Quenching Circuit Discriminator Architecture Impact on a Sub-10 ps FWHM Single-Photon Timing Resolution SPAD

Frédéric Nolet *, Valérie Gauthier, Samuel Parent, Frédéric Vachon, Nicolas Roy, Nicolas St-Jean, Serge A. Charlebois and Jean-François Pratte

Interdisciplinary Institute for Technological Innovation, Department of Electrical and Computer Engineering, Université de Sherbrooke, Sherbrooke, QC J1K 0A5, Canada
* Correspondence: frederic.nolet@usherbrooke.ca; Tel.: +1-819-821-8000-62010

Abstract: In the field of radiation instrumentation, there is a desire to reach a sub-10 ps FWHM timing resolution for applications such as time-of-flight positron emission tomography, time-of-flight positron computed tomography and time-resolved calorimetry. One of the key parts of the detection chain for these applications is a single-photon detector and, in recent years, the first single-photon avalanche diode (SPAD) with a sub-10 ps timing resolution was presented. To reach such a timing resolution, the SPAD was read out by an operational amplifier operated in open-loop as a comparator. This paper presents a comparison between comparators and inverters to determine which type of leading-edge discriminator can obtain the best single-photon timing resolution. Six different quenching circuits (QCs) implemented in TSMC 65 nm are tested with SPADs of the same architecture and in the same operation conditions. This allows us to compare experimental results between the different QCs. This paper also presents a method to measure the SPAD signal slope, the SPAD excess voltage variation and simulations to determine the added jitter of different leading-edge discriminators. For some discriminator architectures, a cascode transistor was required to increase the maximum excess voltage of the QC. This paper also presents the impact on the single-photon timing resolution of adding a cascode transistor for a comparator or an inverter-based discriminator. This paper reports a 6.3 ps FWHM SPTR for a SPAD read out by a low-threshold comparator and a 6.8 ps FWHM SPTR for an optimized 1 V inverter using a cascode transistor for a higher excess voltage.

Keywords: single-photon avalanche diode; quenching circuit; single-photon timing resolution; SiPM; photon-to-digital converter; time of flight; radiation instrumentation

1. Introduction

In the field of radiation instrumentation, there is a trend to reach a sub-10 ps FWHM timing resolution for applications such as time-of-flight positron emission tomography (ToF PET) [1,2], ToF computed tomography (ToF CT) [3] and time-resolved calorimetry. In ToF PET, a coincidence timing resolution (CTR) better than 10 ps FWHM allows a reconstructionless imaging of the PET radiotracer activity distribution with a precision level of 1.5 mm [2,4]. To reach a 10 ps FWHM CTR, major efforts worldwide are currently underway to improve the timing resolution of the whole PET detection chain. On the scintillator side, improved scintillation emission must be developed [5–7] combined with a depth of interaction (DOI) correction [8,9]. On the photodetector side, a recent study showed that reaching a single-photon timing resolution (SPTTR) of 10 ps FWHM at the detector level would allow a CTR in the order of tens of ps FWHM combined with prompt emission scintillators [1].

Single-photon avalanche diode (SPAD) arrays are some of the most promising candidates to reach a CTR of 10 ps FWHM. In recent years, SPAD-based detectors reached a 200 ps FWHM CTR in a commercial ToF PET scanner and a sub-100 ps FWHM for experimental measurements in a laboratory [10–13]. Silicon photomultipliers (SiPMs) are analog...
SPAD arrays connected to a common output node through resistors. The achievable SPTR for that type of detector has improved over the last few years to attain a 70 ps FWHM [6]. Another SPAD array architecture is the photon-to-digital converter (PDC), also known as the digital SiPM [14]. In this type of photodetector, each SPAD is read out individually by a CMOS quenching circuit and connected to digital processing circuits, such as a time-to-digital converter (TDC), counter, time estimator, etc. [15–19]. The digital approach improves the timing resolution since it minimizes the analog’s typical array-wide contributions, such as the high-output capacitance that impacts the rising edge slope of the output signal [20–22], the timing skew coming from the SPAD-to-output distance mismatch [18,19,23] and the high constraints on the array electronic readout, such as a very low electronic noise and very high bandwidth [4,13]. This improvement comes at the cost of a higher design complexity, cost and integration complexity since 3D integration is required to obtain a similar photodetection efficiency [14].

