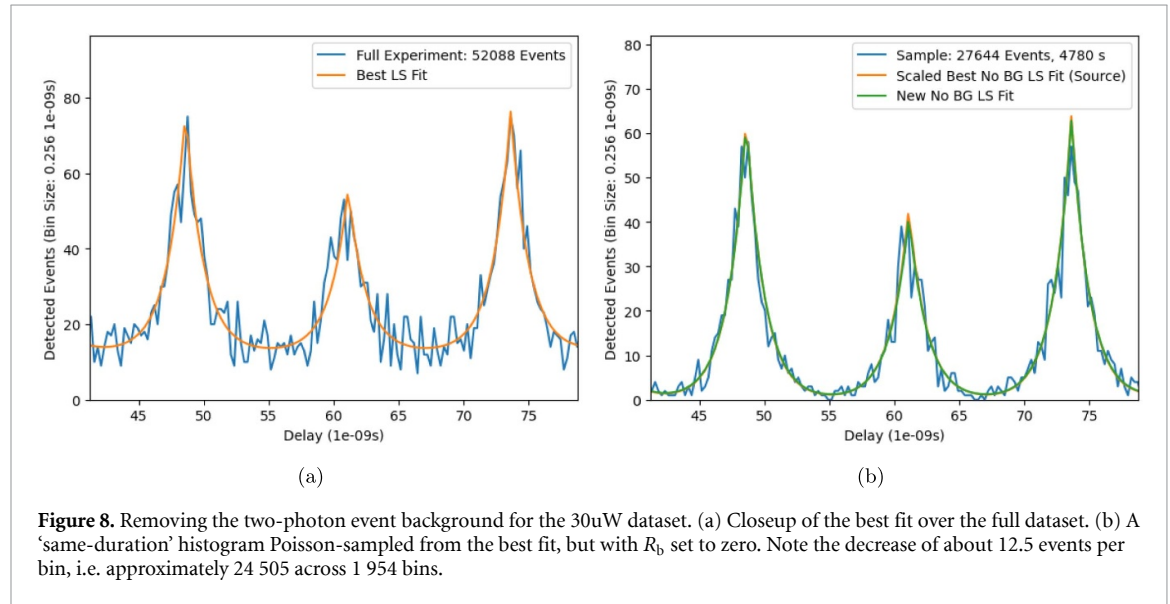
Specifically, consider a steady accumulation of observed events in the vein of the eight datasets listed in table 1. After every 1 000 events, construct a histogram from only those 1 000 events, maintaining independent sampling, and then fit it. Subsequently, average the parameters of that fit with those derived from all previous samples of size 1 000. This concept is called an expanding average, and the value should generally improve over time. Of course, it is still possible for this procedure to produce and sum long sequences of atypical histograms, e.g. ones in figure 4 that are far from the mean, so a rigorous analysis requires the random but iterative Poisson-sampling of histograms, size 1×10^3, to *itself* be redone numerous times, e.g. in the form of 1 500 independent sequences generated for this work. Given all this, the distributions of those 1 500 expanding averages, plus and minus one standard deviation, are displayed in figure 7 for several datasets.

These results show, especially considering the whiskers that denote the uncertainty in *g* when fitting a single histogram of large size, that decent estimates of quality might be possible earlier on by averaging small-sample fits. For instance, fitting the combined 52 088 events in the 30uW dataset provides a lower degree of confidence in *g* than averaging the parameter over 52 independent fits of 1 000 events. This competitive advantage is true as early as 30 000 events. Moreover, even where there is no clear-cut difference, the rapid tapering of confidence bounds still suggests that an expanding average can be a viable early estimator of SPS quality. However, distributions provide no guarantee; the standard deviations of both expanding averages and large-sample fitting, depicted in figure 7, are probabilistic. Some example trajectories, e.g. for 8uW, veer substantially away from the distribution mean even at high numbers of samples. Elsewhere, e.g. for 1p2uW, the distribution itself does not align perfectly with the 'best fit', possibly suggesting a minor systematic error in fitting for histograms of size 1×10^3, just as is present for size 1×10^2. Essentially, the stochastic nature of two-photon event observations induces many complex challenges that stifle certainty for early estimates.

