Comprehensive characterization of InGaAs-InP avalanche photodiodes at 1550nm with an active quenching ASIC

ZHANG, Jun, et al.

Abstract

We present an active quenching application-specific integrated circuit (ASIC), for use in conjunction with InGaAs-InP avalanche photodiodes (APDs), for 1550-nm single-photon detection. To evaluate its performance, we first compare its operation with that of standard quenching electronics. We then test four InGaAs-InP APDs using the ASIC, operating both in the free-running and gated modes, to study more general behavior. We investigate not only the standard parameters under different working conditions but also parameters such as charge persistence and quenching time. We also use the multiple trapping model to account for the afterpulsing behavior in the gated mode, and further propose a model to take account of the afterpulsing effects in the free-running mode. Our results clearly indicate that the performance of APDs with an on-chip quenching circuit significantly surpasses the conventional quenching electronics and makes them suitable for practical applications, e.g., quantum cryptography.

Reference


DOI : 10.1109/JQE.2009.2013210
Abstract—We present an active quenching application-specific integrated circuit (ASIC), for use in conjunction with InGaAs–InP avalanche photodiodes (APDs), for 1550-nm single-photon detection. To evaluate its performance, we first compare its operation with that of standard quenching electronics. We then test four InGaAs–InP APDs using the ASIC, operating both in the free-running and gated modes, to study more general behavior. We investigate not only the standard parameters under different working conditions but also parameters such as charge persistence and quenching time. We also use the multiple trapping model to account for the afterpulsing behavior in the gated mode, and further propose a model to take account of the afterpulsing effects in the free-running mode. Our results clearly indicate that the performance of APDs with an on-chip quenching circuit significantly surpasses the conventional quenching electronics and makes them suitable for practical applications, e.g., quantum cryptography.

Index Terms—ASIC, avalanche photodiodes (APDs), quantum cryptography, single-photon detection, SPAD, telecom wavelengths.

I. INTRODUCTION

INGLE-PHOTON detectors are the key components in numerous photonics-related applications such as quantum cryptography [1], optical time-domain reflectometry [2], [3], and integrated circuit testing [4]. We can classify single-photon detection into four classes: photomultiplier tubes [5], semiconductor avalanche photodiodes (APDs) [6], [7], superconducting detectors [8], and novel proposals such as using a single-electron transistor consisting of a semiconductor quantum dot [9].

In the telecommunication regime (1550 nm), InGaAs/InP APDs are currently fabricated with separate absorption, charge, and multiplication layers [7] to ensure the lattice matching and preserve a low electric field in the InGaAs absorption layer with a narrower bandgap ($E_g = 0.75$ eV for In$_{0.25}$Ga$_{0.75}$As), minimizing the induced leakage currents, while a high electric field in the InP multiplication layer, enhancing the impact ionization effect. The middle charge layer can efficiently control the electric field profiles of the absorption and multiplication layers. The parameters of APDs are affected by many factors such as the crystalline quality of the semiconductor device, imperfections of design and fabrication, quenching circuit, and operational conditions. Therefore, actual performance of these APDs is always compromised and optimized for different applications.

In the past decade, efforts have been made to characterize and further improve APD performance on the single-photon level at 1550 nm [10]–[22]. Recently, integration of the quenching electronics for InGaAs/InP APDs to an application-specific integrated circuit (ASIC) [23], [24] has been implemented. The measured results on some key parameters of APDs demonstrate that active quenching ASICs can efficiently improve the noise-efficiency performance, and it has been shown that these APDs can work in a free-running mode [24]. However, full characterization of APDs with the ASIC is still necessary to better understand the improvements they provided. In this paper, we fully test four InGaAs/InP APDs at 1550 nm with an active quenching ASIC operating both in the free-running and gated modes, and compare the improvements with conventional electronics.

