Correlation Avoidance in Single-Photon Detecting Quantum Random Number Generators by Dead Time Overestimation

Balázs Solymos[1*], Ágoston Schranz[1,2] and Miklós Telek[1,2]

[1]Department of Networked Systems and Services, Faculty of Electrical Engineering and Informatics, Budapest University of Technology and Economics, Műegyetem rkp. 3., Budapest, H-1111, Hungary.

[2]HUNREN-BME Information Systems Research Group, Budapest, Hungary.

*Corresponding author(s). E-mail(s): solymosb@hit.bme.hu;
Contributing authors: aschranz@hit.bme.hu; telek@hit.bme.hu;

Abstract

In the case of quantum random number generators based on single photon arrivals, the physical properties of single-photon detectors, such as time-tagger clocks and dead time, influence the stochastic properties of the generated random numbers. This can lead to unwanted correlations among consecutive samples.

We present a method based on extending the insensitive periods after photon detections. This method eliminates the unwanted stochastic effects at the cost of reduced generation speed. We calculate performance measures for our presented method and verify its correctness with computer simulations and measurements conducted on an experimental setup. Our algorithm has low complexity, making it convenient to implement in QRNG schemes, where the benefits of having uncorrelated output intervals exceed the disadvantages of the decreased rate.

Keywords: quantum random number generation, lasers, single-photon detection, probability

1 Introduction

Provably secure randomness is an essential resource for many applications like Monte Carlo simulations or the cryptographic protocols of the present [1] and even the quantum cryptographic protocols of the future [2]. Conventional pseudorandom number generators are based on complex but deterministic algorithms, unavoidably leading to some undesirable deterministic features in the long run. In contrast, quantum random number generators (QRNGs) [3, 4] exploit the inherent unpredictability of quantum mechanical phenomena to provide a provably secure entropy source. Optical QRNG schemes make use of the quantum nature of light, leading to many possible architectures, such as generators based on the superposition of single-photon paths [5, 6], photon number counting [7, 8], photon arrival times [9–11], quantum phase fluctuations [12], amplified spontaneous emission [13], or even Raman scattering [14].

Using the arrival time of photons is an attractive choice due to the simplicity of the required hardware. The source of randomness in these generators is the light emission process, whose weak optical signal is detected by a single-photon detector. Bits are then generated from the measured arrival times of the individual photons. Ideally, the measured raw data samples should be independent and come from a well-defined, known distribution. However, in a real-world scenario, there are various imperfections we also have to deal with. The finite precision of time measurement introduces unwanted correlations [15], which can be remedied by restarting the time-tagger clock at each detection [9, 16] at the cost of more complicated hardware. Another major factor is the dead time of photon detectors [17], further changing the measured interval distribution.

In this work, we introduce a method to deal with the effect of non-restartable time-tagger clocks and detector dead time simultaneously, at the cost of reduced bit generation speed. Compared to the standard practice of reducing input rates to limit the unwanted correlations due to these effects, our proposed method also allows generator operation in regimes with higher input rates, thus facilitating improved output performance regarding the bit generation rate. The paper is organized as follows: Section 2 describes the basic operation principle of time-of-arrival generators and contains a brief analysis of the measured interval distributions in the non-ideal cases. We introduce our method in Section 3 and evaluate its performance in Section 4. Measurement data presented in Section 5 supports the validity of our method. Finally, Section 6 concludes the paper.

2 Principle of QRNG operation

A whole family of QRNGs operates based on the following concept: a single-photon detector (SPD) detects photons emitted by a suitably attenuated continuous-wave (CW) laser, and a time-tagger card (time-to-digital converter, TDC) assigns time stamps to detections based on its continuously running internal clock signal. We assume the photons to arrive according to a homogeneous Poisson point process (PPP) with rate λ, valid for coherent light sources [18]. We refer to λ as the *input photon rate* of our detection system; it is proportional to the optical power and its value already includes the losses from the $\eta_{\mathrm{d}} < 100\%$ detection efficiency of the SPD. Let S_i denote the ith photon arrival time, and $T_i = S_i - S_{i-1}$ the exponentially distributed time elapsed between S_i and S_{i-1}, where S_0 is the starting time of the measurement. These times are physically measured by counting the clock signal's leading edges between S_i and S_{i-1}, yielding integer values. These integers are the *discretized time differences* (DTDs), discrete random variables denoted by D_i. DTDs undergo well-defined mathematical operations based on the applied random bit generation scheme (e.g., [9]), outputting random bits, which form uniformly distributed, uncorrelated sequences in the ideal case. Such generators are commonly referred to as time-of-arrival (ToA) QRNGs. Our method offers a tool for correlation avoidance of the DTDs that can be used with all such devices; independent of the concrete bit generation algorithm.

Let us denote the time-tagger's resolution—the clock signal's period—by τ. There is a non-zero γ_i time between S_i and the previous leading clock edge, that is, $\gamma_i = S_i - \lfloor S_i/\tau \rfloor \tau$, where $\lfloor \cdot \rfloor$ denotes the floor function, representing the greatest integer less than or equal to its argument. Consequently, $\gamma_i \in [0, \tau)$. We call the random variable γ_i the *phase* of the ith photon detection.

It has been previously known that non-zero phases introduce correlations between the DTDs and, correspondingly, between the random bits generated [9]. In our previous work [15], we have derived a detailed stochastic model of a particular ToA bit generation method, quantitatively analyzing the effects of these phases. We have shown that by increasing the product of the input photon rate of the SPD and the timing resolution ($\lambda\tau$), the correlation coefficients between bits deviate from zero, while the bit-pair and other bit-tuple probabilities deviate from the uniform values. On the other hand, keeping $\lambda\tau$ close to zero severely limits the achievable bit generation rates.

2.1 Distribution and correlation of the observed variables

Bit generation schemes are based on the D_i DTDs since they are the physical observables measured in the setup. According to Ref. [15], focusing only on the first arrival, we can write the following for the distribution of these variables and the corresponding phases, for $x, y \in [0, \tau)$:

$$
\begin{aligned}
F_n(x,y) &\triangleq \Pr(D_1 = n, \gamma_1 < y \mid \gamma_0 = x) \\
&= \begin{cases} \Pr(x + T_1 < y) & \text{if } n = 0, \\ \Pr(n\tau \leq x + T_1 < n\tau + y) & \text{if } n > 0, \end{cases} \\
&= \begin{cases} \chi_{\{y>x\}} \left(1 - \mathrm{e}^{-\lambda(y-x)}\right) & \text{if } n = 0, \\ \mathrm{e}^{\lambda x} \left(1 - \mathrm{e}^{-\lambda y}\right) \mathrm{e}^{-\lambda n \tau} & \text{if } n > 0, \end{cases}
\end{aligned}
\tag{1}
$$

and

$$
f_n(x,y) \triangleq \frac{\mathrm{d}}{\mathrm{d}y} \Pr(D_1 = n, \gamma_1 < y \mid \gamma_0 = x) = \begin{cases} \chi_{\{y>x\}} \lambda \mathrm{e}^{-\lambda(y-x)} & \text{if } n = 0, \\ \lambda \mathrm{e}^{-\lambda(y+n\tau-x)} & \text{if } n > 0, \end{cases} \tag{2}
$$

where χ_A is the indicator of the set A.[1] We note that if $\gamma_0 = 0$ then $F_n(0,\tau) = \Pr(D_1 = n \mid \gamma_0 = 0) = \left(1 - \mathrm{e}^{-\lambda\tau}\right) \mathrm{e}^{-\lambda\tau n}$ results in a geometric distribution [16], retaining the memoryless property of the underlying exponential distribution. This means that successive DTDs, D_i and D_{i+1}, would be uncorrelated after eliminating the effects of non-zero phases.

The conditional and unconditional joint distributions of successive DTDs $D_1, \ldots, D_N$, i.e.,

$$
\Pr(D_1 = n_1, \ldots, D_N = n_N \mid \gamma_0 = x) \quad \text{and} \quad \Pr(D_1 = n_1, \ldots, D_N = n_N),
$$

can also be calculated based on (2). The joint distributions indicate that the $D_1, \ldots, D_N$ variables are correlated [15]. Thus, using the $D_1, \ldots, D_N$ sequence for random bit generation might result in correlated bit sequences.

In Ref. [15], we only focused on the correlations between the random bits generated from the physical process but skipped the numerical analysis of correlations between DTDs. To derive the correlation between successive samples, D_i and D_{i+1}—which is equivalent to the lag-1 autocorrelation coefficient in DTD sequences—, we refer back to our previous work, where we have shown that if the first phase of the process, γ_0, is uniformly distributed between 0 and τ, then every other γ_i has a uniform marginal distribution (Ref. [15], Theorem 1).

[1] Here we have used the fact that the T_i times elapsed between events of the PPP are exponentially distributed, with a cumulative distribution function $F_T(t) = \Pr(T < t) = \chi_{\{t \geq 0\}}(1 - \mathrm{e}^{-\lambda t})$.

Without loss of generality, set $i = 1$ and $i+1 = 2$ and compute the correlation ρ_{D_1,D_2} based on

$$\rho_{D_1,D_2} = \frac{\mathbb{E}(D_1 D_2) - \mathbb{E}(D_1)\,\mathbb{E}(D_2)}{\sqrt{\left(\mathbb{E}\left(D_1^2\right) - \mathbb{E}(D_1)^2\right)\left(\mathbb{E}\left(D_2^2\right) - \mathbb{E}(D_2)^2\right)}}. \tag{3}$$

According to (2), for $n_1 > 0$ and $n_2 > 0$, we have

$$\begin{aligned}\Pr(D_2 = n_2, D_1 = n_1 \mid \gamma_0 = x_0) &= \int_{x_2=0}^{\tau}\int_{x_1=0}^{\tau} f_{n_2}(x_2,x_1)\cdot f_{n_1}(x_1,x_0)\,\mathrm{d}x_1\mathrm{d}x_2\\ &= \int_{x_2=0}^{\tau}\int_{x_1=0}^{\tau} \lambda \mathrm{e}^{-\lambda(x_2+n_2\tau-x_1)}\lambda \mathrm{e}^{-\lambda(x_1+n_1\tau-x_0)}\mathrm{d}x_1\mathrm{d}x_2\\ &= \lambda\tau\left(1-\mathrm{e}^{-\lambda\tau}\right)\mathrm{e}^{-\lambda(n_1\tau+n_2\tau-x_0)}.\end{aligned} \tag{4}$$

Furthermore, using the uniform distribution of γ_0, the expectation of the product D_1D_2 becomes

$$\begin{aligned}\mathbb{E}(D_1D_2) &= \int_0^{\tau}\frac{1}{\tau}\,\mathbb{E}(D_1D_2 \mid \gamma_0 = x)\mathrm{d}x\\ &= \int_0^{\tau}\frac{1}{\tau}\sum_{i=1}^{\infty}\sum_{j=1}^{\infty} ij\Pr(D_2 = i, D_1 = j \mid \gamma_0 = x)\mathrm{d}x = \frac{\mathrm{e}^{-\lambda\tau}}{\left(1-\mathrm{e}^{-\lambda\tau}\right)^2}.\end{aligned} \tag{5}$$