In theory, some of this uncertainty could be better constrained if there was a way to identify and eliminate the background, i.e. the model component in equation (2) that is considered to be uniform across the delay domain. After all, if the best-fit 'ground truths' are to be believed, all the analysed datasets involve substantial counts of two-photon events, or at least triggered 'detections', that are not due to SPS emission. The ratios are listed in table 5 and range from 0.23 to 0.47. So, to investigate the effect of eliminating the background, the generative procedure in this work can be adjusted. After acquiring a 'best fit' of the five parameters in table 3,

Table 5. Background (BG) rates for the best least-squares fit applied to each full dataset. Each rate indicates the portion of events that are associated with BG detections, i.e. a BG ratio.
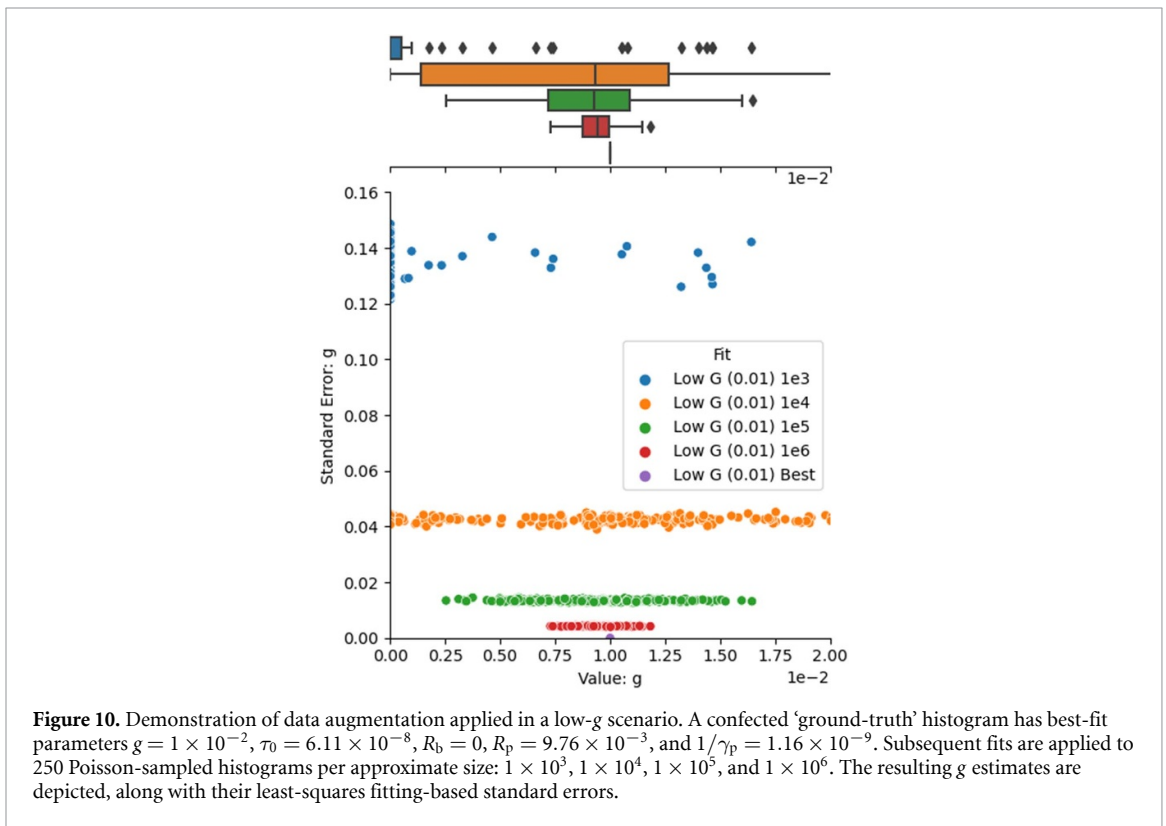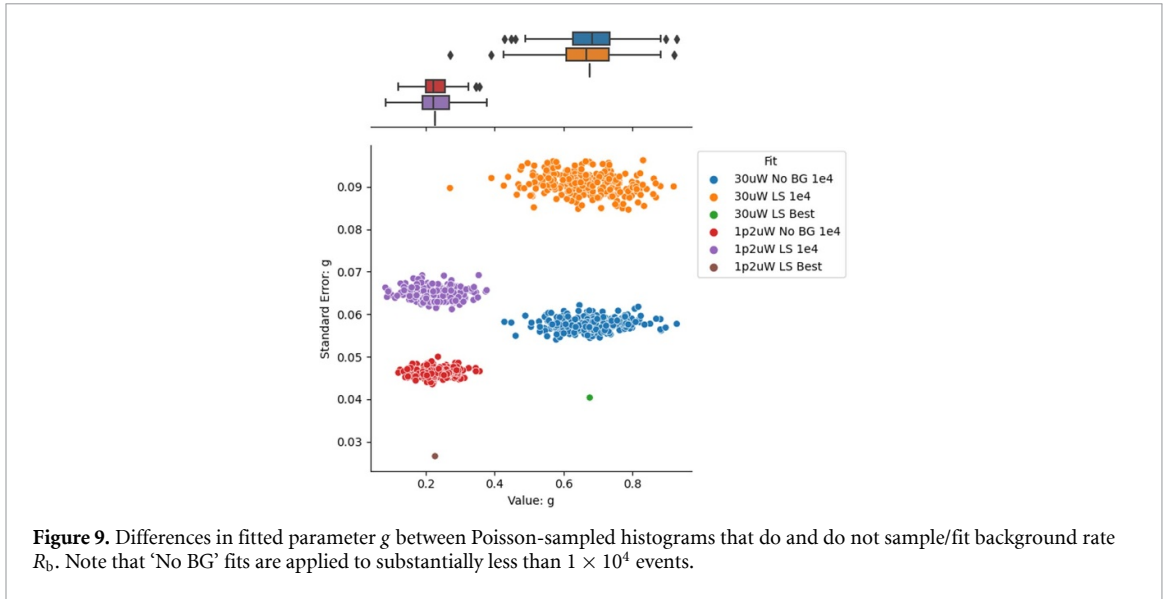
Dataset	Best Fit BG Rate R_b (Hz)	BG Events	Total Events	BG Ratio
1p2uW LS	2.54×10^{-4}	23 787	65 037	0.37
2p5uW LS	7.24×10^{-4}	14 868	51 534	0.29
4uW LS	9.13×10^{-4}	16 748	61 246	0.27
8uW LS	1.14×10^{-3}	6 507	21 611	0.30
10uW− LS	1.20×10^{-3}	7 028	28 817	0.24
10uW+ LS	4.46×10^{-3}	10 549	45 198	0.23
20uW LS	2.18×10^{-3}	20 317	55 469	0.37
30uW LS	2.62×10^{-3}	24 505	52 088	0.47