To achieve an SPTR of 10 ps FWHM with a PDC, the first objective is to design its basic cell, a SPAD combined with a QC. On the SPAD side, multiple developments were made to optimize the SPAD SPTR, such as shallow trench isolation (STI) (27 ps FWHM) [24], edge covering (20 ps FWHM) [23], reducing the size of the SPAD (12.1 ps FWHM for a diameter of 25 µm, 16 ps FWHM for 50 µm and 27.2 ps FWHM for 100 µm) [25] and very small and thin SPADs (7.8 ps FWHM) [20]. On the quenching circuit side, a study showed that the threshold of the discriminator has an impact on the SPAD SPTR [26]. In the current state-of-the-art, the best timing jitter is obtained at a low threshold (<100 mV) using a comparator [20,23,27] instead of a classic inverter, but it requires a higher static power consumption. Nevertheless, a study showed promising results with a standard CMOS inverter down to 7.5 ps FWHM for a 25 µm SPAD [28].

For the low-threshold comparator, there is a certain range of threshold values where the SPAD is at its optimal SPTR value [20,23]. In a recent study [20], we reported a sub-10 ps FWHM SPAD over a threshold range of 0 to 1.5 V. A wide range of threshold values with a fixed SPTR raises the following question:

Considering a SPAD with a fixed SPTR over a large range of comparator threshold voltages, do other discriminators such as inverters with similar electronic jitter react the same and obtain a similar SPTR or are there contributions other than the threshold and electronic jitter that should be mitigated on the discriminator part?

This paper analyzes the SPTR of a SPAD optimized for timing jitter combined with different discriminators within the quenching circuit. This paper presents six different discriminators tested with SPADs of the same architecture and in the same operating conditions. This paper focuses on the measured SPAD signal slope, the added jitter of the SPAD excess voltage variation combined to the time propagation of the QC and the impact on the timing jitter of a cascode transistor to increase the excess voltage. Section 2 introduces the architecture and design considerations, Section 3 presents the materials and methods, Section 4 presents the results obtained with the different architectures, and the results are discussed in Section 5.

2. Architecture

Each quenching circuit studied reads out a SPAD with the same architecture and is implemented in CMOS TSMC 65 nm. The implemented SPAD has a 20 µm diameter and is composed of a standard $p^+$ anode in an n-well cathode with a p-well guard ring [20]. The SPAD shape is a 32-sided polygon. The maximum operating excess voltage of the SPAD is around 1.7 V (around 17% of the breakdown voltage). An improved version of the SPAD architecture of [20] is also implemented on the new test chip. The silicide block layer is used over the SPAD active area to prevent the default drain/source silicide process on the SPAD’s anode. Adding this blocking layer enhances the light transmission to the SPAD and thus its PDE [29]. The contacts’ location in relation to the SPAD anode and guard ring is also updated to avoid having contacts overlaying the SPAD’s high field region.
The purpose of this modification is to reduce the SPAD dark noise, at the expense of a slightly lower photosensitive area (17.6 µm active diameter instead of 20 µm).

On the quenching circuit side, every QC used in this study is a passive quenching with an active recharge and is composed of five sections: the recharge transistor, the quenching transistors, a discriminator, two monostable circuits for the programmable hold-off and recharge time and the 50 ohms driver (Figure 1). Each QC has the exact same design except for the discriminator and an added cascode transistor to increase the maximum excess voltage by 3.3 V. As shown in Figure 1, there are six different discriminator combinations across five different quenching circuits. The discriminator in Figure 1A is an operational amplifier operated in open-loop as a comparator. Full details of the schematics can be found in [20,30]. In B, the discriminator is a 3.3 V inverter, where the NMOS transistor is 19 times bigger than the PMOS transistor to lower the inverter threshold to obtain a better timing response on the rising edge of the signal. In C, the discriminator is a 1 V inverter combined with a cascode transistor. The NMOS transistor is 10 times bigger than the PMOS transistor to maximize the timing response on the rising edge of the signal. All the inverters were optimized to obtain a similar timing response to the comparator through simulations (shown in Section 4.3).

