

Review

Recent Developments in Single-Photon Avalanche Diode (SPAD) Technologies: From Device Engineering to Optimized Photonic Performance in Quantum Communication

Masoud Abrari, Seyyedeh Tahereh Sajjadian, Parsa Nedaei and Majid Ghanaatshoar *

Laser and Plasma Research Institute, Shahid Beheshti University, Tehran 1983969411, Iran; m_abrari@sbu.ac.ir (M.A.); t.sajjadian@mail.sbu.ac.ir (S.T.S.); p.nedaei@mail.sbu.ac.ir (P.N.)

* Correspondence: m-ghanaat@sbu.ac.ir

Abstract

The ability to detect individual photons with exquisite temporal precision has transformed single-photon avalanche diodes (SPADs) into indispensable tools across photonics, with quantum communication emerging as one of their most demanding frontiers. In recent years, device engineering breakthroughs, including refined junction geometries, advanced avalanche quenching schemes and scalable array integration, have redefined the limits of SPAD performance. Parallel advances in fabrication precision and CMOS-compatible architectures have not only expanded spectral sensitivity from the visible to the near-infrared but also enabled systematic suppression of dark count rates, afterpulsing, and timing jitter. These developments have directly impacted the feasibility and robustness of quantum key distribution and time-correlated single-photon counting, where detection efficiency and noise suppression determine system fidelity. This review unifies recent progress in SPAD technology, weaving together innovations in device design, fabrication strategies, and parameter optimization to reveal their collective influence on next-generation quantum-enabled photonic systems. Looking forward, the convergence of hybrid material platforms, on-chip photonic–electronic co-integration, and intelligent quenching control is poised to elevate SPAD performance to meet, and potentially exceed, the stringent requirements of future quantum communication infrastructures.

Keywords: single-photon avalanche diodes (SPADs); device engineering; quantum communication

Received: 9 February 2026

Revised: 28 June 2026

Accepted: 30 June 2026

Published: 4 July 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\) license](https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Single-photon detection, as one of the most significant achievements in optical technology, has played a fundamental role in expanding the frontiers of knowledge and industry. The ability to register a single photon is not only a benchmark for the sensitivity and precision of optical systems, but also provides a foundation for the advancement of quantum science and emerging technologies. At the fundamental level, single-photon detectors enable precise investigations of light–matter interactions on the quantum scale, allowing researchers to study processes such as spontaneous emission, photon correlations, and quantum superposition with unprecedented accuracy. On the applied side, the

importance of single-photon detection becomes even more pronounced. In classical photonics, this technology underpins improvements in systems such as high-sensitivity biomedical imaging, the detection of extremely weak optical signals in astronomy, and the development of time-of-flight methods for 3D imaging. In the quantum domain, single-photon detectors have become indispensable for implementing secure communication protocols based on the principles of quantum mechanics. Technologies such as quantum key distribution, optical information processing, and quantum metrology all rely on detectors capable of determining the presence or absence of a single photon with high reliability. Thus, single-photon detection represents not only an advanced technological capability but also a fundamental necessity for transitioning from classical photonic systems to the quantum infrastructures of the future. Its importance across both scientific and industrial domains has driven extensive efforts to improve efficiency, reduce noise, and enhance scalability, ensuring that single-photon detectors remain a top priority in interdisciplinary research.

Single-photon avalanche diodes (SPADs) have emerged as one of the most important single-photon detection technologies in recent decades, gaining a prominent role in both photonics and quantum science. Their unique feature lies in the ability to initiate an electron avalanche with the absorption of a single photon, thereby generating a measurable electrical signal. This capability is achieved by biasing the diode in Geiger mode, where the electric field is sufficiently strong that even a single carrier can trigger a cascade of impact ionization events. In contrast to linear detectors, which provide an output proportional to light intensity, SPADs deliver a binary output that records the presence or absence of a photon as a distinct pulse. This characteristic makes SPADs ideal for applications where timing precision, high sensitivity, and reliable operation under extremely low-light conditions are crucial.

Another key advantage of SPADs over many competing technologies is their compatibility with CMOS processes. This feature has enabled not only their use in research applications but also their potential for large-scale production, integration with readout circuitry, and implementation in large arrays. Such scalability has paved the way for the development of single-photon cameras, high-speed LiDAR systems, and reliable quantum detectors. As a result, SPADs have gradually evolved beyond a purely laboratory technology and are now recognized as a cornerstone in the transition from classical photonics to the quantum photonic systems of the future.

SPADs trace their origins to research on avalanche breakdown in p-n junctions during the early 1960s. Revolutionary work by R. J. McIntyre at RCA and Ronald Haitz at the Shockley laboratory first proposed that solid-state diodes could detect single photons by operating in a Geiger mode above breakdown [1,2]. Haitz's studies of microscopic breakdown sites (so-called microplasmas) in silicon laid the groundwork, modeling the first diode capable of self-sustained avalanche pulses in response to single photons. By the early 1970s, experiments had confirmed that avalanche photodiodes (APDs) could behave like Geiger counters for photons, enabling single-photon detection in practice [3]. The concept of a solid-state photomultiplier had been born, offering a compact alternative to bulky photomultiplier tubes for single-photon detection.

Through the 1980s, improvements in materials and circuit techniques greatly enhanced SPAD performance. A major advance was the introduction of epitaxial silicon SPADs in the late 1980s, which used a thin high-purity layer to better control the electric field profile. This reduced dark counts and timing jitter by minimizing extraneous carrier diffusion. Around the same time, Perkin-Elmer developed the SLiK silicon SPAD structure, which further suppressed noise pulses and became a staple of early commercial single-photon counters. By the early 1990s, research teams (notably in Milan and EPFL) had

refined SPAD designs to achieve picosecond-scale timing resolution and lower afterpulsing, making SPAD detectors reliable tools in fields like fluorescence lifetime spectroscopy and high-energy physics [4,5].

As microelectronics advanced, the 1990s brought a push to fabricate SPADs in standard CMOS processes to allow on-chip integration of detectors and readout circuitry. Earlier SPAD arrays used custom fabrication and could not be monolithically integrated with logic, limiting them to small modules or multi-anode assemblies. Attempts in 2003 demonstrated the first SPAD devices built in standard CMOS, initially as single pixels and soon as small arrays [6]. Shortly thereafter, fully integrated 2D SPAD arrays were achieved. A 4×8 SPAD array with on-chip counting was reported by Rochas at EPFL [7]. An intermediate innovation was the silicon photomultiplier (SiPM), which is essentially an array of many micro-SPAD pixels summed together to act as one sensor. The SiPM concept, first developed by Bondarenko and collaborators in Russia [8], demonstrated that large-area photon counters could be built from tiled SPAD microcells operating in unison. By the end of the 2000s and into the 2010s, CMOS SPAD arrays grew to hundreds or thousands of pixels, with improving pixel fill factors (FFs) and ever-lower noise. These developments firmly established SPADs as a mainstream technology in optoelectronics.

Today's state-of-the-art SPADs are high-performance devices and are increasingly described as "quantum-ready" components, owing to their ability to resolve single quanta of light with extreme precision. Continued refinements in fabrication have yielded SPADs with dark count rates well below 10 Hz and timing jitter on the order of tens of picoseconds. Modern CMOS SPAD sensors can integrate on-chip time-to-digital converters (TDCs) for timestamping photons with <100 ps resolution and can operate with dead times of only a few nanoseconds. Large arrays (spanning > megapixels) have been reported, often using advanced 3D integration and microlens techniques [9]. These advances make SPAD arrays attractive for quantum imaging and sensing, where high spatio-temporal resolution at the single-photon level is required.

SPAD technology has thus become a key enabler in quantum information science. Single-photon avalanche detectors are widely used in quantum cryptography, such as quantum key distribution (QKD) links, where they serve as the receivers that register each photon of the secret key stream. Likewise, SPAD-based devices have been employed for quantum random number generation, exploiting the inherent randomness of photon detection to produce cryptographic random bits. In 2025, SPADs were truly quantum-ready. They are compact, scalable, and capable of deployment as fundamental building blocks in quantum communication networks, quantum sensors, and other quantum-enabled devices of the future.

Alongside the development of SPAD technology, several other approaches have been proposed for single-photon detection, each with its own advantages and limitations. Photomultiplier tubes (PMTs) were among the earliest single-photon detection devices [10]. Utilizing the photoelectric effect and electron multiplication in a vacuum, they offered exceptionally high optical sensitivity. Although PMTs provide low intrinsic noise and high gain, their large size, requirement for very high operating voltages, fragility, and incompatibility with integrated systems make them unsuitable for many modern applications [11]. In contrast, transition-edge sensors (TESs), which operate based on the superconducting phase transition, are capable of detecting the precise energy of individual photons and can even perform photon-number resolution [12]. However, their requirement for extremely low operating temperatures and slow recovery times restricts their use to specialized laboratory environments with advanced cryogenic systems. Visible Light Photon Counters (VLPCs) also provide efficient single-photon detection within the visible range. Nevertheless, they typically require cryogenic cooling and face challenges in terms of stability and integration with modern electronic circuits [13]. Hybrid photodetectors

(HPDs), which combine the features of PMTs and semiconductor diodes, have attempted to merge the advantages of both technologies. However, their structural complexity, high-voltage requirements, and limited performance range have confined their application to specific areas [14]. Compared with these technologies, SPADs offer significant advantages, such as room-temperature operation, compact size, CMOS compatibility, high temporal resolution, and the ability to be integrated into large arrays. These attributes make SPADs a practical, scalable, and versatile solution for a wide range of applications, from medical imaging to advanced quantum systems.

In recent years, advances in quantum information and communication technologies have created a growing demand for devices capable of detecting single photons with extremely high spatial and temporal precision. Owing to their unique characteristics, SPADs have emerged as practical and reliable detectors in many of these quantum systems. In QKD, the receiver must detect individual photons carrying encrypted information with high accuracy and speed. SPADs meet many of QKD's stringent requirements by offering high detection efficiency, picosecond-level timing resolution, and room-temperature operation. The use of SPAD arrays in such systems not only enables scalability and higher data rates but also facilitates integration with other optical and electronic circuits. Their ability to operate at room temperature, compatibility with standard CMOS manufacturing, and flexibility in architectural design have made SPADs one of the few accessible options for implementing portable and cost-effective quantum systems. SPADs now play a pivotal role not only in laboratory research but also in the development of real-world infrastructure for quantum communication and information processing.

Given the broad range of applications and the rapid pace of technological progress in single-photon detection, there is a growing need for a comprehensive and structured review of the current status and future directions of SPAD technology. This article is organized around three key thematic axes, each playing a critical role in shaping the next generation of optical and quantum devices. First, the device engineering aspects are examined, covering novel SPAD architectures, CMOS-based fabrication technologies, large-scale array implementation, and the challenges of optical–electronic integration. The second one addresses performance optimization, where strategies for reducing dark noise and afterpulsing, enhancing detection efficiency, improving timing precision, and managing dead time are analyzed and compared. Finally, the third axis explores the role of SPADs in quantum technologies, including their use in quantum key distribution, single-photon imaging, and quantum photonic systems. Emerging trends and possible future directions for SPAD development are also discussed. By drawing upon authoritative sources and the latest state-of-the-art research, this review aims to provide a comprehensive, analytical, and application-oriented overview of the evolving landscape of SPAD technology, establishing a scientific foundation for the advancement of next-generation photonic systems.

2. Operating Principles of SPADs

Avalanche diodes are a type of p-n junction diode specifically designed to operate under high reverse bias voltages. The electrical behavior of a diode depends on the applied voltage. In forward bias, a positive voltage is applied to the p-region and a negative voltage to the n-region, reducing the internal electric field and allowing carriers to cross the junction, resulting in a conduction current. In reverse bias, a positive voltage is applied to the n-region and a negative voltage to the p-region, increasing the internal electric field and resulting in a very small leakage current. However, if the reverse voltage exceeds a critical threshold, breakdown takes place and the current increases significantly.

Two main breakdown mechanisms in reverse bias are the tunneling effect and avalanche multiplication. In tunneling, electrons can penetrate the energy bandgap. In avalanche multiplication, at high reverse voltages carriers gain enough energy to ionize atoms and generate new electron–hole pairs. This process can become self-sustaining, resulting in a large current. In this operation, absorption of even a single photon can generate a large electric current [15]. Before discussing the operation in detail, it is essential to provide a concise overview of the avalanche process. When the electric field in a semiconductor exceeds a certain value, charge carriers can acquire sufficient kinetic energy to generate additional electron–hole pairs through the avalanche process, as schematically illustrated in Figure 1a. Consider an electron in the conduction band (labeled 1 in Figure 1b). If the electric field is sufficiently strong, this electron can gain enough kinetic energy before colliding with a valence electron. Upon collision, the high-energy conduction electron transfers part of its energy to the valence electron, promoting it to the conduction band and thereby generating an electron–hole pair (labeled 2 and 2'). The newly generated carriers are then accelerated by the electric field, colliding with other valence electrons and producing additional electron–hole pairs (e.g., 3 and 3' and 4 and 4'), and the process continues in a self-sustaining chain reaction. These, in turn, carry on the process by creating additional electron–hole pairs.

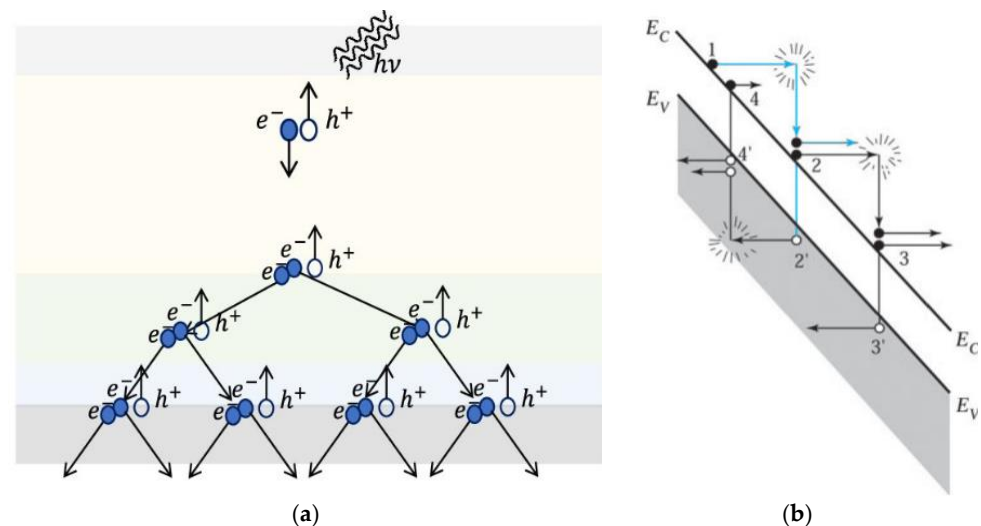


Figure 1. (a) Avalanche phenomenon triggered by impact ionization in avalanche photodiodes [16]. (b) Band diagram showing reverse bias avalanche multiplication, adapted from Ref. [17].

2.1. Operational Modes of Avalanche Photodetectors

The operation mode of avalanche photodetectors varies with the reverse bias level in relation to the breakdown voltage, yielding two possible modes: linear and Geiger. In linear mode, an avalanche photodiode is biased at a constant reverse voltage below its breakdown threshold. Under this condition, the primary photocurrent generated in the absorption region undergoes controlled avalanche multiplication by a factor, M , so that the number of signal photoelectrons, n_s , is proportional to the number of exciting photons, n_p . This proportionality is expressed as $n_s \sim \eta M n_p$, where η and M are the quantum efficiency and the avalanche multiplication factor of APD, correspondingly [15]. As a result, the output current of a linear APD scales linearly with the incident optical power, allowing accurate measurement of signal amplitude. Owing to this property, linear-mode APDs are widely used in optical receivers, where they enhance weak optical signals above the noise level of subsequent electronic amplifiers [18]. In Geiger mode, the SPAD is biased above its avalanche breakdown voltage, allowing a single photon to trigger a self-sustaining avalanche in the p–n junction. Geiger mode is the key operating principle that enables the

detection of single photons in devices such as SPADs. In this mode, the biased voltage creates a very strong electric field within the depletion region. When a single photon is absorbed, it generates an electron–hole pair. Under the influence of the strong electric field, these carriers can initiate an avalanche multiplication process through impact ionization. As a result, the initial charge generated by a single photon is rapidly amplified into a large current pulse that can be easily detected by external electronics. The significance of Geiger mode lies in its ability to convert an extremely weak optical event into a strong and measurable electrical signal. Since the avalanche process produces a nearly identical output pulse regardless of the number of initial carriers, the detector operates in a binary manner, where each detected photon generates a distinct output “click.” Another major advantage of Geiger mode is its extremely high internal gain, typically ranging from 10^5 to 10^6 . This high gain eliminates the need for substantial external amplification and allows reliable single-photon detection even under very low light conditions. In addition, the large avalanche signal improves the signal-to-noise ratio and helps distinguish genuine photon events from electronic noise. Geiger mode also offers excellent timing performance. Because the avalanche builds up within picoseconds to nanoseconds, the arrival time of a photon can be measured with high precision. This capability is crucial for applications requiring accurate timing information.

A typical I–V characteristic of a SPAD is shown in Figure 2. In the absence of charge carriers, the device remains inactive in region 1. Only the arrival of a photon or a thermally generated carrier can trigger the avalanche process. When an avalanche event begins, the current through the diode rises abruptly, moving the device into region 2. This regime, however, is unstable; the large and persistent avalanche current rapidly generates heat. Therefore, a proper operation is essential. To achieve this, the reverse voltage must be lowered below the breakdown voltage to stop the avalanche. Afterward, the SPAD can enter region 3 and then be reset to region 1.

