OPTICA

Check for updates

Single photon multiclock lock-in detection by picosecond timestamping

LUKAS A. JAKOB,¹ WILLIAM M. DEACON,¹ OLIVER HICKS,² ILYA MANYAKIN,^{1,3} Oluwafemi S. Ojambati,¹ Michael Traxler,⁴ and Jeremy J. Baumberg^{1,*}

¹Nanophotonics Centre, Cavendish Laboratory, University of Cambridge, Cambridge CB3 0HE, UK

²Cambridge Consultants, Cambridge CB4 0DW, UK

³National Physical Laboratory, Hampton Road, Teddington TW11 0LW, UK

⁴GSI Helmholtzzentrum für Schwerionenforschung GmbH, Planckstraße 1, 64291 Darmstadt, Germany

*Corresponding author: jjb12@cam.ac.uk

Received 24 August 2021; revised 3 November 2021; accepted 8 November 2021; published 20 December 2021

Extracting signals at low single-photon count rates from large backgrounds is a challenge in many optical experiments and technologies. Here, we demonstrate a single-photon lock-in detection scheme based on continuous photon times-tamping to improve the SNR by more than two orders of magnitude. Through time-resolving the signal modulation induced by periodic perturbations, 98% of dark counts are filtered out and the <1 count/s contributions from several different nonlinear processes identified. As a proof-of-concept, coherent anti-Stokes Raman measurements are used to determine the vibrational lifetime of few molecules in a plasmonic nanocavity. This detection scheme can be applied to all single-photon counting experiments with any number of simultaneous modulation frequencies, greatly increasing SNR and resolving physical processes with picosecond time resolution while keeping the photon dosage small. The open instrumentation package provided here enables low-cost implementation.

Published by The Optical Society under the terms of the Creative Commons Attribution 4.0 License. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

https://doi.org/10.1364/OPTICA.441487

1. INTRODUCTION

Transient or time-resolved optical experiments often require elaborate experimental setups designed to measure very low signal intensities over ultrashort time scales; hence, they often suffer from poor signal-to-noise ratios (SNRs) [1]. Typically, the linear response dominates over any perturbation, giving small induced changes in the signal. Increasing the strength of a repetitive perturbation (for instance, the optical pulse intensity) to enhance the nonlinear signal is often problematic since this can damage the samples, preventing stroboscopic measurement. For probing single nanostructures or individual quantum systems, the strong perturbation needed to obtain clear signals unfortunately often induces irrevocable structural changes such as bond cleavage [2], atomic displacements [3], reshaping [4], or ablation [5].

A technique commonly applied to extract such weak signals from a noisy background is lock-in detection. By introducing a modulation to the sample, the amplitude and phase of the emerging signal can be determined using phase-sensitive heterodyne detection while noise at other frequencies is rejected [6]. For instance, in all-optical experiments such as four-wave-mixing in semiconductor optical amplifiers [7] or stimulated emission from single nanocrystals [8], the pump pulse train is amplitudemodulated at high frequency f_{mod} . In scanning near-field optical microscopies, this modulation is provided by the vibration of a tip above the sample [9]. Optical lock-in detection is also used for

2334-2536/21/121646-08 Journal © 2021 Optica Publishing Group

stimulated-emission-depletion microscopy to enhance contrast in fluorescence microscopy [10], and in many other scenarios.

Several versions of phase-sensitive detection have been implemented for single-photon counting detectors [11–16]. However, these rely on binning photon counts in successive time intervals to create a continuous signal for traditional lock-in analysis instead of analyzing each photon individually [Fig. 1(a)]. The detection challenge is to identify ~ 1 cts/s (counts per second) caused by the linear or nonlinear optical processes of interest under a large background. Here, we present a new approach to single-photon lock-in detection utilizing continuous picosecond photon timestamping of each individual photon to temporally resolve the evolution of an optical signal undergoing modulation. By recording the arrival time of each signal photon at the detector alongside timestamps from synchronized reference clocks [Fig. 1(c)], the time dynamics of the signal is extracted. For each photon, the phase (τ_i) and frequency (f_i) of each reference clock at the photon arrival time (t_0) is accurately determined from a linear fit to the timestamp data of the previous N clock detections [Fig. 1(d)]. In the post-experimental data analysis, experimental noise and background signals can be easily removed through temporal gating, implementing lock-in amplification individually for every photon and each reference clock, suitable to increase the SNR of single-photon counting experiments by many orders of magnitude (here with a factor >100).



Fig. 1. Comparison of single-photon detection techniques. (a) Previous single-photon lock-in techniques: photons are counted in time bins to recover an analog re-synthesised signal for analysis with a standard phase-sensitive lock-in algorithm. (b) Time-correlated single-photon counting: the time interval Δt between a trigger pulse and photon detection events is recorded to resolve the time decay of an optical signal, typically fluorescence. (c) and (d) Multiclock single-photon lock-in technique developed here: continuous timestamping of photon arrival times (t_0) and multiple reference signals with frequency f_i allows identification of the phase τ_i of every individual photon compared to all the reference frequencies.