![Figure 1. Overview of the different implementations of the QCs in CMOS 65 nm. The quenching and recharge are controlled through monostable circuits and the SPAD signal is read by either an adjustable threshold comparator (A), an inverter (B,C) or both (D,E).](image)

QCs D and E use three discriminators each that can read the same SPAD, at the same time. This allows for a direct comparison between different architectures and the ability to assess if connecting multiple discriminators will significantly impact the timing resolution. It should be noted that the node capacitance will be slightly higher than for circuits A, B and C, which could have an impact on the measured SPTR. In circuit D, the comparator can be directly compared with two 3.3 V inverters. In circuit E, all three circuits are cascoded, in order to compare a cascoded comparator with two 1 V inverters. Circuits D and E allow us to determine if the cascoding circuit significantly impacts the SPTR of the comparator-based quenching circuit. Table 1 summarizes the size, maximum excess voltage...
and measured QC timing jitter of each architecture. The QC timing jitter was measured following the method described in [20] for a 2 V signal from a Berkeley Nucleonics pulse generator with a slope of 1 V/ns. This measurement is performed to make sure that no discriminator has a very high electronic jitter, limiting the SPTR measurements. The jitter value of the discriminator is not directly translatable to the SPTR measurement and, since it is expected that the signal from the SPAD is steeper, the discriminator jitter contribution should be even lower for the SPTR measurement.

Table 1. Discriminator summary.

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Cascoded Unbalanced 1 V Inv.</th>
<th>Cascoded Standard 1 V Inv.</th>
<th>Cascoded Comp. 3.3 V Inv.</th>
<th>Unbalanced 3.3 V Inv.</th>
<th>Standard 3.3 V Inv.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1 circuits</td>
<td>A &amp; D</td>
<td>C &amp; E</td>
<td>E</td>
<td>E</td>
<td>B &amp; D</td>
</tr>
<tr>
<td>Discriminator size (µm²)</td>
<td>104</td>
<td>24</td>
<td>6</td>
<td>104</td>
<td>88</td>
</tr>
<tr>
<td>Cascode size (µm²)</td>
<td>63</td>
<td>63</td>
<td>63</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Discriminator and cascode size (µm²)</td>
<td>104</td>
<td>87</td>
<td>69</td>
<td>167</td>
<td>88</td>
</tr>
<tr>
<td>Inverter NMOS W/L ratio</td>
<td>40</td>
<td>2</td>
<td>48</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>Inverter PMOS W/L ratio</td>
<td>4</td>
<td>2</td>
<td>2.5</td>
<td>2.5</td>
<td></td>
</tr>
<tr>
<td>Maximum excess voltage</td>
<td>3.3</td>
<td>4.3</td>
<td>4.3</td>
<td>6.6</td>
<td></td>
</tr>
<tr>
<td>QC Jitter (ps FWHM)</td>
<td>4</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td></td>
</tr>
</tbody>
</table>

3. Materials and Methods

3.1. Discriminator’s Input Signal Slope

One of the contributions to the SPTR is the noise over the signal slope at the discriminator’s threshold. The signal underdrive, overdrive and slope at the discriminator’s threshold at the interface between the SPAD and the quenching circuit are of interest to determine how these parameters limit the SPTR. A previous study with an external quenching with an adjustable threshold and an oscilloscope showed that it is possible to measure the input signal slope parameters [31]. In this study, an integrated comparator with an adjustable threshold was used to get rid of the external capacitance that directly impacts the slope of the signal to measure these parameters for an integrated SPAD-QC pair.

Figure 2 illustrates the optical setup and the different steps to measure the SPAD signal slope. The measurement is a time difference histogram between the signal of a Becker and Hickl PHD-400 reference diode (3) and the output of the SPAD and QC (4). The SPAD is set at a fixed point of operation, in this case 1.7 V of excess voltage, resulting in a fixed time between the signal of the reference diode (3) and the signal of the SPAD (1). The time between the signal of the reference diode (3) and multiple points of the signal of the SPAD (1) is measured by sweeping the threshold (2) at different values (2a to 2c). If the threshold is set close to the baseline (2a), the delay of the QC signal will be shorter (4a). The mean of the histogram is calculated to extract the time at which the signal crossed the threshold.