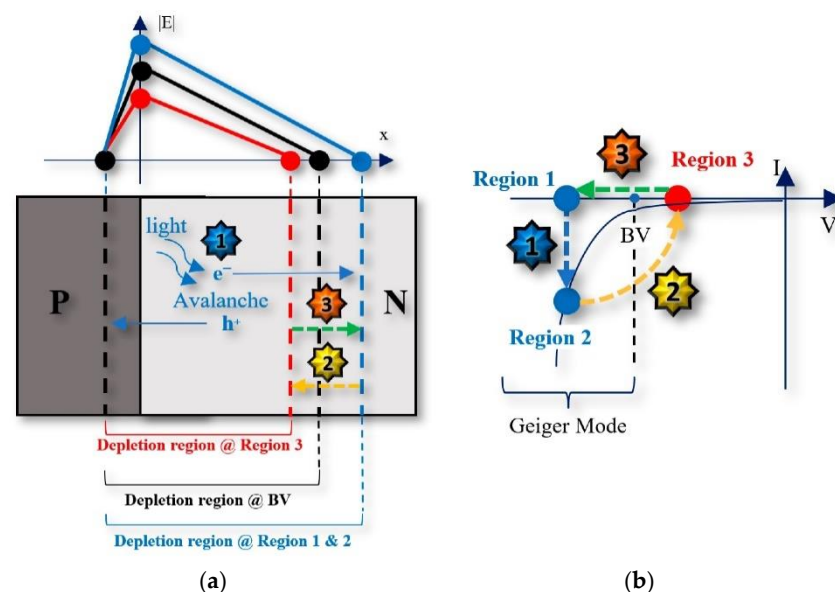


Figure 2. (a) Simple SPAD illustration. (b) Current–voltage (I–V) characteristic curve [19].

2.2. Quenching

The mechanism that reduces the reverse voltage below the breakdown voltage is known as quenching. It is a process used to stop the avalanche current generated after the absorption of a single photon. The quenching plays a critical role in enabling photon detection by biasing the SPAD into its active Geiger mode. Upon the absorption of a single photon, the SPAD initiates an avalanche breakdown, generating a transient photocurrent.

This photocurrent is converted into a voltage signal through a current-to-voltage (I-V) conversion stage, which subsequently deactivates the stop counting mechanism, thereby completing the photon counting process. Following detection, the quenching circuit reduces the bias voltage beneath the breakdown threshold, terminating the avalanche and resetting the SPAD to a nonconductive (cut-off) state for the next detection cycle [20]. The circuit design must enable fast avalanche detection, rapid quenching and rapid restoration of the bias voltage to its initial level. The quenching circuit is designed with the goal of achieving a fast response time, a high count rate, and low noise, and it includes optimized selection between passive and active quenching methods.

2.2.1. Passive Quenching

Passive quenching is a widely used method due to its simplicity and low cost, making passive quenching circuits (PQCs) a popular choice in many applications. In a PQC the avalanche is controlled by a load resistor or a transistor which lowers the bias across the SPAD below the breakdown voltage. The passive quenching operation consists of two main phases: quenching and reset. During the reset step, the SPAD bias voltage is restored above the breakdown level to re-enable Geiger-mode operation. During the quenching phase, the junction capacitance discharges, leading to a voltage drop across the diode. The value of the quenching resistor (R_Q) must be carefully optimized to balance fast avalanche suppression with circuit stability. In this phase, the avalanche current decreases exponentially, reaching a peak value of approximately $I_{D,peak} = V_{EX}/R_D$, where R_D is the internal resistance of the diode, determined by the sum of the space charge layer resistance and the ohmic resistance of the neutral regions crossed by the avalanche current, and V_{EX} is the excess bias voltage, i.e., the voltage above the breakdown voltage that ensures Geiger-mode operation of the SPAD [21]. The selection of the resistor can have a significant impact on the performance of the SPAD. This resistor needs to be high enough in value to limit the avalanche current and prevent self-sustaining avalanches. In general, the key parameters are directly affected by the design and components of the quenching circuit [22–24].

Although the series resistors are essential for quenching the avalanche current, their high value resistance significantly delays the recharge of the depletion region, thus imposing a fundamental limitation on the SPAD recovery time. The use of bias-dependent adaptive resistors has been proposed as a replacement for fixed resistors in quenching methods for SPADs. For instance, by using a resistor composed of Pt/Al₂O₃/Ag in conjunction with a commercial silicon SPAD, the avalanche pulse width reduced to approximately 30 nanoseconds, which is about 10 times faster than conventional passive quenching methods [25].

2.2.2. Active Quenching

One intrinsic drawback of the PQC is the hold-off time of the SPAD. During the reset stage, the SPAD voltage increases and another photon may trigger an undesired avalanche. This unwanted avalanche can keep the comparator output continuously high, resulting in a reduction in the recorded counts. To overcome PQC limitation, active quenching is proposed. In an active quenching circuit (AQC) [26–28], the avalanche current passes through a resistor, where it is sensed by a comparator. The comparator monitors the voltage developed across the resistor and compares it with a predefined reference level. When the voltage exceeds this reference, the comparator generates a digital signal, indicating that an avalanche event has occurred. Following this detection, the quenching transistor promptly reduces the SPAD bias voltage to a value below the breakdown threshold, thereby interrupting the avalanche before it becomes self-sustaining. In contrast to passive quenching, the active approach intentionally and immediately lowers the bias voltage as

soon as the avalanche is detected, ensuring that the SPAD is quickly and accurately reset for subsequent photon detection. Because active quenching dynamically modulates the bias, the recovery time is significantly shorter than that of the passive method. Several parameters influence the time required for this process, such as the sensitivity of the AQC, the physical distance between the SPAD and the AQC, and the maximum value of the avalanche current [23].

In recent years, to improve the speed and stability of quenching in SPAD detectors the use of fast active quenching circuits (FAQCs) based on heterojunction bipolar transistor (HBT) technologies such as SiGe and InGaP/GaAs has gained significant attention [29]. These circuits are designed for rapid quenching, low noise and cost-effective implementation in continuous detection mode.

In this architecture, the avalanche current is sensed at the SPAD cathode using a high-speed SiGe HBT comparator, while a relatively high quenching voltage (~ 6 V) is applied to the SPAD anode to overcome the low voltage tolerance of HBTs. The output of the comparator drives a D flip-flop followed by an RF amplifier based on InGaP/GaAs. This amplifier simultaneously performs the quenching and applies the hold-off time.

Active quenching active reset (AQAR) circuits have also played a significant role in improving the performance of SPAD-based pixels [30]. Ultrafast AQAR designs implemented in 28 nm fully depleted silicon on insulator (FD-SOI) CMOS technology have achieved quenching times below 630 ps and over 90% reductions in afterpulsing probability [31]. In addition, the use of the innovative dual-node avalanche detection method, which senses the avalanche at both the body and gate of transistors, has significantly enhanced detection sensitivity. These circuits are compatible with monolithic 3D SPAD structures, enabling the first full integration of AQAR with SPAD on a single die. Apart from the inherently higher noise in these structures, the AQAR circuit has ensured stable device operation even at high excess bias voltages.

2.3. Performance Metrics

The performance of SPADs is governed by several parameters that critically affect their operation. Photon detection efficiency (PDE), which represents the probability that a photon is detected by the SPAD, is one of these parameters. Thermal, tunneling, and trapping processes in the semiconductor material also result in the generation of electron–hole pairs that trigger avalanches, and these events give rise to a finite dark count rate (DCR). Another important parameter is afterpulsing, which is caused by the delayed release of carriers that become trapped during the primary avalanche. Afterpulsing phenomena can initiate secondary avalanches and add correlated noise to the detector output. A fundamental metric that characterizes SPAD performance is the temporal uncertainty between the arrival of a photon and the generation of its corresponding output pulse, a fluctuation commonly referred to as timing jitter. This timing variation limits the precision of single-photon detection systems and plays a crucial role in applications that require accurate time measurements. After an avalanche, the dead time specifies the period during which the SPAD cannot detect any subsequent photons. This limitation directly affects the maximum count rate of the detector and is a critical factor in high-intensity photon applications. Optimizing these parameters is essential for high-performance photon detection and guides SPAD design for advanced applications.

2.3.1. PDE

PDE is a metric used to evaluate the performance of a SPAD in converting incoming photons into a detectable electrical signal. In other words, PDE represents the probability that a photon entering the active area will trigger an avalanche process, resulting in a transient current. The PDE depends on three factors:

Photon absorption probability (η_{abs}): Each photon must first be absorbed in the active region. Otherwise, it may be absorbed or reflected by dielectric layers, metals, or interfaces. Therefore, photon absorption depends on the wavelength, thickness and material properties of the layers.

Carrier collection efficiency (P_{coll}): After absorption, electrons and holes are generated and must reach the region capable of triggering an avalanche. The geometry of the active area and the electric field play a crucial role in efficient carrier collection.

Avalanche triggering probability (P_{trigg}): Even if the carriers reach the appropriate region, an avalanche must be initiated, and the likelihood of this depends on the excess bias voltage and the SPAD's internal structure. Therefore, the PDE can be approximated as the product of these three factors, as $\text{PDE} = \eta_{\text{abs}} \cdot P_{\text{coll}} \cdot P_{\text{trigg}}$ [32].

2.3.2. DCR

In SPADs, avalanches can arise not only from photon absorption but also from unwanted events occurring in complete darkness. The DCR is the count of pulses produced without any incident photons, arising from thermal or tunneling effects that lead to false avalanches [33]. At high temperatures, thermal excitation makes the main contribution to the DCR, whereas at low temperatures, but in a strong electric field, the main contribution to the DCR will be made by tunneling excitation caused by excess bias. These unwanted events are closely linked to lattice defects in the diode. The thermal contribution is primarily governed by trap-assisted carrier generation and tunneling via trap states. The trap-assisted mechanism is commonly described by the Shockley–Read–Hall (SRH) model.

In the high electric field of the depletion region, carriers may undergo band-to-band tunneling from the valence to the conduction band, a process that is largely independent of temperature. As shown in Figure 3, DCR can be reduced by lowering the temperature and optimizing device fabrication, including defect control and electric field design.

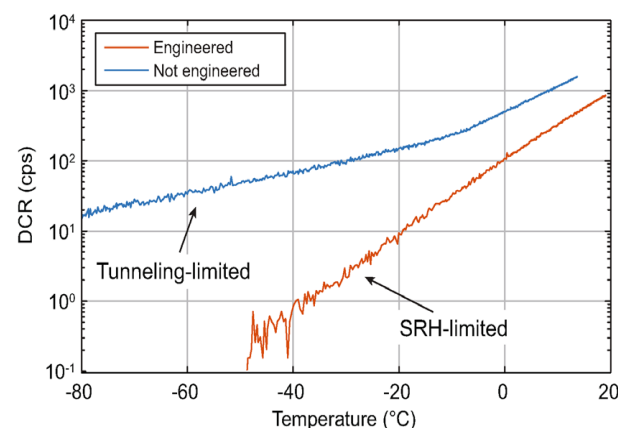


Figure 3. DCR dependence on temperature for SPADs with and without engineering of the electric field [34].

2.3.3. Afterpulsing

Besides the DCR, afterpulsing can induce unwanted events in SPADs. During an avalanche, some carriers may be trapped and later released, which can result in a train of pulses following single-photon absorption [35]. This phenomenon originates from the presence of trap levels released during the avalanche process. During this process, a portion of the charge carriers become trapped in regions associated with crystal defects or at the interface. The release of these trapped carriers occurs randomly and at unpredictable

times since it is governed by the finite lifetimes of the traps. The lifetime of each trap depends on the nature of the defect, the trap depth within the energy bandgap, and the thermal conditions of the device. Consequently, the gradual release of these carriers can induce afterpulsing noise in the output signal, thereby degrading the timing performance of the detector [36]. Afterpulsing is usually modeled using a statistical approach. The total afterpulsing probability can be derived from $p_{ap} = \int_0^\infty \eta_{ap}(\tau) d\tau$, where τ is the statistical time between trapping and release of the trapped carrier and η_{ap} is the afterpulsing probability density. Afterpulsing minimization is a key factor in SPADs for quantum applications. Low afterpulsing leads to faster quenching recovery and therefore the higher speed of the device. To mitigate this effect, several strategies are commonly employed. One effective approach is the use of an appropriate dead time after each avalanche event, allowing sufficient time for trapped carriers to detrapp before the detector is rearmed. Lowering the operating temperature also significantly reduces trap-assisted carrier generation and release, thereby decreasing afterpulse probability. In addition, the implementation of active quenching circuits helps to rapidly extinguish the avalanche, minimizing the number of carriers injected into traps. Careful optimization of the excess bias voltage further reduces carrier trapping by limiting avalanche charge. Finally, improvements in material quality and device fabrication can reduce defect density, thereby suppressing the origin of trap states. Together, these techniques effectively reduce false counts and improve the quantum bit error rate in quantum communication systems [37].

The configuration of the quenching circuit plays a critical role in reducing noise effects and suppressing afterpulsing in SPADs. The most effective approach is the use of an active quenching circuit (AQC), where the avalanche event is rapidly detected and the bias voltage is quickly reduced below the breakdown level. This fast and controlled quenching significantly limits the avalanche charge, thereby reducing carrier trapping and afterpulse probability. An improved configuration is the active quenching with active reset, which not only quenches the avalanche but also applies a well-defined dead time before resetting the detector, allowing trapped carriers to be released safely. In contrast, passive resistive quenching (PRQ) is simpler but slower in extinguishing the avalanche, leading to higher charge flow and increased afterpulsing, making it less suitable for low-noise quantum applications. Advanced designs often combine a fast discriminator/comparator, feedback control, and programmable dead time circuits to optimize both timing performance and noise suppression.

2.3.4. Timing Jitter

Timing jitter in SPADs represents the temporal uncertainty between photon absorption and the generation of the corresponding output pulse [38]. This uncertainty primarily arises from fluctuations in avalanche buildup variations in carrier transit time. This effect determines the temporal resolution of the detector and directly affects applications such as quantum communication. This mechanism is usually characterized by the FWHM of temporal counts. Timing jitter mainly depends on the drifting of carriers in the depletion region. The speed and uniformity of carrier passage through this region determine the time uncertainty between photon absorption and the onset of the avalanche process.

The excess bias voltage across the junction can influence the timing jitter by changing the avalanche build-up time. Higher excess voltage leads to a faster formation of a stable avalanche. This acceleration occurs because the amplified electric field increases the impact ionization coefficients for both electrons and holes, allowing the carrier population to multiply at a significantly faster rate. Consequently, the statistical fluctuations in the avalanche build-up time, often referred to as the deterministic component of the jitter, are

minimized [39]. In conclusion, minimizing timing jitter and enhancing temporal resolution can be effectively achieved through an increased excess bias voltage combined with an optimized electric field design [21].

2.3.5. Dead Time

Another significant source of noise in SPADs is the dead time or recovery time that is defined as the interval immediately following an avalanche event during which the detector is unable to register any new photon. After quenching, there is usually a dead time during which the SPAD cannot detect another photon until the necessary bias is restored. This limitation directly affects the performance of quantum communication systems, particularly QKD. Photons arriving during the detector's dead time are not only unregistered but can also introduce errors, reducing the effective photon counting rate and potentially creating security vulnerabilities in the key exchange process. Therefore, in the design and optimization of SPADs for quantum applications, minimizing dead time and managing intrinsic noise are of critical importance to maintain the maximum photon transmission rate and ensure both the accuracy and security of the system [40]. Dead time can be described by the sum of the passive quench and reset timings [41] as $T_{\text{dead}} = T_{\text{quench}} + T_{\text{reset}}$, where T_{quench} is the time to quench the avalanche current and T_{reset} is the time to recharge the SPAD to its operating voltage.

3. Device Engineering and Architectures

3.1. Single-Pixel Designs

The design and optimization of semiconductor devices are crucial for the advancement of SPAD technology. Over the years, a variety of SPAD architectures have been developed to meet the diverse requirements of applications such as quantum communication, time-of-flight (ToF) systems, imaging and LiDAR [42]. Design choices, including the use of single-pixel devices versus pixel arrays, CMOS integration and the implementation of quenching circuits, have a significant impact on the overall performance, functionality and applicability of SPAD-based systems in these fields. Moreover, the choice of SPAD type directly affects key performance parameters, such as PDE, DCR, and afterpulsing effects. In QKD, PDE and DCR are the critical parameters for achieving long-distance transmission without amplification units. Increasing the PDE in SPADs generally leads to a higher DCR, showing the natural trade-off between PDE and DCR, which remains one of the main challenges in SPAD design. By optimizing the multiplication layer and reducing the active area, these designs have achieved high PDE while simultaneously suppressing DCR and afterpulsing near room temperature. These advancements demonstrate that careful engineering of SPAD geometrical and structural parameters plays a crucial role in optimizing both sensitivity and noise performance for QKD applications [43].

To improve the performance of single SPADs, advanced structures such as the double epitaxial structure are employed. In this structure, by using multiple semiconductor layers with different doping concentrations, precise control of the electric field, reduction in the breakdown voltage, and prevention of premature breakdown are achieved. The double-epitaxial SPAD structure consists of an n-type substrate and a p-type epitaxial layer composed of two regions with different doping levels. Four main regions are formed during fabrication: the shallow n region, which acts as the cathode; the enrichment region, designed to reduce the breakdown voltage and shape the electric field; the sinker, providing a low-resistance current path; and the isolation region, which electrically isolates the device from adjacent pixels. Careful design of these regions optimizes device performance and prevents premature breakdown.