The signal-processing technology to convert the arrival time of an electronic pulse to a digital timestamp is well established and widely used in high-energy physics [17,18]. Field programmable gate array (FPGA) boards can perform logical operations with digital electronic signals in real time with bandwidths exceeding 100 MHz. Time-to-digital converters (TDC) implemented with FPGAs can now achieve timing precision <10 ps [19] while it is possible to parallelize devices using 264 channels or more on one board [20].

Combining picosecond timestamping with pulsed optics holds enormous potential to improve existing and enable new applications. Although the proposed experimental design can increase the SNR in all single-photon experiments, it is particularly advantageous in the areas of quantum correlation, time-of-flight spectroscopy, and scanning near-field microscopy. A particular new capability provided by continuous photon timestamping is the ability to compare photon signals to multiple reference clocks at the same time. For instance, here this allows single-photon lock-in synchronization simultaneously to: (1) the optical pulse repetition rate, (2) the laser power modulation (here, two different on-off periodic modulations), and (3) extra triggers (here, the pulse delay scan). This retrieves the maximum possible information content of each detected photon. In comparison, start-stop photon detection schemes for time-correlated signal photon counting (TCSPC) synchronize only to optical pulses, greatly limiting their use [Fig. 1(b)] [21]. We also note some similarities to lidar technology, where multiple laser repetition rates are used to avoid distance ambiguity. However, these measurements are performed in quick succession; instead here, we resolve a signal modulating with several frequencies at once in our setup.

In this paper, coherent anti-Stokes Raman spectroscopy (CARS) is used as an example of a typical nonlinear experiment with low single-photon count rates that benefits from our scheme [22,23]. We thus briefly describe the measurements, while noting the general applicability of the technique. In CARS, molecular vibrations v are excited by coherently pumping with two laser pulses [pump ω_p and Stokes ω_S , Figs. 2(a) and 2(b)], tuned so that their frequency difference matches the vibration, $\omega_p - \omega_S = v$. Subsequently, anti-Stokes scattering to $\omega_x + \nu$ of a time-delayed separate probe pulse at ω_x is used to identify the relaxing molecular vibrations and measure their vibrational lifetime. To access a domain that can observe CARS from single molecules [22], nanoscale optical confinement is needed and is accessed here using plasmonic nanocavities [24]. These are based on a nanoparticleon-mirror geometry (NPoM) in which an ordered molecular monolayer (here, biphenyl-4-thiol, BPT) is sandwiched in a nmthick gap between Au facets [25]. A background of nonresonant four-wave-mixing dominates the detected signal at $\omega_x + \nu$ [3,26] while additional noise is introduced by electronic dark counts and stray light. Our detection scheme isolates <1 cts/s rates of this nonlinear CARS signal using picosecond photon timestamping through a low-cost open-architecture FPGA board.

2. EXPERIMENTAL SETUP

The basis for detecting a nonlinear optical signal above other contributions is to identify and subtract the linear components and separate the multitude of other nonlinear signals. Here, this is achieved by modulating the two laser beams exciting the sample [Fig. 2(a)]. For CARS experiments, each of the three laser beams contributes to the total recorded count rate [Fig. 2(c)]. Pump and



Fig. 2. (a) Schematic experimental setup of CARS. Three pulsed lasers with a repetition period $T_{\rm rep}$ are focused on a nanoparticle-on-mirror (NPoM) nanocavity sample, with both Stokes and probe modulated periodically with period $T_{\rm mod}$. The emerging signal detected by a SPAD contains photons from the nonlinear CARS process, which must be isolated from other contributions. (b) Excitation scheme of CARS, with pump and Stokes exciting a molecular vibration, probed by anti-Stokes scattering of the probe beam with time delay Δt . (c) Modulation of the detected SPAD signal. Only when all three beams are turned on can the nonlinear three-pulse CARS signal be observed (green).

probe pulses generate broadband contributions across the detected spectral window through electronic anti-Stokes Raman scattering. With pump and Stokes beams exciting the sample, a two-pulse CARS process is also possible but cannot be time-resolved since the pump then both excites and probes. Only all three pulses together excite emission of the desired three-pulse CARS signal, enabling investigation of the ultrafast vibrational dynamics. In systems with a high signal and stability, these contributions can be separated by sequential acquisition of spectra without fast laser modulation. However, with a fixed photon budget on samples such as nanoscopic structures and single molecules, a new approach is required to increase the SNR and avoid damage.