3.2. Overdrive Variation Jitter

The time propagation delay is the time it takes for the signal to go from the input to the output of the discriminator. A variation in one of the input signal parameters (underdrive, overdrive or slope) will cause a variation of the time propagation delay. This variation adds to the total jitter. In this case, it is the excess voltage amplitude that varies, which translates to the overdrive parameter. To estimate the jitter contribution, two measurements are required. First, the SPAD excess voltage must be measured using the method described in [20]. Second, the time propagation delay must be measured as a function of the variation of the overdrive.
Experimentally, the time propagation delay is measured by sweeping the signal input amplitude (the SPAD excess voltage) while the threshold is fixed. In simulation, the time propagation delay can also be estimated by reproducing the Figure 2 setup. For the measurement, the threshold of the different QCs is fixed to a certain value (100 mV in the case of the comparator) and the SPAD excess voltage is swept from 0 V to 2 V to cover the SPTR measurements at 1.7 V of excess voltage. The time propagation delay is simulated across the 5 corners of the technology, which allows a range of variation for each discriminator. Using the method described in [20], the time propagation delay is then combined with the measured SPAD excess voltage variation to estimate the added jitter.

![Figure 2](image_url)

**Figure 2.** Optical setup used to measure the SPAD signal slope at the input of the QC (1). The setup is used to measure a time delay between a reference, provided by the laser and the PHD-400 diode (3), and the output signal of the quenching circuit (4). Using a constant SPAD biasing operating point, the QC threshold (2) is swept from few mVs to the excess voltage to measure the time delay difference caused by the comparator triggering at different levels of the SPAD signal.

### 3.3. Single-Photon Timing Resolution

During the SPAD and QC SPTR measurements, the SPAD is illuminated by a Spectra Physics Mai Tai 80 MHz Ti:Sapphire pulsed laser (pulse width < 100 fs) as presented in [20]. The light is attenuated to ensure the measurements are made in a photon-starved environment and the wavelength is 410 nm (an appropriate wavelength for LYSO-based modules in PET). The SPTR measurements were all performed at an excess voltage of 1.7 V to compare the different quenching circuits. For each discriminator’s architecture, the measurements were conducted on 4 different ASICs, with multiple QCs of the same architecture per the ASIC. To extract the FWHM, the FW(1/10)M and the FW(1/100)M of each data set, a fit combining two exponentially modified Gaussians (EMGs) is applied (Figure 3). Since the signal variation is very steep and thus could change the FWHM evaluation, the EMGs are used instead of a moving average or a filter. Other studies report using a Gaussian fit to analyze SPTR data [22] but this only fits the first part of the signal. Other studies use EMGs instead to include the right side of the histogram in the fit [32].

For our measurements, a second EMG (in blue) was used to fit the discontinuity at about \( t = 150 \) ps. The origin of the discontinuity has yet to be identified but is not an uncommon artefact and is often found in the measurements of the SPTR in similar analysis [33]. The timing jitter measurements are acquired with a Teledyne Lecroy SDA 6000A oscilloscope, with a 40 GSa/s and a 6 GHz vertical bandwidth.
Figure 3. For each data set, a fit combining two exponentially modified Gaussians (EMGs) is applied to extract the value of the FWHM, the FW(1/10)M and the FW(1/100)M.

3.4. Cascade Architecture Impact on the SPTR

The gate bias of the cascode transistor impacts the slope and the amplitude of the signal sent to the discriminator. The highest amplitude voltage at the input of the discriminator is the cascode gate bias voltage minus the threshold of the transistor. Since the cascode source is routed to 1 V transistors (QCs C and E) for the 1V inverters, the cascode bias is limited to ensure the proper operation of the transistors.

The setup presented in Figure 2 is reproduced in a Cadence simulation using a verilog model to emulate the laser source with its parameters, a SPAD model [34,35] and the simulated circuit, to estimate the impact of the cascode transistor at different gate bias points. The cascode bias is varied from 0 V to 1.4 V to account for the cascode threshold of about 0.4 V. The SPAD excess voltage is varied from 0 V to 2 V. The impact of the cascode is simulated across the 5 corners of the technology and for the different values of the junction capacitance of the SPAD identified in the signal slope section. The range of the simulated values is then compared to the measured SPTR of the different discriminator architectures.

4. Results

4.1. SPAD Characteristics

The SPAD characteristics are summarized in Table 2. The SPAD breakdown voltage is 9.9 V. The single SPAD PDE is higher than 15% in the wavelength range of 400 nm to 500 nm with a peak of 27% at 420 nm. The addition of the silicide block layer over the SPAD active area increased the peak PDE from 8% to 27% when compared to the previously developed SPAD [20]. The contacts’ location modification reduced the DCR from 2.8 k to 0.6 k (cps/µm^2) and the afterpulsing from 10% to 1%. These modifications did not impact the timing resolution of the SPAD.