3.2. Array SPADs

In some applications where pixel size is not a critical factor, combining multiple small SPADs into a compact array can offer better performance than a single SPAD. In such configurations, SPADs are arranged in a two-dimensional matrix. It is known that the dead time between detection events can limit data throughput, so SPADs and their readout circuits are integrated into arrays where diodes can operate simultaneously (free-running mode) or within specific time windows (time-gated mode). However, a common issue in arrays is crosstalk, which can cause false counts in neighboring pixels. The crosstalk is classified as electrical, where free carriers from the primary avalanche diffuse into adjacent depletion regions, triggering new avalanches, or optical, where secondary photons generated during the avalanche propagate to neighboring pixels and initiate new avalanches. To suppress optical crosstalk, opaque materials can be placed between pixels to prevent the propagation of secondary photons [44]. Recent studies have shown that optical–electrical crosstalk rather than optical propagation plays the dominant role in SPAD arrays. To address this issue, a carrier extraction structure (CES) has been proposed and experimentally demonstrated. This structure employs p-doped channels in the inter-pixel region to electrically guide photo-generated holes toward the surface, preventing them from reaching neighboring pixels. CES application achieves effective crosstalk suppression without introducing etching damage or reducing device fabrication yield. Experimental results indicate that the use of CES can achieve over 90% crosstalk conquest. Specifically, reductions of 91.5% and 96.2% were observed in arrays with 100 μm and 50 μm pixel pitches, respectively. These findings not only clarify the dominant crosstalk mechanisms in planar InGaAs/InP SPADs but also provide a practical pathway for fabricating high-density, low-crosstalk arrays, which are essential for high-resolution 3D imaging and quantum photonic technologies. Figure 4 shows a cross-sectional schematic of the reported planar InGaAs/InP SPAD arrays comprising the CES.

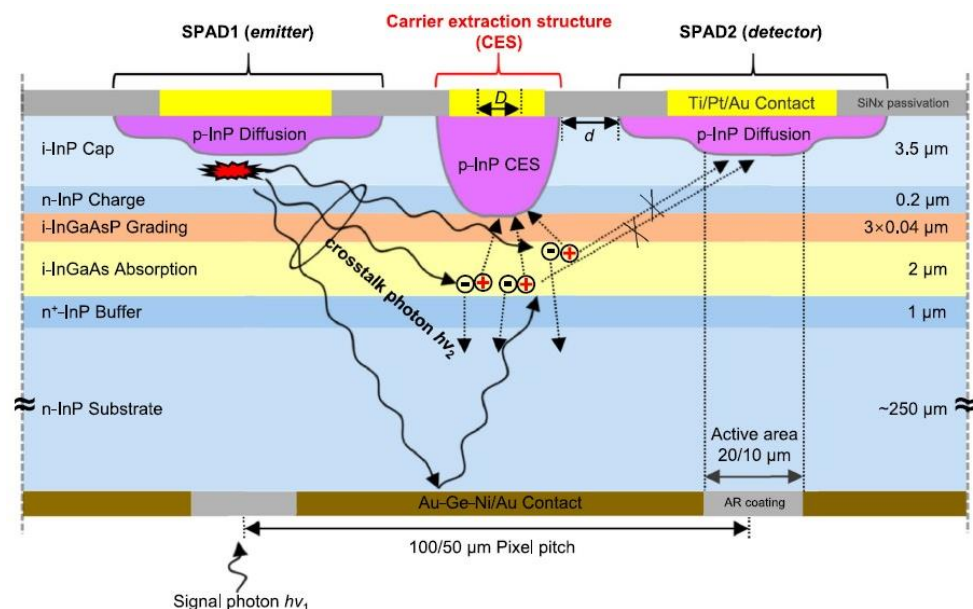


Figure 4. Planar InGaAs/InP SPAD array with CES [45].

3.3. CMOS SPADs

For the purpose of reducing power consumption, miniaturizing hardware components, and lowering system costs, researchers have been focusing on developing integrated single-photon imaging systems using high-voltage CMOS technologies. Since the early 2000s, when the first CMOS SPADs were demonstrated, significant changes have

been made to their structure and design. With advances in CMOS technology, it is now possible to fabricate high-density compact pixel arrays that not only offer lower power consumption and reduced manufacturing costs but can also be integrated with other digital and processing circuits. The CMOS-based SPAD imaging system consists of an array of pixels, each containing a photon sensor and control circuit. Researchers are interested in how CMOS technology can be used to design complex SPAD-based systems. In this case, integrating the SPAD with the electronic circuitry on the same chip reduces parasitic effects. CMOS SPADs are suitable when an array of detectors is required, as in quantum imaging systems [46], weak-value quantum measurement [47], and quantum random number generators (QRNGs) [48].

CMOS SPADs are categorized into two groups: non-isolated and isolated devices [49]. In non-isolated devices, the central cathode is surrounded by the anode and is in direct contact with the substrate (Figure 5a), enabling the formation of a deep depletion region that enhances NIR sensitivity but at the cost of degraded timing performance. In contrast, isolated SPADs are typically designed by introducing controlled n^+ diffusion at the center of the junction below a shallow p^+ boron implantation in an n-doped well, allowing the active area to be precisely defined (Figure 5b).

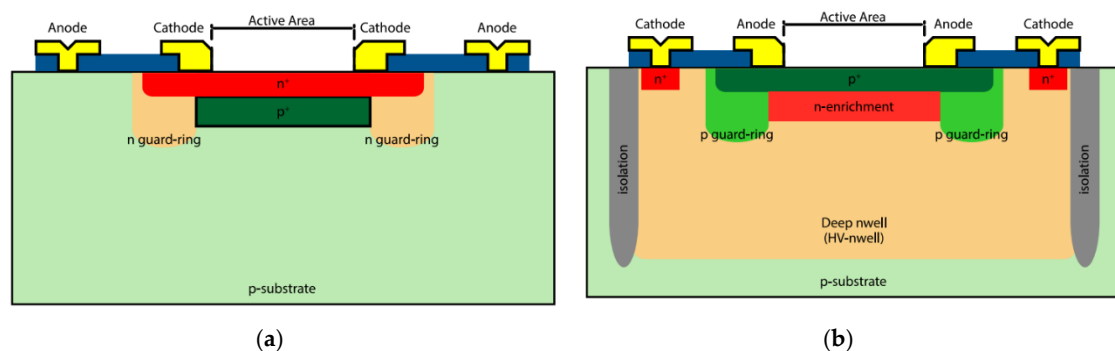


Figure 5. Schematics of (a) non-isolated and (b) isolated CMOS SPADs [21].

A CMOS SPAD is composed of an active region that enables photon absorption and avalanche multiplication, guard rings to prevent edge breakdown, and isolation layers to suppress substrate noise. Key performance metrics, including PDE, DCR and timing resolution, are largely determined by the shape of the active area, the type of junctions, and the design of the guard rings. The active region can be circular, rectangular, octagonal, or polygonal, while the isolation layers are often buried to minimize noise. Most CMOS SPAD detectors are based on a p^+/n -well junction which is intrinsically isolated from the p -type substrate while a deep n -well is commonly used to separate the guard ring. This configuration allows simultaneous access to both the anode and cathode, enabling direct connection to the readout circuit at low voltage. For the structure in Figure 6a, only an additional deep n -well implantation is required, and higher sensitivity in the red and NIR regions can be achieved with deeper junctions, such as a p -well/deep n -well (Figure 6b). Furthermore, n -in- p structures with n^+/p -well or deep n -well/ p -epi/ p^+ sub-junctions (Figure 6c,d) offer higher PDE in the red and NIR regions, but they require a quenching resistor and capacitive decoupling for the readout circuit, making their use in a single pixel less straightforward.

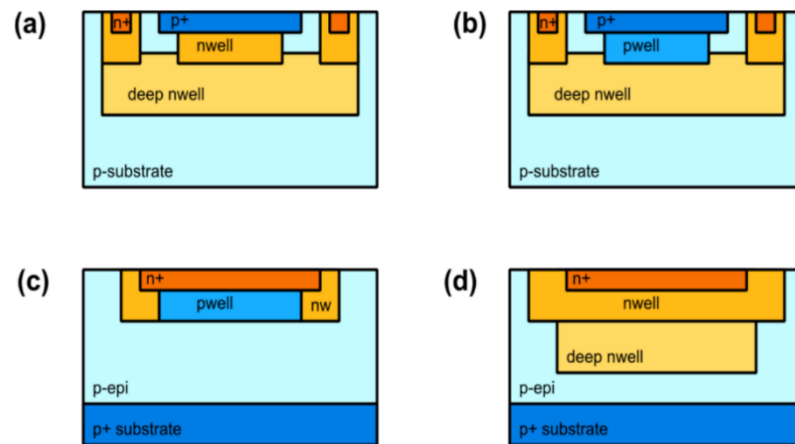


Figure 6. Cross sections of different CMOS SPAD devices: (a) p⁺/n-well; (b) p-well/deep n-well; (c) n⁺/p-well; (d) deep n-well/p-epi/p⁺ substrate [49].

The guard ring is one of the key structural components of a SPAD, designed to prevent premature edge breakdown and to improve the electric field distribution, thereby enhancing device stability and ensuring uniform avalanche multiplication. Various types of guard ring structures have been developed in SPAD designs, each optimized to control edge breakdown, shape the electric field distribution, and improve device performance across different CMOS technologies [50,51]. Another CMOS SPAD design employs an n⁺/p-well junction configuration often combined with a deep p-well isolation layer [52]. This structure effectively suppresses substrate noise and minimizes tunneling leakage, making it suitable for dense array integration. The p-type isolation provides better electrical separation from the substrate, reducing crosstalk and enhancing array uniformity. Such designs typically offer higher PDE in the near-NIR region and improved compatibility with low-voltage readout circuits, making them suitable for advanced CMOS image sensors and 3D-stacked SPAD architectures, as illustrated in Figure 7. Table 1 provides a comprehensive overview of the different junction types and guard ring structures implemented in the CMOS technology used for SPAD design. The table details the specifications of each junction and guard ring type along with the corresponding CMOS process, enabling a clear comparison of the structural and fabrication variations among the samples.

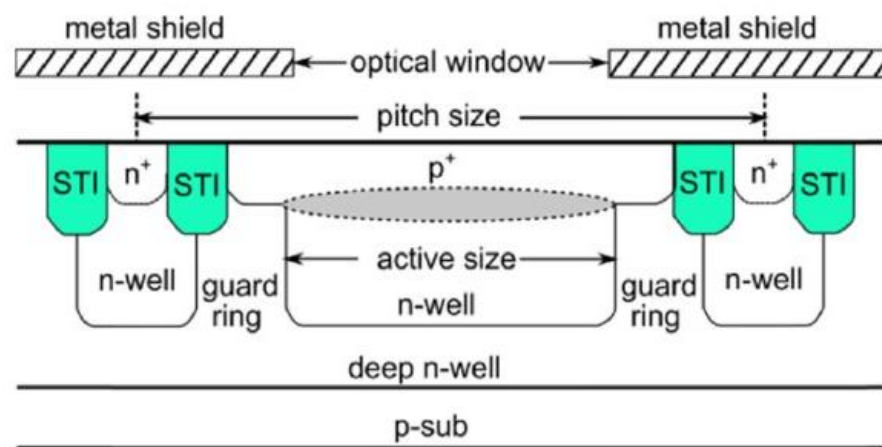


Figure 7. Cross section of the p⁺/n-well SPAD with a deep n-well guard ring. The shadowed areas indicate the avalanche region [52].

Table 1. Recent advances in CMOS SPAD fabrication across different technology nodes.

Tech. (nm)	Junction Type	Guard Ring	Ref.
CMOS 65	p ⁺ /n-well	p-well	[18]
CMOS 90	n ⁺ /p-substrate	n-well	[53]
CMOS 130	p-well/DNW	Deep virtual	[54]
CMOS 180	PSD/n-well/DNW	STI	[55]
CMOS 180	p ⁺ /shallow n-well	Shallow p-well	[51]
CMOS 150	p ⁺ /n-well	DNW	[52]
CMOS 180	p-well/epi/DNW	STI	[56]
CMOS 350	n ⁺ /p-substrate	n-tub	[57]
CMOS 350	p ⁺ /n-tub	p-tub	[58]
CMOS 140 BSI	p ⁺ /n-well	p-well	[59]

DPW (deep p-well); DNW (deep n-well); BSI (back side illuminated); STI (shallow trench isolation); PSD (p-type shallow doping).

3.4. Junction Engineering (Avalanche Region Engineering)

In the design of SPADs for quantum systems, precise engineering of the p-n junction plays a pivotal role in defining and optimizing key performance metrics. The junction must be carefully engineered to ensure a uniform electric field in the depletion region and sufficiently strong to initiate avalanche multiplication while simultaneously preventing an increase in statistical noise such as DCR and afterpulsing. Nowadays, custom SPADs are usually divided into two categories: reach-through and planar SPADs. Reach-through SPADs have vital applications in quantum communication due to their high PDE in the NIR, low DCR, and low-cost commercial modules. In this structure, the multiplication and drift regions are separated within a thick quasi-intrinsic p-type substrate. Photons are absorbed from the back side and generate carriers that drift toward the high-field multiplication region, where avalanche triggering occurs. In this structure, p⁺/n geometry electrons are the principal carriers that trigger the avalanche, resulting in a high PDE. The planar SPADs are based on a front-illuminated n⁺/p structure. These devices feature a shallow avalanche region and a thin quasi-intrinsic p absorption layer formed via double epitaxy on an n⁺ substrate [60,61].

Recently, resonant cavity enhanced (RCE) structures have been designed to improve the performance of GeSn-based SPADs operating at wavelengths of 1550 and 2000 nm. These devices consist of a PIPIN GeSn/Si (p⁺⁺ GeSn, i GeSn, p Si, i Si and n⁺⁺ Si) heterostructure embedded within an optical cavity incorporating a distributed Bragg reflector (DBR), which enhances optical absorption and photon detection efficiency. By optimizing the structure, a single-photon detection efficiency of up to about 80% at room temperature has been achieved [62].

CMOS SPADs exhibit limited sensitivity in the NIR region; however, the performance of this approach can be improved. In this regard, Gobinath et al. [63] presented an improved current-assisted SPAD fabricated via a 350 nm CMOS process, featuring a 14 µm thick epitaxial layer that provides enhanced NIR sensitivity. The main structure is a p⁺/p-epi/n-well junction for the central avalanche region. In this work, two avalanche region geometries are compared. In CA-SPAD-1, which has a square geometry due to the sharp corners, the electric field distribution increases, therefore decreasing PDE and increasing timing jitter. In CA-SPAD-2 (Figure 8a), in order to make the field uniform in the p-n junction, the anode and cathode have a cylindrical design.

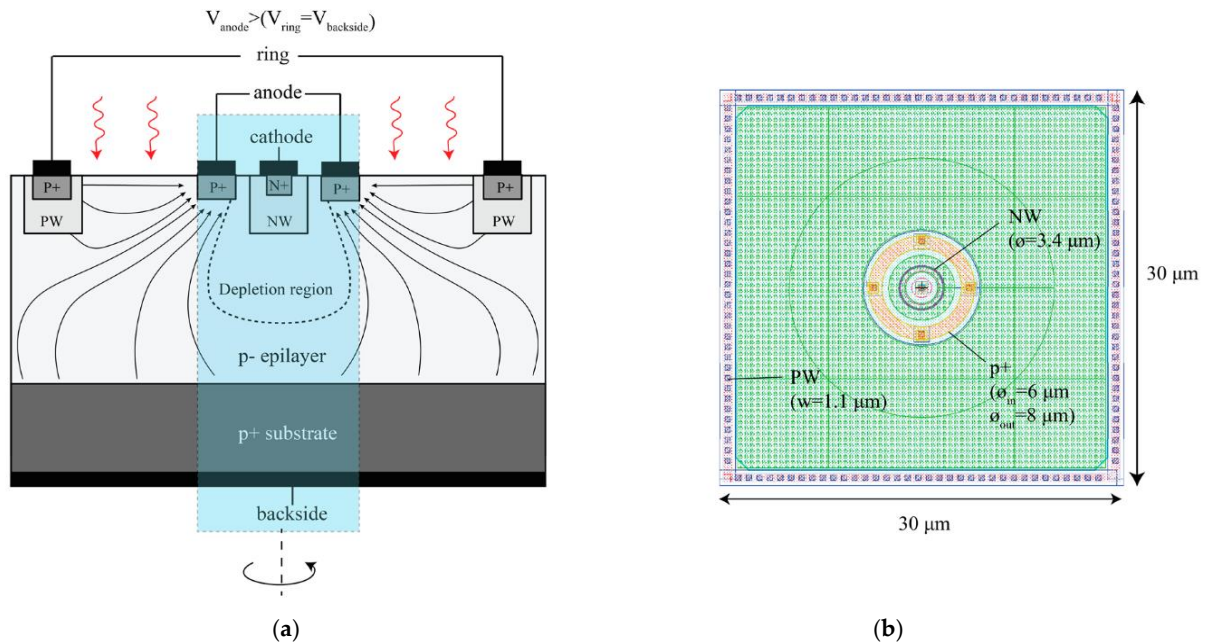


Figure 8. (a) Cross section of the CA-SPAD-2 device. (b) Top-view image of the device [63].

A third ring-shaped electrode was employed to generate a drift field within the epilayer. As shown in Figure 8b, the ring has a square geometry because it causes the maximum fill factor when making arrays and the ring shape is not critical for the drift field. In addition to the SPADs previously introduced and used in quantum applications, a new generation of SPADs exist that are back illuminated (BI) [64]. A BI SPAD based on 40 nm CMOS image sensor (CIS) technology is presented in Figure 9. It is based on a p-well/deep n-well junction with a wide absorption region, providing a high photon detection probability (PDP) over a broad wavelength range. This SPAD achieves the highest PDP and lowest timing jitter, so it has potential for key roles in photonic and quantum applications.