A widefield microscope guides these pulses onto the nanocavity sample (Fig. 3), which are generated at 820 nm (Stokes), 726 nm (pump), and 722 nm (probe) from an $f_{rep} = 80$ MHz pumped Spectra-Physics optical parametric oscillator (OPO). The three lasers are spectrally tuned so that pump and Stokes pulses can resonantly drive the molecular vibration while the Stokes pulse is off-resonant to vibrations excited by the probe pulse (see Supplement 1, Fig. S2a). With ultrafast pulses of 500 fs duration, the spectral resolution of the exemplar experiment here is 50 cm^{-1} . Stokes and probe beams are each modulated with electro-optic modulators (EOMs) driven by function generators (HP 33120A) producing a square wave output at variable frequency f_{mod} . The two function generators are phase locked and operate with a fixed phase difference of 90°. For CARS experiments, probe beam pulses are delayed by Δt with respect to the other two laser beams using a delay stage running in a continuous loop around $\Delta t = 0$ $(f_{\text{delay}} = 1 \text{ Hz})$. The Stokes pulses are temporally aligned to the pump pulses (by optimizing the two-pulse CARS signal) initially, after which this coincident relative delay is fixed. All three beams are spatially overlapped and co-focused on the sample.

In the detection path, the laser light is blocked using spectral filters and a single-photon avalanche diode (SPAD) from Micro Photon Devices (\$PD-100-CTD) detects the signal photons. Each arriving photon leads to an electronic pulse with a timing accuracy of 35 ps. Typical count rates in our experiment range from 1 to 1000 cts/s obscured by at least 100 cts/s of dark counts and stray light.

In addition to the single-photon counts, three reference signals are recorded. A fast photodiode monitors the pulse repetition rate of the Ti:Sapphire pump laser (Spectra-Physics Mai Tai, photodiode integrated into laser head) producing a digitally conditioned $f_{\rm rep} = 80$ MHz pulse train. Together with a synchronized TTL square wave at $f_{\rm mod}$ and TTL trigger pulses at the beginning of each delay stage loop ($f_{\rm delay}$), all electronic signals are passed to the FPGA board (Fig. 3). This board continuously converts the



Fig. 3. Experimental setup for CARS experiments with photon time tagging. Three laser beams are produced by an 80 MHz pulsed pump laser and an OPO. Two of these beams are 50 kHz modulated by EOMs driven with square waves. The probe beam delay is continuously varied by moving a mirror, giving trigger synchronization pulses at the beginning of each loop. All beams are spatially overlapped and focused onto the sample in a microscope. A SPAD detects the spectrally filtered signal photons. A reference signal for the pulsed laser is provided by a fast photodiode (PD). The electronic signals shown as V(t) are passed to a central FPGA board, where the signal and clock timestamps are recorded digitally.



Fig. 4. Time-to-digital conversion with an FPGA. (a) Timestamps with ps precision are recorded by combining a fine TDC and a coarse counter. Data is streamed to the computer via gigabit Ethernet (GbE) when triggered by a signal photon. (b) In the fine TDC, an electronic signal is carried along a tapped delay line connected to a tap register. When the signal is stopped by the FPGA clock, the register entry is recorded.

arrival times of all electronic signals to digital timestamps that are streamed to a computer and saved. This time-to-digital conversion is achieved by combining a fine time-to-digital converter (TDC) and a 32-bit coarse counter [Fig. 4(a)]. The fine TDC consists of a tapped delay line that provides 30 ps accuracy with a range of 5 ns. The coarse counter simply counts increments of the 200 MHz internal FPGA clock and hence covers the nanosecond to second regime.

In a tapped delay line (as commonly applied in time-to-digital conversion [18]), each element is connected to a register that records the state of the carried signal and hence the current position of the signal within the carry line [Fig. 4(b)]. To determine the photon arrival time, an incoming electronic signal launches a pattern that travels along the delay line until it receives a stop signal that causes the position of the signal to be read from the register. Since the stop signal is provided by the FPGA clock, the arrival time of the signal is the time the signal spent propagating along the delay line before the latest clock. The propagation distance is converted to time in post-experimental data processing.

Since the laser pulse reference clock is at 80 MHz, recording every pulse arrival time would require a prohibitive data transfer rate >25 GB/s. Therefore, a trigger system was implemented to only record reference timestamps when a signal photon is detected: Every time a photon arrives at the SPAD, a defined (and tunable) number of most recent timestamps from all reference clocks are sent to the host PC by the readout controller alongside the photon timestamp. This allows the user to control the number of reference timestamps recorded for each signal event, and thus optimize the SNR by maximizing the accuracy of the laser reference given the maximum data transfer rate at each SPAD count rate. As we show below, even a stable repetitive laser system experiences random variations of <0.1% in cavity length, which, if not tracked, greatly reduce the timing precision. Here, ten timestamps per reference clock were sufficient to reach the optimum time resolution and thus minimize the required data size.

The TDC was implemented on a Digilent Arty Z7 development board specifically programmed for this application. We provide the FPGA software online [27] under an open source license to allow for low-cost implementation. Here, we demonstrate the function of this system with three reference clocks and one signal channel. However, the FPGA board supports up to eight input signals that can be assigned to record either the reference or signal channels, allowing for more complex experimental setups.