The DTDs' expected values $\mathbb{E}(D_1) = \mathbb{E}(D_2)$ and second moments $\mathbb{E}\left(D_1^2\right) = \mathbb{E}\left(D_2^2\right)$ can be calculated using Ref. [15, Eq. (12)], yielding

$$\mathbb{E}(D_1) = \mathbb{E}(D_2) = \sum_{n=1}^{\infty} n\cdot\Pr(D_1 = n) = \frac{\left(1-\mathrm{e}^{-\lambda\tau}\right)^2}{\lambda\tau \mathrm{e}^{-\lambda\tau}}\sum_{n=1}^{\infty} n\cdot \mathrm{e}^{-\lambda\tau n} = \frac{1}{\lambda\tau} \tag{6}$$

and

$$\mathbb{E}\left(D_1^2\right) = \mathbb{E}\left(D_2^2\right) = \sum_{n=1}^{\infty} n^2\cdot\Pr(D_1 = n) = \frac{\left(1-\mathrm{e}^{-\lambda\tau}\right)^2}{\lambda\tau \mathrm{e}^{-\lambda\tau}}\sum_{n=1}^{\infty} n^2\cdot \mathrm{e}^{-\lambda\tau n} = \frac{\left(1+\mathrm{e}^{-\lambda\tau}\right)}{\lambda\tau\left(1-\mathrm{e}^{-\lambda\tau}\right)}. \tag{7}$$

Finally, the correlation between D_1 and D_2, purely a function of the product $\lambda\tau$, is

$$\rho_{D_1,D_2} = \frac{\frac{\mathrm{e}^{-\lambda\tau}}{\left(1-\mathrm{e}^{-\lambda\tau}\right)^2} - \frac{1}{(\lambda\tau)^2}}{\frac{\left(1+\mathrm{e}^{-\lambda\tau}\right)}{\lambda\tau\left(1-\mathrm{e}^{-\lambda\tau}\right)} - \frac{1}{(\lambda\tau)^2}} = \frac{(\lambda\tau)^2\mathrm{e}^{-\lambda\tau} - \left(1-\mathrm{e}^{-\lambda\tau}\right)^2}{\lambda\tau\left(1-\mathrm{e}^{-2\lambda\tau}\right) - \left(1-\mathrm{e}^{-\lambda\tau}\right)^2}. \tag{8}$$

The correlation tends to zero as $(\lambda\tau) \to 0$ or $(\lambda\tau) \to \infty$, its value is negative in between (see Fig. 1). It is monotonically decreasing until obtaining its minimum of -0.2233 around $\lambda\tau = 3.5749$. Thus, increasing $\lambda\tau$ from zero increases the magnitude of correlations between successive DTDs,[2] and the resulting sequence of random variables will always contain systematic correlations. Although the standard practice of reducing the optical power (limiting $\lambda\tau$) is a valid approach to decrease correlations, it also severely limits the capabilities of the QRNGs. For example, only allowing $|\rho_{D_1,D_2}| < 10^{-4}$ means that $\lambda\tau$ has an upper bound of 0.0346, which can limit certain architectures in terms of bit generation rates [19, Sec. 3.3]. Therefore, finding a different way of eliminating correlations whilst allowing higher $\lambda\tau$ values can prove beneficial.

2.2 Dead time

An additional limitation is imposed by the inability of physical devices to observe all successive photon arrivals. Detectors usually have a dead time, an insensitive time interval of length ζ after a detected photon arrival, during which they cannot register any new arrivals. This means that after a photon detection at S_i, no photons arriving before $S_i + \zeta$ are recognized. Consequently, for the observed photon arrivals $S_i > S_{i-1} + \zeta$ holds for $\forall i > 0$. Our model assumes that photon arrivals during the dead time interval are undetected, and such arrivals do not reset the dead time.

Similarly to the previous case free of dead time, we can compute the distribution of the DTDs $D_1, \ldots, D_N$ as follows. Assume that $\zeta = k\tau + \delta$ is constant with $k \in \mathbb{N}$ and $0 \leq \delta < \tau$, meaning that $\Pr(D_1 < k) = 0$. Then, for $n \geq k$, the conditional distribution is [15]

$$\begin{aligned} F_n(x,y) &= \Pr(D_1 = n, \gamma_1 < y \mid \gamma_0 = x) \\ &= \begin{cases} \Pr(x + T_1 + \delta < y) & \text{if } n = k, \\ \Pr\left((n-k)\tau \leq x + T_1 + \delta < (n-k)\tau + y\right) & \text{if } n > k, \end{cases} \\ &= \begin{cases} \chi_{\{x+\delta<y\}}\left(1 - \mathrm{e}^{-\lambda(y-x-\delta)}\right) & \text{if } n = k, \\ \chi_{\{\tau<x+\delta<\tau+y\}}\left(1 - \mathrm{e}^{-\lambda(y-x-\delta+\tau)}\right) & \\ \quad + \chi_{\{x+\delta<\tau\}}\mathrm{e}^{-\lambda(\tau-x-\delta)}\left(1 - \mathrm{e}^{-\lambda y}\right) & \text{if } n = k+1, \\ \left(\mathrm{e}^{-\lambda((n-k)\tau-x-\delta)}\right)\left(1 - \mathrm{e}^{-\lambda y}\right) & \text{if } n > k+1, \end{cases} \end{aligned} \tag{9}$$

[2]This statement is valid until the global minimum is reached at $\lambda\tau = 3.5749$; however, values of $\lambda\tau > 1$ are impractical. They represent a domain in which, on average, more than one photon arrives within a clock period. This practically means a good-quality SPD with high photon rate tolerance connected to a low-resolution TDC. This domain is irrelevant in the present discussion.

and for $n \geq k$, the conditional density is

$$f_n(x,y) = \frac{\mathrm{d}}{\mathrm{d}y} F_n(x,y) = \begin{cases} \chi_{\{x+\delta<y\}} \lambda \mathrm{e}^{-\lambda(y-x-\delta)} & \text{if } n = k, \\ \chi_{\{x+\delta<\tau\}} \lambda \mathrm{e}^{-\lambda(y-x-\delta+\tau)} + \chi_{\{\tau<x+\delta<\tau+y\}} \lambda \mathrm{e}^{-\lambda(y-x-\delta+\tau)} & \text{if } n = k+1, \\ \lambda \mathrm{e}^{-\lambda(y+(n-k)\tau-\delta-x)} & \text{if } n > k+1. \end{cases} \tag{10}$$

Along the lines of the dead time free case, we compute the distribution of D_1 and the joint distribution of D_1 and D_2 from (10), utilizing the uniform distribution of γ_0, as

$$p_{n_1} \triangleq \Pr(D_1 = n_1) = \frac{1}{\tau} \int_{x_0=0}^{\tau} \int_{x_1=0}^{\tau} f_{n_1}(x_1, x_0) \, \mathrm{d}x_1 \mathrm{d}x_0, \tag{11}$$

$$p_{n_1,n_2} \triangleq \Pr(D_2 = n_2, D_1 = n_1) = \frac{1}{\tau} \int_{x_0=0}^{\tau} \int_{x_1=0}^{\tau} \int_{x_2=0}^{\tau} f_{n_2}(x_2, x_1) \cdot f_{n_1}(x_1, x_0) \, \mathrm{d}x_2 \mathrm{d}x_1 \mathrm{d}x_0. \tag{12}$$

The distributions allow us to calculate the expected values $\mathbb{E}(D_1 - k)$, $\mathbb{E}\left((D_1 - k)^2\right)$ and $\mathbb{E}((D_1 - k)(D_2 - k))$, along with the correlation $\rho_{D_1,D_2} = \rho_{D_1-k,D_2-k}$:

$$\mathbb{E}(D_1 - k) = \sum_{n_1=1}^{\infty} n_1 p_{n_1} = \frac{1+\lambda\delta}{\lambda\tau}, \tag{13}$$

$$\mathbb{E}\left((D_1 - k)^2\right) = \sum_{n_1=1}^{\infty} n_1^2 p_{n_1} = \frac{1 + \lambda\delta + \mathrm{e}^{-\lambda\tau}(2\mathrm{e}^{\lambda\delta} - 1 - \lambda\delta)}{\lambda\tau\left(1 - \mathrm{e}^{-\lambda\tau}\right)}, \tag{14}$$

$$\mathbb{E}((D_1 - k)(D_2 - k)) = \sum_{n_1=1}^{\infty} n_1 \sum_{n_2=1}^{\infty} n_2 p_{n_1,n_2}, \tag{15}$$

$$\rho_{D_1,D_2} = \mathrm{corr}(D_1, D_2) = \frac{\mathbb{E}((D_1 - k)(D_2 - k)) - \mathbb{E}^2(D_1 - k)}{\mathbb{E}\left((D_1 - k)^2\right) - \mathbb{E}^2(D_1 - k)}, \tag{16}$$

where we provided closed-form expressions for the former two and computed the latter two numerically.

Figure 1 depicts the correlation of consecutive DTDs as a function of the photon arrival rate for selected values of the dead time. We note that the correlation is independent of the integer part of the dead time, k, and only its fractional part, δ, affects the values. The figure verifies that the correlation tends to zero as the photon arrival rate decreases to zero, but for higher photon arrival rates the correlation strongly depends on the dead time.

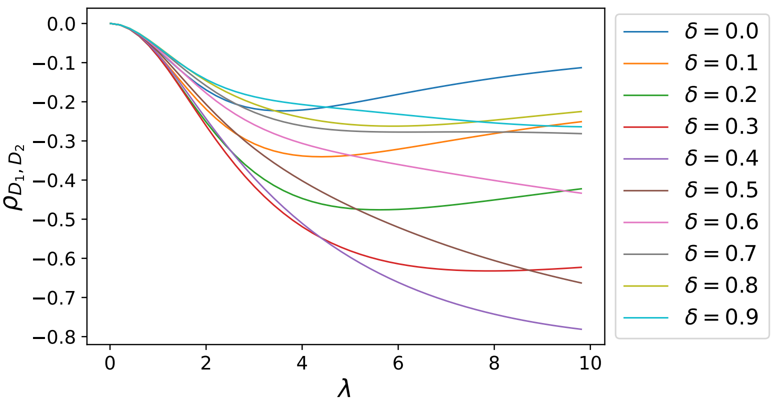

Fig. 1 Correlation of consecutive DTDs as a function of the input photon rate λ and fractional dead time δ, with $\tau = 1$.

Note that the presence of dead time reduces the measured rate of photon detections. When $S_i > S_{i-1} + \zeta$, the mean time between photon observations is

$$\mathbb{E}(S_i - S_{i-1}) = \mathbb{E}(T_i) = \frac{1}{\lambda} + \zeta = \frac{1 + \lambda\zeta}{\lambda}. \tag{17}$$

As a consequence, the average rate at which the D_i samples are obtained is

$$\begin{aligned} \lambda_{\mathrm{d}} &= \lim_{c\to\infty} \frac{\text{observed photon arrivals in } [0, c\tau)}{c\tau} \\ &= \frac{1}{\mathbb{E}(S_i - S_{i-1})} = \frac{1}{\mathbb{E}(T_i)} = \frac{\lambda}{1+\lambda\zeta}. \end{aligned} \tag{18}$$

3 Dead time overestimation

To eliminate the correlation between successive D_i values, we introduce an approach called the *overestimation* of dead time. The approach is based on the following observation. The conditional distribution in (9) is such that for $n > k+1$ the conditional characteristic function

$$\begin{aligned} \bar{F}_n(x,y) &= \Pr(D_1 = n, \gamma_1 < y \mid \gamma_0 = x, D_1 > k+1) \\ &= \frac{\Pr(D_1 = n, \gamma_1 < y \mid \gamma_0 = x)}{\sum_{j=k+2}^{\infty} \Pr(D_1 = j \mid \gamma_0 = x)} \\ &= \frac{\left(\mathrm{e}^{-\lambda((n-k)\tau - x - \delta)}\right)\left(1 - \mathrm{e}^{-\lambda y}\right)}{\sum_{j=k+2}^{\infty} \left(\mathrm{e}^{-\lambda((j-k)\tau - x - \delta)}\right)\left(1 - \mathrm{e}^{-\lambda\tau}\right)} \\ &= \mathrm{e}^{-(n-(k+2))\lambda\tau}\left(1 - \mathrm{e}^{-\lambda y}\right) \end{aligned} \tag{19}$$

is independent of x and δ, and satisfies

$$\begin{aligned}&\Pr(D_1 = n, \gamma_1 < y \mid \gamma_0 = x, D_1 > k+1)\\ &= \underbrace{\Pr(D_1 = n \mid \gamma_0 = x, D_1 > k+1)}_{e^{-(n-(k+2))\lambda\tau}\left(1-e^{-\lambda\tau}\right)} \cdot \underbrace{\Pr(\gamma_1 < y \mid \gamma_0 = x, D_1 > k+1)}_{\frac{1-e^{-\lambda y}}{1-e^{-\lambda\tau}}},\end{aligned} \tag{20}$$

that is, D_1 and γ_1 are independent when $D_1 > k+1$. This also means that D_2, which depends on γ_1, will be independent of D_1 as long as $D_1 > k+1$.