Table 2. SPAD characteristics summary.

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>This Work</th>
<th>[36]</th>
<th>[20]</th>
<th>[37]</th>
<th>[38]</th>
<th>[39]</th>
<th>[33]</th>
<th>[24]</th>
<th>[28]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Technology (nm)</td>
<td>65</td>
<td>55</td>
<td>65</td>
<td>65</td>
<td>90</td>
<td>150</td>
<td>160</td>
<td>180</td>
<td>180</td>
</tr>
<tr>
<td>V_{BD} / V_{EX} (V)</td>
<td>9.9/1.7</td>
<td>31/7</td>
<td>9.9/1.5</td>
<td>9/0.4</td>
<td>15/2.4</td>
<td>18/3</td>
<td>26/3-9</td>
<td>11/0.8</td>
<td>22/8</td>
</tr>
<tr>
<td>Diameter (µm)</td>
<td>17.6</td>
<td>8.8</td>
<td>20</td>
<td>8</td>
<td>6.4</td>
<td>10</td>
<td>10-80</td>
<td>14</td>
<td>14</td>
</tr>
<tr>
<td>DCR (cps/µm^2)</td>
<td>0.6 k</td>
<td>2.6</td>
<td>2.8 k</td>
<td>15.6 k</td>
<td>3.1</td>
<td>0.4</td>
<td>0.13</td>
<td>4 k</td>
<td>0.2</td>
</tr>
<tr>
<td>Afterpulsing (%)</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;10</td>
<td>&lt;1</td>
<td>0.85</td>
<td>&lt;1</td>
<td>&lt;1.26</td>
<td>N/A</td>
<td>&lt;0.1</td>
</tr>
<tr>
<td>@ Hold-off (µs)</td>
<td>0.1 µs</td>
<td>4.5 ns</td>
<td>0.1 µs</td>
<td>5 µs</td>
<td>15 ns</td>
<td>150 ns</td>
<td>50 ns</td>
<td>N/A</td>
<td>3 ns</td>
</tr>
<tr>
<td>Peak PDE (%)</td>
<td>27</td>
<td>62</td>
<td>8</td>
<td>5.5</td>
<td>44</td>
<td>31</td>
<td>71</td>
<td>N/A</td>
<td>55</td>
</tr>
<tr>
<td>SPTR (ps FWHM)</td>
<td>6.3</td>
<td>30</td>
<td>7.8</td>
<td>235</td>
<td>51</td>
<td>42</td>
<td>28</td>
<td>27</td>
<td>7.5</td>
</tr>
</tbody>
</table>
4.2. Discriminator’s Input Signal Slope

In Figure 4a, each point is a time delay histogram for different threshold values. The measured slope of the first 1000 mV of the input signal is about 5 V/ns. Using this measurement, we adjust the parameters of the SPAD model, such as the junction capacitance, to obtain a better match between the measurements and simulations of the timing response of the quenching circuits (Figure 4b) [34,35]. The updated SPAD model is then used to obtain a more realistic timing response when designing a new quenching circuit. The set of five curves, with different junction capacitance, is used to simulate the time propagation delay of the quenching circuit as a function of the overdrive in the next section.

![Figure 4](image_url)

Figure 4. SPAD signal at the discriminator input reported from the comparator threshold value. In (a) is the measured SPAD signal slope using the optical setup described in Section 3.1. In (b) is a comparison between the measured signal and multiple simulations using the same setup and different SPAD junction capacitances in the SPAD SPICE model.

Figure 5 shows the extracted SPTR as a function of the SPAD signal slope from the same data. The measurement shows that a steeper slope leads to a better SPTR, since the contribution of the noise over the slope jitter will be smaller. However, the improvement saturates after 4 V/ns, suggesting that the noise over the slope jitter contribution becomes negligible compared to other timing jitter contributions. In the case of this SPAD at an excess voltage of 1.7 V, this condition (4 V/ns or higher slope) is achieved with a threshold of 1 V or lower. Therefore, one could expect that this SPAD combined with any QC with similar intrinsic timing jitter should obtain a similar SPTR as long as its threshold is below 1 V.

