Experimental observation of nonclassical effects on single-photon detection rates

K. J. Resch, J. S. Lundeen, and A. M. Steinberg

Department of Physics, University of Toronto, Toronto, Ontario, Canada M5S 1A7

It is often asserted that quantum effects can be observed in coincidence detection rates or other correlations, but never in the rate of single-photon detection. We observe nonclassical interference in a singles rate, thanks to the intrinsic nonlinearity of photon counters. This is due to a dependence of the effective detection efficiency on the *quantum statistics* of the light beam. Such measurements of detector response to photon pairs promise to shed light on the microscopic aspects of silicon photodetectors, and on general issues of quantum measurement and decoherence.

Although quantum electrodynamics is one of the most well-established and most accurately tested of physical theories, the vast majority of optical physics still relies on the classical theory of the electromagnetic field. This is because, as was recognized early in the development of quantum optics [1], classical and quantum theories of light yield precisely the same predictions for a wide variety of phenomena. In particular, the two theories are equivalent for calculations of intensities (i.e., mean photon number) and effects linear in the intensity; only by studying higher-order correlations (e.g., Hanbury-Brown-Twiss-style coincidence measurements [2]) can one identify clear signatures of nonclassical effects. Over the past 20 years or so, many experiments have applied coincidence-counting and autocorrelation techniques to access these higher-order correlation functions and demonstrate a variety of fascinating quantum effects [3,4]. Since quantum and classical theories predict identical intensity measurement outcomes, such papers often begin by contrasting the featureless rate of single-photon detection with the quantum phenomena observable *only* in coincidence rates. (A striking exception is to be found in [4], where a quantum effect is used to alter the first-order coherence of two beams, leading to a modulation in intensity that can be observed in the rate of single detection events, or equivalently at a classical square-law detector.) Here, we show experimentally that the rate of photodection at a single detector can in fact be sensitive to quantum-statistical effects even under conditions of *constant* intensity, due to the frequently overlooked nonlinearity of typical photon counters such as avalanche photodiodes.

The standard model of photodetection used in quantum optics dates back to Glauber [1], and is appropriate for a low-efficiency detector whose photocurrent is, in the classical limit, proportional to the incident intensity, or the photon number. Real photon counters rely on extremely nonlinear processes such as electron-hole-pair avalanches to amplify the signal from a single photon to a detectable current. As a result, nearly all such devices (a rare exception, still not in widespread use, is the Rockwell SSPM [5,6]) are incapable of distinguishing between a single photon and two or more photons within a certain (device-dependent) ''dead time.'' For an ideal, 100%-efficient detector this would imply, within the dead time, a photocount rate (which we term the ''singles rate,'' although a single detection event need not

correspond to a single photon) proportional not to $\langle n \rangle \equiv \langle a^\dagger a \rangle$, but rather to the step function $\langle \Theta(n-0.5) \rangle$, which is equal to 1 for any photon number $n \geq 1$. The violently nonlinear character of this response implies that even a single avalanche photodiode is sensitive not only to the mean photon number, but also to higher-order terms. If one restricts oneself to considering photon numbers no higher than 2, for example, the response function may be Taylor expanded as $\frac{3}{2}\langle n \rangle - \frac{1}{2}\langle n^2 \rangle$. This expansion can be extended to include orders up to n^j where j is the maximum number of photons considered in the subspace, or written under normal ordering as $\langle a^\dagger a \rangle - \frac{1}{2} \langle a^\dagger a^\dagger aa \rangle + \frac{1}{6}\langle a^\dagger a^\dagger a^\dagger aaa \rangle - \cdots$. (It is this response, the rate of single detection events, which we refer to as the ''singles rate''; it is important to bear in mind that this can differ from the arrival rate of single *photons*.) For a real detector with a finite quantum efficiency, this model can be generalized. However, no precise theoretical model for such a detector exists, nor has the effect been experimentally studied prior to now. Our naive model suggests that if the single-photon detection efficiency is η, the probability of a detection event in the presence of two photons should be $\eta + (1-\eta)\eta = 2\eta - \eta^2$. This classical-minded model neglects the possibility of any cooperative effects between the photons, along with the possibility of stimulated-emission processes in the detector. Thus, measurements of the ''two-photon detection probability'' should help one construct accurate microscopic models of photon counters. In particular, it may be possible to probe issues in quantum measurement theory in this way. At very early times, the detection event is dominated by coherent processes such as electron-hole pair production. As the avalanche process amplifies this first pair, the system is coupled to a large number of degrees of freedom and effective decoherence occurs. Once coherence is lost, our model should correctly describe the effect of additional photons impinging on the detector. If, however, a second photon arrives during the coherent phase of the detection process, it is reasonable to expect a deviation from this prediction.