Thus, the correlation of the consecutive D_i values comes from the small samples; i.e., when $D_i = k$ or $D_i = k+1$, then D_i and D_{i+1} are correlated. We can exploit this property in the overestimation algorithm to avoid unwanted correlations.

In the following sections, unless the unit of time is specified explicitly, we assume τ and ζ to have arbitrary, unspecified time units, whilst λ is measured in [counts]/[unit of time].

3.1 Overestimation method

Let us overestimate the dead time with an interval covering m clock cycles, where $m \in \mathbb{Z}^+$ such that $\zeta = k\tau + \delta \leq m\tau$. We refer to m as the overestimation parameter. After a detection event, we start an $m\tau$ long safety interval from the next rising clock edge. If a photon is detected after the dead time is over but before this safety interval has ended, we discard the detection event from any further calculations and extend the safety interval by $m\tau$, counted from the following rising edge.

Suppose the safety interval is eventually over because no early detection extends it further. In this case, we continue using our bit generation method as if the previous detection happened at the end of the safety interval. That is, we count the next time difference between the end of the safety interval and the next detection time, then digitize it. See an example in Fig. 2. This approach can be thought of as an algorithm taking the $\mathbb{D} = \{D_1, D_2, \ldots,\}$ DTDs as input and outputting the $\mathbb{V} = \{V_1, V_2, \ldots,\}$ *virtual DTDs* (vDTDs). The algorithm (described in Algorithm 1) has the added benefit of placing the starting points of measurable intervals right to the beginning of a clock cycle, essentially realizing the ideal case of $\gamma_{i-1} = 0$, yielding geometrically distributed vDTDs.

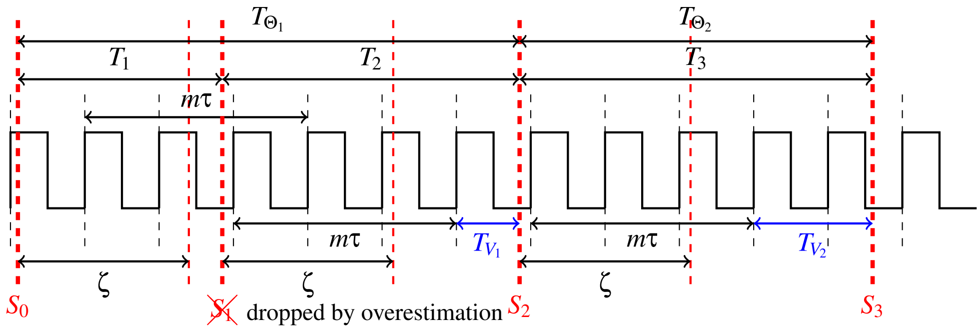

Fig. 2 Example of the overestimation method with overestimation parameter m and dead time ζ ($m = 3, \zeta = 2.3, \tau = 1$). The square signal represents the measurement clock. Thick red dashed lines at S_0, S_1, S_2, and S_3 denote actual photon detection times, and lighter red lines show the end of the corresponding dead times. T_1, T_2, T_3 are the intervals responsible for the $D_1 = 2$, $D_2 = 4$, $D_3 = 5$ DTDs without overestimation. The photon detected at S_1 arrives before the safety interval is over, which is therefore dropped by the overestimation algorithm. T_{V_1} and T_{V_2} note the resulting virtual intervals considered in our method, responsible for $V_1 = 0$, $V_2 = 1$ virtual DTDs, while T_{Θ_1} and T_{Θ_2} are the intervals responsible for $\Theta_1 = 6$ and $\Theta_2 = 5$, with $\beta_1 = \{2,4\}$ and $\beta_2 = \{5\}$ respectively. (For the notation β_ℓ, Θ_ℓ, T_{Θ_ℓ}, refer to Sec. 3.2.)

Algorithm 1 Algorithm of the overestimation method

Require: m ▷ Overestimation parameter

while True **do**
  get D ▷ Obtain last DTD at a new detection
  if $D > m$ **then** ▷ Check if safety interval is over
    $V \leftarrow D - (m+1)$ ▷ Generate V virtual DTD
  end if
end while

Let $\mathbb{S} = \{S_0, S_1, \ldots,\}$ be the observed photon arrival times with dead time ζ (that is, $\forall i$: $S_i > S_{i-1} + \zeta$) and $\mathbb{D} = \{D_1, D_2, \ldots,\}$ be the sequence of measured DTDs associated with $\mathbb{S}$. Let $\mathbb{V} = \{V_1, V_2, \ldots,\}$ be the virtual DTD sequence generated by Algorithm 1 from $\mathbb{D}$.

Theorem 1. *The virtual DTD sequence generated by Algorithm 1, $\mathbb{V}$, is composed of i.i.d. elements with geometric distribution:* $\Pr(V_\ell = n) = (1 - \mathrm{e}^{-\lambda\tau})\mathrm{e}^{-\lambda\tau n}$.

Proof. For the distribution of DTDs D_i greater than m, we can write

$$\begin{aligned}
&\Pr(D_i = n \mid \gamma_{i-1} = x_{i-1}, D_i > m) \\
&= \frac{\left(1-\mathrm{e}^{-\lambda\tau}\right)\mathrm{e}^{-\lambda((n-k)\tau - x_{i-1} - \delta)}}{\sum_{j=m+1}^{\infty}\left(1-\mathrm{e}^{-\lambda\tau}\right)\mathrm{e}^{-\lambda((j-k)\tau - x_{i-1} - \delta)}} = \frac{\mathrm{e}^{-\lambda n\tau}\left(1-\mathrm{e}^{-\lambda\tau}\right)}{\mathrm{e}^{-\lambda\tau(m+1)}} \\
&= \left(1-\mathrm{e}^{-\lambda\tau}\right)\mathrm{e}^{-\lambda(n-(m+1))\tau},
\end{aligned} \tag{21}$$

where γ_{i-1} is the arrival phase of S_{i-1}. Using the $V \leftarrow D-(m+1)$ assignment rule in line 4 of Algorithm 1, we have

$$\begin{aligned}
&\Pr(V_\ell = n \mid \gamma_{i-1} = x_{i-1}) \\
&= \Pr(D_i = (m+1)+n \mid \gamma_{i-1} = x_{i-1}, D_i > m) \\
&= \left(1-\mathrm{e}^{-\lambda\tau}\right)\mathrm{e}^{-\lambda(n+m+1-(m+1))\tau} = \left(1-\mathrm{e}^{-\lambda\tau}\right)\mathrm{e}^{-\lambda n\tau}
\end{aligned} \tag{22}$$

for the distribution of the V_ℓ variable, which is independent of the phase γ_{i-1}. □

Note that without dead time, the choice of $V \leftarrow D-1$ assignment rule in line 4 of Algorithm 1 would be sufficient since it removes the first fractional clock period, which is responsible for the correlation of successive samples in this case. Additionally, removing m full-length clock periods does not affect the discrete distribution of samples [16]. Using this scheme comes at a cost, as the time used to overestimate the dead time cannot be used for bit generation, leading to a decreased bit generation rate.

One could reason that we could have the same effect by simply reducing the optical power intensity (the photon rate λ) to a regime where correlations and distortions in the distributions vanish. We argue that our algorithm is a better choice than power reduction, both from a philosophical and a numerical point of view.

First, it is true that by decreasing the optical power, the probability $\Pr(D_i \leq k+1)$ decreases, consequently reducing the number of DTDs causing correlations. However, this probability is never exactly zero—unless λ is set to zero, preventing bit generation. Algorithm 1, on the other hand, removes every problematic DTD, yielding a theoretically correlation-free sequence of virtual DTDs.

Second, reducing the input rate also reduces the available number of measurement samples for bit generation per unit time. Consequently, power reduction limits achievable output bit generation speeds. [3]

[3] The power reduction approach is disadvantageous even in terms of the achievable min-entropy rate, as the maximum of the min-entropy per unit time often lies in a parameter regime corresponding to a higher $\lambda\tau$ product than what the power reduction approach would still allow. See Sec. 4.4 for the discussion about entropy rates.

3.2 Virtual DTD generation rate

For the performance assessment of Algorithm 1, let us define the u-long subsequence of $\mathbb{D}$, $\beta_\ell = \{D_i, D_{i+1}, ..., D_{i+u-1}\}$, responsible for generating the ℓth vDTD, V_ℓ. According to the algorithm, β_ℓ starts with an uninterrupted run of zero or more DTDs smaller than or equal to m, and ends with a single element greater than m ($D_{i-1} > m$ and $D_{i+u-1} > m$, but $D_t \leq m$ $\forall t \in (i, i+u-2)$). Note that the set of all such subsequences, $\{\beta_\ell\}$, is a partition of $\mathbb{D}$, since $\forall i:\ D_i \in \bigcup_\ell \beta_\ell$ and $(D_i \in \beta_x \wedge D_i \in \beta_y) \Rightarrow (\beta_x = \beta_y)$.

The number of elapsed clock signal edges between generating $V_{\ell-1}$ and V_ℓ is $\Theta_\ell = \sum_{k=0}^{u-1} D_{i+k}$, where u is the length of β_ℓ and Θ_ℓ is the sum of β_ℓ's elements.