II. SETUP AND THE PRINCIPLE OF THE ASIC

The schematic setup for testing APDs is shown in Fig. 1. A digital delay pulse generator (DG 535, Stanford Research Systems, Inc.) provides synchronous signals for the whole system. One of its periodic outputs drives a 1550-nm laser diode (LD) to produce short optical pulses with $\sim$200 ps full-width at half-maximum (FWHM). The optical pulses are split into two parts by a 10/90 asymmetric fiber beam splitter (BS). Ninety percent of the signal is monitored by a power meter (IQ 1100, EXFO Company) to regulate the precise variable attenuator (Var. ATT, IQ 3100, EXFO Co.) in real time and stabilize the intensity of the output from the attenuator that goes to the pigtailed APD. The pins of the APD and ASIC are soldered together on a small
printed circuit board, while the body of APD is fixed on the top of four-layer thermoelectric cooler and actively stabilized with a closed-loop control.

The schematic diagram of the ASIC, fabricated with a 0.8-μm complementary metal oxide semiconductor (CMOS) process, is shown in the gray area of Fig. 1. The amplitude of the gate signals from the complex programmable logic device (CPLD) is first converted to the power supply voltage VDD (+5 V) of the chip. Two pulses are then generated to control the p-type MOS (PMOS) and n-type MOS (NMOS) switches, respectively, which have extremely fast rise and full times. There is a very short delay between the two control pulses to avoid the simultaneous conduction of the two switches. The timing is such that at the beginning of the gate, the pMOS switch is closed while the nMOS switch is opened, to charge the voltage at the quenching point (Q) up to VDD, and then, the pMOS switch is reopened.

We also observe avalanche signals due to signal reflections or electronic noise. These signals can be reduced. This brings a lot of benefits such as superior signal integrity, minimized parasitic capacitance, and reducing fake avalanche signals due to signal reflections or electronic noise. We know, due to the ASIC, the sizes of the electronics are greatly decreased and the electronic cables and the lengths of wires are reduced. This brings a lot of benefits such as superior signal integrity, minimized parasitic capacitance, and reducing fake avalanche signals due to signal reflections or electronic noise.

The three curves manifestly exhibit the performance improvements provided by the new quenching electronics. The improvement of a factor of 3 in the \( P_{DC}-P_{DE} \) performance for #3 APD shown in Fig. 3(a) is better than expected. As we know, due to the ASIC, the sizes of the electronics are greatly decreased and the electronic cables and the lengths of wires are reduced. This brings a lot of benefits such as superior signal integrity, minimized parasitic capacitance, and reducing fake avalanche signals due to signal reflections or electronic noise. We also observe \( P_{DC}-P_{DE} \) performance improvements on other APDs, for instance, for #2 APD shown in Fig. 3(a), the ratio is always about 1 (no improvement) when \( P_{DE} < 13\% \) and slowly increases to about 2 when \( P_{DE} \approx 25\% \).

We see, in Fig. 3(b), a significant improvement in the \( P_{AP} \) between the two cases as expected. The \( P_{AP} \) is generally proportional to the total number of carriers generated during an avalanche, and hence motivates small and rapidly quenched avalanches. The results here clearly illustrate that the circuit response and quenching time of the new system for the avalanche

III. PERFORMANCE TESTS OF APDS

We have tested four commercial APDs: #1 (JDSU0131T1897); #2 (JDSU0122E1711); and #3 (Epitaxx9951E9559) from JDSU; as well as #4 (PLI-DOI61910-040W059-076) from Princeton Lightwave, Inc., and compared the different performance characterizations of these APDs with the ASIC quenching system.

A. Integrated Versus Conventional Quenching Electronics

First, we perform the key parameter measurements on the same (#3) APD using the new (ASIC) and old (conventional nonintegrated circuit) [11], [14] quenching electronics under the same settings \( (T = 223\ K) \). Fig. 3 shows the comparison results for dark count \( (P_{DC}) \) per nanosecond versus single-photon detection efficiency \( (P_{DE}) \) probabilities, afterpulse probability \( (P_{AP}) \) and jitter, respectively. Using the double-gate method [14] (we discuss this in the latter section) as shown in Fig. 2, these parameters can be related to

\[
P_{DC} = \frac{C_{DC}}{f\tau_{AB}} \quad P_{AP} = \frac{C_{AP}}{C_{DE}/\tau_{CD}} \quad P_{DE} = \frac{1}{\mu} \ln \left( \frac{1-C_{DC}/f}{1-C_{DE}/f} \right) \tag{1}
\]

where \( C_{DC} \) \( (C_{AP}, C_{DE}) \) is the observed dark count (afterpulse, detection) rate, \( \tau_{AB} \) \( (\tau_{CD}) \) is the effective width of detection (afterpulse) gate in nanoseconds and \( \mu \) is the mean photon number per optical pulse with repetition frequency of \( f \). During the experiment, the conditions are \( f = 10 \ kHz \), \( \tau_{AB} = \tau_{CD} = 100 \ ns \), and \( \mu = 1 \), and these are fixed unless specifically mentioned in this paper.