Figure 8. Removing the two-photon event background for the 30uW dataset. (a) Closeup of the best fit over the full dataset. (b) A 'same-duration' histogram Poisson-sampled from the best fit, but with R_b set to zero. Note the decrease of about 12.5 events per bin, i.e. approximately 24 505 across 1 954 bins.

visually depicted by figure 8(a), the background rate R_b can be set to zero, subsequently allowing the generation of a 'No BG' Poisson-sampled histogram, such as is exemplified in figure 8(b).

Crucially, without a background rate, the average event rate in table 1 and the total number of events in a sampled histogram are scaled by a factor of one minus the 'BG Ratio' listed in table 5. So, for instance, a 30uW 'No BG' histogram of size 1×10^4 actually contains around 5 300 events. Even so, despite the substantially fewer events detected over the same measurement duration, figure 9 shows that four-parameter fits applied to the 'No BG' histograms—R_b in table 2 is set to zero and not allowed to vary—considerably tighten their standard errors. Background elimination is thus a worthwhile pursuit in acquiring better estimates of SPS quality. However, as the figure demonstrates, this would not impact the variance intrinsic to the Poisson process that describes event detection; a significant degree of uncertainty is unavoidable.

As a final reminder, the data augmentation technique in this paper is generally applicable, provided that certain physical assumptions hold. Accordingly, consider a state-of-the-art low-g SPS device [33–35] with the following ground-truth emission profile: $g = 1 \times 10^{-2}$, $\tau_0 = 6.11 \times 10^{-8}$, $R_b = 0$, $R_p = 9.76 \times 10^{-3}$, and $1/\gamma_p = 1.16 \times 10^{-9}$. As figure 10 shows, one can Poisson-sample histograms to quantify uncertainty for this device in much the same manner as for the FI-SEQUR quantum dot. However, there are caveats. With such a tiny MPE peak, sufficient detections must be gathered to even acknowledge its existence. For example, fitting synthetic histograms of 1×10^3 events predominantly and falsely suggests a perfect $g = 0$; any exceptions are statistical outliers. Moreover, even when the number of two-photon events is sufficiently large to resolve g, e.g. when the 250 histograms of size 1×10^6 statistically indicate $g = 9.368 \times 10^{-3} \pm 0.847 \times 10^{-3}$, one must also consider whether the standard error of the fitting algorithm itself washes out the g value of interest. The unfortunate takeaway is that acquiring sufficient confidence in more extreme values of g requires a much greater signal-to-noise ratio, e.g. by increasing CPS or integration times. Whatever the case, the power of the data augmentation technique is that it can quantify these expectations for experimentalists.