Similar to λ_d, we define λ_v, the *virtual count rate* at which the vDTDs are generated, as

$$\lambda_\mathrm{v} = \lim_{c\to\infty} \frac{\text{number of vDTDs } V_\ell \text{ generated in } [0, c\tau)}{c\tau}. \tag{23}$$

Theorem 2. *The virtual count rate λ_v can be expressed as*

$$\lambda_v = \frac{\mathrm{e}^{-\lambda((m+1)\tau-\zeta)}\left(\mathrm{e}^{\lambda\tau}-1\right)}{\tau(\lambda\zeta+1)}. \tag{24}$$

Proof. Consider the $\{Z_0, Z_1, \ldots\}$ sequence, where for $i \geq 0$

$$Z_i = \begin{cases} 0 \text{ if } D_i \leq m, \\ 1 \text{ if } D_i > m. \end{cases} \tag{25}$$

The sum $S_N = \sum_{i=0}^{N} Z_i$ then gives the number of vDTDs generated by Algorithm 1 from an original N-long $\{D_1, \ldots, D_N\}$ DTD sequence. We can then write

$$\begin{aligned} &\Pr(Z_i = 1 \mid \gamma_{i-1} = x_{i-1}, D_{i-1} = n_{i-1}) = \Pr(D_i > m \mid \gamma_{i-1} = x_{i-1}, D_{i-1} = n_{i-1}) \\ &= \Pr(D_i > m \mid \gamma_{i-1} = x_{i-1}) = \sum_{n=m+1}^{\infty} \mathrm{e}^{-\lambda(n\tau-\zeta-x_{i-1})}\left(1-\mathrm{e}^{-\lambda\tau}\right), \end{aligned} \tag{26}$$

and

$$\Pr(Z_i = 0 \mid \gamma_{i-1} = x_{i-1}, D_{i-1} = n_{i-1}) = 1 - \Pr(Z_i = 1 \mid \gamma_{i-1} = x_{i-1}, D_{i-1} = n_{i-1}).$$

Consequently, Z_i only depends on γ_{i-1}, in the sense that

$$\Pr(Z_i = 1 \mid \gamma_{i-1} = x_{i-1}) = \Pr(Z_i = 1 \mid \gamma_{i-1} = x_{i-1}, D_{i-1} = n_{i-1}, \ldots, D_1 = n_1, \gamma_0 = x_0).$$

That is, the $\{Z_1,\ldots,Z_N\}$ sequence is dependent on an underlying $\{\gamma_0,\gamma_1,\ldots,\gamma_{N-1}\}$ phase sequence. According to (9), the consecutive γ_i values form a Markov chain, since $\Pr(\gamma_i < x_i \mid \gamma_{i-1} = x_{i-1}) = \Pr(\gamma_i < x_i \mid \gamma_{i-1} = x_{i-1},\ldots,\gamma_0 = x_0)$. The stationary phase distribution satisfies

$$f(y) = \int_{x=0}^{\tau} f(x)g(x,y)\mathrm{d}x, \tag{27}$$

where $g(x,y)$ can be obtained from (10) using that the conditional phase density at the first photon arrival after the dead time is

$$g(x,y) = \frac{\mathrm{d}}{\mathrm{d}y}\Pr(\gamma_1 < y \mid \gamma_0 = x) = \sum_{n=0}^{\infty} f_n(x,y). \tag{28}$$

The solution of (27) is $f(y) = \chi_{\{0\le y<\tau\}}\frac{1}{\tau}$.

Due to the ergodicity of the γ_i Markov chain, as N tends to infinity, the number of samples in the $\{\gamma_0,\gamma_1,\ldots,\gamma_{N-1}\}$ phase sequence which fall into the $(x,x+\Delta)$ interval is proportional to $f(x)\cdot\Delta$.

Using this, the ratio of DTDs longer than m can be written as

$$\begin{aligned}\mathcal{S} &\triangleq \lim_{N\to\infty}\frac{S_N}{N} = \int_{x=0}^{\tau}\frac{1}{\tau}\Pr(D_i > m \mid \gamma_{i-1} = x)\,\mathrm{d}x = \int_{x=0}^{\tau}\frac{1}{\tau}\sum_{n=m+1}^{\infty}\mathrm{e}^{-\lambda(n\tau-\zeta-x)}\left(1-\mathrm{e}^{-\lambda\tau}\right)\mathrm{d}x \\ &= \sum_{n=m+1}^{\infty}\frac{(\mathrm{e}^{\lambda\tau}-1)^2\mathrm{e}^{-\lambda(n+1)\tau-\zeta)}}{\lambda\tau} = \frac{(\mathrm{e}^{\lambda\tau}-1)\mathrm{e}^{-\lambda((m+1)\tau-\zeta)}}{\lambda\tau}.\end{aligned} \tag{29}$$

The expected virtual count rate can then be calculated as

$$\lambda_{\mathrm{v}} = \mathcal{S}\cdot\lambda_{\mathrm{d}} = \frac{(\mathrm{e}^{\lambda\tau}-1)\mathrm{e}^{-\lambda((m+1)\tau-\zeta)}}{\lambda\tau}\cdot\frac{\lambda}{1+\lambda\zeta} = \frac{\mathrm{e}^{-\lambda((m+1)\tau-\zeta)}\left(\mathrm{e}^{\lambda\tau}-1\right)}{\tau(\lambda\zeta+1)}, \tag{30}$$

where λ_{d} is the original rate with dead time, as obtained in (18). □

Let $\Theta = \lim_{\ell\to\infty}\Theta_\ell$ be the stationary number of leading clock edges between generating consecutive V_ℓ values. Theorem 2 defines its mean as $\mathbb{E}(\Theta) = 1/(\lambda_{\mathrm{v}}\tau)$. The expected time for generating a vDTD with Algorithm 1, T_Θ, can then be written as

$$\mathbb{E}(T_\Theta) = \tau\cdot\mathbb{E}(\Theta) = \frac{1}{\lambda_{\mathrm{v}}} = \frac{\tau(\lambda\zeta+1)}{\mathrm{e}^{-\lambda((m+1)\tau-\zeta)}\left(\mathrm{e}^{\lambda\tau}-1\right)}. \tag{31}$$

The vDTD sample generation rate computed according to Theorem 2 is depicted in Fig. 3.

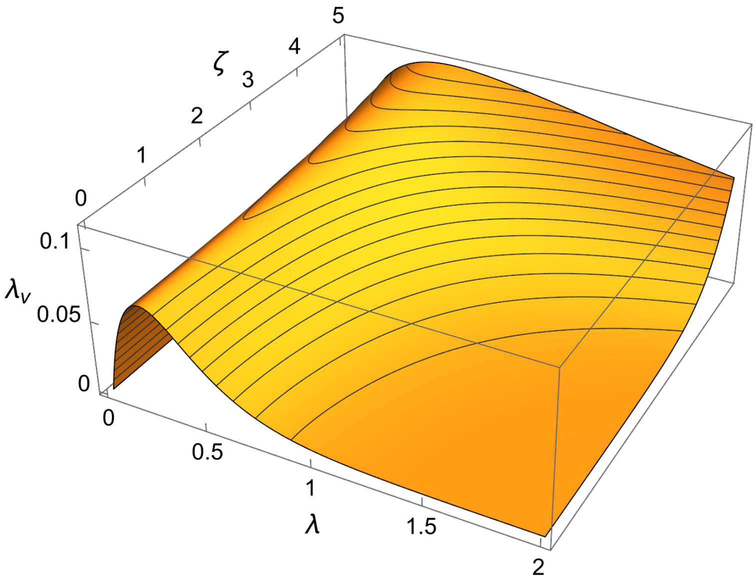

Fig. 3 Virtual count rate, λ_v, as a function of the input photon rate λ and dead time ζ, with $\tau = 1$, $m = 5$.

3.3 Computation of further performance indices

Theorem 2 calculates the mean number of non-discarded detections. The analysis approach of this section allows the computation of more detailed performance indices of Algorithm 1.

To compute the distribution of Θ_1 based on (9), we introduce $\hat{\Theta}(z,x_0) = \mathbb{E}\left(z^{\Theta_1} \mid \gamma_0 = x_0\right)$, the z-transform of Θ_1; $F_\mathrm{d}(z,x_0,x_1) = \sum_{n=0}^{m} z^n f_n(x_0,x_1)$ describing the discarded arrivals; and $F_\mathrm{a}(z,x_0,x_1) = \sum_{n=m+1}^{\infty} z^n f_n(x_0,x_1)$ describing the non-discarded (accepted) arrivals. Based on these functions, $\hat{\Theta}(z,x_0)$ can be obtained as

$$\begin{aligned}\hat{\Theta}(z,x_0) = &\int_{x_1} F_\mathrm{a}(z,x_0,x_1)\mathrm{d}x_1 + \int_{x_1}\int_{x_2} F_\mathrm{d}(z,x_0,x_1)F_\mathrm{a}(z,x_1,x_2)\mathrm{d}x_2\mathrm{d}x_1 + \dots \\ &= \sum_{i=1}^{\infty}\int_{x_1}\dots\int_{x_i} F_\mathrm{d}(z,x_0,x_1)\dots F_\mathrm{d}(z,x_{i-2},x_{i-1})F_\mathrm{a}(z,x_{i-1},x_i)\mathrm{d}x_i\dots\mathrm{d}x_1. \quad (32)\end{aligned}$$

The cumulative distribution function (CDF) of the initial phase distribution after a non-discarded photon arrival is provided in the second term of (20). Its density function (obtained by a derivation according to the function parameter) is

$$f_\mathrm{init}(x) = \frac{\lambda \mathrm{e}^{-\lambda x}}{1-\mathrm{e}^{-\lambda\tau}}, \quad (33)$$

for $0 \le x \le \tau$. The distribution of Θ_1 is obtained in z-transform domain as

$$\hat{\Theta}(z) = \mathbb{E}\left(z^{\Theta_1}\right) = \int_x f_{\text{init}}(x)\hat{\Theta}(z,x)\mathrm{d}x. \tag{34}$$

We note that the mean "time" between observations, which we computed directly in the previous section, is

$$\mathbb{E}(\Theta) = \frac{\mathrm{d}}{\mathrm{d}z}\hat{\Theta}(z)\bigg|_{z=1}.$$

Unfortunately, the infinite number of integrals in (32) makes the numerical analysis of $\hat{\Theta}(z)$ computationally challenging but can be efficiently approximated using the following Erlangization approach.

3.4 Approximation based on an Erlang clock

Following the pattern of Ref. [15, eq. (50)], we map $f_n(x_0,x_1)$, as introduced in (10), into matrices of size $\hat{N}\times\hat{N}$:

$$\begin{aligned}
\{\mathbf{A}_n\}_{ij} &= \Pr(J_1 = j, D_1 = n \mid J_0 = i) \\
&= \begin{cases} \Pr\big(\Omega = n\hat{N} + j - i - L\big) & \text{if } n\hat{N} + j \ge i + L, \\ 0 & \text{otherwise,} \end{cases} \\
&= \begin{cases} q(1-q)^{n\hat{N}+j-i-L} & \text{if } n\hat{N} + j \ge i + L, \\ 0 & \text{otherwise,} \end{cases}
\end{aligned} \tag{35}$$

where $\hat{N}$ is the order of the Erlang clock, $q = \frac{\lambda\tau}{\lambda\tau+\hat{N}}$ and the discretized version of dead time is $L = \lfloor \hat{N}\zeta/\tau \rfloor$, an integer. Furthermore, $J_i \in \{1,\ldots,\hat{N}\}$ denotes the phase of the grid process at S_i, while Ω denotes the number of phase changes.

To compute the number of intervals associated with discarded and non-discarded arrivals, we introduce $\mathbf{A}_\mathrm{d}(z) = \sum_{n=0}^{m} \mathbf{A}_n z^n$ and $\mathbf{A}_\mathrm{a}(z) = \sum_{n=m+1}^{\infty} \mathbf{A}_n z^n$.

The Erlang clock based approximate of $\hat{\Theta}(z,x_0)$ is obtained by considering that an accepted photon arrival is preceded by an arbitrary number of dropped photon arrivals, thus

$$\Theta(z) = \sum_{i=0}^{\infty} \mathbf{A}_\mathrm{d}^i(z)\mathbf{A}_\mathrm{a}(z) = \left(\mathbf{I} - \mathbf{A}_\mathrm{d}(z)\right)^{-1}\mathbf{A}_\mathrm{a}(z), \tag{36}$$

$\mathbf{I}$ denoting an identity matrix of appropriate size. From this, the distribution of Θ can be obtained by inverse z-transform and its kth factorial moment as

$$f_k = \mathbb{E}\left(\Theta(\Theta-1)\dots(\Theta-k+1)\right) = \frac{\mathrm{d}^k}{\mathrm{d}z^k}\, v_{\mathrm{init}}\Theta(z)\mathbb{1}\Big|_{z=1}, \tag{37}$$

where $\mathbb{1}$ is a column vector of ones, $v_{\mathrm{init}} = \frac{\hat{v}}{\hat{v}\mathbb{1}}$, and $\{\hat{v}\}_i = q(1-q)^{i-1}$ is the discretized version of f_{init}, introduced in (33). E.g., the squared coefficient of variation (SCV) of Θ can be obtained from the factorial moments as

$$C_\Theta^2 = \frac{E\left(\Theta^2\right) - E(\Theta)^2}{E(\Theta)^2} = \frac{f_2 + f_1 - f_1^2}{f_1^2}. \tag{38}$$

4 Numerical investigations

In this section, we validate the obtained analytical results against simulations for some performance indices.