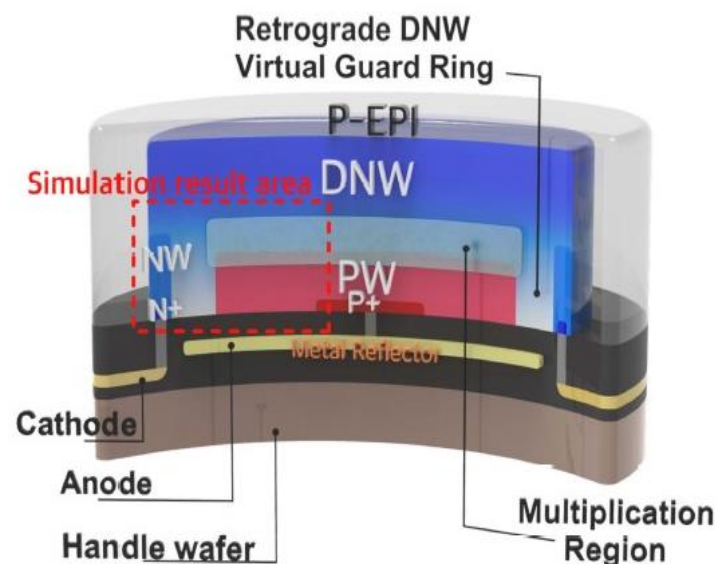


Figure 9. Cross section of the proposed back-illuminated SPAD [64].

Recent advances in semiconductor fabrication have enabled the realization of SPADs in advanced bipolar-CMOS-DMOS (double metal oxide semiconductor) (BCD) technologies [65], allowing compact and cost-effective integration for mobile and 3D sensing applications. Figure 10 illustrates the structure of the proposed SPAD implemented in 55 nm

BCD technology. The device employs a shallow junction formed between a p^+ region and a high-voltage n-well (HVNW), with a deep n-well layer placed underneath. The low-doped HVNW helps to suppress tunneling noise and enhance the PDE, while the DNW further improves the quantum efficiency at longer wavelengths, extending the spectral response of the detector.

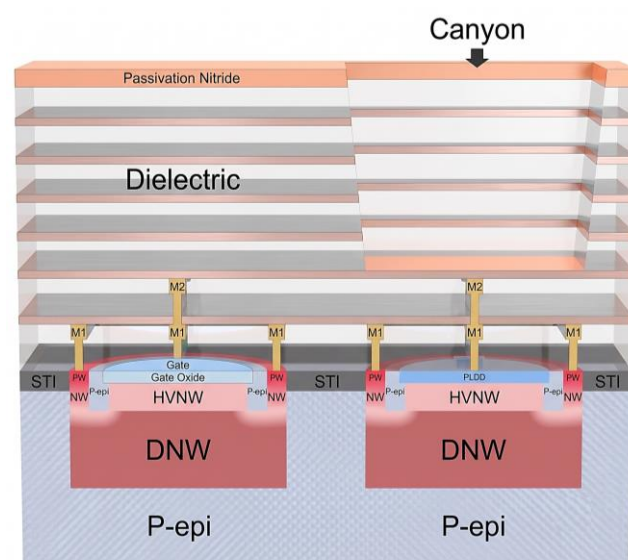


Figure 10. Cross section of a BCD-SPAD [66].

4. Fabrication Strategies and Material Platforms

4.1. Advances in Fabrication Technologies

The fabrication technology of SPADs has advanced remarkably in recent years, enabling various approaches to achieve improved single-photon detection performance. Depending on the target spectral range, application type, and material constraints, diverse fabrication routes have been proposed and implemented.

The study by Itzler et al. [67] represents one of the first systematic efforts to implement InGaAs/InP technology in Geiger mode and laid the foundation for SPAD designs based on the separate absorption–charge–multiplication (SACM) architecture. MOCVD epitaxial growth was employed to form the absorption, charge, and multiplication layers, while the introduction of graded InGaAsP transition layers reduced valence-band discontinuity and carrier trapping at the heterointerfaces. A buried p–n junction was created through a two-step zinc diffusion process to suppress edge breakdown and peripheral leakage, and a SiN_x passivation layer was applied to ensure long-term device stability. The authors demonstrated that even small variations in the thickness or doping concentration of the epitaxial layers could lead to substantial changes in the internal electric field profile, breakdown voltage, and dark count rate.

In the study by Takai et al. [68], a novel SPAD structure compatible with $0.18\ \mu\text{m}$ CMOS technology was introduced, aimed at enhancing sensitivity in the NIR region for automotive LiDAR applications. In this design, two customized layers of an n-SPAD and a p-SPAD were added to the standard CMOS process to create a high-field region that is fully depleted (Figure 11). This configuration increased the depletion width and improved carrier collection without requiring excessively high bias voltages. In addition, the use of a p-epitaxial layer with a graded doping profile enhanced NIR photon absorption and improved the PDE. The design achieved a high PDE of 19.4% at 870 nm without a reduction in the fill factor (33.1%), thereby overcoming the long-standing trade-off between PDE and FF. From a fabrication standpoint, this work represents an important step toward the

convergence of single-photon detectors with standard CMOS processes. The incorporation of custom layers within an industrial framework without requiring fundamental process modifications demonstrates a mature level of semiconductor design. Optimization of doping concentration, precise control of the depletion region, and suppression of edge breakdown at low bias voltages were among the key achievements of this study. Furthermore, the addition of a surface p^+ layer to collect surface-generated carriers significantly reduced the DCR and stabilized performance over a wide temperature range (25–130 °C).

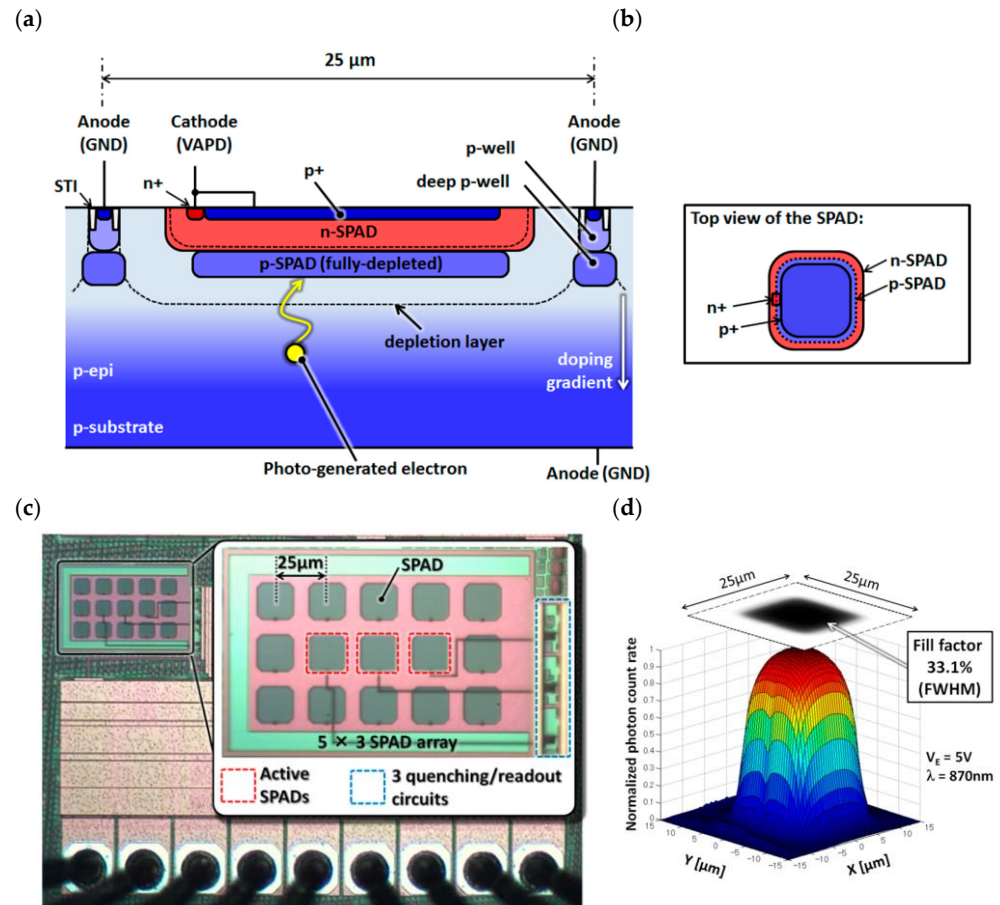


Figure 11. Structure and performance of an NIR-enhanced SPAD fabricated using 0.18 μm CMOS technology. The cross-sectional view (a) illustrates the custom n-SPAD and p-SPAD layers together with the p-epitaxial layer featuring a graded doping profile, which collectively form a strong electric field and a fully depleted region. (b) The top view of the SPAD. The microscopic image (c) shows the 3×5 SPAD array along with the quenching and readout circuits. The 3D plot (d) presents the photon count rate map at 870 nm, corresponding to a fill factor of 33.1% [68].

In the study by Fang et al. [69], an advanced architecture for InGaAs/InP single-photon detectors was proposed, aiming to achieve exceptionally high detection efficiency at a wavelength of 1550 nm. The device was designed based on the SAGCM structure and fabricated through MOCVD epitaxial growth with precise control of layer thickness and doping concentration. The key innovation lies in the introduction of a dielectric–metal ($\text{SiO}_2/\text{metal}$) reflective layer above the photon-sensitive region, which through constructive interference between the semiconductor–dielectric and dielectric–metal interfaces provides an effective reflectivity of about 80% at 1550 nm. This design enhances optical absorption and consequently improves the PDE by approximately 20% compared with conventional structures. Additionally, the dual-step Zn diffusion and careful electric field

tuning within the charge and multiplication layers ensured field uniformity and minimized edge breakdown. The study demonstrates a successful integration of epitaxial optimization, electric field engineering, and readout circuit co-design. In particular, the development of a monolithic integrated readout circuit (MIRC) using low-temperature co-fired ceramic (LTCC) technology marked an important step in reducing parasitic capacitance and suppressing afterpulsing effects. These advancements enabled a detection efficiency of 60% at room temperature with a dark count rate of 340 kcps, while operation at a lower temperature (253 K) achieved 40% efficiency with a dark count rate of 3 kcps. This work exemplifies the synergy between optoelectronic design and precise fabrication in InGaAs/InP systems.

The development of SPADs based on standard and advanced CMOS technologies is one of the main research topics. For example, the use of 28 nm fully depleted FD-SOI technology with n-well/p-sub structures and low-voltage, over-deep n-well configurations has led to a significant reduction in DCR and an improvement in PDE. Precise design of the junction layout and depletion region also allows effective control over dead time and afterpulsing [70].

To extend the spectral sensitivity of SPADs into the near-infrared region, researchers have explored more advanced technologies such as Ge-on-Si. In the study by Wanitzek et al. [71], a novel approach was proposed to reduce noise and enhance the performance of Ge-on-Si single-photon detectors. The work focuses on the design of a double-mesa structure, demonstrating that the addition of an extra etching step significantly reduces the electric field at the diode edges, thereby lowering both the dark current and DCR in linear and Geiger modes. The epitaxial layers were grown using MBE, featuring a two-step Ge growth process to achieve high crystalline quality and eliminate threading dislocations. The final structure consists of a Ge absorption layer (300–400 nm thick), an intrinsic Si multiplication layer, and a p-type Si charge layer with precisely controlled doping ($\sim 1.4 \times 10^{17} \text{ cm}^{-3}$), maintaining a balance between the electric field and avalanche probability (Figure 12). Furthermore, a highly doped p⁺ Ge/Si contact layer was employed to improve surface passivation and reduce contact resistance.

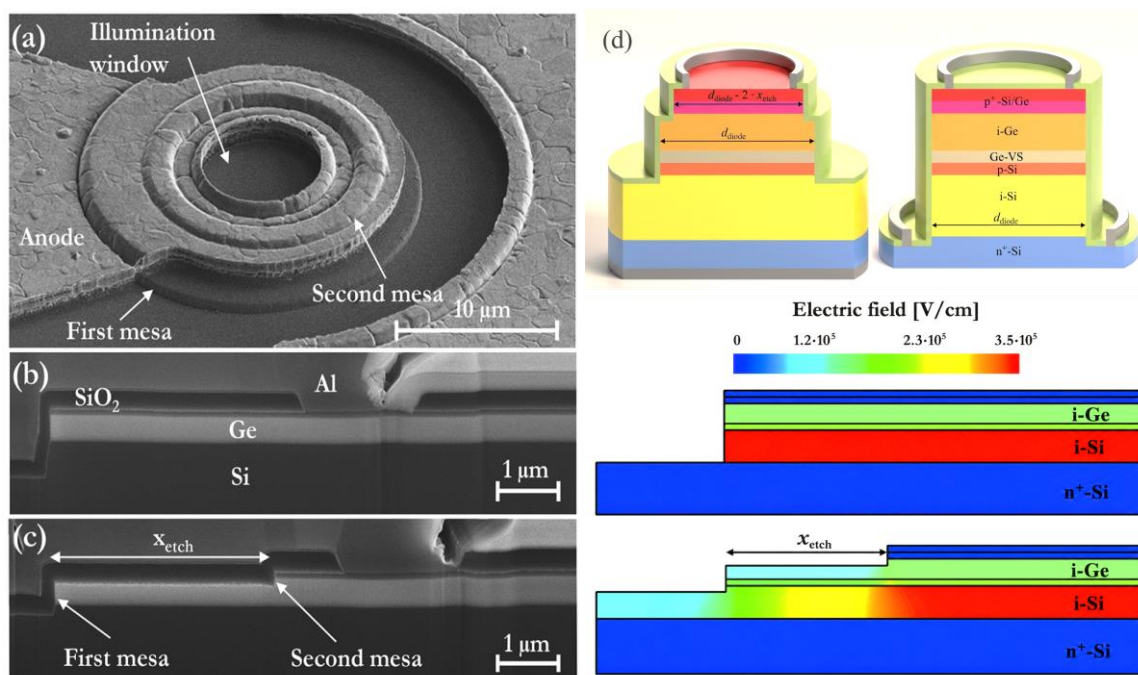


Figure 12. Comparison of single-mesa and double-mesa structures in Ge-on-Si SPADs. (a) SEM image of the final device featuring two successive mesa etching steps, including the central illuminated region, inner mesa, and outer mesa. (b,c) Cross-sectional SEM images of the single- and double-mesa structures. (d) Three-dimensional schematic of the active layers comprising i-Ge, i-Si, and the Ge-VS interface layer for both configurations, along with the simulated electric field distribution. The double-mesa design significantly suppresses the electric field around the diode periphery and eliminates lateral leakage paths [71].

The double-mesa process involves deep etching using inductively coupled plasma reactive-ion etching (ICP-RIE) in an HBr environment. Ti/Al metal contacts were sputtered to prevent aluminum diffusion into the semiconductor. This simple yet effective design achieved more than a 260-fold reduction in dark current under linear mode and a 100-fold decrease in DCR under Geiger mode at 110 K. Moreover, the thin absorption layer in the double-mesa structure enabled a fast response time of 81 ps and a PDE of 7.3%, highlighting the advantage of geometrical design in suppressing peripheral electric fields and improving avalanche timing. This approach demonstrates that geometric and edge engineering in Ge-on-Si architectures can serve as an efficient alternative to complex doping-control methods for noise reduction and performance enhancement in infrared SPADs.

Further progress has been made using III–V semiconductors such as InGaAs/InAlAs to fabricate SPADs with high sensitivity in the telecommunication band. Precise epitaxial growth and engineering of the absorption region, combined with the introduction of built-in electric fields, have improved both response speed and detection sensitivity [72]. Although these devices require cooling and high bias voltages, they demonstrate superior performance in LiDAR and long-wavelength optical communication systems.

In innovative design efforts, III–V nanowires have been employed to reduce detector volume and enhance electro-optical control (Figure 13). A notable example is an InGaAs–GaAs nanowire-based structure that enables free-running operation and exhibits a strong suppression of afterpulsing [73].

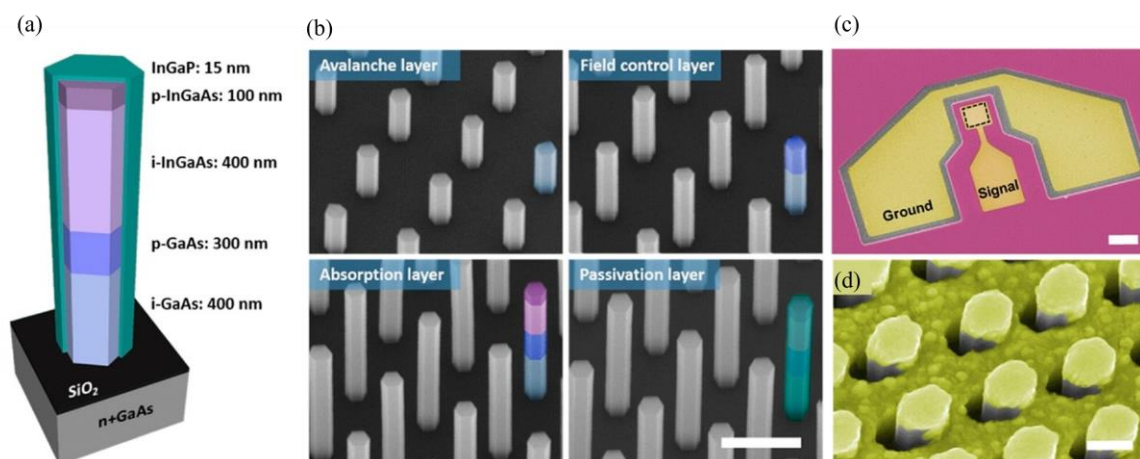


Figure 13. Structure and microstructure of the InGaAs–GaAs nanowire-based single-photon detector designed for free-running operation and suppression of afterpulsing. (a) Layered schematic of the nanowire, comprising an intrinsic InGaAs absorption region, a p-GaAs field region, and an In-GaP control layer grown on an n⁺-GaAs/SiO₂ substrate. (b) Microscopic images showing different growth stages and layer depositions. (c) Electrode layout of the device with signal and ground lines for microprobe contact. (d) Microscopic view of the ordered nanowire array, demonstrating high geometric uniformity and strong potential for large-scale integration. Reprinted with permission from [73]. Copyright 2018 American Chemical Society.