The three curves manifestly exhibit the performance improvements provided by the new quenching electronics. The improvement of a factor of 3 in the \( P_{DC}-P_{DE} \) performance for #3 APD shown in Fig. 3(a) is better than expected. As we know, due to the ASIC, the sizes of the electronics are greatly decreased and the electronic cables and the lengths of wires are reduced. This brings a lot of benefits such as superior signal integrity, minimized parasitic capacitance, and reducing fake avalanche signals due to signal reflections or electronic noise. We also observe \( P_{DC}-P_{DE} \) performance improvements on other APDs, for instance, for #2 APD shown in Fig. 3(a), the ratio is always about 1 (no improvement) when \( P_{DE} < 13\% \) and slowly increases to about 2 when \( P_{DE} \approx 25\% \). The \( P_{DC}-P_{DE} \) performance improvement ratio strongly depends on the APD devices and operational conditions. Although the reasons of the significant improvement for #3 APD are not clear yet, one possibility could be different gate heights and discrimination approaches between the two quenching systems, as it is the noise that is improved here, for a given excess bias voltage.

We see, in Fig. 3(b), a significant improvement in the \( P_{AP} \) between the two cases as expected. The \( P_{AP} \) is generally proportional to the total number of carriers generated during an avalanche, and hence motivates small and rapidly quenched avalanches. The results here clearly illustrate that the circuit response and quenching time of the new system for the avalanche...
discrimination are faster than the old system. We will come back to this in more details in the following sections.

Timing jitter (time resolution) is another key parameter. It is defined as the temporal uncertainty of detection output for an avalanche with fixed arrival time of photons. Time jitter strongly depends on device fabrication and $P_{DE}$, corresponding to excess bias ($V_{eb}$) on the APD. Larger $V_{eb}$ can generate higher electric fields, which will shorten the trapping time of the carriers in the absorption and grading layers, and also the buildup time of avalanche, hence reducing the jitter. To measure this, we use a time-correlated single-photon counting (TCSPC) board (SPC-130, Becker & Hickl GmbH) with a time resolution of 6 ps FWHM and minimum time slot of 815 fs to measure the jitter properties. A synchronized signal from a pulse generator is used as the TCSPC’s “stop” while the detection output signal is used as “start.” The measured jitter is the overall jitter of the system, including the jitter (< 60 ps) and width (∼200 ps) of arrived optical pulses, the APD’s intrinsic jitter owing to the stochastic process of carrier dynamics, as well as from the associated electronics. The jitter performance is shown in Fig. 3(c) and we only see a minor improvement when $P_{DE} < 10\%$. We expect the electronic jitter to be slightly better as the ASIC can efficiently reduce the propagation time and jitter of the signals. At higher $P_{DE}$, we do not observe the improvement, and the negligible difference between the two cases is due to contributions from the associated external electronics, e.g., CPLD and discriminator that are used with the new system but not the old one. However, varying degrees of improvement have been observed on other APDs even at higher $P_{DE}$.

B. $P_{DC}$, $P_{DE}$, and Thermal Activation Energy

In order to illustrate the universal improvements afforded by this new quenching system, we use the new system operating in the gated mode to repeat the measurement on different APDs and temperature settings, as shown in Figs. 4 and 5.

The $P_{DC}$–$P_{DE}$ behavior of #1, #2, and #4 APDs are very similar, with $P_{DC} \sim 1.6 \times 10^{-6}$ ns$^{-1}$ and $P_{DE} = 10\%$ at 223 K, as shown in Fig. 4, but much better than #3 APD. Fig. 5 shows the $P_{DC}$ behavior of #2 APD from 210 to 238 K, and we see a reduction in $P_{DC}$ to $4.5 \times 10^{-7}$ ns$^{-1}$ for the same $P_{DE}$.