This experiment relies on a polarization-based version [7,8] of the two-photon interferometer originally devised by Hong, Ou, and Mandel [9]. Our experimental setup is shown in Fig. 1. Light from a mode-locked 800-nm titanium:sapphire laser is frequency doubled in a 0.2-mm β-barium borate (BBO) crystal. We separate the fundamental beam from
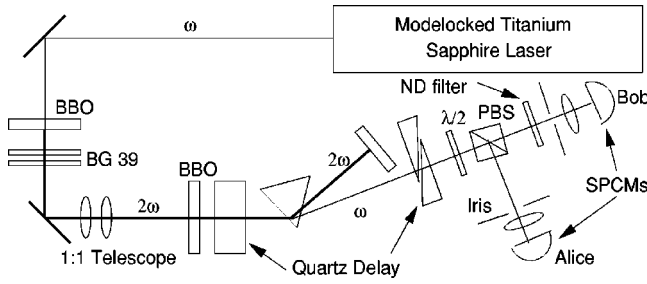
FIG. 1. Experimental setup for polarization-based two-photon interferometer: BG 39, colored glass filter; $\lambda/2$, half-wave plate; PBS, polarizing beam splitter; Alice/Bob, single-photon counting modules. Different transmission ND filters were used for each data accumulation run.

the second harmonic with colored glass filters (BG 39). The second harmonic serves as the pump laser for the production of type-II degenerate down-conversion in a second, 0.1-mm, BBO crystal. The down-converted photon pairs exit the nonlinear crystal collinear with each other and the pump laser, and have orthogonal polarizations—one photon's polarization is horizontal (H), and the other is vertical (V). Due to BBO's birefringence, the V photon trails the H photon by an average of 7 fs. The pump laser is separated from the down-converted photons by using a fused-silica prism. A 9-mm-thick quartz plate and a pair of translatable birefringent quartz prisms are used to delay the photons relative to one another inside the interferometer. The photon-pair polarization is rotated 45° by a half-wave plate and the pair goes through a cube polarizing beam splitter (PBS). The output beams of the PBS travel through adjustable irises and are focused onto 180-μm avalanche photodiodes (APDs): (EG&G model SPCM–AQR–13). In this way, each photon has a 50% chance of reaching each detector; see Fig. 1.

When the birefringent delay is set so that the photon wave packets arrive at the PBS simultaneously, there is a drop in the coincidence rate. The half-wave plate and the PBS serve to measure the photons in the $|\pm 45°\rangle$ basis. There are two possible Feynman paths that lead to a coincidence event: photon 1 can be transmitted at the PBS and photon 2 reflected, or vice versa. The amplitudes for these two processes interfere destructively due to the different relative phases of $|H\rangle$ and $|V\rangle$ in the states $|\pm 45°\rangle = (|H\rangle \pm |V\rangle)/\sqrt{2}$, leading to a coincidence null, provided the photons arrive at the beam splitter within their coherence length. This interference has no effect on the intensity of the light at either detector. If the photons arrive at the beam splitter outside each other's coherence length, then they are distinguishable, in principle, based on their arrival time at the detectors. In this case no interference is observed.

Theoretically, this coincidence dip can obtain a visibility of 100% for a narrow-band pump, and slightly lower values when the finite pulse duration and crystal length are taken into account [7]. However, for the purpose of this experiment, we obtained a higher photon collection efficiency and higher photon counting rates by opening the irises in front of the detectors at the cost of lowering the coincidence interference visibility. In addition to using large irises, we employed a slightly imbalanced 1:1 telescope in the pump laser beam

to again increase the collection efficiency [10].

To our knowledge, there is no microscopic physical model for the response of an APD when hit by two photons at the same time. In order to give a physically intuitive explanation for the origin and magnitude of an effect on the singles rate, we analyze the Hong-Ou-Mandel (HOM) interferometer using our naive model of detection. This model neglects the specifics of silicon APDs, and coherent effects between the multiple electron-hole excitation pathways. When the HOM interferometer is not balanced and there is no interference, each photon pair entering the device will lead to 0, 1, or 2 photons reaching a given detector—with probabilities 0.25, 0.5, and 0.25, respectively. When interference is occurring, half of the photon pairs entering the device will impinge, as a pair, on the detector and the other half will leave through the other exit port (for the ideal case of 100% coincidence visibility).