Figure 9. Differences in fitted parameter *g* between Poisson-sampled histograms that do and do not sample/fit background rate R_b. Note that 'No BG' fits are applied to substantially less than 1×10^4 events.

Figure 10. Demonstration of data augmentation applied in a low-*g* scenario. A confected 'ground-truth' histogram has best-fit parameters $g = 1 \times 10^{-2}$, $\tau_0 = 6.11 \times 10^{-8}$, $R_b = 0$, $R_p = 9.76 \times 10^{-3}$, and $1/\gamma_p = 1.16 \times 10^{-9}$. Subsequent fits are applied to 250 Poisson-sampled histograms per approximate size: 1×10^3, 1×10^4, 1×10^5, and 1×10^6. The resulting *g* estimates are depicted, along with their least-squares fitting-based standard errors.

5. Discussion

One of the primary motivators of this work is the publication of a MAP method for SPS quality estimation [38] that promised up to two orders of magnitude shorter data acquisition times. Such a claim is understandably alluring to experimentalists, as performing statistical measurements on quantum emitters is resource expensive and time consuming.

Unfortunately, this statement must be interpreted carefully. First of all, experimental context matters. For instance, the laser pulses in the MAP publication [38] appear to have a period of 1 μs, as opposed to the 12.5 ns in this work. This detail means that sequential peaks in the data are well separated and distinct; inaccuracies in fitting the decay factor γ_p are unlikely to affect confidence in *g* as much as they do for the compressed comb structures in this work.

Admittedly, a reduced frequency of excitation pulses does decrease the number of photons detected within a delay window of interest, so there is a trade-off. Indeed, in the case of many practical applications, it

is appealing to drive the SPS as often as possible to maximise the single-photon generation rate. Hence, assessing SPS quality under regimes of higher-frequency excitation retains its appeal. Of course, there is a fundamental limit here, in that the SPS cannot be excited until it has time to relax; the spontaneous-emission lifetime for the FI-SEQUR system is around one to two nanoseconds, and the peaks cannot be brought closer to each other than that. However, if decay factor γ_p could be confidently fitted by a quick supplementary experiment, e.g. involving distinct excitation pulses, then perhaps overlapping peaks would no longer be a major factor of uncertainty, enabling many more valuable observations in the same amount of time without any of the drawbacks.

Notably, time is a poor proxy for the pace of accumulating information if the rates of detecting two-photon events vary dramatically between experiments. A brief measurement can correspond to many two-photon events if the count rates are high. Such a reason is why this work aligned synthetic histograms across datasets by the number of events observed rather than the time passed. Indeed, for the sake of fair comparison between different methods of SPS quality estimation, one has to consider histograms of equal information content: an equivalent number of events observed across the given time domains. If so, maybe being able to sample so many repeated peaks proves minorly beneficial when fitting FI-SEQUR data, in that R_p becomes accurately characterised. However, since g relies heavily on the MPE peak, early quality estimates may actually benefit more from a relatively high proportion of observations existing in that central zone. Essentially, the 50-second integration time presented in the MAP article [38] may provide many more *informative* events within the relevant domain than most FI-SEQUR datasets do for the same amount of time; see figure 2(b) and extrapolate from the ten-second histogram.

The signal-to-background ratio is another complicating factor. Table 5 in this work shows that the FI-SEQUR datasets contain numerous 'false' detection events that, if eliminated, would significantly improve estimates; see figure 9. In contrast, the experimental data considered in the MAP publication [38] appears to have no background. Of course, the feasibility of independent background characterisation within an experimental context, let alone its elimination, is another matter to discuss elsewhere. The point here is that, second for second, the correlation measurement setup in the MAP paper appears to observe many more two-photon events under adjacent peaks than the FI-SEQUR experiments detailed here.