4.1 Simulations

We created simulation runs, each consisting of 1 million consecutively generated intervals, with a custom-built Python program. For sample interval generation, we utilized Python's built-in pseudorandom "random" library[4] to simulate photon emission times for particular λ and τ parameters. We also simulated the effect of a constant ζ dead time (emissions in the dead time period are not registered as detections) and then used these intervals as the input for a Python function implementing Algorithm 1 to generate simulated vDTD distributions and calculate various statistics of the simulation results. We obtained every data point by taking the mean of 20 independent simulation runs. In figures, the standard deviation of the statistic is also denoted with a blue error bar based on the 20 samples—although this value is mostly too small for graphical visibility.

First, we verified the validity of simulations using the lag-1 correlations in (16), as well as the mean value of DTDs in (13). The dead time in the simulation had zero integer part ($k = 0$) and a fractional part δ varying between 0 and 0.9. The clock resolution was set to $\tau = 1$, and we swept the value of λ between 0 and 10. The results in Fig. 4 show excellent agreement between theory and simulations.

[4]Although pseudorandom number generators cannot provide truly random numbers, the output they produce is still suitable for initial investigations, as this output is expected to mimic the statistical properties of truly random sequences, without the indeterministic features.

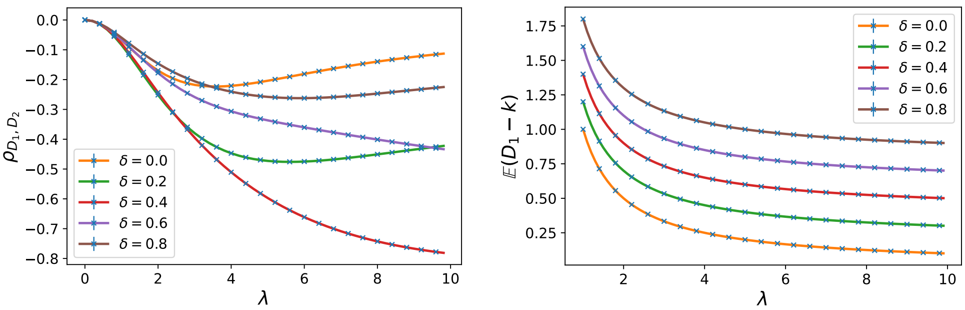

Fig. 4 Comparison of theoretically calculated (solid lines) and simulated results (markers) for the correlation between successive DTDs (left) and the mean value of DTDs (right), as a function of λ and selected fractional dead times δ. The simulation uses $\tau = 1$, and the step size for λ is 0.1, but only every fourth data point is shown here for better visibility.

Theoretically obtained and simulated results also align for further performance measures, such as the virtual count rate. Figure 5 shows two cases; the results support the validity of the theoretical model presented in Theorem 2. Using these simulations, we also checked the validity of results when using the approximation method based on an Erlang clock, as presented in Section 3.4. We found that this approximation already has a decent accuracy with relative errors[5] in the order of 10^{-2} for $\hat{N} = 100$ and 10^{-3} for $\hat{N} = 1000$ Erlang phase parameters, while allowing for the approximation of arbitrary performance indices. An example of simulated and approximated results for C^2_Θ can be seen in Fig. 6.

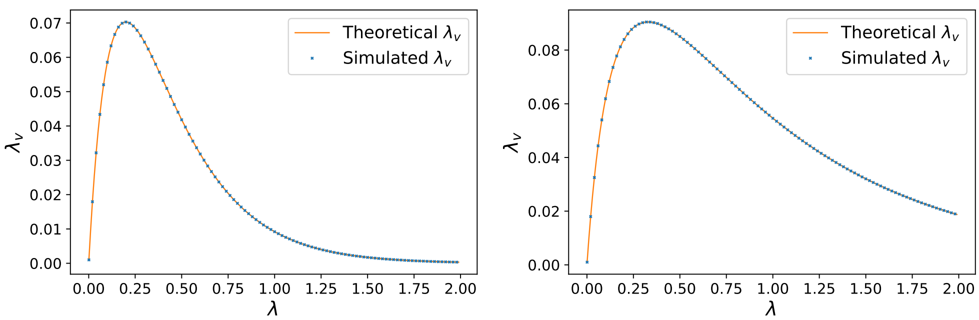

Fig. 5 Theoretically derived and simulated results for the virtual count rate λ_v as a function of the λ input photon rate, for different dead times ($\zeta = 1.8, 4.2$, left to right) with $m = 5, \tau = 1$. The simulation step size for λ is 0.05, but only every second data point is shown here for better visibility.

[5]The relative error is defined as the difference in percentage between the approximate and theoretical values when the latter is taken to be 100%.

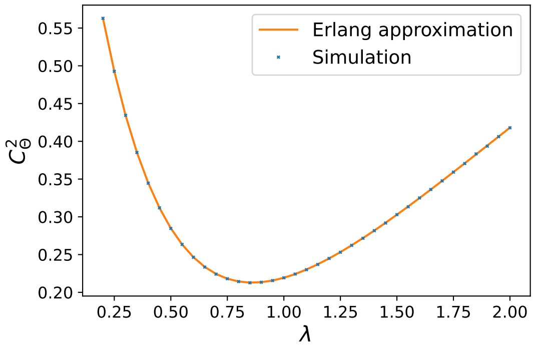

Fig. 6 Simulated and approximated results for the SCV, C^2_Θ, for different input photon rates λ and fixed dead time $\zeta = 0.7$. The approximation uses $\hat{N} = 1000$ Erlang phases with $\tau = 1$, $m = 5$.

4.2 Performance cost

To demonstrate the performance cost of Algorithm 1, we compare the DTD and vDTD generation rates. Comparing λ_d and λ_v indicates that for $\lambda\tau \ll 1$, the difference in output rates is not substantial, but when $\lambda\tau \sim 1$, the performance cost of using Algorithm 1 becomes apparent, as seen in Fig. 7. We can further define the $\lambda_\mathrm{v}/\lambda_\mathrm{d}$ ratio to quantify this performance loss:

$$\frac{\lambda_\mathrm{v}}{\lambda_\mathrm{d}} = \frac{\mathrm{e}^{-\lambda((m+1)\tau-\zeta)}\left(\mathrm{e}^{\lambda\tau}-1\right)}{\tau(\lambda\zeta+1)} \cdot \frac{1+\lambda\zeta}{\lambda} = \frac{\left(\mathrm{e}^{\lambda\tau}-1\right)\mathrm{e}^{-\lambda((m+1)\tau-\zeta)}}{\lambda\tau}. \tag{39}$$

Eq. (39) indicates that the critical defining factor for performance loss is the difference $m\tau - \zeta$ (which we will call the *accuracy of overestimation*), corresponding to how much we overestimate ζ with $m\tau$. While $m\tau$ needs to be strictly greater than ζ for Algorithm 1 to provide uncorrelated vDTDs, it is beneficial to choose $m\tau$ as close to ζ as possible. This effect is illustrated in Fig. 8.

4.3 Maximally achievable virtual count rate

When generating vDTDs with Algorithm 1, increasing the λ input photon rate beyond a certain point decreases the final virtual count rate as the probability of detections corresponding to smaller D_i values rises. Thus, finding the optimal input λ corresponding to the maximally achievable output λ_v is important.

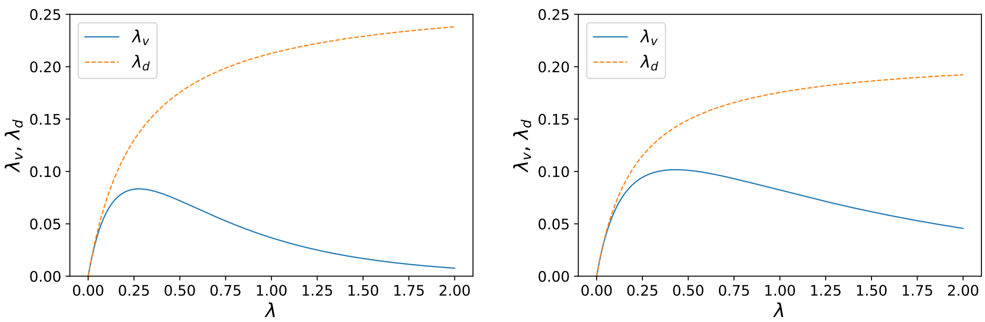

Fig. 7 Comparison of achievable output rates at λ input photon rates for different dead times ($\zeta = 3.7, 4.7$, left to right) with (λ_v) and without (λ_d) using Algorithm 1 with $\tau = 1$, $m = 5$.

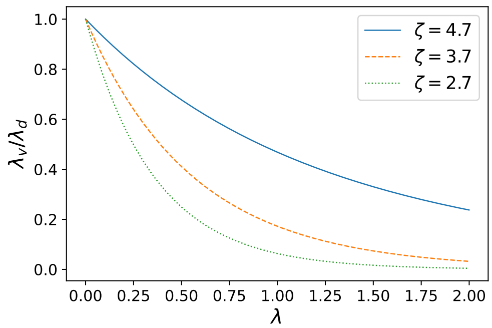

Fig. 8 The performance cost ratio λ_v/λ_d as a function of λ input photon rate for different dead times ($\zeta = 2.7, 3.7, 4.7$) and $\tau = 1, m = 5$.

Using Eq. (30), we can find this maximum by solving

$$\begin{aligned}\frac{\partial \lambda_v}{\partial \lambda} &= \frac{\partial}{\partial \lambda} \frac{e^{-\lambda((m+1)\tau-\zeta)}\left(e^{\lambda\tau}-1\right)}{\tau(\lambda\zeta+1)} \\ &= \frac{\left(e^{\lambda\tau}-1\right)[\zeta-(m+1)\tau]e^{-\lambda((m+1)\tau-\zeta)}}{\tau(\lambda\zeta+1)} \\ &\quad - \frac{\zeta\left(e^{\lambda\tau}-1\right)e^{-\lambda((m+1)\tau-\zeta)}}{\tau(\lambda\zeta+1)^2} + \frac{e^{\lambda\tau-\lambda((m+1)\tau-\zeta)}}{\lambda\zeta+1} = 0\end{aligned} \tag{40}$$

for λ. Unfortunately, this equation has no algebraic solution but can still be solved numerically. Solutions for an example parameter set are compared to simulation results in Fig. 9.