For silicon-based SPADs, several approaches have been proposed to increase light absorption at longer wavelengths. One effective strategy involves engineering a low-doped epitaxial absorption layer, expanding the drift region, and optimizing the junction geometry. Such a design enhances responsivity in the 800–900 nm range while improving field uniformity and reducing noise [74].

In the case of superconducting detectors, such as superconducting nanowire single-photon detectors (SNSPDs), the implementation of novel designs including Peano fractal patterns, Bragg mirrors, and multi-angle optical traps has improved absorption, eliminated polarization sensitivity, and enhanced uniformity. These architectures are particularly advantageous for quantum applications that demand extremely high sensitivity and low error rates [75].

Ensuring uniform performance and controlling structural defects play a crucial role in SPAD scalability. Optimization of doping layers, reduction of carrier traps, and the use of precise lithographic masks have been shown to minimize dark counts and improve inter-pixel uniformity. Such improvements are especially critical in advanced CMOS designs for large-scale (megapixel) SPAD arrays [41].

4.2. Material Platforms and Evolution in Single-Photon Detector Technologies

The choice of material fundamentally determines the operational limits, detection efficiency, and long-term scalability of single-photon detectors. While fabrication strategies define how precisely these devices can be engineered, the intrinsic properties of the employed materials, such as bandgap, lattice quality, defect density, and carrier ionization behavior, ultimately dictate performance boundaries. Over the past two decades, material development for SPADs and SNSPDs has progressed along two distinct but complementary paths: semiconducting materials for avalanche-based photon multiplication and superconducting thin films for cryogenic single-photon sensing. In the semiconductor domain, silicon has remained the dominant platform for visible detection due to its CMOS compatibility and excellent crystal quality, while III–V compounds and Ge-on-Si heterostructures have extended the detection range into the near-infrared. In parallel, superconducting materials such as NbN, NbTiN, and amorphous WSi have enabled single-photon detection at telecommunication wavelengths with nearly unity efficiency.

The following subsections provide an overview of the principal materials employed in these technologies, tracing their evolution from conventional silicon to low-dimensional materials. Each subsection highlights the rationale behind material selection, intrinsic advantages and drawbacks, and the historical context that has guided the transition toward more advanced material systems.

4.2.1. Silicon-Based SPADs

Silicon has been the foundational material for SPADs since their introduction in the late 1990s, largely due to its exceptional material maturity, intrinsic crystal quality, and full compatibility with CMOS fabrication technologies. The indirect bandgap of silicon (1.12 eV) limits its spectral response, yet it conveniently covers most visible and near-infrared applications in imaging, ranging, and fluorescence detection. From a material perspective, silicon offers very low defect and trap densities, allowing for sub-picoampere dark currents even at room temperature. The small ratio between the hole and electron ionization coefficients ($k \ll 1$) results in reduced avalanche noise and enables stable, low-jitter photon multiplication. Furthermore, the presence of a high-quality native oxide (SiO_2) simplifies electrical isolation and surface passivation, which are crucial for large-area arrays. These properties, combined with abundant wafer availability and well-established doping control, have made Si the dominant choice for visible-range SPADs and a key enabler of monolithically integrated photon counting imagers. Rawat et al. [16] designed a

low-voltage Si avalanche photodiode epitaxially grown on a SOI substrate, featuring a submicron π -layer ($\approx 0.8 \mu\text{m}$) and an engineered doping contrast to initiate avalanche breakdown below 10 V (Figure 14). To overcome the limited absorption of silicon in the near-infrared region, a photon-trapping microhole (PTMH) array was introduced and passivated with SiO_2 , effectively extending the optical path and suppressing surface leakage. The resulting devices exhibited a dark current of only $\sim 30 \text{ pA}$ and a five-fold enhancement in EQE at 850 nm, emphasizing how microstructured Si architectures and precise doping engineering can simultaneously enhance absorption efficiency and reduce parasitic leakage within CMOS-compatible material platforms.

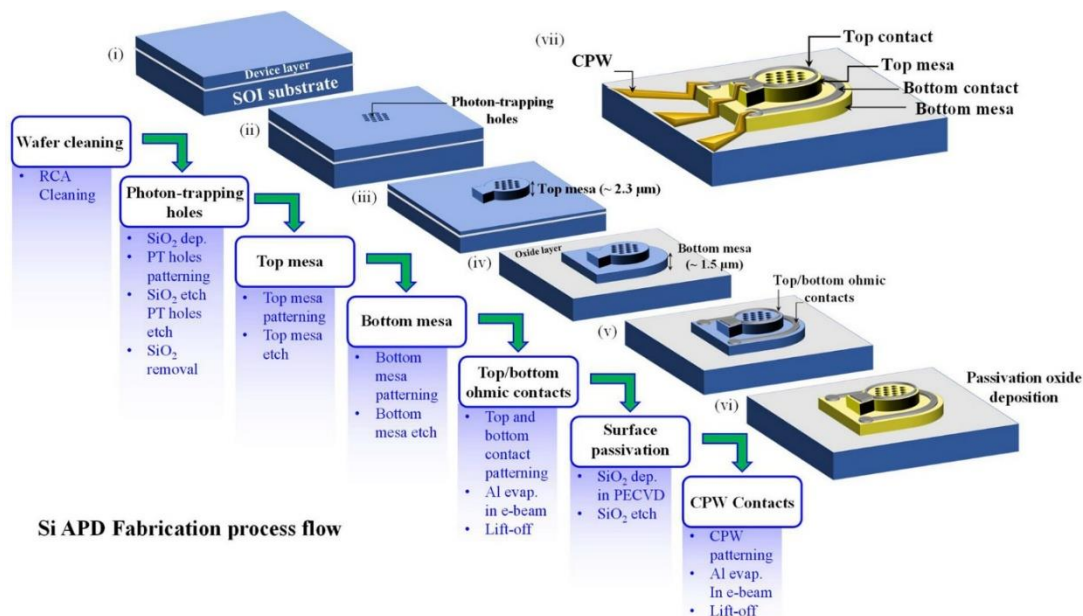


Figure 14. CMOS-compatible fabrication steps for Si-APDs. PTMHs are defined through stepper-based optical lithography. Both the mesa and PTMH sidewalls are coated with SiO_2 for surface passivation and suppression of off-state leakage. Coplanar waveguide (CPW) metal contacts are integrated to lower contact resistance and support high-speed device operation [16].

Despite its extensive industrial foundation, silicon exhibits intrinsic limitations rooted in its band structure and absorption characteristics. As the absorption coefficient drops sharply beyond $1 \mu\text{m}$, detection efficiency declines for telecommunication wavelengths, making Si-based SPADs unsuitable for long-range or fiber-optic systems. Additionally, its relatively large phonon-scattering cross section imposes constraints on carrier lifetime and the temperature dependence of the breakdown process. Nevertheless, these limitations are purely material-driven and have inspired the development of alternative compound semiconductors and hybrid platforms. Overall, silicon remains the benchmark against which other SPAD materials are evaluated, combining structural perfection, low intrinsic noise, and unmatched manufacturability.

4.2.2. III–V Compound SPADs

The introduction of III–V compound semiconductors into SPAD technology marked a critical step toward extending photon detection beyond the intrinsic absorption limit of silicon. Materials such as InGaAs, InP, GaAs, and GaAsP possess direct bandgaps and strong absorption coefficients in the near-infrared region, enabling single-photon detection up to $1.55 \mu\text{m}$, which is the essential wavelength for fiber-optic communication and quantum key distribution. Among them, the InGaAs/InP heterostructure has emerged as the benchmark platform owing to its favorable band alignment and ability to separate

absorption and multiplication regions. The InGaAs layer, with an adjustable bandgap around 0.75 eV, efficiently absorbs near-infrared photons, while the wider-bandgap InP multiplication region (1.34 eV) suppresses tunneling and impact ionization noise. To mitigate lattice mismatch ($\sim 3\%$) and strain accumulation between these layers, composition-graded buffers or quaternary alloys such as InGaAsP are often inserted, providing smoother band transitions and reducing defect-related dark counts.

Tang et al. [45] reported a planar InGaAs/InP SPAD array with a carrier-extraction structure grown by MOCVD, where a $2\text{ }\mu\text{m}$ $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ absorption layer and a $1\text{ }\mu\text{m}$ InP multiplication layer were separated by triple InGaAsP grading layers to relieve the valence-band discontinuity (Figure 15). A p-type Zn-diffused InP channel was introduced between pixels to laterally extract photogenerated holes and suppress optical-electrical crosstalk by over 90%. This material engineering strategy eliminates the need for deep isolation trenches and preserves the pristine InP surface, maintaining a dark current below 1 nA and negligible PDE degradation. The work highlights how controlled Zn diffusion and heterointerface engineering in lattice-matched InGaAs/InP epitaxy can yield high-density SPAD arrays with minimized defect-assisted tunneling and improved reliability for short-wave infrared (SWIR) photon detection.

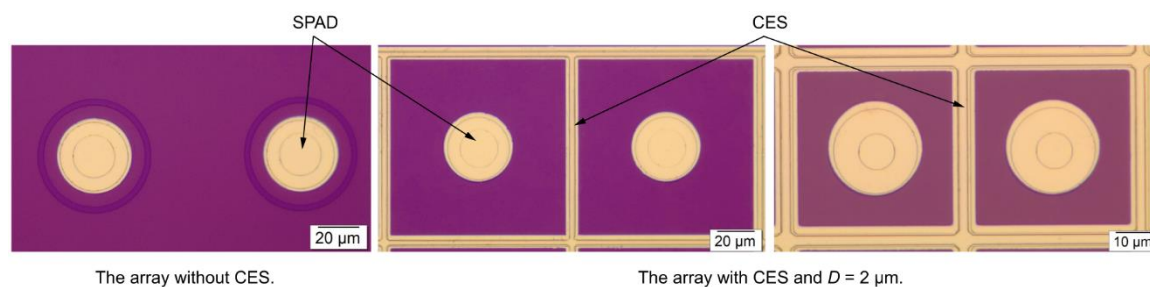


Figure 15. Top-view layout of neighboring pixels in the array: a $100\text{ }\mu\text{m}$ pitch configuration without CES (left), a $100\text{ }\mu\text{m}$ pitch configuration incorporating CES (center), and a $50\text{ }\mu\text{m}$ pitch configuration with CES (right) [45].

Xie et al. [76] modeled an SAGCM InP/InGaAs SPAD composed of an undoped $\text{In}_{0.53}\text{Ga}_{0.47}$ absorption layer, an InP multiplication layer, and an intermediate InGaAsP grading adjacent to a charge layer to smooth the bandgap transition and control the electric field. The wide-bandgap InP region effectively suppresses tunneling-induced dark counts, while the lattice-matched InGaAs layer ensures high absorption at 1550 nm . The model reproduced experimental data with $<12\%$ error and confirmed that trap-assisted tunneling (TAT) dominates dark noise at low temperatures ($<245\text{ K}$). The study demonstrates the material-level balance between high photon detection efficiency and minimized dark count rate through optimized heteroepitaxy and electric field engineering.

Farrell et al. [73] demonstrated a nanowire-based InGaAs-GaAs absorption and multiplication avalanche photodiode (SAM-APD) structure, grown by selective-area MOCVD, where a 400 nm GaAs multiplication layer and a 400 nm $\text{In}_{0.26}\text{Ga}_{0.74}\text{As}$ absorption layer were axially integrated and passivated by a thin InGaP shell. The nanowire geometry allowed strain relaxation and minimized defect-assisted tunneling, achieving a dark count rate below 10 Hz and negligible afterpulsing (<0.25) at 77 K in free-running mode. The study highlights how InGaAs-GaAs heteroepitaxy can provide lattice- and surface-quality advantages over conventional planar InGaAs/InP SPADs while maintaining efficient NIR absorption.

Early devices fabricated via MOCVD and MBE techniques demonstrated proof-of-concept single-photon operation at $1.3\text{--}1.55\text{ }\mu\text{m}$, although with relatively high dark currents and limited uniformity. Telesca et al. [77] investigated InGaAs/InP SPADs featuring

a double-zinc diffusion structure to suppress charge persistence noise. The front-illuminated SAGCM devices used lattice-matched $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ absorption and InP multiplication layers grown by MOCVD, while two sequential Zn diffusions defined the p^+ regions with controlled depth ($d_s \approx 100$ nm). By tuning the shallow-diffusion radius ($\Delta r = 4\text{--}10$ μm), the electric field distribution in the InGaAs layer was optimized to drain thermally-generated holes before they reach the active area. Material-level optimization yielded a 70% reduction in charge persistence and a lower DCR (~ 7 kcps at 225 K) without altering the epitaxial stack. This study emphasizes how precise dopant diffusion profiling within InP-based heterostructures can improve reliability and enable operation above 250 K without increasing tunneling currents.

From a materials standpoint, the III–V family offers high absorption efficiency and tunable bandgaps, but these advantages come with intrinsic challenges. The presence of heterointerfaces and polar bonds increases the likelihood of trap-assisted tunneling and carrier afterpulsing, while the mechanical softness and thermal mismatch of InP-based epilayers complicate large-area integration. Furthermore, due to higher intrinsic carrier generation rates, these devices typically require thermoelectric cooling (-50 °C to -100 °C) to suppress dark noise. Recent material refinements, including double-mesa etching to confine the electric field, controlled doping in avalanche-control layers, and epitaxial re-growth to heal dislocations have substantially improved reproducibility and PDP. The hybrid bonding of InGaAs-based epitaxial stacks onto silicon readout wafers now offers a compromise between the superior infrared sensitivity of III–V compounds and the scalability of CMOS platforms. As a result, III–V SPADs continue to serve as the material cornerstone for high-performance long-wavelength single-photon detection.

4.2.3. Ge-on-Si SPADs (Hybrid Silicon Platforms)

To extend single-photon sensitivity into the SWIR range while maintaining CMOS process compatibility, researchers have developed Ge-on-Si SPADs that leverage the strong absorption of germanium (bandgap ≈ 0.66 eV) together with the mature fabrication infrastructure of silicon. The lattice mismatch of approximately 4.2% between Ge and Si presents a fundamental materials challenge, as it induces threading dislocations and interface defects that act as generation centers for dark carriers. Early Ge-on-Si devices suffered from dark count rates exceeding 10^5 cps even at cryogenic temperatures, primarily due to these defects. However, advances in epitaxial growth, such as two-step low-temperature Ge deposition, cyclic annealing, and the use of defect-filtered buffer layers ($\text{Si}_{1-x}\text{Ge}_x$ graded buffers or aspect ratio trapped Ge) have significantly improved crystal quality and reduced dislocation densities below 10^7 cm^{-2} . In this hybrid configuration, the Ge absorption layer, typically $0.3\text{--}0.5$ μm thick, is grown atop a lightly doped Si substrate or an intrinsic Si layer that serves as the avalanche region. The wider-bandgap Si effectively confines the electric field, limiting tunneling noise while allowing efficient carrier multiplication. Structural innovations such as double-mesa etching and planar passivation oxides have been introduced to improve edge uniformity and reduce trap-assisted breakdown. These material refinements have enabled Ge-on-Si SPADs to achieve photon detection efficiencies of 15–20% at 1310–1550 nm and timing jitters below 100 ps at 77 K, with cost and integration advantages unattainable by InGaAs/InP devices. As growth techniques continue to minimize dislocation-related leakage, Ge-on-Si platforms are becoming a promising route toward large-area and CMOS-compatible infrared SPAD arrays for 3D sensing and quantum communication applications.

Warburton et al. [78] demonstrated one of the first epitaxial Ge-on-Si SPADs using a SACM structure, where a 1 μm intrinsic Si layer served as the multiplication region and an undoped Ge absorption layer was grown by reduced-pressure CVD directly on Si. The

Ge layer was seeded at low temperature ($\sim 350^\circ\text{C}$) and subsequently thickened and cyclically annealed to reduce the threading dislocation density to $\sim 10^6\text{--}10^7\text{ cm}^{-2}$. Despite the $\sim 4.2\%$ lattice mismatch, this approach enabled efficient absorption at 1310–1550 nm while maintaining CMOS compatibility. The devices exhibited a 4% detection efficiency at 1310 nm and a dark count rate of $\sim 10^6\text{--}10^7\text{ Hz}$ at 100 K, with a noise equivalent power (NEP) of $1 \times 10^{-14}\text{ W}\cdot\text{Hz}^{-1/2}$. This work confirmed that Ge-on-Si heteroepitaxy could provide infrared sensitivity without III–V materials, though surface leakage from mesa sidewalls and interface defects remained the main sources of dark current. Vines et al. [79] realized planar Ge-on-Si SACM SPADs grown by RP-CVD, using a $\sim 1\text{ }\mu\text{m}$ undoped Ge absorber atop a $\sim 1.5\text{ }\mu\text{m}$ intrinsic Si multiplication region with a B-implanted Si charge sheet, a planar geometry that keeps high fields away from Ge sidewalls and suppresses surface-generation leakage. Wanitzek et al. [71] introduced Ge-on-Si SPADs employing a double-mesa geometry to suppress peripheral electric fields and trap-assisted tunneling at the Ge/SiO₂ interface. The heterostructure was grown by MBE, consisting of a 500 nm intrinsic Si multiplication layer, a 100 nm p-type charge layer ($\approx 1.4 \times 10^{17}\text{ cm}^{-3}$), and a 300 nm unintentionally doped Ge absorber grown on a low-temperature Ge virtual substrate (Ge-VS) to annihilate threading dislocations. The double-mesa configuration effectively reduced the dark current by $>260\times$ and the dark count rate by $\approx 100\times$ compared to a single-mesa device at 110 K, reaching DCR $\approx 9.5 \times 10^5\text{ Hz}$, PDE $\approx 7.3\%$, and a record low timing jitter of 81 ps. These results demonstrate how interface-field control and defect-filtered Ge growth can markedly enhance SWIR sensitivity and enable near-cryogenic operation of CMOS-compatible Ge-on-Si SPADs.