Essentially, this work was unable to confirm the suggested advantage of the Poisson-likelihood approach [38] compared to least-squares fitting. Uncertainties in g, due to both fitting errors and stochastic effects, remain unavoidable for now. Of course, if the claim was instead referring to the efficiency of computationally estimating g, regardless of its error, then the reformulation provided by equation (5) of this work will find appreciation for its added algorithmic utility.

6. Conclusion

Determining the second-order correlation function that describes the emission statistics of a light source requires the acquisition of two-photon events. In the context of intended SPS quantum dots, such measurements are costly in both time and material resources. Consequently, there exists a research endeavour focussed on estimating SPS quality as quickly and accurately as possible. However, this work cautions that recent optimism in the literature, promoting new techniques capable of 'fast estimates in under a minute' and 'a one-to-two order of magnitude speed-up', may need to be tempered.

A central concern is that novel estimation approaches may be promoted without validation on a sufficient amount of data. One singular experiment, i.e. one instance of how a histogram binning observed events evolves, is simply not enough. Hence, assuming detections adhere to Poisson statistics, this work centres on a generative method similar to bootstrapping that allows for the computational synthesis/analysis of new data from existing datasets. Specifically, having applied this process of data augmentation to eight datasets studying a fibre-coupled InGaAs/GaAs epitaxial quantum dot, this work assesses estimates of the SPS figure of merit g, a.k.a. $g^{(2)}(0)$, and discusses the impact of stochastic variability in the measurements.

The major contributions of this work are as follows:

- The reformulation of well-known theoretical equations describing the expected delay-domain histograms of detected two-photon coincidences, in the context of an SPS excited by a pulsed laser, thus allowing data to be fit in a computationally efficient manner.
- Proof of principle that data augmentation is a valuable tool in quantifying the unavoidable systematic errors in g that result from the variability of Poisson processes. This work cautions that neglecting these errors, a common oversight in the literature, can lead to unwarranted overconfidence in early g estimates and pre-mature declarations of SPS devices as state-of-the-art.

- A comparison between the standard least-squares approach and a recently proposed Poisson-likelihood method, finding no significant/systematic advantage for the latter and even a potential vulnerability to numerical instabilities during optimisation.
- An investigation of whether expanding averages over small-size histograms is competitive against *g* estimation based on a single large-size histogram. The results suggest expanding averages show some promise for early estimation, but deeper analysis is required to determine their general utility rigorously.
- The finding that suppressing background counts would boost the fitting accuracy of *g* but has no significant discernible effect on the error from the stochastic variability in detections.

Ultimately, this research serves as another example of how experimental physics can benefit from applying data-driven approaches commonly found in machine learning, such as data augmentation. In this case, provided the real-world data adheres well to the assumptions underlying a theoretical model, experimental observations can be supplemented with numerical 'extrapolations' from which bootstrapped statistics can be deduced. Only with this improved understanding can methods for assessing SPS quality be judged appropriately for speed *and* accuracy.

Data availability statement

The data that support the findings of this study are openly available at the following URL/DOI: https://github.com/UTS-CASLab/sps-quality.

Acknowledgment

Conflict of interest

The authors have no conflicts to disclose.

Author contributions

David Jacob Kedziora: Conceptualisation (supporting); Formal Analysis (lead); Methodology (lead); Software (lead); Visualisation (lead); Writing—original draft (lead); Writing—review & editing (equal). **Anna Musiał**: Conceptualisation (equal); Data Curation (lead); Investigation (equal); Writing—review & editing (equal). **Wojciech Rudno-Rudziński**: Investigation (equal); Writing—review & editing (equal). **Bogdan Gabrys**: Conceptualisation (equal); Methodology (supporting); Project Administration (lead); Supervision (lead); Writing—review & editing (equal).

ORCID iDs

David Jacob Kedziora ⬤ https://orcid.org/0000-0002-0673-182X
Anna Musiał ⬤ https://orcid.org/0000-0001-9602-8929
Wojciech Rudno-Rudziński ⬤ https://orcid.org/0000-0002-3475-7517
Bogdan Gabrys ⬤ https://orcid.org/0000-0002-0790-2846