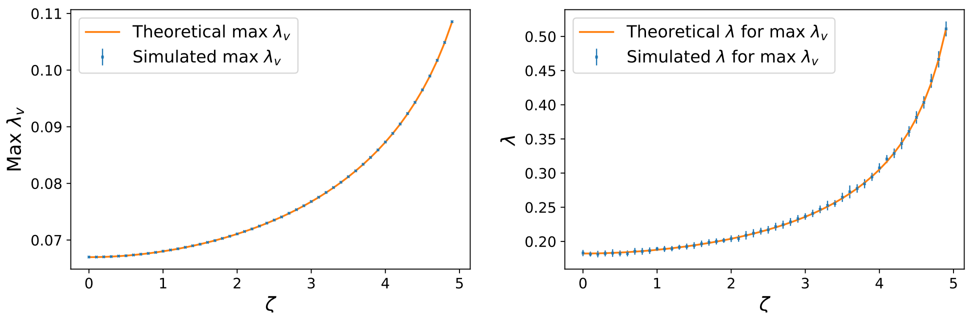

Fig. 9 Maximally achievable virtual count rates (λ_v) and the corresponding input rates (λ) for different dead times (ζ) with fixed $m = 5$ and $\tau = 1$ parameters.

The accuracy of the overestimation ($m\tau - \zeta$) also has a critical effect on maximum achievable rates. This reinforces the importance of choosing $m\tau$ close to ζ.

Note that compared to the practice of reducing the λ_d input rate for correlation mitigation, the maximal λ_v output virtual count rates provided by our method exceed the typical power limited λ_d input rates (see e.g. the end of Sec. 2.1) as long as $m\tau - \zeta$ is chosen properly.

4.4 Entropy of the output counts

Due to Algorithm 1, the vDTDs are independent and identically geometrically distributed with

$$\Pr(V = v) = p_v = p(1-p)^v,\ v \in \mathbb{Z}^+ \tag{41}$$

probabilities where $p = 1 - \mathrm{e}^{-\lambda\tau}$. Consequently, the min-entropy of a vDTD is

$$H_\infty(V) = \min_v(-\log_2 p_v) = -\log_2\left(1 - \mathrm{e}^{-\lambda\tau}\right) \tag{42}$$

and its (Shannon) entropy is

$$\begin{aligned} H(V) &= -\sum_v p_v \log_2 p_v = \frac{-(1-p)\log_2(1-p) - p\log_2 p}{p} \\ &= \frac{\lambda\tau \cdot \log_2(\mathrm{e}) \cdot \mathrm{e}^{-\lambda\tau} - \log_2(1-\mathrm{e}^{-\lambda\tau}) \cdot (1-\mathrm{e}^{-\lambda\tau})}{1-\mathrm{e}^{-\lambda\tau}}. \end{aligned} \tag{43}$$

The min-entropy of a random variable provides the upper bound of uniform bits that can be extracted from the variable [20] and can never exceed its Shannon entropy, making it a more efficient measure when assessing random number generators. The other main factor determining the achievable raw entropy generation speed is the rate at which measurement samples are obtained. When using Algorithm 1 this rate is the λ_v virtual count rate, as it determines the speed at which Algorithm 1 generates vDTDs. The (min-)entropy rates, defined as the (min-)entropy generated per unit time, are the products of the (min-)entropy per random variable and the rate at which random variables are generated. Their values can be calculated as $h(V) = \lambda_v \cdot H(V)$ and $h_\infty(V) = \lambda_v \cdot H_\infty(V)$, respectively.

4.5 Handling non-constant dead time

The dead time ζ may not be constant in real systems. We also consider the case when ζ is a random variable to model this effect.

4.5.1 Finite support ζ distributions

We first show that the virtual count rate is monotonic in ζ, then provide limits for λ_v assuming finite-support dead time distributions.

Monotonicity of λ_v in ζ

λ_v is monotonic in ζ, since

$$\frac{\partial \lambda_v}{\partial \zeta} = \frac{\partial}{\partial \zeta} \frac{e^{-\lambda((m+1)\tau-\zeta)}\left(e^{\lambda\tau}-1\right)}{\tau(\lambda\zeta+1)} = \frac{\lambda\zeta^2 e^{-\lambda((m+1)\tau-\zeta)}\left(e^{\lambda\tau}-1\right)}{\tau(\lambda\zeta+1)^2} > 0, \tag{44}$$

because $\lambda > 0$, $\zeta \geq 0$, and $\tau > 0$ by definition, which also makes $e^{\lambda\tau} > 1$, therefore Eq. (44) holds true for all valid ζ.

Bounded ζ

For the case of finite-support ζ distributions, we can use the upper bound of the distribution to set m adequately. In contrast, due to the monotonicity in ζ, we can use the lower bound of ζ to calculate the worst-case performance characteristics of Algorithm 1 for the chosen m. More precisely, given an upper bound ζ_U and lower bound ζ_L for ζ, we can substitute $\zeta = \zeta_L, m = \lfloor \zeta_U/\tau \rfloor + 1$ into our previous formulae to get worst-case results in terms of the achievable λ_v. Since we set our m overestimation parameter according to ζ_U, and λ_v is maximal when $m\tau - \zeta$ is minimal, the constant $\zeta = \zeta_U$ distribution corresponds to the best case scenario, yielding a maximal λ_v for the given m. Substituting these into Eq. (24), we obtain

$$\begin{aligned} &e^{-\lambda\left[\left(\left\lfloor \frac{\zeta_U}{\tau} \right\rfloor + 2\right)\tau - \zeta_L\right]}\left(e^{\lambda\tau}-1\right) \cdot \frac{1}{\tau(\lambda\zeta_L+1)} \leq \lambda_v \quad \text{and} \\ &\lambda_v \leq e^{-\lambda\left[\left(\left\lfloor \frac{\zeta_U}{\tau} \right\rfloor + 2\right)\tau - \zeta_U\right]}\left(e^{\lambda\tau}-1\right) \cdot \frac{1}{\tau(\lambda\zeta_U+1)}. \end{aligned} \tag{45}$$

This way, even if we do not know the exact value or distribution of ζ, we can still give a lower and upper estimate for the achievable virtual count rates.

4.5.2 Unbounded dead time distributions

For a fixed value of m, a particular sample from an arbitrary ζ distribution can fall into two categories:

$$A_1 : \ \zeta \leq m\tau,$$
$$A_2 : \ \zeta > m\tau,$$

where A_1 and A_2 are mutually exclusive and complete. Due to the law of total probability, the stationary distribution of the vDTDs can be written as

$$\Pr(V = v) = \Pr(V = v \mid \zeta \leq m\tau) \cdot \Pr(\zeta \leq m\tau) + \Pr(V = v \mid \zeta > m\tau) \cdot \Pr(\zeta > m\tau), \tag{46}$$

where the first part of the sum corresponds to A_1 and the second part to A_2. In the case of A_1, the corresponding distribution of V is the same as in Sec. 3.1 since $\zeta \leq m\tau$, and in this case, $\Pr(V = v \mid \zeta \leq m\tau)$ is independent of ζ and equal to (22). In the case of A_2, $\Pr(V = v \mid \zeta > m\tau)$ is no longer independent of ζ; therefore, V is no longer ensured to be uncorrelated and may show unwanted correlations. However, the probability of potentially correlated samples is $\Pr(\zeta > m\tau)$, and can be adjusted by the choice of m. Larger m values result in a lower sample generation rate, λ_v, but a lower probability of correlated samples, and the opposite holds for smaller m values. The proper choice of m can set an appropriate trade-off.

5 Measurements and experimental results

We tested Algorithm 1 with the physical setup presented in detail in Ref. [19]. A green semiconductor laser (Thorlabs LP520-SF15) working in CW conditions is the source of photons, with a wavelength of 519.9 nm. After passing through several tunable attenuators to set the desired photon rate, the light is detected by a low-noise photomultiplier (PicoQuant PMA-175 NANO), and its output pulses are time-tagged by a time-to-digital converter (PicoQuant TimeHarp 260). Figure 10 shows the block diagram of the experimental setup.

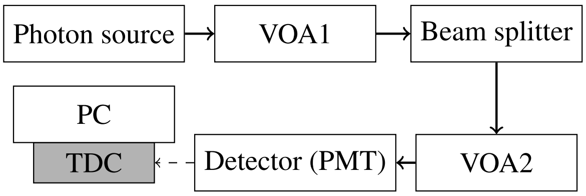

Fig. 10 Experimental setup used for measurements. VOA: variable optical attenuator; PMT: photomultiplier tube; TDC: time-to-digital converter card; PC: computer. (Beam splitter functions as an additional 20 dB attenuator.)

The maximum photon rate tolerated by the photomultiplier is around 5 Mcps (million counts per second). The highest resolution of the detection system is $\tau = 250\,\mathrm{ps}$, while the total dead time is reported to be typically around 2 ns. According to our measurement results, while 2 ns can be considered a lower limit for the dead time, there are cases where the system exhibits behaviour corresponding to larger values of ζ. Therefore, we cannot consider ζ to be constant.

At first glance, correlation coefficients predicted by e.g. (8) look negligible for the parameter set we use. However, our previous research showed that even seemingly low correlations between DTDs become noticeable once the samples are used for random bit generation. Earlier, we conducted measurements on the same experimental setup and increased the detection rate to around $3.72 \cdot 10^6$ cps. The NIST Statistical Test Suite [21], one of the primary tools of randomness assessment, failed the generated bit sequence on the *Runs* test at a significance level of 0.01, showing that consecutive bits feature a non-zero correlation [19].

We collected measurement data of $2 \cdot 10^9$ observed photon arrival times with a mean detection rate of $\lambda_\mathrm{d} \approx 1.05 \pm 0.01$ Mcps. Rescaling after accounting for the typical dead time of the system according to (18) results in an input photon rate of $\lambda = 1.052$ Mcps.

We also created time-binned versions of the original, unbinned measurement data to investigate possible $\lambda\tau$ statistics beyond our experimental setup's range of operational limits. To do so, we used data recorded with the device's own τ time resolution and created lower resolution versions of the same experiment—as if we used a longer, $\tau' = K_\mathrm{b} \cdot \tau$ clock period, where K_b is a positive integer. The binning method is presented in Algorithm 2.

We obtained additional *binned datasets* corresponding to $K_\mathrm{b} = 2,\ 5,\ 10,\ 100,\ 1000$. We applied Algorithm 1 to the unbinned and binned raw datasets. We refer to the output of Algorithm 1 as *overestimated data*.

Algorithm 2 Binning algorithm

Require: $\mathbb{D}$ ▷ Original DTD samples

Require: K_b ▷ Integer

$\mathbb{D}' = [\,]$ ▷ Array of binned DTDs

$c_1 := 0$ ▷ Carry

for $i = 1$ to length($\mathbb{D}$) **do**

  get $D_i = \mathbb{D}[i]$

  $\mathbb{D}'[i] = \lfloor (D_i + c_i)/K_\mathrm{b} \rfloor$

  $c_{i+1} = D_i + c_i \pmod{K_\mathrm{b}}$

end for

For the unbinned data ($K_\mathrm{b} = 1$), we set $m = 1000$ as a safe overestimation parameter,[6] and $m' = 500,\ 200,\ 100,\ 10,\ 1$ for the binned data with $K_\mathrm{b} = 2,\ 5,\ 10,\ 100,\ 1000$, respectively, following the rule $m' = 1000/K_\mathrm{b}$.[7]

We evaluated the *raw* and *overestimated* (both unbinned and binned) datasets in the following ways:

1. By calculating the autocorrelation of (v)DTD sequences.
2. By counting single (v)DTD occurrences. As the distribution of values (the histogram) is expected to be geometrically distributed, we fit it to the expected form. We then calculated the goodness of fit and checked the fitting parameters.
3. By counting the relative frequencies of consecutive (v)DTDs' value pairs. Measured pair statistics are compared to the expected value of the ideal, independent case—calculated as the product of relative frequencies of single (v)DTDs—via hypothesis testing.

The results of the evaluation methods are detailed below.