The origin of the dark counts is mainly due to the defect concentration in the semiconductor device. There are two main mechanisms for the generation of dark carriers: thermal generation and tunneling generation. The thermal generation means that a carrier is transferred from the valence band to the conduction band either directly or via the midgap defects, owing to the thermal excitation. Tunneling generation means that a
versus is the detrapping rate at time to the falling, and temperature (s, the /61/84 as we make our detection gates longer. The for #1 APD with different . When is can be expressed as values for is the Boltzmann constant and behavior. The first approach measures the total noise behavior as a function of . When is large enough, say, 100 , the is negligible and the measured noise is primarily due to dark counts. After subtracting , the quantity of noise left can be attributed to afterpulsing. This method was used in [24] while straightforward, generally overestimates AP.

The other approach, the double-gate method [14], as used in our setup is illustrated in Fig. 2. If there is a click during the detection (AB) gate, the CPLD will also generate an afterpulse (CD) gate after AB’s reset pulse with a delay of to the falling edge of the AB gate. This corresponds to the deadtime. The CPLD also generates a reset pulse for the CD gate only when an afterpulse detection is registered during this gate. This method directly measures AP.

Assuming a Poisson distribution, can be expressed as

\[ P_{AP} = \left[ \frac{1 - e^{-R_{AP}(\tau_d)\tau_{BD}}} {\tau_{CD}} \right] \tag{3} \]

where is the detrapping rate at time and is the avalanche probability. We use a multiple trapping model (multiple detrapping times) to describe \( R_{AP}(\tau_d) \) [19], [20],

\[ R_{AP}(\tau_d) = \sum_i N_i e^{-\tau_d/\Delta t_i} \tag{4} \]

where \( N_i \) is the number of trapped carriers at the end of the detection gate with a detrapping time constant of \( \Delta t_i \). There are single trapping models that use a single detrapping time constant \( \Delta t \) but, in many cases, this does not correspond to the measured results. The multiple trapping model effectively fits the measured results but some physical questions remain, e.g., why only two detrapping time parameters are needed for modeling one AP while three parameters are required for another, etc. In fact, quantitative description and modeling for AP behavior is still an intractable problem.

To illustrate the suitability for free-running operations, we look at the AP as we make our detection gates longer. The results for #1 APD are plotted in Fig. 7 and fitted using the multiple trapping model. is fixed at 100 ns while \( \tau_{AB} \) and the photon’s arrival positions are altered. If the active quenching was slow, then the arrival position, or time, of the photon’s appearance in the AB gate would be reflected in the \( P_{AP} \) behavior. A photon creating an avalanche at the start of a long gate would generate more carriers, increasing the chances for subsequent

\[ P_{DC} \propto T^2 e^{-E_a(T)/kT} \tag{2} \]

where is the Boltzmann constant and is a function of temperature with slow variations. In Fig. 6, four curves of log(\( P_{DC}/T^2 \)) versus \( 1/kT \) for #1 APD with different values are plotted. We evaluate the difference of \( E_a \) values for two small temperature ranges (216–223 and 233–238 K). The fitting values are displayed in Fig. 6. The results clearly show that generally higher temperatures induce larger \( E_a \) values and suggest that the thermal generation mechanism around 238 K dominates while the TAT mechanism is more significant around 216 K[26].

C. Afterpulsing

During an avalanche process, due to a photon detection, dark count effects, or afterpulsing itself, a carrier can be trapped by a defect in the multiplication layer. This carrier may excite another avalanche—an afterpulse, during subsequent gates. This process severely limits the APD performance for high-frequency operation due to the need to apply long, typically \( \sim 10 \mu s \), deadtimes where the APD is inactive. There are two methods to measure the behavior. The first approach measures the total noise behavior as a function of . When is large enough, say, 100 , the is negligible and the measured noise is primarily due to dark counts. After subtracting , the quantity of noise left can be attributed to afterpulsing. This method was used in [24] while straightforward, generally overestimates AP.