Na et al. [80] introduced a CMOS-compatible GeSi SPAD that operates at room temperature, fabricated on an SOI substrate using a 250 nm Si multiplication layer, a p-type charge layer, and a 450 nm epitaxial Ge absorber grown by UHV-CVD between 300 and 600°C followed by $600\text{--}800^\circ\text{C}$ annealing. A boron-doped Si screening layer was introduced beneath the Ge sidewalls to terminate high electric fields and prevent trap-assisted tunneling at defective interfaces. This material-level field control reduced the dark current density to $\approx 4.5\text{ }\mu\text{A cm}^{-2}$ and enabled room-temperature Geiger-mode operation with DCR $\approx 8\text{ pA}$, PDP $\approx 12\%$ at 1310 nm, jitter $\approx 188\text{ ps}$, and after-pulsing $< 1\%$ at 0.75 V excess bias. The study marks a key advance in defect-filtered Ge/Si heteroepitaxy and dopant engineered field screening, showing that high-quality Ge-on-Si SPADs can now match InGaAs/InP performance without cryogenic cooling.

Wu et al. [81] numerically designed a CMOS-compatible Ge-on-Si SPAD employing a charge-focusing SACM structure designed to minimize electric fields at the Ge/Si interface and suppress leakage through defective corners of the Ge absorber. The device incorporates a 200 nm lightly p-doped Ge layer epitaxially grown atop a high-voltage p-well (HVPW) Si layer with a tailored DPW junction that laterally confines the avalanche region. Simulations show that reducing the DPW width from $18\text{ }\mu\text{m}$ to $2\text{ }\mu\text{m}$ concentrates the avalanche current at the DPW/N-buried-layer junction while keeping the high field away from the Ge edges, thereby lowering the dark count rate and afterpulsing without modifying the CMOS process. This material-engineered field distribution enables low-noise and low-jitter NIR detection within standard Si photonic platforms.

Despite the clear CMOS compatibility and afterpulsing advantages of Ge-on-Si SPADs, several intrinsic material limitations still prevent them from matching the room-temperature performance of mature InGaAs/InP devices. First, the Ge/Si lattice mismatch introduces threading dislocations and interface defects, which act as generation/recombination centers and increase dark current and DCR. Second, Ge has a narrow bandgap, so thermally generated carriers become significant at elevated temperatures and strongly limit room-temperature Geiger-mode operation. Third, the absorption edge of Ge is temperature-dependent; lowering the temperature widens the germanium bandgap, which

reduces absorption near 1550 nm, but operating at higher temperatures enhances absorption while also raising thermal noise. Finally, although the Si multiplication layer can reduce trap-assisted afterpulsing compared with InP multiplication layers, efficient 1550 nm detection still requires careful control of Ge thickness, Ge/Si interface quality, charge layer design, and electric field distribution. Therefore, the main intrinsic challenges of Ge-on-Si SPADs are the trade-off between 1550 nm absorption and thermally induced DCR, lattice mismatch defects, and interface carrier transport.

4.3. Defect Control and Uniformity Enhancement in SPADs

While the choice of semiconductor material defines the fundamental performance limits of a single-photon avalanche diode, the actual device efficiency, noise level, and reproducibility are often governed by defects and inhomogeneities introduced during epitaxy, doping, and post-processing. Structural imperfections such as dislocations, impurity clusters, and interface traps alter local electric field profiles, facilitate trap-assisted tunneling, and increase the DCR. Therefore, understanding and controlling these defects is as crucial as selecting the material itself. Over the past two decades, advances in epitaxial growth, thermal annealing, and surface passivation have significantly improved the crystalline quality and electrical uniformity of SPADs across silicon, III–V, and Ge-on-Si platforms. The following sections summarize the principal defect mechanisms, their material-specific manifestations, and the engineering strategies developed to mitigate them and ensure consistent large-scale fabrication.

4.3.1. Sources of Defects in SPAD Materials

Defects in SPAD materials originate from both bulk crystal imperfections and interfacial processes occurring during epitaxy and device fabrication. In silicon-based SPADs, dislocations and oxygen-related precipitates inherited from Czochralski-grown wafers can introduce deep-level traps that promote trap-assisted tunneling and increase dark counts under high electric fields [76]. For III–V semiconductors such as InGaAs and InP, the higher ionicity of the lattice and lower formation energy of native defects make them more susceptible to antisite defects, vacancies, and compositional inhomogeneity, especially at the heterointerface between absorption and multiplication layers. In heteroepitaxial systems like Ge-on-Si, the lattice mismatch ($\sim 4.2\%$) leads to the formation of threading dislocations that propagate through the Ge layer and act as generation/recombination centers, severely degrading the dark current and afterpulsing behavior. Defects and impurities near the depletion region act as generation/recombination centers following SRH statistics, often exacerbated by the absence of high-temperature annealing and gettering steps in standard deep-submicron processes, which increase trap-assisted tunneling and DCR [82]. Additionally, interface states at semiconductor/oxide boundaries (e.g., Si/SiO₂ or InP/SiO₂) and residual contamination from plasma etching or diffusion sources introduce charge traps that distort the local electric field, causing premature breakdown or excess jitter. High-quality materials with minimal defects and dislocations are essential, since local field enhancement around such imperfections can trigger premature avalanche or microplasma formation [15]. Trap-related phenomena such as carrier release following an avalanche event can further induce afterpulsing noise, which is particularly severe in III–V and Ge-based SPADs due to high trap densities [24]. These defects, though microscopic in scale, collectively determine macroscopic performance metrics such as DCR, PDP, and timing uniformity, emphasizing the need for defect-aware material processing in every SPAD platform.

4.3.2. Role of Fabrication Process in Defect Formation

The density and distribution of defects in SPADs are critically determined by the fabrication sequence, including epitaxial growth, ion implantation, thermal annealing, and surface passivation. During epitaxy, deviations in substrate temperature or precursor flow can alter stoichiometry and cause vacancy clusters or interstitial impurities, which subsequently act as recombination centers. In silicon-based SPADs, rapid thermal annealing (RTA) is routinely employed to activate dopants and repair implantation-induced lattice damage; however, incomplete recovery of point defects or oxygen precipitates can still contribute to dark current. High-temperature annealing effectively neutralizes the majority of radiation-induced defect complexes, restoring up to 85% of the original DCR level in irradiated CMOS SPAD arrays [83]. As noted by Wu et al. [84], suboptimal parameters in customized doping and RTA steps can increase defect densities and lead to wide variations in DCR across SPAD chips, emphasizing the importance of process control during epitaxial recrystallization. Sicre et al. [85] identified that hot-carrier stress primarily degrades the top Si/SiO₂ interface in CMOS SPADs through Si–H bond dissociation. These stress-induced interface traps act as mid-gap generation centers, contributing to long-term DCR drift and reliability degradation.

For III–V SPADs, precise control over MBE or MOCVD growth conditions, especially V/III ratio and temperature gradients, is essential to minimize compositional grading and dislocation propagation across heterointerfaces. In Ge-on-Si systems, cyclic annealing and defect-filtered buffer layers are used to annihilate threading dislocations and suppress mismatch-related leakage paths. As demonstrated by Wu et al. [81], adopting a charge-focusing SACM design effectively reduces the electric field at the Ge–Si interface and the Ge sidewall edges, thereby mitigating defect-induced leakage and afterpulsing commonly observed in lattice-mismatched Ge-on-Si SPADs.

The diffusion of dopants during high-temperature steps must also be optimized, since over-diffusion of Zn or B in InP or Si platforms can locally distort electric fields and enhance tunneling probabilities. Post-growth surface treatments play an equally important role: plasma damage, native oxide regrowth, or contaminant adsorption can introduce interface traps that degrade timing uniformity and breakdown stability. Therefore, surface passivation (for instance, using SiO₂, Al₂O₃, or InGaP capping layers) has become an indispensable step to stabilize surface charge and minimize leakage. As highlighted by Zou and Hu [75], the SiO₂ defect layer embedded within the optical microcavity plays a critical role in both optical tuning and defect suppression. Uniform ion-beam deposition of this layer is essential to prevent microcracks and interface traps at the NbTiN/SiO₂ boundary, which could otherwise increase dark noise and destabilize superconducting current. According to a study carried out by Na et al. [80], defects in Ge with energy levels close to the band edges are responsible for the generation of dark counts, implying that DCR can be further minimized through optimization of the fabrication process. In their device, this optimization mainly involved refining the UHV-CVD growth and annealing sequence to suppress threading dislocations, introducing a boron-doped Si screening layer to reduce sidewall field concentration, and improving surface passivation quality using ALD-deposited Al₂O₃, all of which contributed to the record low dark current. Xun et al. [86] showed that STI regions, where Si–SiO₂ interface charges accumulate, can increase the baseline DCR by more than 40% even before irradiation, highlighting the critical role of surface states in defect-mediated noise generation. Overall, the correlation between process parameters and defect dynamics underscores that SPAD fabrication is a delicate balance between achieving precise junction control and maintaining crystalline perfection, where each step (from epitaxy to passivation) directly impacts device reliability and noise performance.

4.3.3. Engineering Strategies for Defect Mitigation

To minimize the detrimental effects of bulk and interfacial defects, a variety of structural and process-level engineering strategies have been developed for SPADs. Among the most effective is the implementation of guard ring and junction isolation structures, which redirect the electric field away from edges and surface defect regions that tend to initiate premature breakdown. In CMOS-compatible Si-SPADs, STI combined with p^+ guard rings significantly reduce the contribution of surface leakage paths, while optimized field shaping ensures uniform avalanche triggering across the active area. As demonstrated by Jiang, Scott, and Deen [87], placing a polysilicon gate above the n-well guard ring effectively displaces the defective Si–SiO₂ interface associated with STI away from the active region, thereby reducing dark noise originating from interface traps. As evidenced by Jouni et al. [88], proton irradiation induces clustered vacancy-related defects in the depletion region of CMOS SPADs, forming sub-mid-gap energy levels that enhance trap-assisted tunneling and Poole–Frenkel emission. These clusters dominate dark count generation under high electric fields, although controlled thermal annealing at moderate temperatures can partially restore the crystalline order and reduce defect-mediated leakage.

In compound semiconductor SPADs, particularly InGaAs/InP devices, charge control layers with tailored doping profiles are introduced between the absorption and multiplication regions to confine the high-field zone and prevent carrier trapping at the heterointerface. Similarly, the use of double diffusion profiles (as in Zn-diffused InP or B-diffused Si junctions) enables smoother electric field gradients and suppresses trap-assisted tunneling near abrupt doping transitions. For epitaxial platforms such as Ge-on-Si, the adoption of gettering techniques, including thermal-cycle oxidation and sacrificial SiO₂ layers, effectively removes metallic impurities and reduces threading dislocation density before final device patterning. Additionally, the combination of chemical surface passivation (e.g., Al₂O₃ by ALD or sulfur-based wet treatments for III–V) with dielectric encapsulation has proven essential in neutralizing interface traps and stabilizing long-term dark current. These integrated strategies collectively ensure that the electric field is both spatially uniform and temporally stable during avalanche events, these being key prerequisites for achieving low DCR, minimal afterpulsing, and high pixel-to-pixel reproducibility in SPAD arrays.

Lattice mismatch at the Ge/Si and Ge/SiO₂ interfaces generates a high density of trap-assisted tunneling centers, which dominate dark count behavior in Ge-on-Si SPADs. Sidewall damage introduced by ICP-RIE etching further increases defect-related leakage [37]. To suppress these effects, the use of thermally grown GeO₂ passivation instead of PECVD SiO₂, as demonstrated by Coughlan et al. [89], has proven effective in reducing dark current by minimizing interfacial trap density. In addition, adopting a two-step Ge growth or wafer-bonding approach can mitigate dislocation formation at the Ge/Si boundary, while multimesa geometries confine electric fields away from defective edges, collectively enhancing the stability and noise performance of devices.

4.3.4. Material-Dependent Approaches to Defect Management

The strategies for defect control in SPADs must be tailored to the intrinsic material properties and growth mechanisms of each semiconductor platform. In silicon-based SPADs, where the crystal quality of the substrate is exceptionally high and the lattice is covalently bonded, the dominant defects arise from impurities, implantation damage, and oxide-interface traps rather than extended dislocations. Consequently, silicon devices benefit most from thermal annealing, oxidation gettering, and surface passivation. These processes eliminate point defects and neutralize charge at Si/SiO₂ boundaries. The mature CMOS environment also allows precise junction formation and reproducible isolation,

making defect-related noise more predictable and manageable. As reported by Villa et al. [90], large-diameter silicon SPADs often exhibit slightly reduced sensitivity at their central regions due to increased series resistance and localized self-quenching, a phenomenon that can mimic nonuniform defect-related response across the active area.

In contrast, III–V SPADs such as InGaAs/InP and GaAsP systems face fundamentally different challenges. Their ionic and polar lattice structures, combined with heteroepitaxial growth, make them prone to lattice mismatch, vacancy complexes, antisite defects, and compositional grading, especially at absorption/multiplication interfaces. As a result, controlling these defects relies heavily on epitaxial perfection/optimized V/III ratio, growth temperature, and interface grading to suppress threading dislocations and interface states. Post-growth diffusion and charge control layers are engineered to stabilize the field and mitigate tunneling paths caused by these crystalline imperfections. Between these two extremes lies the Ge-on-Si hybrid platform, where the 4.2% lattice mismatch introduces a unique combination of dislocation-driven and interface-related defects. Here, cyclic annealing, defect-filtered buffer layers, and careful passivation become indispensable to bridge the crystalline and amorphous regimes. Overall, while silicon SPADs rely on mature process cleanliness, III–V and Ge-based SPADs depend on epitaxial precision and heterointerface engineering, reflecting how defect management strategies evolve directly from the chemistry and lattice structure of the underlying materials.

For quantum-level single-photon detection, the ability of a material platform to overcome defect-related limitations is primarily determined by several intrinsic and process-dependent material properties. High crystalline quality and low dislocation density are essential because extended defects act as generation–recombination centers and promote trap-assisted tunneling under the high electric fields required for Geiger-mode operation. A low concentration of deep-level traps is also critical, since these traps can capture avalanche carriers and later release them, leading to afterpulsing and temporal instability. In addition, a suitable bandgap and low intrinsic carrier generation rate help suppress thermally induced dark counts, while favorable carrier ionization characteristics enable efficient avalanche triggering without excessive multiplication noise. Stable and passivable interfaces are also equally important, particularly in CMOS, III–V, and Ge-on-Si platforms, where interface states at semiconductor/oxide or heterojunction boundaries can distort the local electric field and trigger premature breakdown. Moreover, lattice compatibility, controlled doping, chemical stability during processing, and the ability to form uniform electric field profiles are necessary to minimize local field enhancement and pixel-to-pixel nonuniformity. Therefore, materials capable of supporting quantum-level SPAD performance need to efficiently absorb photons within the target spectral range while simultaneously ensuring defect tolerance, stable interfaces, low trap density, and consistent avalanche characteristics. Vines et al. [79] showed that the Ge-on-Si platform can reduce defect-related limitations in SWIR single-photon detection by assigning absorption and multiplication to materials with appropriate intrinsic properties. In this structure, Ge enables SWIR photon absorption, while the high-quality Si multiplication layer helps suppress deep-level trap-assisted afterpulsing that is commonly observed in InGaAs/InP SPADs. The favorable Ge/Si band alignment also facilitates electron transport and reduces carrier accumulation at the heterointerface. This confirms that high crystalline quality, low trap density, suitable band alignment, and stable heterointerfaces are key material properties for reducing DCR and afterpulsing in quantum-level single-photon detection.

4.4. Emerging Materials and Heterogeneous Integration for Next-Generation SPADs

The continuing demand for higher detection efficiency, broader spectral coverage, and scalable integration has driven the search for new material systems and device archi-

structures beyond conventional silicon, III–V, and Ge-on-Si SPAD platforms. While traditional semiconductors have matured through decades of optimization, their intrinsic trade-offs (between breakdown voltage, dark-count noise, and spectral response) limit further progress, especially in near-infrared and high-speed applications. Recent advances in materials science and nanoscale processing have enabled the exploration of emerging candidates such as two-dimensional (2D) transition metal dichalcogenides, wide-bandgap semiconductors, perovskites, and hybrid CMOS-III–V architectures. These materials offer tunable band structures, superior carrier transport, and the potential for heterogeneous monolithic integration, paving the way toward SPAD technologies with unprecedented performance and functional diversity.

4.4.1. 2D and Van Der Waals Materials

Two-dimensional semiconductors and van der Waals heterostructures have recently emerged as promising materials for next-generation SPADs due to their atomic-scale thickness, defect-free interfaces, and tunable bandgap properties. Materials such as graphene, MoS₂, WS₂, and WSe₂ exhibit high carrier mobility and mechanical flexibility while enabling efficient light absorption in the visible-to-near-infrared spectrum. Unlike bulk semiconductors, 2D materials can be stacked without lattice-matching constraints, allowing direct integration with silicon or III–V substrates through van der Waals bonding. This heterointerface eliminates many of the defect-related issues that typically arise in epitaxial growth, such as threading dislocations and interface traps, thus potentially reducing DCR and enhancing timing uniformity. Furthermore, their atomically smooth surfaces and controllable electronic properties provide a new platform for engineering ultra-thin avalanche regions, quantum-tunneling junctions, and waveguide-integrated SPADs suitable for on-chip quantum photonic circuits.