3. RESULTS

With this setup, data files are acquired containing a list of arrival timestamps. For the first demonstrations here, two reference clock signals are chosen: f_{rep} and f_{mod} . In general, any periodic reference signal can be chosen as a clock enabling a plethora of different applications. As an initial calibration experiment, a single laser pump beam modulated at $f_{mod} = 1$ kHz is focused on the sample and spontaneous Stokes scattering from the nanocavity is recorded by the SPAD (see Fig. S1 in Supplement 1 for the spectrum detected).

From the periodic clock timestamps, the frequency f of the clock at the time of arrival of each photon at the detector is determined. For each of the detected photons, the clock frequency slightly varies due to fluctuations in the period, as seen in the histograms of Figs. 5(a) and 5(b). Over this 100 s data set, the average pulse repetition rate of the laser was $f_{\rm rep} = 79.64$ mHz with a standard deviation of 25 kHz. For the EOM modulation of the laser, a frequency of 1.00001 kHz is found ± 0.78 MHz as expected for this source. These results vary little between successive data sets, demonstrating the accuracy of our global clock.

The point in the clock cycle τ when each signal photon was detected is now also calculated. Extracting this value for all photons detected during a measurement allows us to reconstruct the signal modulation through the clock cycle, of a length set by the inverse of the clock frequency. This projects all counts into a single clock cycle; hence, repetitive but extremely low count rates can be analyzed by simply increasing the integration time.

We first apply this temporal reconstruction to the periodically pulsed laser. Using the laser repetition rate to wrap the signal within the $T_{rep} = 12.5$ ns long clock period places most counts near a specific time that depends on a system electronic delay from optical paths, cables, and latency [Fig. 5(c)]. All photons emerging from the sample due to excitation by the periodic laser pulses are detected within a $\Delta \tau \sim 200\,$ ps wide window because Raman is a prompt process. Even though the photons are emitted within the 500 fs optical pulse width, this peak is broadened by the detection electronics. In contrast, electronic dark counts and stray light photons are uncorrelated with the excitation pulses and hence give a constant background signal spread over the whole period. Exploiting this property of dark counts allows us to distinguish them from photons emitted by the sample and remove them from the data. This results in a >98% reduction in dark counts and thus an increase of $T_{\rm rep}/\Delta\tau > 60$ in the SNR for low count rate experiments [Fig. 5(c)]. The approach to remove dark counts in the time domain is similar to previous reports of time-gated singlephoton counting, which has been demonstrated both with actively quenched detectors [28-30] and digitally by TCSPC [31].

The same concept of temporal reconstruction is now applied to determine which lasers excited the sample when each signal photon



Fig. 5. Analysis of timestamp data from spontaneous Raman scattering. (a) and (b) Histograms of the clock frequency f for (a) the laser pulses and (b) the EOM modulation, over 100 s. (c) Reconstruction of the laser pulse shape discriminating signal (green) and dark counts (gray). Inset: 0.2 ns time window $\Delta \tau$ including 60% of signal counts (dashed). Bin size 25 ps. (d) Reconstruction of the laser modulation cycle, recording Raman scattering from a nanocavity with single laser pump turned off and on (laser power 2 μ W). Bin size 2 μ s.

was generated. This is crucial to separate the linear and nonlinear components of the signal. To demonstrate this functionality in our calibration case, the single laser pump is switched on and off at frequency 1 kHz and 50% duty cycle while recording Raman scattering from the nanocavity sample (average laser power on sample 2 μ W). Consequently, during the laser modulation cycle, two temporal regions with constant count rates are observed, with the background for laser-off being simply identified [Fig. 5(d)].

When the power of the laser is reduced by nearly a hundredfold (50 nW on the sample), the modulation depth of the signal is greatly reduced since the dark count rate is now larger than the Raman signal [Fig. 6(a)]. Dark counts outside the 200 ps window centered on the laser pulse in Fig. 5(c) are now removed, reducing the count rate outside the laser window to close to zero [Fig. 6(b)] with only a residual 2% of unfiltered dark counts remaining. This dark count rejection increases the modulation visibility by 400%, but in measurements with even lower laser power or more stray light/dark counts, this enhancement can exceed 4000%. Analyzing the distribution of the noise in the modulation reveals two Poisson distributions for the on and off states of the laser [Fig. 6(c)]. These arise from the single-photon counting statistics in the experiment and show that the SNR is now only limited by the photon shot noise, which can thus be improved by increasing the integration time to collect more photons.

To demonstrate the ability of the setup to detect a small nonlinear signal upon a large background, CARS experiments are then carried out. The detected CARS spectra are shown in Fig. S2b in Supplement 1. Since two beams are now modulated, each at $f_{\rm mod} = 50$ kHz (to improve the SNR) but with one phase shifted by 90°, and with the addition of a third beam of constant intensity, the signal modulation shows four distinct windows of different



Fig. 6. Demonstration of dark count removal. (a) Reconstruction of laser-modulated spontaneous Raman scattering from a nanocavity at low average laser power (50 nW) over 100 s, with a Raman signal smaller than the dark count rate. (b) Reconstruction of modulated signal after dark count removal. Bin size for (b) and (c) 2 μ s. (c) Distribution of noise in (b), with fitted Poisson distributions (lines) indicating the single-photon statistics of the experiment.

heights [Fig. 7(a)], as expected from the experimental design [Fig. 2(c)]. Dark counts in the CARS experiments throughout the entire modulation period are removed as detailed above. The measured count rate is highest when all three beams illuminate the sample and lowest when only the pump beam is on. Analyzing the different windows allows contributions from each individual laser combination to be determined [yellow, orange, and red in Fig. 7(a)]. Subtracting these values from the count rate when all lasers illuminate the sample allows the nonlinear three-pulse CARS count rate to be extracted [green in Fig. 7(a)].