Thus for $\eta = 1$, the singles-counting rate drops from 75% of the pair-production rate to 50%, a dip with 33% visibility. For a detector with $\eta < 1$ that obeys our model, the rate drops from $[\eta/2 + (2\eta - \eta^2)/4]$ to $[(2\eta - \eta^2)/2]$, for a visibility $V_s = \eta/(4 - \eta)$. For a realistic coincidence visibility less than 100%, we expect the singles visibility to fall off linearly with the coincidence visibility, V_c: $V_s = V_c \eta/(4 - \eta)$.

In order to investigate the singles effect and its dependence on the collection efficiency, four runs were performed: three with different neutral density (ND) filters in front of one detector (Bob), and one with no ND filter. By lowering the transmission of the ND filters, the overall collection efficiency can be lowered. The three ND filters allowed for 80%, 57%, and 27% transmission of light to give us a broad range of efficiencies. The irises in front of the detectors had 8-mm diameters. The data were accumulated by scanning the quartz delay over the interference dip many times. An individual scan counted singles and coincidences at 280 evenly spaced delay settings for 1 second per point. Typically 100 such 5-min scans were repeated for up to 20 h, in alternating directions to accumulate good statistics and minimize systematic errors due to laser drift. Over the 15 hours of data taking for the experiment with no ND, Alice and Bob's singles rates dropped by 10%.

For the experiment with no ND filter and with 8-mm irises, the singles rates for Alice and Bob were 13 600 s^{-1} and 13 800 s^{-1}, respectively; the coincidence rate was 510 s^{-1}. By removing the down-conversion crystal, the backgrounds at Alice and Bob were measured to be 1415 s^{-1} and 2006 s^{-1}, respectively. Not including potential background from the crystal itself, these background measurements are an overestimate. The crystal has imperfect transmission and will attenuate both the pump laser and the on-axis background. The systematic error introduced by performing the measurement in this way could be as large as 20%, which is much higher than the Poissonian counting statistical error. Based on the singles rates at the beginning of the data collection, the background level was measured to be 9% for Alice and 12% for Bob.

The coincidence and singles counts were binned and summed based on their delay position to create the raw ex-
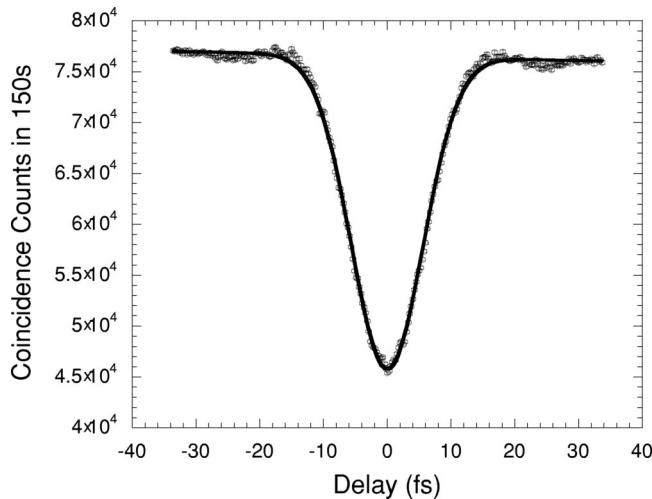
FIG. 2. Typical plot of coincidence rates versus the relative time delay. The Gaussian fit yields a visibility of $(39.39\pm0.05)\%$.
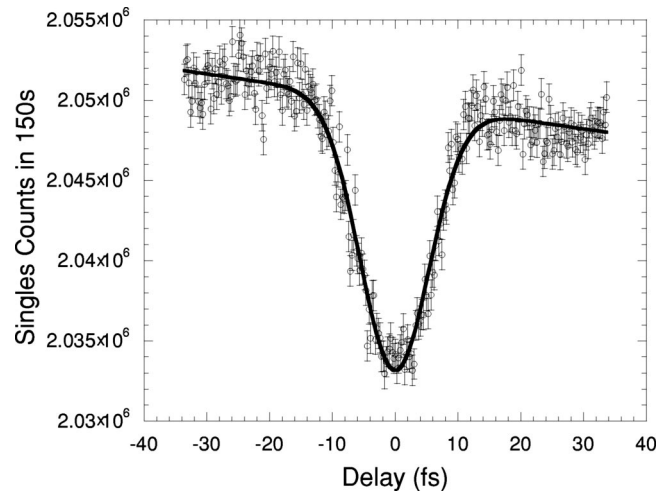