Recent demonstrations of low-dimensional SPAD prototypes have validated the potential of 2D semiconductors, particularly MoS₂, as active materials for photon counting devices. A notable example is the work by Roy et al. [91], who demonstrated a dual-gated bilayer graphene/MoS₂ van der Waals heterostructure capable of number-resolved single-photon detection at cryogenic temperatures. In this architecture, the MoS₂ layer acts as the photogeneration region, while the bilayer graphene channel (electrostatically gapped by a transverse field) serves as both the multiplication and readout elements. The study establishes a concrete experimental foundation for integrating 2D materials such as MoS₂ into SPAD-type architectures with ultralow noise and high single-photon resolution. The device achieved an optical gain on the order of 10^{10} , a DCR as low as 0.07 Hz, and a dynamic range of ≈ 40 dB. These results confirm that van der Waals stacking can provide ultraclean lattice-mismatch-free interfaces, enabling avalanche-like carrier multiplication without the excess noise typical of bulk APDs.

4.4.2. Hybrid CMOS-III–V Platforms

Heterogeneous integration of III–V semiconductors with silicon CMOS technology has emerged as a compelling route toward high-efficiency, broadband, and low-noise SPADs. III–V materials such as InGaAs, InP, and GaAs provide strong absorption and high carrier multiplication in the near-infrared range (1000–1700 nm), while CMOS offers mature fabrication, low-cost scalability, and advanced quenching/readout electronics. Direct monolithic growth of III–V layers on Si remains challenging due to lattice and thermal expansion mismatch, which induce threading dislocations and interface traps that elevate dark counts. To overcome these limitations, hybrid bonding and wafer-level transfer techniques such as direct oxide bonding, adhesive polymer bonding, and micro-bump interconnects enable the precise alignment of epitaxial III–V photodiodes with CMOS circuits without introducing high-density defects. This approach combines the superior optical

properties of III-V absorbers with the reproducibility and integration density of silicon technology, establishing a pathway toward compact and high-performance SPAD arrays for quantum communication and LiDAR systems. Ren et al. [92] demonstrated a 4×1 InGaAs/InP SPAD array flip-chip integrated on a silicon photonic SOI platform, achieving >10% on-chip PDE and HOM (Hong-Ou-Mandel) interference with 45% visibility, validating the feasibility of hybrid III-V/CMOS integration for scalable quantum SPAD systems.

4.4.3. Wide-Bandgap and Perovskite Materials

Beyond silicon and III-V semiconductors, wide-bandgap materials such as silicon carbide (SiC) and gallium nitride (GaN) have gained attention as alternative platforms for next-generation SPADs operating under high-field and high-temperature conditions. Their large bandgaps (3.2 eV for 4H-SiC and 3.4 eV for GaN) enable extremely low DCR, high breakdown voltage, and remarkable radiation hardness, which are crucial for spaceborne photon counting and harsh environment imaging. In particular, GaN-based avalanche photodiodes have already demonstrated sub-nanosecond timing jitter and ultraviolet sensitivity, suggesting the feasibility of GaN-SPADs once reliable single-photon triggering and quenching circuits are realized. Similarly, SiC offers a high carrier saturation velocity and chemical robustness, making it an ideal candidate for high-temperature single-photon detection where silicon devices fail.

The recent progress summarized by Lei et al. [93] demonstrates that GaN-based avalanche photodiodes (APDs) can already achieve high multiplication gains (10^5 – 10^6), ultraviolet responsivity, and sub-nanosecond avalanche build-up, approaching the dynamic range required for Geiger-mode operation. The transition from linear APDs to SPADs in GaN mainly depends on suppressing trap-assisted tunneling and surface leakage currents, both originating from deep-level defects and threading dislocations formed during heteroepitaxial growth. Through epitaxial lateral overgrowth, field-plate termination, and optimized mesa passivation, the dark current has been reduced by more than an order of magnitude, suggesting that controlled Geiger triggering could soon be achieved. Moreover, the article highlights that integrating GaN with 2D interfacial layers or low-defect AlGaN buffer stacks could further lower the avalanche threshold and stabilize breakdown, paving the way for GaN-based SPADs capable of room-temperature, ultraviolet single-photon detection.

A recent breakthrough in wide-bandgap single-photon technology was reported by Yu et al. [94], who realized a free-running ultraviolet SPAD based on 4H-SiC. The device employed a beveled-mesa structure to achieve uniform electric field distribution and suppress edge leakage, enabling ultralow dark current and exceptional stability under high bias. Using a combined passive quenching and active-reset readout circuit with tunable hold-off control, the researchers significantly reduced afterpulsing to only 0.3%, while maintaining a PDE of $\approx 10\%$ at 266 nm. The demonstrated uniformity of the active area and stable gain characteristics confirm that 4H-SiC can support Geiger-mode avalanche operation without thermal runaway or defect-induced noise. These results establish SiC as a viable wide-bandgap candidate for UV-sensitive radiation-hard SPADs, expanding single-photon detection into regimes where silicon and III-V devices face intrinsic limitations.

Parallel to these developments, hybrid organic–inorganic perovskites such as MAPbI₃ have emerged as tunable photoactive materials combining high absorption coefficients with adjustable bandgaps (1.5–2.3 eV). Recent perovskite photodetectors have shown low dark currents, high detectivity, and room-temperature operation, indicating strong noise advantages relative to Si in some regimes [95]; however, true perovskite APD/SPAD operation remains an open research direction. Representative works demonstrate high-performance perovskite photodiodes and platforms approaching photon

counting sensitivity, but not Geiger-mode APDs. The soft-lattice structure of perovskites, while susceptible to defect formation and ion migration, also allows strain-engineered tuning of carrier lifetime and gain uniformity, which is promising for flexible and room-temperature SPAD architectures. Together, wide-bandgap and perovskite materials represent complementary pathways; one focusing on robustness and radiation tolerance, and the other on low-cost tunable optoelectronic integration.

4.4.4. Quantum and Neuromorphic SPADs

The continuous expansion of material platforms, from silicon and III-V compounds to 2D van der Waals semiconductors, wide-bandgap crystals, and hybrid perovskites has redefined the design space of single-photon avalanche diodes. Each emerging class contributes a distinct advantage: atomically thin 2D materials provide defect-free heterointerfaces and ultrafast charge transport; hybrid CMOS-III-V integration enables broadband and scalable architectures; wide-bandgap semiconductors such as GaN and SiC deliver radiation-hard and high-temperature operation; and perovskites introduce solution processability and mechanical flexibility. Among wide-bandgap semiconductors, GaN, Al-GaN, InGaN, 4H-SiC, ZnO, and β -Ga₂O₃ are promising candidates for neuromorphic photonic and optoelectronic computing because of their wide bandgaps, low leakage currents, high thermal stability, and strong UV/visible photoresponse [96–100]. GaN and SiC materials are particularly attractive for robust and high-speed neuromorphic photodetectors, while ZnO and β -Ga₂O₃ can provide defect-assisted persistent photoconductivity, which is useful for emulating synaptic plasticity. In parallel, halide perovskites such as MAPbI₃, FAPbI₃, CsPbBr₃, CsPbI₃, and low-dimensional Ruddlesden–Popper perovskites offer strong light absorption, tunable bandgaps, solution processability, and ion-migration-mediated memory effects [101–107]. These properties make them suitable for artificial photonic synapses, neuromorphic vision sensors, and adaptive detector systems where optical input, memory retention, and event-driven signal processing can be integrated at the material level. Together, these advances suggest a shift from material-limited SPAD performance toward a regime of function-tailored design, where structural and electronic parameters can be independently optimized through heterogeneous integration. The interplay between materials engineering, interface control, and circuit co-design now determines the achievable trade-off among efficiency, noise, and timing precision.

Looking ahead, these developments are expected to catalyze the emergence of quantum and neuromorphic SPADs capable of intelligent photon detection beyond binary counting. Integration with quantum photonic circuits will enable on-chip time-correlated measurements, entangled-photon detection, and scalable sources. Simultaneously, coupling SPADs with adaptive electronics or memristive synapses may lead to neuromorphic imaging systems that mimic biological perception, performing local learning and event-driven processing directly at the detector level. As the boundary between materials and architecture continues to blur, the convergence of wide-bandgap robustness, 2D interface perfection, and CMOS-level integration marks a decisive step toward the next generation of intelligent, energy-efficient, and quantum-ready SPAD technologies.

5. Performance Optimization

Optimization of SPADs demands a delicate balance among key performance parameters such as DCR, PDE, timing jitter, afterpulsing probability, and operational speed. Unlike conventional photodiodes, SPADs operate in the Geiger regime, where stochastic avalanche triggering amplifies both desired and parasitic events. Thus, every improvement in sensitivity or detection efficiency must be accompanied by strategies to suppress unwanted noise and timing fluctuations. This section focuses on the major performance-lim-

iting factors in SPADs and the engineering approaches developed to overcome them, including defect and interface control to reduce DCR and optical as well as electronic optimization to enhance PDE across visible and NIR spectra. Additionally, the inherent trade-offs among the functional parameters, particularly between efficiency and DCR, are analyzed in the context of different material systems and architectures, providing a comprehensive view of how SPAD performance can be systematically tuned from the single-device to array level.

5.1. Origins and Types of Noise in SPADs

Noise in SPADs arises from both intrinsic stochastic processes and extrinsic material or circuit imperfections. Unlike conventional photodiodes that operate in a linear regime, SPADs function above the breakdown voltage, where each absorbed photon or thermally generated carrier can trigger a full avalanche event. Consequently, every fluctuation in carrier generation or avalanche dynamics manifests as an electrical pulse indistinguishable from a true photon event, making noise suppression a central challenge for performance optimization.

The dominant noise sources in SPADs can be classified into several categories. Thermal generation noise, primarily associated with defect-related trap states in the depletion region, gives rise to the DCR. Afterpulsing noise originates from carriers trapped during an avalanche that are later released, retriggering secondary avalanches correlated in time with the initial event. Timing noise (jitter) stems from variations in avalanche build-up time, carrier diffusion, and electronic readout fluctuations, directly degrading temporal resolution. In addition, electronic noise from quenching and amplification circuits and optical crosstalk in densely packed arrays further distort the photon statistics, particularly in high-speed or high-density architectures.

Each of these noise mechanisms scales differently with material properties, device geometry, and operational mode. Thermal and trap-assisted processes dominate in narrow-bandgap or defect-rich semiconductors, while timing and electronic noise become more pronounced at high gating frequencies. Therefore, noise reduction in SPADs cannot rely on a single approach but requires a coordinated optimization across materials, device structure, and circuit level, ranging from epitaxial purity and interface passivation to quenching circuit design and optical isolation. The following sections detail the mechanisms and mitigation strategies for DCR, afterpulsing, and jitter, which together define the ultimate limit of sensitivity and timing accuracy in modern SPAD systems.

5.2. Reduction in Dark Count Rate

The dark count rate represents one of the most critical noise parameters in SPADs, defining the fundamental limit of detection sensitivity. DCR originates from spurious avalanches triggered by thermally or field-generated carriers rather than photons. Its magnitude depends strongly on the material bandgap, defect density, trap states, surface leakage, and electric field profile within the multiplication region. In silicon-based SPADs, dark counts mainly arise from SRH generation at mid-gap traps and surface states at the Si/SiO₂ interface, whereas in InGaAs and Ge-on-Si, band-to-band tunneling and trap-assisted tunneling dominate due to smaller bandgaps and higher field strengths. To suppress DCR at the material and process level, multiple engineering approaches have been developed. In silicon devices, advanced thermal annealing and oxidation gettering eliminate implantation-induced point defects, while surface passivation with high-quality oxides or nitrides reduces interface trap density.

Wu et al. [108] demonstrated that both room-temperature and high-temperature annealing significantly reduce the DCR of irradiated CMOS SPADs. The DCR dropped by up to 40% after four weeks of self-annealing at room temperature and further decreased

during 120–140 °C annealing, consistent with the removal of phosphorus vacancy and oxide-related defect complexes. Finazzi et al. [109] demonstrated that laser annealing can effectively mitigate radiation-induced dark-count degradation in InGaAs/InP SPADs. By irradiating devices with 3 MeV protons (4×10^8 p/cm²) and subsequently performing 1550 nm laser annealing at 200 mW for 60 s, they achieved up to a 14-fold reduction in DCR without affecting PDE uniformity. The simulations confirmed localized heating (~455 K) sufficient to heal lattice defects in the InGaAs layer, revealing that near-infrared laser annealing is a rapid and efficient post-irradiation recovery strategy for III–V SPADs. For InGaAs/InP SPADs, graded heterointerfaces and optimized doping in charge-control layers minimize electric field peaks that drive tunneling. Pellegrini et al. [110] demonstrated that introducing a multistep InGaAsP graded layer and a properly doped InP charge-control layer between the absorption and multiplication regions effectively smooths the band discontinuity and suppresses tunneling-induced dark counts in InGaAs/InP SPADs. Furthermore, emerging fabrication techniques such as hydrogen passivation, ALD of conformal dielectrics, and cryogenic epitaxy are increasingly being used to suppress deep-level defect states across all platforms.

At the device and operational level, DCR can be further minimized by controlling the electric field and thermal activation. Guard-ring and beveled-mesa structures confine the high field region away from surface defects, reducing edge-related leakage. Liu et al. [111] experimentally demonstrated that the width of the virtual guard ring strongly affects dark noise in CMOS SPADs. When the guard-ring width was reduced from 2 µm to 1 µm, the electric field strength and current density near the device edge increased, causing a 2.7-fold rise in DCR and a two-fold rise in afterpulsing. The study concluded that maintaining a wider guard-ring region provides a better trade-off between fill factor and noise, effectively suppressing field-driven leakage and trap-assisted tunneling.

Temperature optimization (typically cooling to 200–250 K) for InGaAs SPADs lowers thermal generation rates exponentially. As demonstrated by Itzler et al. [112], gated-mode operation combined with temperature optimization between 200 K and 250 K effectively suppresses thermally generated dark counts in InGaAs/InP SPADs, yielding DCR-hold-off curves that collapse onto a universal temperature-dependent form. These advances have reduced DCR in modern SPADs by several orders of magnitude compared to early designs, enabling sub-Hz dark-count levels in cryogenic silicon devices and sub-kHz performance in near-infrared III–V detectors. Achieving further improvement now hinges on atomic-level defect control and interface perfection, linking material purity directly to system-level noise stability.

5.3. Suppression of Afterpulsing

Afterpulsing intensity depends strongly on the density and lifetime of deep-level traps, as well as on the hold-off time and excess bias during avalanche recovery. Therefore, most suppression strategies aim either to reduce trap occupancy or to accelerate trap release without compromising detection efficiency. At the circuit level, the most effective approach is the implementation of fast active-quenching and adjustable hold-off circuits. These circuits terminate the avalanche within a few nanoseconds and keep the diode reverse bias voltage below breakdown for a programmable duration, allowing trapped carriers to recombine. Stipčević et al. [113] developed an ultra-fast active-quenching circuit that terminates the avalanche within a few tens of nanoseconds, thereby preventing trap recharging and reducing afterpulse probability in silicon SPADs used for quantum communication. Their design demonstrated that shortening the quenching and hold-off phases can effectively suppress time-correlated avalanches without compromising detection efficiency. Liu et al. [114] introduced a hybrid quenching circuit combining fast active and high-frequency gated-passive modes, which suppressed higher-order afterpulses in

InGaAs(P)/InP SPADs by up to 85% at 223 K while maintaining gate-free detection efficiencies of 4–10%, highlighting that rapid avalanche termination is an effective route for minimizing correlated dark events. Lai et al. [115] developed a power-law-based correction model that quantitatively compensates for afterpulsing-induced errors in InGaAs SPAD systems, significantly improving count-rate accuracy and signal-to-noise ratio without hardware modification.

At the device and process level, reducing trap density through material engineering is equally crucial. Yu et al. [94] demonstrated that employing a wide-bandgap 4H-SiC material with a beveled-mesa geometry can intrinsically suppress trap formation and surface-related defects. Their device achieved an afterpulsing probability as low as 0.3% and a dark count rate of 133 kcps at 266 nm, confirming that process-level geometry control and high-purity semiconductor materials are effective routes to minimizing trap-mediated noise in SPADs. Restelli et al. [116] demonstrated that limiting avalanche charge through sub-nanosecond gating in InGaAs/InP SPADs drastically suppresses afterpulsing, with a two-orders-of-magnitude reduction when the gate duration was shortened from 1.5 ns to 0.5 ns. This highlights that controlling the temporal and charge dynamics of the avalanche process is an effective means of minimizing trap-related correlated noise.

The operational strategy of gated-mode detection provides an intrinsic reduction mechanism. By limiting the active window to a few nanoseconds and synchronizing the bias pulse with the optical signal, trap filling and re-triggering are minimized. Combining this technique with optimized quenching circuits has become the standard configuration for low-afterpulse SPAD operation in both visible and near-infrared ranges.