References

[1] OIDA 2020 OIDA quantum photonics roadmap: every photon counts (available at: https://opg.optica.org/abstract.cfm?URI=OIDA-2020-3OIDA) 3
[2] Sasaki M *et al* 2011 Field test of quantum key distribution in the Tokyo QKD network *Opt. Express* **19** 10387
[3] Liao S-K *et al* 2017 Satellite-to-ground quantum key distribution *Nature* **549** 43–47
[4] Schimpf C, Reindl M, Huber D, Lehner B, Silva S F C D, Manna S, Vyvlecka M, Walther P and Rastelli A 2021 Quantum cryptography with highly entangled photons from semiconductor quantum dots *Sci. Adv.* **7** 16
[5] 2022 40 years of quantum computing *Nat. Rev. Phys.* **4** 1–1
[6] Deutsch I H 2020 Harnessing the power of the second quantum revolution *PRX Quantum* **1** 020101
[7] Zhong H-S *et al* 2020 Quantum computational advantage using photons *Science* **370** 1460–3
[8] Lemos G B, Borish V, Cole G D, Ramelow S, Lapkiewicz R and Zeilinger A 2014 Quantum imaging with undetected photons *Nature* **512** 409–12

[9] Tenne R, Rossman U, Rephael B, Israel Y, Krupinski-Ptaszek A, Lapkiewicz R, Silberberg Y and Oron D 2018 Super-resolution enhancement by quantum image scanning microscopy *Nat. Photon.* **13** 116–22

[10] Casacio C A, Madsen L S, Terrasson A, Waleed M, Barnscheidt K, Hage B, Taylor M A and Bowen W P 2021 Quantum-enhanced nonlinear microscopy *Nature* **594** 201–6

[11] Segawa T F and Igarashi R 2023 Nanoscale quantum sensing with nitrogen-vacancy centers in nanodiamonds–a magnetic resonance perspective *Prog. Nucl. Magn. Reson. Spectrosc.* **134–135** 20–38

[12] Tong Y, Zhang W, Qin X, Xie Y, Rong X and Du J 2023 A customized control and readout device for vector magnetometers based on nitrogen-vacancy centers *Rev. Sci. Instrum.* **94** 014709

[13] Ménoret V, Vermeulen P, Moigne N L, Bonvalot S, Bouyer P, Landragin A and Desruelle B 2018 Gravity measurements below 10-9 g with a transportable absolute quantum gravimeter *Sci. Rep.* **8** 12300

[14] Stray B *et al* 2022 Quantum sensing for gravity cartography *Nature* **602** 590–4

[15] Kimble H J, Dagenais M and Mandel L 1977 Photon antibunching in resonance fluorescence *Phys. Rev. Lett.* **39** 691–5

[16] Kurtsiefer C, Mayer S, Zarda P and Weinfurter H 2000 Stable solid-state source of single photons *Phys. Rev. Lett.* **85** 290–3

[17] Zwiller V, Blom H, Jonsson P, Panev N, Jeppesen S, Tsegaye T, Goobar E, Pistol M-E, Samuelson L and Björk G 2001 Single quantum dots emit single photons at a time: antibunching experiments *Appl. Phys. Lett.* **78** 2476–8

[18] Buckley S, Rivoire K and Vučković J 2012 Engineered quantum dot single-photon sources *Rep. Prog. Phys.* **75** 126503

[19] Kuhn A, Hennrich M and Rempe G 2002 Deterministic single-photon source for distributed quantum networking *Phys. Rev. Lett.* **89** 067901

[20] Blinov B B, Moehring D L, Duan L-M and Monroe C 2004 Observation of entanglement between a single trapped atom and a single photon *Nature* **428** 153–7

[21] Chandrasekaran V, Tessier M D, Dupont D, Geiregat P, Hens Z and Brainis E 2017 Nearly blinking-free, high-purity single-photon emission by colloidal InP/ZnSe quantum dots *Nano Lett.* **17** 6104–9

[22] Zhu C *et al* 2022 Room-temperature, highly pure single-photon sources from all-inorganic lead halide perovskite quantum dots *Nano Lett.* **22** 3751–60

[23] Yuan Z, Kardynal B E, Stevenson R M, Shields A J, Lobo C J, Cooper K, Beattie N S, Ritchie D A and Pepper M 2002 Electrically driven single-photon source *Science* **295** 102–5

[24] Michler P (ed) 2017 *Quantum Dots for Quantum Information Technologies* (Springer)

[25] Doherty M W, Manson N B, Delaney P, Jelezko F, Wrachtrup J and Hollenberg L C 2013 The nitrogen-vacancy colour centre in diamond *Phys. Rep.* **528** 1–45

[26] Hensen B *et al* 2015 Loophole-free Bell inequality violation using electron spins separated by 1.3 kilometres *Nature* **526** 682–6