The other approach, the double-gate method [14], as used in our setup is illustrated in Fig. 2. If there is a click during the detection (AB) gate, the CPLD will also generate an afterpulse (CD) gate after AB’s reset pulse with a delay of to the falling edge of the AB gate. This corresponds to the deadtime. The CPLD also generates a reset pulse for the CD gate only when an afterpulse detection is registered during this gate. This method directly measures AP.

Assuming a Poisson distribution, can be expressed as

\[ P_{AP} = \left[ \frac{1 - e^{-R_{AP}(\tau_d)\tau_{BD}}} {\tau_{CD}} \right] \tag{3} \]

where is the detrapping rate at time and is the avalanche probability. We use a multiple trapping model (multiple detrapping times) to describe \( R_{AP}(\tau_d) \) [19], [20],

\[ R_{AP}(\tau_d) = \sum_i N_i e^{-\tau_d/\Delta t_i} \tag{4} \]

where \( N_i \) is the number of trapped carriers at the end of the detection gate with a detrapping time constant of \( \Delta t_i \). There are single trapping models that use a single detrapping time constant \( \Delta t \) but, in many cases, this does not correspond to the measured results. The multiple trapping model effectively fits the measured results but some physical questions remain, e.g., why only two detrapping time parameters are needed for modeling one AP while three parameters are required for another, etc. In fact, quantitative description and modeling for AP behavior is still an intractable problem.

To illustrate the suitability for free-running operations, we look at the AP as we make our detection gates longer. The results for #1 APD are plotted in Fig. 7 and fitted using the multiple trapping model. is fixed at 100 ns while \( \tau_{AB} \) and the photon’s arrival positions are altered. If the active quenching was slow, then the arrival position, or time, of the photon’s appearance in the AB gate would be reflected in the \( P_{AP} \) behavior. A photon creating an avalanche at the start of a long gate would generate more carriers, increasing the chances for subsequent

\[ P_{DC} \propto T^2 e^{-E_a(T)/kT} \tag{2} \]

where is the Boltzmann constant and is a function of temperature with slow variations. In Fig. 6, four curves of log(\( P_{DC}/T^2 \)) versus \( 1/kT \) for #1 APD with different values are plotted. We evaluate the difference of \( E_a \) values for two small temperature ranges (216–223 and 233–238 K). The fitting values are displayed in Fig. 6. The results clearly show that generally higher temperatures induce larger \( E_a \) values and suggest that the thermal generation mechanism around 238 K dominates while the TAT mechanism is more significant around 216 K[26].

C. Afterpulsing

During an avalanche process, due to a photon detection, dark count effects, or afterpulsing itself, a carrier can be trapped by a defect in the multiplication layer. This carrier may excite another avalanche—an afterpulse, during subsequent gates. This process severely limits the APD performance for high-frequency operation due to the need to apply long, typically \( \sim 10 \mu s \), deadtimes where the APD is inactive. There are two methods to measure the behavior. The first approach measures the total noise behavior as a function of . When is large enough, say, 100 , the is negligible and the measured noise is primarily due to dark counts. After subtracting , the quantity of noise left can be attributed to afterpulsing. This method was used in [24] while straightforward, generally overestimates AP.

The other approach, the double-gate method [14], as used in our setup is illustrated in Fig. 2. If there is a click during the detection (AB) gate, the CPLD will also generate an afterpulse (CD) gate after AB’s reset pulse with a delay of to the falling edge of the AB gate. This corresponds to the deadtime. The CPLD also generates a reset pulse for the CD gate only when an afterpulse detection is registered during this gate. This method directly measures AP.

Assuming a Poisson distribution, can be expressed as

\[ P_{AP} = \left[ \frac{1 - e^{-R_{AP}(\tau_d)\tau_{BD}}} {\tau_{CD}} \right] \tag{3} \]

where is the detrapping rate at time and is the avalanche probability. We use a multiple trapping model (multiple detrapping times) to describe \( R_{AP}(\tau_d) \) [19], [20],

\[ R_{AP}(\tau_d) = \sum_i N_i e^{-\tau_d/\Delta t_i} \tag{4} \]

where \( N_i \) is the number of trapped carriers at the end of the detection gate with a detrapping time constant of \( \Delta t_i \). There are single trapping models that use a single detrapping time constant \( \Delta t \) but, in many cases, this does not correspond to the measured results. The multiple trapping model effectively fits the measured results but some physical questions remain, e.g., why only two detrapping time parameters are needed for modeling one AP while three parameters are required for another, etc. In fact, quantitative description and modeling for AP behavior is still an intractable problem.