5.1 Autocorrelation of (v)DTD sequences

First, we calculated the autocorrelation coefficients of every dataset, denoted as a_1 and a_1^o for raw and overestimated data, respectively. The unbinned raw dataset shows correlation coefficients in the order of 10^{-5}. The half-width of the 95% confidence interval for zero

[6]Examining the measurement data, we conclude that $\zeta < 1000\tau$ with high enough certainty that this choice of m can be considered safe, faithfully overestimating the dead time.

[7]The binning algorithm rescales the necessary overestimation parameter by $1/K_\mathrm{b}$, as the dead time of the underlying process is unchanged. If $\zeta < m\tau$, then $\zeta < (m/K_\mathrm{b}) \cdot (K_\mathrm{b} \cdot \tau)$ holds trivially. The choice of $m' = m/K_\mathrm{b}$ yields a comparable dataset to the unbinned set overestimated by m; using the original overestimation parameter for the binned sequence would result in a greatly reduced λ_v.

correlation is

$$\frac{\sqrt{2}\cdot\mathrm{Erf}^{-1}(0.95)}{\sqrt{2\cdot 10^9}} = \frac{1.96}{\sqrt{2\cdot 10^9}} = 4.38\cdot 10^{-5}$$

for $2\cdot 10^9$ samples, where $\mathrm{Erf}^{-1}(\cdot)$ is the inverse error function. Obtaining such small correlation coefficients is expected even without overestimation when $\lambda\tau \ll 1$—recall that correlations become noticeable as the product increases. Table 1 lists the lag-1 coefficients of raw and overestimated datasets. The only coefficient exceeding 10^{-4} in absolute value is the lag-1 coefficient for the dataset with the largest $\lambda\tau$, using $K_\mathrm{b} = 1000$, which shows a significant and sudden increase, leaping above 10^{-3} in magnitude.

After overestimation, lag-1 coefficients remained in the order of 10^{-5}, within the 95% confidence interval for zero correlation—even without considering the slight growth of the confidence interval due to the reduced number of samples in the overestimated datasets.[8] All of the overestimated sequences show lower magnitude autocorrelation coefficients than their unprocessed counterparts. The difference is most notable for the sequence with binning parameter 1000, which was originally heavily correlated. When overestimated, the sequence performs significantly better. Note that sequences have similar values after being passed through the algorithm—this is expected since all of them are discretized from the same realization of the underlying PPP, and all use the same overestimation parameter after adjusting for dead time, $m'\cdot K_\mathrm{b}$.

Table 1 Lag-1 autocorrelation coefficients of raw (a_1) and overestimated (a_1^o) datasets. Overestimation successfully reduced the absolute values of correlation coefficients for all data.

K_b / m'	$\lambda\tau'$	a_1	a_1^o
1 / 1000	$2.630\cdot 10^{-4}$	$4.324\cdot 10^{-5}$	$-7.811\cdot 10^{-6}$
2 / 500	$5.261\cdot 10^{-4}$	$4.322\cdot 10^{-5}$	$-8.175\cdot 10^{-6}$
5 / 200	$1.315\cdot 10^{-3}$	$4.311\cdot 10^{-5}$	$-7.692\cdot 10^{-6}$
10 / 100	$2.630\cdot 10^{-3}$	$4.273\cdot 10^{-5}$	$-1.109\cdot 10^{-5}$
100 / 10	$2.630\cdot 10^{-2}$	$-1.474\cdot 10^{-5}$	$-1.233\cdot 10^{-5}$
1000 / 1	$2.630\cdot 10^{-1}$	$-5.737\cdot 10^{-3}$	$-1.987\cdot 10^{-5}$

[8] E.g., for the shortest dataset ($K_\mathrm{b} = 1000, m' = 1$) with $1.37\cdot 10^9$ samples, the magnitude of the 95% confidence interval increases to $\sqrt{2}\cdot\mathrm{Erf}^{-1}(0.95)/\sqrt{1.37\cdot 10^9} = 1.96/\sqrt{1.37\cdot 10^9} = 5.29\cdot 10^{-5}$.

5.2 Frequencies of (v)DTD values

Histograms show an even more noticeable contrast between the raw and overestimated cases. We fit the function $y = A \cdot \mathrm{e}^{-Ax} + C$ to the histogram data using the least squares method.[9] Ideally, fitting would yield $A = \lambda\tau'$ and $C = 0$—note that this is a discretized version of the exponential probability density function $f_T(t) = \chi_{\{t \geq 0\}} \lambda \cdot \mathrm{e}^{-\lambda t}$.[10] The histograms and results of the fitting are shown in Fig. 11. Histograms show deviations from a geometric distribution for the raw datasets, noticeable even by visual inspection, while overestimated datasets do not. The fitting error statistics of overestimated datasets are at least 3 orders of magnitude better compared to their raw counterparts, both in the case of *mean square error*s (MSEs) and *coefficient of determination* parameters (R^2; perfect fit is $R^2 = 1$). The resulting A parameters for the overestimated data are also in agreement with the expected $\lambda\tau'$ values,[11] although slightly larger. This is most probably because the expected $\lambda\tau'$ values were calculated with the spreadsheet dead time value of 2 ns, but in reality, the actual dead-time-like imperfections of the measurement setup caused a bigger reduction of the effective rate than what the constant $\zeta = 2$ ns correction accounted for. The fitting results are summarised in Tables 2 and 3.

Table 2 A parameters of curve fitting before and after overestimation

Data	Raw A	Overestimated A	Expected ($\lambda\tau'$)
$K_\mathrm{b} = 1$	$2.578 \cdot 10^{-4}$	$2.638 \cdot 10^{-4}$	$2.630 \cdot 10^{-4}$
$K_\mathrm{b} = 2$	$5.154 \cdot 10^{-4}$	$5.276 \cdot 10^{-4}$	$5.261 \cdot 10^{-4}$
$K_\mathrm{b} = 5$	$1.285 \cdot 10^{-3}$	$1.319 \cdot 10^{-3}$	$1.315 \cdot 10^{-3}$
$K_\mathrm{b} = 10$	$2.553 \cdot 10^{-3}$	$2.636 \cdot 10^{-3}$	$2.630 \cdot 10^{-3}$
$K_\mathrm{b} = 10^2$	$2.440 \cdot 10^{-2}$	$2.609 \cdot 10^{-2}$	$2.630 \cdot 10^{-2}$
$K_\mathrm{b} = 10^3$	$1.751 \cdot 10^{-1}$	$2.396 \cdot 10^{-1}$	$2.630 \cdot 10^{-1}$

[9] We utilized the Scipy python library's "curve_fit" method with initial guiding guesses determined by the expected $\lambda\tau'$ parameter, and 10^5 maximum evaluations.

[10] As shown in Eq. (41) and Ref. [16], sampling exponentially distributed time intervals with parameter λ—using a restartable clock with resolution τ and no dead time—yields geometrically distributed samples. Thus, an equivalent exponential fit is also a valid substitute for this geometric fit. The additional C parameter is introduced because we only considered data in the histograms corresponding to the first part of the distribution that fits into the predetermined amount of histogram bins.

[11] For the $K_b = 100$ and $K_b = 1000$ cases, bigger deviation of the fit parameters are expected due to smaller sample sizes (since the number of histogram bins was also scaled with K_b for comparability of results) and higher impact of the C fitting parameter.

Table 3 MSE and $1-R^2$ values of curve fitting before and after overestimation

Data	Raw		Overestimated	
	MSE	$1-R^2$	MSE	$1-R^2$
$K_b = 1$	$5.445 \cdot 10^{-11}$	$1.242 \cdot 10^{-2}$	$9.242 \cdot 10^{-14}$	$2.053 \cdot 10^{-5}$
$K_b = 2$	$2.278 \cdot 10^{-10}$	$1.299 \cdot 10^{-2}$	$2.447 \cdot 10^{-13}$	$1.359 \cdot 10^{-5}$
$K_b = 5$	$1.780 \cdot 10^{-9}$	$1.624 \cdot 10^{-2}$	$1.111 \cdot 10^{-12}$	$9.866 \cdot 10^{-6}$
$K_b = 10$	$1.164 \cdot 10^{-8}$	$2.658 \cdot 10^{-2}$	$3.814 \cdot 10^{-12}$	$8.472 \cdot 10^{-6}$
$K_b = 10^2$	$2.861 \cdot 10^{-6}$	$6.683 \cdot 10^{-2}$	$6.545 \cdot 10^{-10}$	$1.457 \cdot 10^{-5}$
$K_b = 10^3$	$9.155 \cdot 10^{-4}$	$2.959 \cdot 10^{-1}$	$2.508 \cdot 10^{-6}$	$5.848 \cdot 10^{-7}$

5.3 Frequencies of successive (v)DTD pair values

If the individual (v)DTDs are independent, then the joint probabilities satisfy

$$\begin{aligned} \Pr(D_i = k, D_{i+1} = \ell) &= \Pr(D_i = k) \cdot \Pr(D_{i+1} = \ell) \text{ and} \\ \Pr(V_i = k, V_{i+1} = \ell) &= \Pr(V_i = k) \cdot \Pr(V_{i+1} = \ell)\,. \end{aligned} \tag{47}$$

We can use this for hypothesis testing, where our null hypothesis is that the tested data is from an ideal binomial trial with a probability given by (47), and gather evidence trying to refute this.[12] We applied binomial statistical tests on each of the $\{D_i = k, D_{i+1} = \ell\}$ and $\{V_i = k, V_{i+1} = \ell\}$ pair statistics for $k, \ell \in \{0, 1, \ldots, 19\}$, yielding a p-value for each of the 400 pairs to investigate possible deviations from the expected distribution in the case of consecutive detections.

We set the target of the *comprehensive* significance level per dataset to 0.01. Since we are looking only at the most extreme p-values, we used the Bonferroni correction (due to the multiple comparisons problem) [22] to get *individual* significance levels of $2.5 \cdot 10^{-5}$ that we then compare to each of the 400 p-values. If any p-value is lower than the *individual* significance level, then the whole dataset fails at the *comprehensive* significance level.

The results of the statistical tests show a clear contrast between the raw and the overestimated data in favour of the latter. The raw data scored minimum p-values of $1.6 \cdot 10^{-5}$ without binning ($K_b = 1$), and $5.9 \cdot 10^{-7}$, $3.4 \cdot 10^{-13}$, $9.2 \cdot 10^{-31}$, 0 and 0 for binned sets ($K_b = 2,\ 5,\ 10,\ 100,\ 1000$, respectively), which are orders of magnitude under the individual

[12] Successful rejection of the null hypothesis constitutes a test failure.