![Figure 5](image_url)

Figure 5. SPTR as a function of the calculated SPAD signal slope from Figure 4. The SPTR lowers as a function of the slope with a saturation around 4 V/ns.
4.3. Overdrive Variation Jitter

This section presents the simulated time propagation delay of the different QCs at multiple SPAD excess voltages. Here, we make use of the SPAD model updated from the input signal slope data set (Section 4.2). To determine the time propagation delay of the QC architecture, multiple simulations are performed by varying the SPAD capacitance as well as the different corners for the CMOS circuits. Figure 6a shows the propagation delay of the QCs as a function of the SPAD excess voltage for the CMOS TT (typical–typical corner case) and parameter set #3 (Figure 4).

![Figure 6. In (a) is the QC time propagation delay as a function of the SPAD excess voltage for multiple architectures. In (b) is the added jitter of the propagation delay variation considering a 30 mV FWHM variation of the excess voltage. The region of interest for this paper is at 1.7 V (since it is the excess voltage chosen for the measurement) and is boxed in red.](image)

The added jitter is obtained using the excess voltage derivative of the propagation delay multiplied by a 30 mV FWHM of excess voltage noise. This allows us to estimate the contributions of the different architectures. The value of 30 mV FWHM was reported in a previous study for a SPAD of the same architecture and was measured again on the SPAD under test [20]. Since the SPAD is operated at 1.7 V of excess voltage, this contribution was highlighted in Figure 6b. From this simulation, we expect the SPTR to be similar for the comparator and the 1 V inverter but a little worse for the 3.3 V inverters. The contribution of the time propagation delay was estimated for each SPAD model parameter set and each corner for the CMOS circuit, providing a range for this contribution. To estimate the range of the SPTR for the different architectures, it is possible to add in quadrature the evaluated added jitter to a baseline SPAD SPTR. Considering a baseline SPAD SPTR of 6.5 ps FWHM for each architecture, the estimated SPTR ranges are as follows: the comparator is between 6.5 and 8.7 ps FWHM, the UB 1 V Inv. is between 6.6 and 8.3 ps FWHM, the 1 V Inv. is between 6.9 and 9.0 ps FWHM, the 3.3 V Inv. UB is between 7.2 and 12 ps FWHM, and the 3.3 V Inv. STD is between 9.2 and 14 ps FWHM. These values are represented as the min and max values in Figure 7 and are summarized in Table 3.
4.4. Single-Photon Timing Resolution

The SPTR of the SPAD and the different QC architectures were measured and compared to the estimated jitter from the simulated time propagation delay of each discriminator. As shown in Figure 7, the measured SPTRs of the SPAD combined with the comparator are between 7.0 and 7.8 ps FWHM over 10 different samples. For the 1 V Inverter UB, the values are between 7.5 and 8.4 ps FWHM. For the 1 V Inverter STD, the values are between 8.1 and 9.4 ps FWHM. For the 3.3 V Inverter UB, the values are between 8.5 and 10.1 ps FWHM. These values are summarized in Table 3.

The measured SPTR is in good accordance with the estimated jitter of the time propagation delay variation of the different QC architectures as shown by the blue and red curves.

4.5. Cascode Transistor Impact on the SPTR

The SPTR of the SPAD and the different QC architectures with cascode transistor were measured and compared to the estimated jitter from the simulated time propagation delay of each discriminator.

Figure 8 shows the measured SPTR of the SPAD combined with the comparator for a cascode gate bias between 1 V and 1.4 V. The minimum simulated value is associated with a FF (fast-fast corner case) simulation case and the smallest SPAD junction capacitance (Figure 4b), the typical value is associated with a TT simulation case and the middle SPAD junction capacitance and the maximum simulated value is associated with a SS (slow–slow corner case) simulation case and the highest SPAD junction capacitance value. The measurements show that the SPTR degrades towards a 1 V cascode bias since the amplitude of the signal that goes to the discriminator input (the input signal overdrive) is getting smaller. For this result, two comparators with a cascode transistor were measured.
10 of 15

#1 stops seeing the SPAD signal at 1.02 V and comparator #2 at 1.07 V. The measurements were conducted up to 1.4 V of cascode gate bias to stay within the inverter maximum gate voltage of 1 V (see Figure 1). The SPTR follows the same improvement as shown in simulations. The SPTR obtained at 1.4 V is around 8 ps FWHM and simulation shows that the contribution is negligible from 1.5 V of gate cascode bias, thus it was not possible to investigate it in measurements. Nevertheless, 8 ps FWHM at 1.4 V of cascode gate bias shows that a well-implemented cascode transistor allows for an improvement in the maximum excess voltage of the SPAD while not degrading the SPTR.