FIG. 3. Typical plot of the singles rate versus the relative time delay. The Gaussian fit yields a visibility of $(0.816\pm0.009)\%$. When background is included, the visibility becomes $(0.90\pm0.01)\%$, based on statistical errors only. With the measured efficiency for Alice of $(8.4\pm1.1)\%$, including the systematic error, our model predicts a visibility of $(0.84\pm0.11)\%$.

perimental data. The raw data contained fringes with an 800-nm spacing, which are due to a classical interference effect and have a visibility of approximately 1%. These fringes are caused when the optic axis of the down-conversion crystal is not in the same plane as the optic axis in the quartz delay. Since the period of these fringes is faster than the time scale for the HOM dip (2.6 fs versus 20 fs), these effects can be easily distinguished. The fringes were removed by averaging pairs of data points separated by half the classical fringe period. We fit the data to a Gaussian plus a straight line, to account for residual effects of drift or non-uniform transmission through the quartz prisms. Although the detailed shape of the dip depends on characteristics of the ultrafast pump pulse, and approaches a triangle for a cw pump, the Gaussian fits our data well, and minimizes the number of free parameters. This allows us to achieve the best possible precision in comparing visibilities of different curves (whose shapes we have independently verified to be identical to within measurement uncertainties).

Prior to the data accumulation, the efficiency of detection was measured. This efficiency is the product of the actual intrinsic quantum efficiency of the photodetector and the path efficiency of the interferometer. The path efficiency includes losses due to finite-sized irises, imperfect transmission through optics, and the ND filters, but *not* the effect of the PBS. In order to determine the efficiency η_A of one detector, the iris in front of the other detector is closed to 1–2 mm [11,6]. The efficiency is then $\eta_A=2C_{AB}/(S_B-B_B)$, where C_{AB} is the coincidence rate between Alice and Bob, S_B is the singles rate at Bob, B_B is the background rate at Bob, and the factor of 2 compensates for the PBS.

Coincidence and singles rates are shown in Figs. 2 and 3 as a function of the relative time delay, for the data run with no ND filters. The data have been binned and summed according to the relative time delay, and the fringes have been removed by the averaging method previously mentioned. The coincidence visibility is $(39.39\pm0.05)\%$, and the raw singles visibility at Alice is $(0.816\pm0.009)\%$, where the errors reflect only the statistical fitting uncertainties. Correct-

ing for the $(9.4\pm1.9)\%$ background that we measured at Alice the corrected visibility is $(0.90\pm0.01)\%$. The efficiency was measured to be $(8.4\pm1.1\%)$, the large uncertainty being due to the sensitivity to background when small irises are used. Our model predicts a visibility of $(0.85\pm0.11)\%$ based on the uncertainty in the efficiency measurement. The centers and widths of the coincidence dip and Alice's singles rate dip agree to within just over one standard deviation.
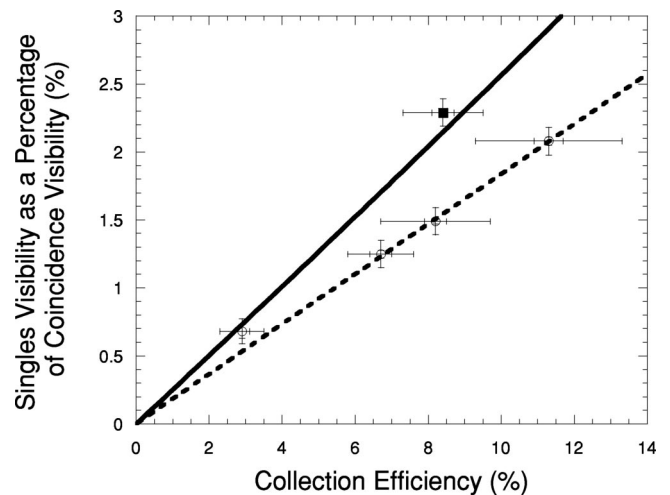

FIG. 4. Plot of the ratio of singles visibility to coincidence visibility versus the collection efficiency. The solid square is a data point from Alice's singles visibility, and the open circles are from Bob's singles visibility. The solid line is the prediction of our naive model. The dashed line is a linear fit with a zero y-intercept to the data for Bob based on statistical errors only. The data are shown with two sets of error bars. The smaller error bars represent statistical uncertainties only, and the larger set includes both the statistical and the systematic errors.