To illustrate the suitability for free-running operations, we look at the AP as we make our detection gates longer. The results for #1 APD are plotted in Fig. 7 and fitted using the multiple trapping model. is fixed at 100 ns while \( \tau_{AB} \) and the photon’s arrival positions are altered. If the active quenching was slow, then the arrival position, or time, of the photon’s appearance in the AB gate would be reflected in the \( P_{AP} \) behavior. A photon creating an avalanche at the start of a long gate would generate more carriers, increasing the chances for subsequent
afterpulses, than in the short gate regime or if the photon arrived at the end of a gate. The overlapping curves show that the $P_{AP}$ behavior does not change for long gates, nor is it dependent on the arrival time, and hence should not change when we move to a free-running regime.

We finally study the temperature dependence of afterpulse. The experimental results and fitting curves are shown in Fig. 8 and the fitting parameters are listed in Table I. When the temperature is varied from 238 to 223 K, the $P_{AP}$ behavior is almost identical due to the close trap lifetime parameters, but when the temperature is at 216 K (#1 APD) or 210 K (#2 APD), there is a distinct increase for the $P_{AP}$. The detrapping life time can be modeled as [27]

$$\Delta t \propto e^{E_{tr} / kT} / T^2$$

where $E_{tr}$ is the trapping activation energy. This formula means that lower temperatures cause larger $\Delta t$ for traps corresponding to larger $P_{AP}$.

Moreover, when $\tau_d \leq 10 \mu s$, the $P_{AP}$ of #2 APD at 210 K, in Fig. 8, is less than at other temperatures, but the $P_{AP}$ of #1 APD at 216 K is not. According to the fitting results at 210 K, there is a trap type with a fast detrapping life time of 615 ns in #2 APD, which causes rapid detrapping at small $\tau_d$, but when $\tau_d$ becomes large, the effect of this trap type is gradually diminished while the other trap types with 2560 and 10135 ns life times start to dominate the detrapping process. Unfortunately, this kind of fast detrapping time is too short and/or too weak to be measured at the other three temperatures and for #1 APD.

In order to validate the aforementioned phenomena, we perform the measurements of $P_{AP}$ behaviors of #4 APD from another manufacturer, whose results are shown in Fig. 9. The $P_{AP}$ increases from 233 to 210 K while the cross point appears between 210 and 203 K, which agrees well with our explanation for the different $P_{AP}$ behaviors. We believe that the $P_{AP}$ models so far are not perfect, and further investigations, including effective models and experiments, are still needed.

D. Free-Running Mode

Free-running operation is very important for many applications such as asynchronous and CW photon counting and quantum cryptography [1], etc. Due to the lower noise characteristics of InGaAs/InP APDs that use this active quenching ASIC, some of us have recently been able to show that this is now also possible for APDs in the telecom regime.

As in the gated regime, the operation in the free-running mode depends on the parameters of $V_{DC}$, $\tau_d$, and $T$. However, unlike the gated mode, the afterpulse parameter in the free-running mode is more difficult to evaluate.

As we said, with respect to Fig. 7, the $P_{AP}$ does not depend on the width of the gate, which is applicable for the free-running mode. Indeed, it may not be obvious how the afterpulse probability evolves when the gate is open for such long times, though it would appear that at worst, the probability continues to decrease over the period of detection. Nonetheless, we have previously seen that for short deadtimes, the afterpulsing dominates [24]. As we have now been able to use the double-gate method to characterize the afterpulsing, in the gated regime, we can use a simple model to describe the detection and noise rates for the free-running mode

$$R = \eta N (1 - \eta N \tau_d) (1 - P_{AP})$$

with $\eta = 1 - e^{-\mu F_{PE}} (1 - P_{DC})$, considering the Poisson distribution. $P_{PE}$ and $P_{DC}$ are the detection efficiency and dark count probability, and $N$ is the input photon number. The term of $(1 - \eta N \tau_d)$ is for deadtime correction. If we put $\mu = 0$, we recover the noise rate. $P_{AP}$ is the total afterpulsing contribution at $\tau_d$, calculated from integrating over the gated afterpulse probability from $\tau_d$ to infinity (in practice, 100 $\mu$s is sufficient). Fig. 10 shows the experimental rates as well as the results of our model as a function of $\tau_d$.