Figure 8. Measured SPTR as a function of cascode bias for the comparator-based QC.

Figure 9 shows the measured SPTR of the SPAD combined with the two different 1 V inverter QCs as a function of the cascode bias, the standard inverter and the unbalanced inverter (Table 1). For the QC with the STD inverter (Figure 9b), the measurements show that the SPTR degrades significantly towards a 1 V cascode bias since the amplitude of the signal that goes through the discriminator is getting smaller and smaller and the circuit stops seeing the SPAD signal at around 1.05 V. The SPTR gets to a stable value once the cascode gate bias is higher than 1.3 V.

The SPTR variation of the 1 V inverter UB is a little different from the other two (Figure 9a). Below 1 V of cascode bias, there is a second phenomenon that occurs. The SPAD signal is capacitively coupled to the discriminator through the cascode for a cascode bias of 800 mV and below until it reaches a breaking point. This phenomenon occurs only for very
small values of the SPAD junction capacitance and hence at a very high slope. For higher capacitance values, as shown by the maximum value of the simulation, the behavior is similar to the other QC architectures. The SPTR for a 1.2 to 1.4 V cascode gate bias is around 8 ps FWHM and simulations show the cascode transistor contribution to be negligible from 1.1 V of cascode bias.

5. Discussion

This paper presents a small area and thin SPAD that can be read out by either inverters or a comparator with an adjustable threshold and obtain a very similar SPTR. This is important for PDCs since there is an array of QCs in a system and an inverter-based QC does not have static power consumption. For more information on the comparison of the power consumption for both architectures, a comparison for a PET scanner is presented in [20]. Nevertheless, the comparator seems to be marginally better than the inverter to reach the ultimate SPTR performance.

Table 4 compares the different architectures considering their size, maximum excess voltage, intrinsic QC jitter and the best SPTR achieved for each circuit at an excess voltage of 1.7 V. The 3.3 V inverter has a higher SPTR than the other QCs, and this is explained by the SPAD operating point, which causes a higher time propagation delay variation due to the higher threshold of the inverter (Figure 6).

Table 4. Results summary per QC architecture.

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Comp.</th>
<th>Cascoded UB 1 V Inv.</th>
<th>Cascoded Std 1 V Inv.</th>
<th>Cascoded Comp.</th>
<th>UB 3.3 V Inv.</th>
<th>Std 3.3 V Inv.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total discriminator size (μm²)</td>
<td>104</td>
<td>87</td>
<td>69</td>
<td>167</td>
<td>88</td>
<td>40</td>
</tr>
<tr>
<td>Maximum excess voltage (V)</td>
<td>3.3</td>
<td>4.3</td>
<td>4.3</td>
<td>6.6</td>
<td>3.3</td>
<td>3.3</td>
</tr>
<tr>
<td>QC jitter (ps FWHM)</td>
<td>4</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Static power consumption (μW)</td>
<td>165</td>
<td>0</td>
<td>0</td>
<td>165</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Best SPTR (ps FWHM) at 1.7 excess voltage</td>
<td>6.3</td>
<td>6.8</td>
<td>8.1</td>
<td>7.0</td>
<td>8.4</td>
<td>9.5</td>
</tr>
<tr>
<td>Best SPTR (ps FW(1/10)M) at 1.7 excess voltage</td>
<td>14.31</td>
<td>17.9</td>
<td>18.8</td>
<td>16.2</td>
<td>25.8</td>
<td>30.2</td>
</tr>
<tr>
<td>Best SPTR (ps FW(1/100)M) at 1.7 excess voltage</td>
<td>32.4</td>
<td>36.6</td>
<td>36.14</td>
<td>44.2</td>
<td>45.6</td>
<td>49.1</td>
</tr>
</tbody>
</table>