Since the expected singles visibility scales linearly with the coincidence visibility, Fig. 4 shows their ratio versus efficiency. Both the singles visibility and the efficiency are corrected for background. There are two sets of error bars plotted on each data point. The smaller set of error bars represents the statistical errors only. The larger set of error bars includes the systematic error due to the uncertainty in the background. The singles visibility at Bob is linearly proportional to the efficiency. The four data points for Bob fall on a straight line (dashed line) with a zero y-intercept with a χ^2 of 5.6 based solely on the statistical errors. The slope differs from that of our theoretical curve (solid line), but by an amount attributable to systematic uncertainty in background. We have made preliminary measurements that suggest that we are missing much of the background contribution to the singles rates at both of our detectors. This is due to fluorescence background created in the down-conversion crystal or in its antireflection coatings, and is difficult to isolate from the signal itself. Refinement of background measurement techniques is an important subject for future work. It should also be mentioned that, since the background measurement is incorporated into the values for the efficiency and for the singles visibility, a higher background will increase both values, but by different amounts.

We have demonstrated a quantum effect on the counting rate of a single-photon counter, showing that its effective efficiency is a function of the quantum statistics of the incident beam. A simple model for this effect is roughly consistent with the data; however, it will be necessary to refine our estimates of various background sources in order to further pursue quantitative tests of this model. By adjusting the correlation properties of the incident beam, we are able to probe the response of silicon avalanche photodiodes to photon pairs (as compared with single photons) on time scales of the order of 5 fs, and for delays up to 60 fs. Since energy and momentum relaxation time scales in silicon are on the order of 100 fs, and the incident photons may be prepared in identical quantum states (with focal spots on the order of 180 μm), a quantum-mechanical description of the microscopic detection process might be expected to be necessary in this regime. At present, the dependence of this effect on detector efficiency agrees qualitatively with a simple classical model, and the agreement can be made quantitative when a 20% systematic uncertainty in background is included. For pairs of photons separated by delays ranging from 20 fs to 60 fs, well outside the interference region, no systematic variation in the counting rate was observed at the 0.1% level. Over the 10-fs width of the quantum interference pattern, the shapes of the singles and coincidence pattern are the same to within 2.5%. Thus, quantum corrections to our model of the detector response to photon pairs appear to be quite small in the regime probed so far. By better characterizing the backgrounds and by extending these studies to other temporal or spatial scales or other detectors, the type of system presented here should provide a potentially powerful tool for studying decoherence and the measurement problem at ultrafast time scales in a variety of physical systems.

[1] R.J. Glauber, in *Quantum Optics and Electronics*, edited by C. DeWitt, A. Blandin, and C. Cohen-Tannoudji (Gordon and Breach Science Publishers, New York, 1965), pp. 63–185.

[2] R. Hanbury Brown and R.Q. Twiss, Nature (London) **177**, 27 (1956).

[3] R.Y. Chiao, P.G. Kwiat, and A.M. Steinberg, in *Advances in Atomic, Molecular, and Optical Physics*, edited by B. Bederson and H. Walther (Academic Press, New York, 1994), Vol. 34.

[4] X.Y. Zou, L.J. Wang, and L. Mandel, Phys. Rev. Lett. **67**, 318 (1991).

[5] A.W. Lightstone, A.D. MacGregor, D.E. MacSween, R.J. McIntyre, C. Trottier, and P.P. Webb, Electron. Eng. (U.K.) **61**, 37 (1989).

[6] P.G. Kwiat, A.M. Steinberg, P.H. Eberhard, and M.D. Petroff, Appl. Opt. **33**, 1844 (1994).

[7] G. Di Giuseppe, L. Haiberger, F. De Martini, and A.V. Sergienko, Phys. Rev. A **56**, R21 (1997).

[8] Y.H. Shih and A.V Sergienko, Phys. Rev. A **50**, 2564 (1994).

[9] C.K. Hong, Z.Y. Ou, and L. Mandel, Phys. Rev. Lett. **59**, 2044 (1987).

[10] C.H. Monken, P.H. Souto Ribeiro, and S. Padua, Phys. Rev. A **57**, R2267 (1998).

[11] D.N. Klyshko, Sov. J. Quantum Electron. **10**, 1112 (1980).