With this method to extract the nonlinear signal, time-resolved CARS measurements can now be performed by delaying the probe pulse compared to the pump and Stokes pulses. This is achieved by mechanically scanning a delay line for the probe pulse. Conventionally slow scans are performed, integrating until a sufficient SNR is achieved at each time point; however, this produces strong artifacts in the delay scan due to transient changes in the emission spectrum (here caused by movement of Au atoms on the nanoparticle facet [32,33]) and damage to the nanostructure. Hence, the delay stage is instead continuously scanned back and forth at $f_{delay} = 1$ Hz and a third reference trigger is introduced into the FPGA from the scanning delay line. This allows each photon detected to also be tagged with the probe time delay at which it was measured, thus building up the entire time-delay curve simultaneously, without any artifacts.

Time-resolved tracks for all contributions identified by the laser modulation are compared in Fig. 7(b). The two-pulse CARS signal induced by pump and Stokes beams (red) stays constant, as well as the electronic Raman scattering from the pump (orange). While electronic Raman scattering from the probe also leads to a constant signal, vibrational pumping by surface-enhanced Stokes scattering of pump and probe photons adds a time-dependent



Fig. 7. Time-resolved CARS experiments. (a) CARS signal during 50 kHz laser modulation period (dark counts removed). Contributions from the different lasers are marked in orange (pump), red (Stokes), and yellow (probe), with the nonlinear CARS signal in green. Bin size 40 ns. (b) Extracted contributions from (a), locked-in to the delay stage sweep as the probe (Pr) pulses are delayed with respect to pump (Pu) and Stokes pulses. (c) Time-resolved CARS signal with exponential decay fit to extract a vibrational lifetime of 1595 fs of the BPT molecules in the nanocavity. For negative delays, the signal follows a Gaussian, matching the 500 fs pulse width. Average laser power on sample per beam is 0.2 μ W.

signal to the yellow contribution in Fig. 7(a). This signal decays equally to both positive and negative delays as the pump or probe excite the sample identically. On the other hand, the three-pulse CARS count rate decreases for both positive and negative delay, but it is not symmetric around 0 ps [Fig. 7(c)]. When the probe pulse arrives before the molecules are excited, the signal vanishes quickly with a rise time corresponding to the pulse length of 500 fs [orange dashed, Fig. 7(c)]. For probe pulses arriving after pump and Stokes, the signal decreases exponentially. From an exponential fit, the lifetime of the 1585 cm⁻¹ vibration of BPT is estimated as 1600 ± 400 fs. This signal is emitted from only an estimated 100 molecules in the nanocavity gap, billions of times fewer than for solution measurements. Previous attempts to measure this without single-photon detection required average laser powers >4 μ W per beam, at least tenfold more than in this photon-counting lock-in mode, and which is enough to perturb and destroy the nanocavity structures. With safe powers of $I = 0.2 \mu W$ per beam employed here, the CARS signal is 1000-times smaller (since it scales as I^{3}) and is below the limits of integrating detectors. As the signal strongly varies from nanoparticle to nanoparticle, further ongoing experiments and theory are required for a full analysis, but are beyond the scope of this article.

4. DISCUSSION

To quantify the improvements made by eliminating dark counts, the window (within which counts are not removed) is centered on the pulse and increased in width [Fig. 5(a) inset]. For each width, the percentage of dark counts that are removed is calculated as well as the photon events preserved [Fig. 8(a)]. A window width



Fig. 8. Characterization of dark count removal while varying filter window width $\Delta \tau$ [see Fig. 5(c)]. (a) Percentage of signal counts preserved and dark counts removed for increasing window widths. (b) Ratio of signal counts to dark counts, compared to unfiltered data (horizontal line at 1.3).

of $\Delta \tau = 200$ ps retains ~60% of real photon events and only 1.6% of all dark counts. To remove a higher percentage of dark counts, the window width must be reduced, thus decreasing the signal counts. The ratio of signal counts to dark counts [Fig. 8(b)] shows that by filtering the dark counts, this ratio can be increased by almost two orders of magnitude from 1.3 in the unfiltered data to >100. Since a compromise between the best SNR and maximum retained signal counts is demanded, we thus choose a window width of 200 ps, which retains the signal counts most efficiently [before the green curve in Fig. 8(a) saturates] and increases the SNR by 60.