To design a QC, its propagation delay variation as a function of overdrive must be taken into account and we showed it can be precisely simulated. As seen in Figure 7, the contribution of the propagation delay variation combined with the excess voltage variation (or overdrive variation) fits the measured SPTR for the different QC architectures. Even though the measured intrinsic jitter of the comparator is higher than the inverter for a standard electrical signal (Table 1), it provides the lowest jitter when coupled to a SPAD. Hence, to design the optimal QC for a specific SPAD, certain SPAD characteristics must be measured beforehand. First, using a comparator with an adjustable threshold, the signal slope of the SPAD should be measured to adjust the SPAD model used in simulation. Second, the SPAD excess voltage variation should be measured using a comparator with an adjustable threshold. This will provide a suitable testbench for the QC simulation and help identify if a comparator is required for the designed SPAD. It is also important to validate if the threshold has an impact on the SPAD SPTR. For instance, most bigger SPADs or externally quenched SPADs do not have similar performances in a wide threshold range [23,27].
Furthermore, if the cascode transistor that increases the excess voltage is designed to minimize its impact on the slope and overdrive of the input signal, the cascode does not degrade the timing resolution. Its impact on the slope and on the overdrive can be estimated through simulations.

A last point that must be considered, when performing very precise SPTR measurements, is the set of oscilloscope parameters that must be optimized to obtain the best possible measurements [40]. The measurements presented in this paper were conducted with a Teledyne Lecroy SDA 6000A oscilloscope, with 40 GSa/s and a 6 GHz bandwidth. To determine if the oscilloscope was limiting the measurements, the measurements with the comparator and the 1 V inverter UB were also conducted with an Agilent MSO-X 90324A, with 80 GSa/s and a 13 GHz bandwidth. A better sampling rate allows us to minimize the quantization noise of the acquisition, but the bandwidth and the noise floor of the scope must also be considered. In our case, simply switching to the Agilent oscilloscope to increase the sampling rate provided the same SPTR results. Reducing the vertical scale parameter allowed us to reduce the noise floor from the oscilloscope to obtain a better SPTR [41]. This lowered the SPTR of every SPAD and QC measured by about 1 ps FWHM (Figure 10). This resulted in an SPTR of 6.3 ps FWHM for a SPAD–comparator pair (Figure 10c) and 6.8 ps FWHM for a SPAD combined with a 1 V inverter UB (Figure 10d).

Figure 10. SPTR for two different quenching circuit architectures, measured with two different oscilloscopes. In (a,b), the measurements were taken with a Teledyne Lecroy SDA 6000A oscilloscope, with 40 GSa/s and a 6 GHz bandwidth; in (c,d), the measurements were conducted with an Agilent MSO-X 90324A, with 80 GSa/s and a 13 GHz bandwidth. The SPTR of the reference QC with a comparator without the cascode is in (a,c) and the SPTR measurement of the cascoded 1 V inverter UB is in (b,d).
6. Conclusions

This paper presents an implementation comparison between six different leading-edge discriminators for a sub-10 ps FWHM SPTR SPAD. This paper reports a 6.3 ps FWHM SPTR for a SPAD read out by an adjustable threshold comparator and a 6.8 ps FWHM SPTR for an optimized 1 V inverter using a cascode transistor for a higher excess voltage. We demonstrated that a comparator is not always required for a small area and thin SPAD. To design an optimal quenching circuit, the following SPAD characteristics must be measured beforehand: the SPAD signal slope and the SPAD excess voltage variation. This provides a suitable SPAD model for the QC simulation and helps identify if a comparator is required through time propagation delay simulation as a function of overdrive variation. It is important to also validate if the threshold has an impact on the SPAD SPTR, as different SPAD architectures can have a different signal slope and hence a different timing response to the same discriminator. Using this method, we were able to estimate the timing measurements of the six different leading-edge discriminators and the measurements are in accordance with the simulations. The results also show that a cascode transistor can be used to increase the maximum excess voltage of the QC with minimum impact on the SPTR.


Funding: The authors acknowledge the financial contributions of the Natural Sciences and Engineering Research Council of Canada (CRSNG), the Arthur B. McDonald Canadian Astroparticle Physics Research Institute (Queen’s U., ON), the Institut quantique (U. de Sherbrooke, QC) and the Canada Foundation for Innovation (CFI).

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

2. Conti, M.; Bendriem, B. The new opportunities for high time resolution clinical TOF PET. *Clin. Transl. Imaging* 2019, 7, 139–147. [CrossRef]


29. Marwick, M.; Andreou, A. Single photon avalanche photodetector with integrated quenching fabricated in TSMC 0.18 µm 1.8 V CMOS process. *Electron. Lett.* **2008**, *44*, 643–644. [CrossRef]


**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.