[27] Bathen M E and Vines L 2021 Manipulating single-photon emission from point defects in diamond and silicon carbide *Adv. Quantum Technol.* **4** 2100003

[28] Sajid A, Ford M J and Reimers J R 2020 Single-photon emitters in hexagonal boron nitride: a review of progress *Rep. Prog. Phys.* **83** 044501

[29] He Y-M *et al* 2015 Single quantum emitters in monolayer semiconductors *Nat. Nanotechnol.* **10** 497–502

[30] Gao T, von Helversen M, Antón-Solanas C, Schneider C and Heindel T 2023 Atomically-thin single-photon sources for quantum communication *npj 2D Mater. Appl.* **7** 4

[31] Brown R H and Twiss R Q 1956 Correlation between photons in two coherent beams of light *Nature* **177** 27–29

[32] Waks E, Santori C and Yamamoto Y 2002 Security aspects of quantum key distribution withsub-Poisson light *Phys. Rev.* A **66** 042315

[33] Miyazawa T *et al* 2016 Single-photon emission at 1.5 μm from an InAs/InP quantum dot with highly suppressed multi-photon emission probabilities *Appl. Phys. Lett.* **109** 132106

[34] Hanschke L, Fischer K A, Appel S, Lukin D, Wierzbowski J, Sun S, Trivedi R, Vučković J, Finley J J and Müller K 2018 Quantum dot single-photon sources with ultra-low multi-photon probability *npj Quantum Inform.* **4** 43

[35] Schweickert L *et al* 2018 On-demand generation of background-free single photons from a solid-state source *Appl. Phys. Lett.* **112** 093106

[36] James D F V, Kwiat P G, Munro W J and White A G 2001 Measurement of qubits *Phys. Rev.* A **64** 052312

[37] Huber D, Reindl M, Huo Y, Huang H, Wildmann J S, Schmidt O G, Rastelli A and Trotta R 2017 Highly indistinguishable and strongly entangled photons from symmetric GaAs quantum dots *Nat. Commun.* **8** 15506

[38] Cortes C L, Adhikari S, Ma X and Gray S K 2020 Accelerating quantum optics experiments with statistical learning *Appl. Phys. Lett.* **116** 184003

[39] Musiał A *et al* 2020 Plug & play fiber-coupled 73 kHz single-photon source operating in the telecom O-band *Adv. Quantum Technol.* **3** 2000018

[40] Harshvardhan G M, Gourisaria M K, Pandey M and Rautaray S S 2020 A comprehensive survey and analysis of generative models in machine learning *Comput. Sci. Rev.* **38** 100285

[41] Maharana K, Mondal S and Nemade B 2022 A review: data pre-processing and data augmentation techniques *Glob. Transit. Proc.* **3** 91–99

[42] Jablonka K M, Ongari D, Moosavi S M and Smit B 2020 Big-data science in porous materials: materials genomics and machine learning *Chem. Rev.* **120** 8066–129

[43] Tao H, Wu T, Aldeghi M, Wu T C, Aspuru-Guzik A and Kumacheva E 2021 Nanoparticle synthesis assisted by machine learning *Nat. Rev. Mater.* **6** 701–16

[44] Srocka N *et al* 2018 Enhanced photon-extraction efficiency from InGaAs/GaAs quantum dots in deterministic photonic structures at 1.3 μm fabricated by *in-situ* electron-beam lithography *AIP Adv.* **8** 085205

[45] Żołnacz K *et al* 2019 Method for direct coupling of a semiconductor quantum dot to an optical fiber for single-photon source applications *Opt. Express* **27** 26772

[46] Newville M, Stensitzki T, Allen D B and Ingargiola A 2014 LMFIT: Non-linear least-square minimization and curve-fitting for Python (https://doi.org/10.5281/zenodo.11813)

[47] Baronnier J, Mahler B, Boisron O, Dujardin C, Kulzer F and Houel J 2021 Optical properties of fully inorganic core/gradient-shell CdSe/CdZnS nanocrystals at the ensemble and single-nanocrystal levels *Phys. Chem. Chem. Phys.* **23** 22750–9

[48] Allemand A, Kulzer F, Mahler B, Dujardin C and Houel J 2022 Optical properties of individual CdS/CdSe/CdS nanocrystals: spherical quantum wells as single-photon sources *Nanotechnology* **33** 275703