Currently, limitations of the timing precision in this setup are observed in the distorted pulse shape. Even though the laser pulses are <1 ps, the reconstructed pulse shape [Fig. 5(c)] has a width of \sim 200 ps. The SPAD has a nominal jitter of only 35 ps and the FPGA board has a timing precision of 30 ps (see Fig. S3), but further inaccuracies are introduced by the fast photodiode (internal to Spectra-Physics Mai Tai pump laser) and amplification of the MHz clock signal by two amplifiers (Mini-Circuits ZFL-1000LN). Additionally, the detector electronic response adds a shoulder to the peak, thus decreasing the percentage of preserved signal counts after filtering out dark counts. Improvements in the detection electronics can thus further increase the fraction of dark counts removed by this technique. For instance, without FPGA noise and reaching the SPAD resolution of 35 ps would give another \sim 10-fold improvement.

Here, we presented a scheme of modulated lasers for the example of CARS. Typically, laser modulation and lock-in detection is also used for stimulated Raman scattering (SRS). Since SRS requires the detection of small changes in the pump laser intensity, the detected powers are far above the single-photon regime. In an optimized configuration, our FPGA setup can only record up to 1 Mcts/s count rates (pW) and is therefore not suitable for SRS without upgrading the electronics to deal with much higher count rates.

To highlight the potential of the presented technology, Fig. 9 compares different photon detection techniques. Depending on the photon count rates, different photodetectors must be selected. For single-photon experiments, SPADs are suitable, while traditional photodiodes (PD) are needed for spectroscopies delivering higher light intensities (>nW). In between, avalanche photodiodes (APD) and photomultiplier tubes (PMT) provide detection of photocurrents with high gain.

The crucial experimental SNR strongly depends on signal detection and amplification. A SPAD delivers one voltage pulse for



Fig. 9. Comparison of different photon detection techniques. SPADs, APDs, and PMTs can detect single-photon count rates whereas PDs operate at higher light intensities. The SNR of the detectors can either be improved by lock-in amplifiers in the high signal regime, or with the photon timestamping setup (SPAD + timetag) presented here for single-photon detection, which can be extended to higher count rates using a combination of several SPADs (multiSPAD + timetag). A detailed description of the comparison is provided in Supplement 1.

every detected photon and thus the only source of noise is photon shot noise and dark events. For photodiodes, where an electrical current induced by light is produced, additional noise sources include the electronic shot noise due to the diode dark current and thermal detector noise as well as noise from amplifiers necessary to record the small output currents.

For photodiodes, a typical way to improve this SNR is to modulate the excitation light source (or sample perturbation) at a fixed frequency and use a lock-in amplifier to record the amplitude of the modulated detected photocurrent. Analogously, our FPGA system can enhance the SNR in single-photon experiments by more than two orders of magnitude. By resolving this signal modulation, it is possible to filter out unwanted background photons and those from other contributions. In the experiment presented here, we demonstrate reliable detection of three-pulse CARS at count rates of 1 cts/s within a background of more than 300 cts/s. This ratio of >10² is remarkable for a single-photon experiment, but it can increase even further for applications with high backgrounds such as stray light.

Improving the SNR even further would be possible by enhancing the timing precision to narrow the electronic pulse as discussed above. Increasing the detector count rates beyond the saturation of a single SPAD can be handled by splitting the light intensity over multiple SPADs and connecting them to different channels of the FPGA. In combination with a spectrometer grating, a SPAD array could even then resolve the spectral dependence of the signal.

Here, we have demonstrated the working principle of our setup using laser pulses, laser modulation, and the delay stage sweep as reference clocks. However, any periodic signal can act as a reference, making our setup attractive for a wide range of experimental research fields. In particular the additional time resolution of the signal during a reference period can enable a plethora of new applications. In a separate study, we have applied this scheme to measure time-resolved perturbations of a plasmonic nanocavity by a mid-infrared laser [34]. Other applications include tracking the light emission from an optoelectronic device induced by an alternating voltage to characterize the response time of the device. In scanning near-field optical microscopy, an oscillating tip above the sample can provide a reference frequency, both drastically increasing the SNR and recording the signal as a function of tip-sample distance. Moreover, with a dispersive fiber, photons can be delayed depending on their color, enabling optical time-of-flight spectroscopy [35] with an 80 MHz lock-in frequency. Finally, the switching of photoactive molecules can be resolved in time with modulated lasers. We suggest that this is of particular interest for single-molecule fluorescence spectroscopy and microscopy.

In conclusion, we developed a method to separate different contributions to a signal by resolving the variation of the detected single photons over the period of multiple reference signals. The technique relies on continuously recording the arrival time of each photon at a single-photon detector with an FPGA board and comparing it to reference clock timestamps. With this setup, we reconstructed the periodic modulation of a sub-ps excitation laser, allowing 98% of stray light and dark counts to be filtered out. The capability of this method was demonstrated in a CARS experiment, where single photon per second count rates of a nonlinear signal were detected. Due to the high flexibility for different reference clock signals from Hz to MHz frequencies, this concept can be applied universally to all single-photon experiments, drastically increasing their SNR.