It is clear that a more complicated model is warranted. However, the physics of these limitations is clear. In the small $\tau_d$ region, we underestimate the rates as we do not take account of cascaded afterpulses, i.e., higher order effects. The more interesting region from 20 to 40 $\mu$s, we are overestimating due to the difficulty in defining an appropriate integration range, which will also change as a function of the photon flux, the intrinsic detection efficiency, and the deadtime. Importantly, we can also conclude that for small $\tau_d$, if $N$ increases, then the $P_{AP}$ value will decrease, since photon clicks will increase while the multiple afterpulsing effects will be relatively less likely.

Our model makes a first attempt to both understand the afterpulsing, and to develop a reliable technique for determining the detector’s characteristics, without resorting to complicated techniques in a double-gate regime, there is still some way to go. Although there is an apparent need for large $\tau_d$ that, in turn, limits the maximum count rate, this is highly dependent on the photon flux to be detected and free-running APDs are certainly highly advantageous for applications with low to moderate count rates.

E. Charge Persistence

Charge persistence (CP) is not normally a problem for synchronous detectors as the photons arrive during the gate. However, what happens if a photon arrives before the gate is applied,
as is possible in the free-running mode, before the APD is activated after a deadtime?

When the detector is “OFF,” i.e., at $V_{1ro}$ below $V_{1p}$, with only a few volts so that primary dark carriers can still be generated and multiplied by the average dc gain but with a small probability. When the gate pulse arrives some of the carriers that remained in the multiplication layer can induce avalanches [28]. This is called “CP,” or sometimes referred as the “twilight effect” [29]. Similarly, when the CP carriers are released before the gate pulses with the time difference less than the effective transit time, they can also create afterpulses [28]. Now let us consider another case, where photons always appear before the gate. Based on the same principle, in this case, the number of dark CP carriers will be increased and the CP effect will be expanded.

We experimentally test this effect and the results are shown in Fig. 11. By varying the time difference between the arrival times of gates ($T_d$) and photons ($T_p$), we observe the changes of the normalized (for $\mu$) noise per gate for #4 APD. The two almost identical behaviors show that the CP effect is proportional to photon numbers and, per photon, can generate noise of about 10% of the dark count level with the time difference less than 1 ns. When the time difference is larger than $\sim 5$ ns, the CP effect is negligible due to the characteristic exponential decrease. Moreover, through using TCSPC, we also observe the detection events at the beginning of the gate are more than those at other

**TABLE I**

<table>
<thead>
<tr>
<th>APD</th>
<th>T(K)</th>
<th>$\Delta t_1$(ns)</th>
<th>$\Delta t_2$(ns)</th>
<th>$\Delta t_3$(ns)</th>
</tr>
</thead>
<tbody>
<tr>
<td>#1</td>
<td>216</td>
<td>1135</td>
<td>5645</td>
<td></td>
</tr>
<tr>
<td>#1</td>
<td>238-223</td>
<td>860</td>
<td>4385</td>
<td></td>
</tr>
<tr>
<td>#2</td>
<td>210</td>
<td>615</td>
<td>2560</td>
<td>10135</td>
</tr>
<tr>
<td>#2</td>
<td>238-223</td>
<td>1020</td>
<td>2165</td>
<td>5075</td>
</tr>
</tbody>
</table>

Fig. 8. $P_{\text{AP}}$ versus $\tau_d$ for #1 (a) and #2 (b) APDs at different $T$.

Fig. 9. $P_{\text{AP}}$ versus $\tau_d$ for #4 APD at different $T$.

Fig. 10. Plot of the detection and noise rates as a function of deadtime for #2 APD at $V_{\text{DC}} = 48.62$ V, $N = 10$ KHz with CW photons and $T = 210$ K, operating in the free-running mode.