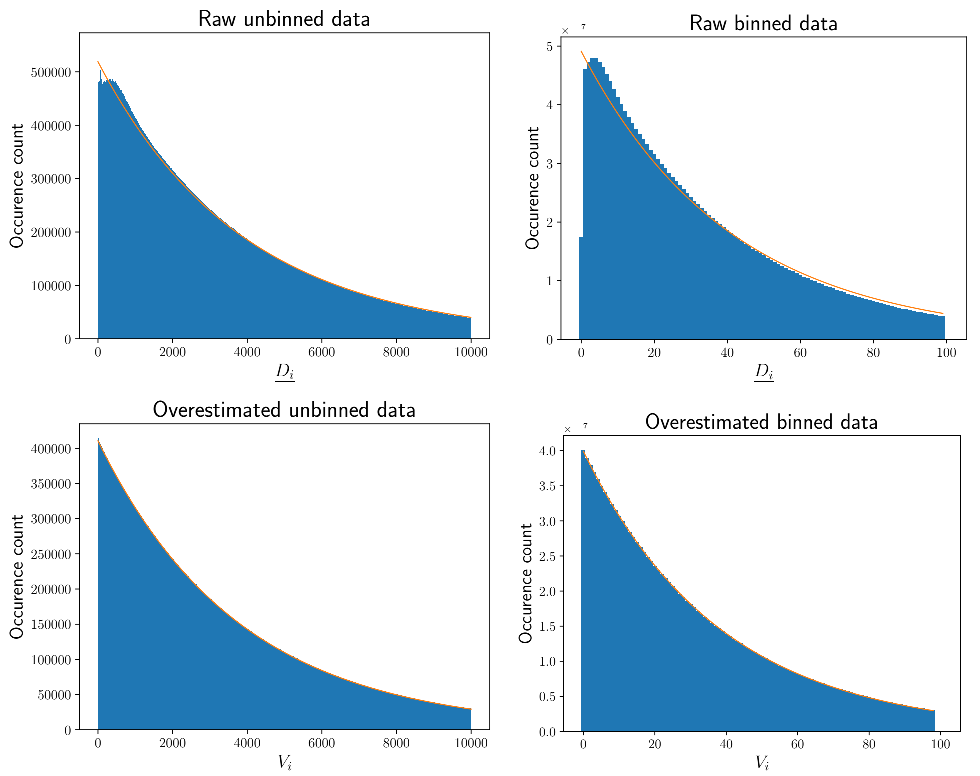

Fig. 11 Histograms and the results of curve fitting for measurement data before and after overestimation. Due to the effect of dead time, we shifted the histogram left before fitting, not including the originally empty bins for smaller D_i values. We denote these shifted values by $\underline{D_i}$. Figures on the left correspond to the original measurement data (unbinned), while figures on the right correspond to a binned case with $K_\mathrm{b} = 100$. The top row shows histograms of the unprocessed data, while the bottom row shows the resulting histograms after using Algorithm 1. The orange lines indicate the results of the attempted curve fitting.

significance level and, therefore, fail the test. The minima of p-values obtained for overestimated datasets range from $6 \cdot 10^{-4}$ ($K_\mathrm{b} = 10$) to 0.01 ($K_\mathrm{b} = 1000$), which, unlike results from the raw data, are all above the individual significance level, passing the test.

5.4 Further measurement results

The statistical tests signified that the overestimation algorithm can transform distorted distributions into distributions very close to exponential/geometric. Newly measured datasets with detected photon rates of ∼400, ∼600, and ∼800 kcps were also evaluated with the previously presented methodology, yielding similar results, emphasizing the gains.

To stress the potentially disadvantageous effect of correlations in measured DTDs, we also utilized the simple bit generation algorithm presented in Ref. [9] and tested the resulting bit sequences with the NIST STS statistical test suite [21]. The sequences with higher $\lambda\tau$ values had failing results for some of the test cases, while bit sequences created from the vDTDs passed all the test cases.

We also calculated the experimental ratio of measured input count rates to the virtual count rates achieved by Algorithm 1. We note that we only have measurement data available corresponding to low values ($\sim 10^{-4}$) of $\lambda\tau$, but the experimental results all stay within the bounds given by (45), using $\zeta_L = 10\tau$ and $\zeta_U = 999\tau$. The experimental output/input rates of Algorithm 1 range from 0.774 (for ~ 1 Mcps input rate) to 0.906 (for ~ 400 kcps input rate), which is a tolerable performance loss for eliminating the correlations within the generated DTD series.

6 Conclusion

We have introduced an algorithm to eliminate the dependencies between bits from single-photon detecting QRNG schemes. Compared to reducing the input optical power to limit operation into a regime with low correlations, our approach also allows generator operation in parameter regimes with higher input rates, potentially facilitating improved bit generation rates. The proposed procedure constructs a purely geometric distribution obtained from the discretized measurements of the underlying arrival process by overestimating the insensitive period after registered photon detections. The algorithm avoids correlations between successive time samples by discarding a period used for overestimation, which contains a random component depending on the arrival of photons with respect to the underlying time resolution grid. This virtually realizes the ideal case of no dead time and zero starting phase, yielding geometrically distributed *virtual discretized time differences* (similarly to a restartable measurement clock without dead time), preserving the memoryless property of the exponentially distributed physical process. Dead time overestimation features a slight compromise by reducing the output rate of detections used for bit generation.

The validity of our analytic results regarding the algorithm's theoretical soundness and performance metrics is supported by both computer simulations and measurements conducted on an experimental setup. The algorithm has low complexity, making it convenient to implement in random number generators where it is desirable to work with uncorrelated time samples before bit assignment or to harness randomness from an exponential/geometric distribution. Although we evaluated our algorithm's performance on collected datasets, its low complexity also makes it easy to implement in continuous operation modes. Depending on the focal points of the actual QRNG scheme, the benefits of dead time overestimation can largely exceed the disadvantages of a decreased effective count rate.

List of abbreviations

- CDF – cumulative distribution function
- cps – count(s) per second
- CW – continuous-wave
- DTD – discretized time difference
- MSE – mean square error
- NIST – National Institute of Standards and Technology
- PC – personal computer
- PMT – photomultiplier tube
- PPP – Poisson point process
- SCV – squared coefficient of variation
- SPD – single-photon detector
- TDC – time-to-digital converter
- ToA – time-of-arrival
- QRNG – quantum random number generator/generation
- vDTD – virtual discretized time difference
- VOA – variable optical attenuator

Declarations

Availability of data and materials

The datasets used and/or analysed during the current study are available from the corresponding author upon reasonable request.

Competing interests

The authors declare that they have no competing interests.

Funding

M.T. was supported by the OTKA K-138208 project of the Hungarian Scientific Research Fund. B.S. was supported by the Ministry of Culture and Innovation and the National

Research, Development and Innovation Office within the Quantum Information National Laboratory of Hungary (Grant No. 2022-2.1.1-NL-2022-00004). Á.S. was supported by the OTKA K-142845 project of the Hungarian Scientific Research Fund and also received funding from the European Union under grant agreement No. 101081247 (QCIHungary project), which has been implemented with the support provided by the Ministry of Culture and Innovation of Hungary from the National Research, Development and Innovation Fund.

Authors' contributions

B.S. provided the original concept of the dead time overestimating algorithm and conducted simulations and measurements. M.T. implemented the scheme and obtained results regarding the performance indices in Mathematica. Á.S. assembled the physical measurement setup. All three authors contributed to developing the theory and writing and proofreading the manuscript.

Acknowledgements

Not applicable.

References

[1] Dodis Y, Ong SJ, Prabhakaran M, Sahai A. On the (im)possibility of cryptography with imperfect randomness. In: 45th Annual IEEE Symposium on Foundations of Computer Science. IEEE; 2004. p. 196–205. Available from: https://doi.org/10.1109/FOCS.2004.44.

[2] Gyöngyösi L, Bacsardi L, Imre S. A survey on quantum key distribution. Infocommunications Journal. 2019;11(2):14–21. https://doi.org/10.36244/ICJ.2019.2.2.

[3] Herrero-Collantes M, García-Escartín JC. Quantum random number generators. Reviews of Modern Physics. 2017;89(1):015004. https://doi.org/10.1103/RevModPhys.89.015004.

[4] Mannalatha V, Mishra S, Pathak A. A comprehensive review of quantum random number generators: concepts, classification and the origin of randomness. Quantum Inf Process. 2023 Dec;22(12). https://doi.org/10.1007/s11128-023-04175-y.

[5] Jennewein T, Achleitner U, Weihs G, Weinfurter H, Zeilinger A. A fast and compact quantum random number generator. Review of Scientific Instruments. 2000 apr;71(4):1675–1680. https://doi.org/10.1063/1.1150518.

[6] Stefanov A, Gisin N, Guinnard O, Guinnard L, Zbinden H. Optical quantum random number generator. Journal of Modern Optics. 2000 mar;47(4):595–598. https://doi.org/10.1080/09500340008233380.

[7] Fürst H, Weier H, Nauerth S, Marangon DG, Kurtsiefer C, Weinfurter H. High speed optical quantum random number generation. Optics express. 2010;18(12):13029–13037. https://doi.org/10.1364/OE.18.013029.

[8] Gras G, Martin A, Choi JW, Bussières F. Quantum Entropy Model of an Integrated Quantum-Random-Number-Generator Chip. Physical Review Applied. 2021 may;15(5). https://doi.org/10.1103/physrevapplied.15.054048.

[9] Stipčević M, Rogina BM. Quantum random number generator based on photonic emission in semiconductors. Review of scientific instruments. 2007;78(4):045104. https://doi.org/10.1063/1.2720728.

[10] Wahl M, Leifgen M, Berlin M, Röhlicke T, Rahn HJ, Benson O. An ultrafast quantum random number generator with provably bounded output bias based on photon arrival time measurements. Applied Physics Letters. 2011;98(17):171105. https://doi.org/10.1063/1.3578456.

[11] Massari N, Tontini A, Parmesan L, Perenzoni M, Gruijć M, Verbauwhede I, et al. A monolithic SPAD-based random number generator for cryptographic application. In: ESSCIRC 2022- IEEE 48th European Solid State Circuits Conference (ESSCIRC). IEEE; 2022. p. 73–76. Available from: https://doi.org/10.1109%2Fesscirc55480.2022.9911498.

[12] Lei W, Xie Z, Li Y, Fang J, Shen W. An 8.4 Gbps real-time quantum random number generator based on quantum phase fluctuation. Quantum Information Processing. 2020 nov;19(11). https://doi.org/10.1007/s11128-020-02896-y.

[13] Williams CR, et al. Fast physical random number generator using amplified spontaneous emission. Optics express. 2010;18(23):23584–23597. https://doi.org/10.1364/OE.18.023584.

[14] Bustard PJ, Moffatt D, Lausten R, Wu G, Walmsley IA, Sussman BJ. Quantum random bit generation using stimulated Raman scattering. Optics Express. 2011 nov;19(25):25173. https://doi.org/10.1364/oe.19.025173.

[15] Schranz Á, Solymos B, Telek M. Stochastic performance analysis of a time-of-arrival quantum random number generator. IET Quantum Communication. 2024 6;5(2):140–156. https://doi.org/10.1049/qtc2.12080.

[16] Schranz Á, Udvary E. Mathematical analysis of a quantum random number generator based on the time difference between photon detections. Optical Engineering. 2020;59(4):044104. https://doi.org/10.1117/1.OE.59.4.044104.

[17] Müller JW. Generalized dead times. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment. 1991;301(3):543–551. https://doi.org/10.1016/0168-9002(91)90021-H.

[18] Glauber RJ. Coherent and incoherent states of the radiation field. Physical Review. 1963 9;131(6):2766–2788. https://doi.org/10.1103/PhysRev.131.2766.

[19] Schranz Á. Optical solutions for quantum key distribution transmitters [Ph. D. dissertation]. Budapest University of Technology and Economics; 2021. Available from: http://hdl.handle.net/10890/16991.

[20] Konig R, Renner R, Schaffner C. The Operational Meaning of Min- and Max-Entropy. IEEE Transactions on Information Theory. 2009 9;55(9):4337–4347. https://doi.org/10.1109/tit.2009.2025545.

[21] Rukhin AL, et al. A Statistical Test Suite for Random and Pseudorandom Number Generators for Cryptographic Applications. Gaithersburg, MD, United States: National Institute of Standards & Technology; 2010. Spec. Pub. 800-22, Rev. 1a. Available from: https://doi.org/10.6028/nist.sp.800-22.

[22] Dunn OJ. Multiple Comparisons among Means. Journal of the American Statistical Association. 1961;56(293):52–64. https://doi.org/10.1080/01621459.1961.10482090.