Funding. Engineering and Physical Sciences Research Council (EP/L015978/1, EP/L027151/1, EP/P029426/1, EP/R020965/1, EP/S022953/1); European Research Council under the Horizon 2020 Framework Programme (PICOFORCE (883703), POSEIDON (861950), THOR (829067)).

Acknowledgment. We thank Cambridge Consultants, who provided the resources and collaboration to develop this FPGA and also enabled an open-source implementation. We also thank NPL for funding (I.M.).

Disclosures. The authors declare no conflicts of interest.

Data Availability. Data underlying the results presented in this paper are available in Ref. [36].

Supplemental document. See Supplement 1 for supporting content.

REFERENCES

- 1. Y. R. Shen, The Principles of Nonlinear Optics (Wiley, 1984).
- A. Lombardi, M. K. Schmidt, L. Weller, W. M. Deacon, F. Benz, B. De Nijs, J. Aizpurua, and J. J. Baumberg, "Pulsed molecular optomechanics in plasmonic nanocavities: from nonlinear vibrational instabilities to bond-breaking," Phys. Rev. X 8, 011016 (2018).
- K. T. Crampton, A. Zeytunyan, A. S. Fast, F. T. Ladani, A. Alfonso-Garcia, M. Banik, S. Yampolsky, D. A. Fishman, E. O. Potma, and V. A. Apkarian, "Ultrafast coherent Raman scattering at plasmonic nanojunctions," J. Phys. Chem. C 120, 20943–20953 (2016).
- C. J. DeSantis, D. Huang, H. Zhang, N. J. Hogan, H. Zhao, Y. Zhang, A. Manjavacas, Y. Zhang, W. S. Chang, P. Nordlander, S. Link, and N. J. Halas, "Laser-induced spectral hole-burning through a broadband distribution of Au nanorods," J. Phys. Chem. C **120**, 20518–20524 (2016).
- M. Huang, F. Zhao, Y. Cheng, N. Xu, and Z. Xu, "Origin of laser-induced near-subwavelength ripples: interference between surface plasmons and incident laser," ACS Nano 3, 4062–4070 (2009).
- M. L. Meade, Lock-in Amplifiers: Principles and Applications, IEE Electrical Measurement Series (1983).
- P. Borri, J. Mørk, and J. M. Hvam, "Heterodyne pump-probe and fourwave mixing in semiconductor optical amplifiers using balanced lock-in detection," Opt. Commun. 169, 317–324 (1999).
- L. Piatkowski, N. Accanto, G. Calbris, S. Christodoulou, I. Moreels, and N. F. van Hulst, "Ultrafast stimulated emission microscopy of single nanocrystals," Science 366, 1240–1243 (2019).

- 9. Y. Inouye and S. Kawata, "Near-field scanning optical microscope with a metallic probe tip," Opt. Lett. **19**, 159–161 (1994).
- G. Marriott, S. Mao, T. Sakata, J. Ran, D. K. Jackson, C. Petchprayoon, T. J. Gomez, E. Warp, O. Tulyathan, H. L. Aaron, E. Y. Isacoff, and Y. Yan, "Optical lock-in detection imaging microscopy for contrast-enhanced imaging in living cells," Proc. Natl. Acad. Sci. USA **105**, 17789–17794 (2008).
- A. Lenhard, S. Zaske, and C. Becher, "Lock-in detection of single photons after two-step frequency conversion," Opt. Lett. 37, 4254–4256 (2012).
- F. T. Arecchi, E. Gatti, and A. Sona, "Measurement of low light intensities by synchronous single photon counting," Rev. Sci. Instrum. 37, 942–948 (1966).
- D. Braun and A. Libchaber, "Computer-based photon-counting lock-in for phase detection at the shot-noise limit," Opt. Lett. 27, 1418–1420 (2002).
- A. Restelli, R. Abbiati, and A. Geraci, "Digital field programmable gate array-based lock-in amplifier for high-performance photon counting applications," Rev. Sci. Instrum. 76, 093112 (2005).
- X. Wang, G. Zhang, R. Chen, Y. Gao, L. Xiao, and S. Jia, "Detection efficiency enhancement of single-photon detector at 1.55 μm by using of single photons lock-in and optimal threshold," Opt. Laser Technol. 44, 1773–1775 (2012).
- W. Chen, X. Wang, B. Wang, Y. Wang, Y. Zhang, H. Zhao, and F. Gao, "Lock-in-photon-counting-based highly-sensitive and large-dynamic imaging system for diffuse optical tomography," Biomed. Opt. Express 7, 499–511 (2016).
- J. Wu, Z. Shi, and I. Y. Wang, "Firmware-only implementation of timeto-digital converter (TDC) in field-programmable gate array (FPGA)," in *Proceedings IEEE Nuclear Science Symposium Conference Record* (2003), Vol. 1, pp. 177–181.
- S. Tancock, E. Arabul, and N. Dahnoun, "A review of new time-to-digital conversion techniques," IEEE Trans. Instrum. Meas. 68, 3406–3417 (2019).
- E. Bayer and M. Traxler, "A high-resolution (< 10 ps RMS) 48-channel time-to-digital converter (TDC) implemented in a field programmable gate array (FPGA)," IEEE Trans. Nucl. Sci. 58, 1547–1552 (2011).
- A. Neiser, J. Adamczewski-Musch, M. Hoek, W. Koenig, G. Korcyl, S. Linev, L. Maier, J. Michel, M. Palka, M. Penschuck, M. Traxler, C. Uğur, and A. Zink, "TRB3: a 264 channel high precision TDC platform and its applications," J. Instrum. 8, C112043 (2013).
- 21. D. O'Connor, *Time-Correlated Single Photon Counting* (Academic, 2012).
- S. Yampolsky, D. A. Fishman, S. Dey, E. Hulkko, M. Banik, E. O. Potma, and V. A. Apkarian, "Seeing a single molecule vibrate through timeresolved coherent anti-Stokes Raman scattering," Nat. Photonics 8, 650–656 (2014).