Fig. 11. Noise, including CP and dark counts per gate, normalized by $\mu$, as a function of time difference between detection gate ($T_d$) and photons ($T_p$). The horizontal line is the dark count level. The results are tested using #4 APD at $P_{\text{FDE}} = 10\%$. 

Authorized licensed use limited to: UNIVERSITY OF BRISTOL. Downloaded on June 2, 2009 at 05:16 from IEEE Xplore. Restrictions apply.
regions. The CP effect will cause nonnegligible noise in the case of high-frequency gating or asynchronous high-flux detection.

F. Quenching Time

Quantifying the quenching time, including the circuit reaction time and gate closing time, as shown in Fig. 12, of an avalanche is very important to understand the avalanche dynamics of APDs. Although an active quenching ASIC should have a faster quenching time than conventional electronics, this has not previously been measured. More generally, these results are also pertinent for rapid gating schemes that use very short gates and hence terminate avalanches very quickly.

The principle for measuring the quenching time is to compare the count rate behaviors for detections and afterpulses (see Fig. 12) using the double-gate measurement electronics. The total number of carriers, during an avalanche, should be proportional to the excess bias on APD and the excess bias duration, or the integral of excess bias over the quenching time. Now, we consider the case where photons arrive at the end of the detection gates, by delaying photons. From phases 1 to 2 in Fig. 12, the count rates of detection and afterpulse are both almost constant, while from phases 2 to 3, the detection rate is still constant but the afterpulse rate decreases first due to the decrease of the integral. The time difference between the two phases can be regarded as the reaction time to detect the onset of the avalanche and send the signal to close the nMOS switch. After phase 3, both of the rates drastically decrease until the end of the closing time.

Fig. 13 shows the results of these measurements on #1 APD. From the slope of the detection rate, we can obtain the closing time of the gate, which is only around 1 ns. Although it is very hard to determine a precise value of the reaction time from the fitting results, the slight shift between the detection and afterpulse rates indicates that the reaction time is much less than the closing time.

IV. CONCLUSION

In summary, we have fully characterized an active quenching ASIC and compared its operation with a conventional electronic circuit. To show that the improvements are universal, we also characterized and compared four different InGaAs/InP APDs. The APDs operating in the gated mode exhibit substantial performance improvements compared with the conventional quenching electronics and allow for free-running operation. We also extract thermal activation energies to identify the dominant mechanism of dark counts, and by employing the multiple detrapping model in the gated mode and proposed model in the free-running mode, the afterpulse behaviors are well-illustrated. Moreover, we have characterized the CP and quenching time. The advantages of low afterpulsing and noise in both those regimes are mostly attributed to the state-of-the-art ASIC.

ACKNOWLEDGMENT

The authors would like to thank Dr. A. Rochas and Dr. M. A. Itzler for useful discussions.

REFERENCES


J. Sel. Topics Quantum Electron. 2001, he cofounded id Quantique, a spin-off committed to commercializing quantum mechanics) and was appointed as a Maître (quantum cryptography and the foundations of fundamental physics. He is recognized as one of the leading figures in the field of quantum communication and cryptography.

Nicolas Gisin received the Ph.D. degree in physics from the University of Geneva, Switzerland, in 1981.

He is the Director of the Group of Applied Physics—Optics, University of Geneva. He has extensive experience in both university and industry where he has been instrumental in launching several startup companies including id Quantique. He has authored or cowritten over 150 academic articles. His current research interests include both applied and fundamental physics. He is recognized as one of the leading figures in the field of quantum communication and cryptography.

Hugo Zbinden was born in 1961. He received the Ph.D. degree in rare-earth solid-state lasers from the University of Bern, Bern, Switzerland, in 1991. He developed high-power diode-pumped Nd:YAG lasers for industrial applications. In 1993, he joined the Group of Applied Physics, University of Geneva, Switzerland, where he has been leading the group’s experimental activities in quantum optics (quantum cryptography and the foundations of quantum mechanics) and was appointed as a Maître d’Enseignement et de Recherche (MER) in 2003. In 2001, he cofounded id Quantique, a spin-off committed to commercializing quantum cryptography and other quantum technologies.