- T. Ichimura, N. Hayazawa, M. Hashimoto, Y. Inouye, and S. Kawata, "Tip-enhanced coherent anti-Stokes Raman scattering for vibrational nanoimaging," Phys. Rev. Lett. 92, 220801 (2004).
- J. J. Baumberg, J. Aizpurua, M. H. Mikkelsen, and D. R. Smith, "Extreme nanophotonics from ultrathin metallic gaps," Nat. Mater. 18, 668–678 (2019).
- F. Benz, C. Tserkezis, L. O. Herrmann, B. De Nijs, A. Sanders, D. O. Sigle, L. Pukenas, S. D. Evans, J. Aizpurua, and J. J. Baumberg, "Nanooptics of molecular-shunted plasmonic nanojunctions," Nano Lett. 15, 669–674 (2015).
- M. Danckwerts and L. Novotny, "Optical frequency mixing at coupled gold nanoparticles," Phys. Rev. Lett. 98, 026104 (2007).
- "Cambridge NanoPhotonics FPGA system," NanoPhotonics Centre, 2021, https://www.np.phy.cam.ac.uk/fpga-system.
- A. Tosi, A. D. Mora, F. Zappa, A. Gulinatti, D. Contini, A. Pifferi, L. Spinelli, A. Torricelli, and R. Cubeddu, "Fast-gated single-photon counting technique widens dynamic range and speeds up acquisition time in time-resolved measurements," Opt. Express 19, 10735–10746 (2011).
- M. Buttafava, G. Boso, A. Ruggeri, A. Dalla Mora, and A. Tosi, "Timegated single-photon detection module with 110 ps transition time and up to 80 MHz repetition rate," Rev. Sci. Instrum. 85, 083114 (2014).
- I. Nissinen, J. Nissinen, A. K. Länsman, L. Hallman, A. Kilpelä, J. Kostamovaara, M. Kögler, M. Aikio, and J. Tenhunen, "A sub-ns time-gated CMOS single photon avalanche diode detector for Raman spectroscopy," in *Proceedings European Solid-State Device Research Conference* (2011), pp. 375–378.
- S. Ly, G. McNerney, S. Fore, J. Chan, and T. Huser, "Time-gated single photon counting enables separation of CARS microscopy data from multiphoton-excited tissue autofluorescence," Opt. Express 15, 16839–16851 (2007).
- F. Benz, M. K. Schmidt, A. Dreismann, R. Chikkaraddy, Y. Zhang, A. Demetriadou, C. Carnegie, H. Ohadi, B. De Nijs, R. Esteban, J. Aizpurua, and J. J. Baumberg, "Single-molecule optomechanics in 'picocavities'," Science 354, 726–729 (2016).
- C. Carnegie, M. Urbieta, R. Chikkaraddy, B. de Nijs, J. Griffiths, W. M. Deacon, M. Kamp, N. Zabala, J. Aizpurua, and J. J. Baumberg, "Flickering nanometre-scale disorder in a crystal lattice tracked by plasmonic flare light emission," Nat. Commun. 11, 682 (2020).
- R. Chikkaraddy, A. Xomalis, L. A. Jakob, and J. J. Baumberg, "Midinfrared-perturbed molecular vibrational signatures in plasmonic nanocavities," arXiv: 2108.10171 (2021).
- M. Sidorova, S. G. Pavlov, A. D. Semenov, M. Gensch, and H.-W. Hübers, "Fiber-dispersive Raman spectrometer with single-photon sensitivity," Opt. Express 29, 20941–20951 (2021).
- L. Jakob, W. Deacon, O. Hicks, I. Manyakin, O. Ojambati, M. Traxler, and J. Baumberg, "Research data supporting "Single photon multiclock lock-in detection by picosecond time stamping" [Dataset]," University of Cambridge, 2021, https://doi.org/10.17863/CAM.78088.