5.4. Timing Jitter Minimization

Timing jitter arises from statistical variations in avalanche build-up and signal propagation, which ultimately limit temporal resolution in SPAD systems. Its reduction depends on precise control of both the avalanche dynamics and the electronic response of the readout circuit. At the device level, minimizing jitter begins with a uniform electric field distribution and well-confined avalanche regions. Optimized doping profiles in the multiplication layer and thin high-field junctions shorten avalanche build-up time, while reducing the impact of carrier diffusion from the absorption layer. The use of low-defect epitaxial materials and controlled passivation further ensures consistent triggering behavior across the active area. At the circuit level, fast active-quenching and low-noise front-end electronics are essential. High-bandwidth preamplifiers and constant fraction discrimination allow rapid and consistent detection of avalanche onset, reducing timing uncertainty caused by signal amplitude variations. Nolet et al. [117] demonstrated that optimizing the discriminator and active-quenching circuitry of SPADs significantly improves timing performance—achieving sub-10 ps FWHM jitter through low-noise, high-bandwidth front-end design. Goll et al. [118] showed that integrating a SPAD with a high-speed bipolar CMOS (BiCMOS) gating circuit effectively reduces parasitic capacitance and avalanche build-up time from 3 ns to 1.6 ns, resulting in lower timing jitter and faster photon detection response. Temperature stabilization also plays a critical role, since avalanche initiation time and gain depend strongly on carrier mobility and breakdown voltage, both of which vary with temperature. Finally, system-level improvements such as synchronization between the optical trigger and bias waveform, as well as the adoption of thin-junction back-illuminated structures, enable sub-50 ps timing precision in state-of-the-art SPADs. Recent work on back-illuminated SPADs has demonstrated that adopting thin-junction backside-illuminated architectures yields a timing jitter below 50 ps FWHM by minimizing carrier transit time and enhancing optical coupling [119].

In CMOS SPADs, the ultimate timing jitter is limited by the stochastic nature of avalanche initiation, carrier transit dispersion, and the earliest detectable avalanche build-up

inside the high-field region. When diffusion tails, field nonuniformity, parasitic capacitance, and discriminator noise are minimized, the remaining timing uncertainty is mainly governed by impact ionization statistics and front-end bandwidth. Recent CMOS-integrated SPADs have already approached this practical limit. Nolet et al. [120] reported a 65 nm CMOS SPAD with an integrated quenching circuit achieving a 7.8 ps FWHM single-photon timing resolution, while a later optimized discriminator architecture from the same group achieved sub-10 ps performance with values down to about 6.3 ps FWHM [117]. These results indicate that state-of-the-art CMOS SPADs are now operating in the few-picosecond regime, and further improvement will likely require simultaneous reduction of avalanche build-up fluctuations, carrier transport dispersion, and electronic timing noise rather than optimization of a single parameter.

In the near-infrared region, timing jitter requires additional consideration because photon absorption often occurs deeper in the device or in a separate narrow-bandgap absorption layer. For Si SPADs, improving NIR sensitivity generally requires a thicker depletion or absorption region, which increases carrier transit-time dispersion and may introduce a long diffusion tail in the timing response. In InGaAs/InP and Ge-on-Si SPADs, where separate absorption and multiplication regions are commonly used, the total timing response is governed not only by avalanche build-up in the multiplication layer but also by the transport of photogenerated carriers from the NIR absorption layer to the high-field multiplication region. Therefore, the temporal response may deviate from a purely Gaussian distribution and can include an exponential tail associated with carrier transport and thermionic emission across heterointerfaces or grading layers. Key strategies for reducing NIR jitter include thinning or carefully optimizing the multiplication layer, maintaining a sufficiently high but controlled electric field, engineering smooth grading or charge layers to avoid carrier pile-up, suppressing diffusion from neutral regions, and using optical enhancement methods such as backside illumination, resonant cavities, or light-trapping structures to increase NIR absorption without excessively increasing carrier transport distance. In this sense, NIR SPAD design requires a balanced optimization of absorption efficiency, electric field distribution, carrier transit time, avalanche build-up time, and noise generation. As a general result, through coordinated optimization of materials, structure, and circuit design, temporal jitter can be effectively reduced without compromising photon detection efficiency or increasing dark noise.

5.5. Enhancing Photon Detection Efficiency in the Visible and NIR Spectrum

Photon detection efficiency depends on optical absorption, carrier collection, and avalanche triggering probability. Improving PDE requires coordinated optimization of optical coupling, junction design, and electric field control. In the visible range, the most effective approach is to increase the number of photons reaching the active region. This can be achieved by using optimized anti-reflection coatings and microlens arrays that guide incident photons toward the junction. Kohnen Poushi et al. [121] showed that applying an optimized anti-reflection coating (ARC) on CMOS SPADs increases photon detection efficiency in the visible range by improving optical transmission and by preventing standing-wave penetration of the shallow region where the avalanche junction is formed. Their analysis also demonstrated that selecting the correct ARC thickness at each wavelength produces measurable PDE enhancement across the visible spectrum. Pavia et al. [122] experimentally demonstrated that integrating plano-convex microlens arrays on a low-fill factor CMOS SPAD imager (5% active area) increases the effective collected photon flux by a factor of four to eight, depending on the optical f-number. Their measurements showed concentration factors of 4.3 at $f/2$ up to 8.1 at $f/11$, indicating that microlenses substantially recover otherwise-lost optical power and therefore provide a direct boost to the effective PDE without altering the device's intrinsic dark noise.

Thin-junction structures also improve PDE because carriers are generated closer to the high-field region. Back-illuminated architectures further enhance absorption by removing metal interconnects from the optical path and reducing substrate thickness. In the near-infrared range, the limits of silicon absorption become more significant. Increasing the depletion depth or using deep-diffused absorption regions helps to capture long-wavelength photons that penetrate deeper into the material. Efficient carrier transport requires an electric field that extends through the absorption region to prevent recombination. In hybrid platforms such as Ge-on-Si and III–V based devices, PDE can be increased by engineering the absorption layer thickness and reducing internal reflections at hetero-interfaces. Vines et al. [79] report one of the first normal-incidence planar Ge-on-Si SPADs that closes much of the performance gap to InGaAs/InP in the SWIR regime. By using a SACM geometry with a selectively implanted Si charge sheet, they confine the high electric field to the buried Si multiplication layer and keep the Ge absorber in a low-field regime. This planar field-engineering strategy sharply reduces surface-related dark generation compared with earlier mesa Ge-on-Si SPADs and enables a single-photon detection efficiency of 38% at 1310 nm and 125 K, together with about a fifty-fold improvement in NEP over mesa structures. Under matched operating conditions, the devices also show substantially lower afterpulsing than commercial InGaAs/InP SPADs, highlighting Ge-on-Si as a viable low-cost platform for large-area SWIR SPAD arrays in eye-safe LiDAR and quantum sensing applications. Through full optical–electrical simulation, McCarthy et al. [123] showed that integrating a photonic crystal nano-hole array into the Ge absorber can substantially increase absorption in Ge-on-Si SPADs. Their optimized design achieved a $2.5\times$ to $3.4\times$ enhancement in absorbed power at 1550 nm, providing a practical route to higher detection efficiency without a proportional rise in DCR when the photonic crystal is placed in an α -Si overlay rather than etched directly into Ge.

PDE also increases with excess bias because the avalanche triggering probability rises. This improvement must be balanced with the dark count rate and afterpulsing, since higher electric fields increase tunneling and trap activity. The best PDE performance is achieved when optical design, carrier dynamics, and field distribution are optimized together. Govdali et al. [124] demonstrated that in waveguide-coupled Si SPADs, increasing the excess bias directly raises the PDE because the avalanche initiation probability increases under stronger electric fields. However, as they reported, the higher field also elevates the dark count rate due to field-dependent carrier generation, so the operating point must be chosen to balance PDE and noise.

For optimum photon detection efficiency in the visible and NIR ranges, SPADs can also be integrated with photonic and quantum circuit elements that improve photon delivery to the active region. In visible-light systems, microlens arrays, anti-reflection coatings, spectral filters, and optimized free-space coupling improve the probability of incident photons reaching the avalanche region. In NIR and integrated quantum photonic platforms, waveguide-coupled SPADs, grating or edge couplers, resonant cavities, photonic crystal structures, and on-chip interferometric routing circuits can be used to enhance optical confinement and increase the interaction length between the guided optical mode and the absorption region. Such optical circuit engineering is particularly useful for QKD, TCSPC (time-correlated single-photon counting), entangled photon detection, and integrated quantum photonic systems, where efficient mode matching, low-loss routing, and synchronized time-gated readout are required to maximize PDE without increasing dark noise or afterpulsing.

5.6. Parameter Trade-Offs in SPAD Operation

Improving PDE in any SPAD architecture inevitably introduces trade-offs with DCR, afterpulsing probability, and timing jitter. Increasing the excess bias raises the avalanche

triggering probability and therefore enhances PDE, but it also strengthens the electric field in the multiplication region, which increases field-assisted carrier generation and exacerbates DCR [125]. Higher bias can additionally enlarge trap occupation, intensifying afterpulsing unless long hold-off times or optimized quenching electronics are used. Optical design choices that guide more photons into the active region, such as waveguide coupling, anti-reflection structures, and photonic crystal enhancements, may also increase the effective absorption volume, but they must be balanced against the risk of raising the number of generation/recombination sites or nonuniform fields that negatively affect DCR. Therefore, SPAD optimization requires simultaneous control of optical coupling, electric field distribution, and carrier dynamics to reach a bias point where PDE is maximized without unacceptable noise growth. Zhang et al. [126] demonstrated that in InGaAs/InP SPADs, PDE increases almost linearly with excess bias, whereas the DCR grows exponentially due to field-enhanced thermal and tunneling processes. Their measurements explicitly confirm an exponential PDE-DCR trade-off, showing that even modest improvements in avalanche probability require a disproportionately higher DCR, thereby constraining the optimal operating bias. Gulinatti et al. [127] demonstrated that achieving a higher PDE at 800 nm (up to 45% and beyond) using first- and second-generation red-enhanced SPADs (RE-SPADs) requires applying relatively high overvoltages (up to 20 V), which directly leads to an increased DCR. Moreover, to prevent voltage drop in the active region and to maintain precise control over the avalanche process, careful electric field engineering is essential. These trade-offs were addressed in the redesign of the second-generation structure by increasing the anode–substrate breakdown voltage, improving pixel isolation and reducing series resistance, thereby enabling simultaneous realization of high PDE and acceptable DCR in SPAD arrays.

5.7. Comparative Summary of SPAD Performance Across Materials and Architectures

To provide a clear overview of the state of the art in SPAD research and development, we have compiled a comparative table summarizing key performance metrics comprising PDE, DCR, timing jitter, afterpulsing probability, dead time, and operating temperature across a wide range of SPAD structures and materials reported in the literature (Table 2). This includes both experimental implementations and simulation-based studies. This consolidated view enables readers to quickly assess the strengths and limitations of various SPAD designs and to identify trade-offs and trends across different platforms. The tabulated data offers valuable insights into how different materials and device architectures impact critical performance parameters under diverse operating conditions.

Table 2. Key performance metrics of modern SPADs across a wide range of structures and materials.

Ref.	Structure/Material (Type)	PDE (λ /Bias)	DCR	Jitter	Afterpulsing	Dead Time	Temp
[128]	Si SPAD (behavioral model, simulation)	–	$\approx 6 \times 10^2$ cps	–	–	$\approx 2 \mu\text{s}$	300 K
[78]	Ge-on-Si SACM SPAD (mesa, experimental)	4% @ 1310 nm (10% excess bias)	$\sim 6 \times 10^6$ cps	300 ps	low	–	100 K
[129]	CMOS 0.5 μm n-well SPAD (lateral-diffusion guard ring, experimental)	–	7.5×10^2 cps	–	–	60 ns	300 K
[130]	Si SPAD in 0.18 μm standard CMOS, 10 μm active area + guard ring, actively quenched, used	PDP peak at 490 nm (value not explicitly given in text); devices used	3 kcounts/s @ 11 V to 90 kcounts/s @ 12 V (dark, single SPAD, RT)	–	–	≈ 5 ns	≈ 300 K

	in 64-SPAD array receiver [exp]	at 860 nm in comms demo					
[131]	Si SPAD in 0.18 μm CMOS, TWELL guard ring (exp)	6.2% @ 480 nm (1.0 V excess)	26 kHz @ 1.0 V excess	187 ps		5–500 ns	300 K
[132]	Si thin-junction SPAD on SOI + Nanostructured Light-Trapping (experimental)	32% @ 850 nm (≈ 3 V excess)	$\sim 4 \times 10^7$ cps (≈ 40 MHz)	25 ps	–	(10 ns gate)	K 300
[133]	Si SPAD in 0.35 μm CMOS (integrated AQC, experimental)	–	$\approx 1 \times 10^5$ cps @ Vex ≈ 3.3 V	–	<1%	9.5 ns	300 K
[79]	Planar Ge-on-Si SPAD (SACM structure, experimental)	38% @ 1310 nm, 125 K	$\approx 1 \times 10^6$ cps @ 125 K, 5% excess bias	≈ 310 ps	Strongly reduced vs. InGaAs/InP ($\approx 20\%$ of InGaAs afterpulsing at same SPDE)	Gated, 50 ns gate (hold-off variable)	78–125 K
[73]	InGaAs–GaAs Nanowire SAM-APD (experimental)	Not measured ($\lambda = 1064$ nm, free-running)	<10 Hz @ 77 K, VOB ≈ 10 V	<113 ps	$\approx 23\%$ @ VOB = 12 V (within 2 μs)	Passive quench (≈ 40 ns pulse separation)	77 K
[134]	Si SPAD + Compact AQRC (0.18 μm CMOS, experimental)	$\approx 34\%$ @ 450 nm (Vex = 3.5 V)	≈ 6.9 kHz @ 3.5 V	–	$\approx 0.75\%$ (hold-off 4 ns, Vex = 3.5 V)	≈ 5 ns	300 K
[76]	InP/InGaAs SPAD (SAGCM, model verified w/ simulation)	~ 10 – 15% @ 1550 nm (Vex ≈ 5 V)	$\sim 10^4$ cps @ 225 K, Vex = 5 V	–	Visible under short gate-off (<10 μs), strong TAT component	Gated, depends on external PQRC	Down to 225 K
[135]	CMOS SPAD Quenching Circuits (CQC & TBQC, simulation)	–	–	–	Reduced via AQT < 1 ns	7 ns (CQC), 21 ns (TBQC)	300 K
[136]	0.35 μm HV-CMOS SPAD, Type B (experimental)	37.4% @ 780 nm (Vex = 6.6 V)	2.88×10^4 cps @ 25 $^\circ\text{C}$	–	$\approx 10\%$ (dead-time-limited)	5.8 ns	25 $^\circ\text{C}$
[70]	28 nm FD-SOI CMOS SPAD (modified DNW + STI-optimized) (simulation)	–	≈ 20 Hz/ μm^2 @ 5% Vex	–	–	–	–
[87]	180 nm CMOS n ⁺ /p-well SPAD_NG/SPAD_DG/SPAD_FG (experimental)	PDP $\approx 7.35\%$ @ 560 nm, Vex = 0.5 V	SPAD_FG: DCR reduced up to $\sim 50\%$ vs NG (2.17 kHz @ -30 $^\circ\text{C}$, 0.5 V Vex)	–	SPAD_FG AP reduced up to $\sim 40\%$	–	-30 to 30 $^\circ\text{C}$
[137]	n/p-well CMOS SPAD (0.18- μm HV-CMOS)	26% @ 520 nm (Vex = 20%)	4–6 cps/ μm^2	224 ps	4.5–7.5%	PQC-limited	300 K
[45]	InGaAs/InP SPAD Array (Planar SAGCM + Carrier Extraction Structure, experimental)	≈ 15 – 20% @ 1550 nm (typical SAGCM range; PDE not the focus of paper)	<1 nA dark current (device-level), DCR not explicitly stated	–	–	–	300 K & 230 K (operational range)
[138]	Double-multiplication-region SPAD in 110 nm CMOS (experimental + TCAD simulation)	78% @ 500 nm (5.5 V), 25.5% @ 850 nm (5.5 V)	3.7 cps/ μm^2 @ 5.5 V (normalized)	165 ps @ 517 nm, 240 ps @ 850 nm	5.4% @ 5.5 V	~ 7 μs (passive quench)	RT to 50 $^\circ\text{C}$
[71]	Ge-on-Si double-mesa SPAD (experimental)	7.3% @ 1310 nm (Vex = 17.6%)	953 kHz @ 110 K	81 ps @ 17.6% Vex	–	Gated 50 ns, 3 kHz	110 K
[80]	Ge/Si SPAD (experimental)	12% @ 1310 nm (Vex = 0.75 V)	19 kHz/ μm^2	188 ps	<1% ($\tau \approx 90$ ns)	Gated (Tg = 5 ns)	300 K

[124]	Si waveguide-coupled SPAD (experimental)	7–10% @ 488 nm (Vex ≈ 20%)	50–107 kHz	~400 ns (circuit-limited)	–	~270 ns (passive quench)	21 °C
[139]	Ge-on-Si SACM SPAD with photonic crystal nano-holes (simulation)	Up to 2.41× SPDE enhancement @ 1550 nm	DCR varies depending on τSRH & sSRH	–	–	–	225 K (target)
[84]	CMOS SPAD (TSMC 18HV & 13HV), electric field-engineered structures (experimental)	–	Low–high (chip-dependent)	150 ps (FWHM, weak T-dependence for modified structure)	High in defective chips (short DT)	100 ns (used to suppress APP)	0–60 °C
[67]	InGaAs/InP SPAD (experimental)	20–45% @ 1550 nm (gated)	3–10 kHz @ 200 K	30–100 ps	Strong, temperature-dependent	Depends on gate (typically μs–ms)	150–220 K
[90]	CMOS SPAD (0.35 μm, experimental)	Peak ≈55% @ 420 nm (Vex = 6 V)	≈5–1 × 10 ⁵ cps (diameter-dependent, RT)	56–4470 ps (diameter-dependent, integrated QC)	<1% with proper hold-off	20–150 ns (diameter-dependent)	300 K
[140]	Nanostructured Si SPAD (simulation)	≈80% absorption, PDE ≈ 30–60% (λ = 400–1000 nm, Vex = 4–11 V)	–	≈15 ps FWHM	–	–	300 K
[69]	InGaAs/InP SAGCM + dielectric–metal reflector (experimental)	60% @ 1550 nm (1.25 GHz SWG)	340 kcps	–	14.8%	~3 ns	300 K
[141]	110 nm CIS, deep PW/DNW SPAD, no-STI, virtual guard ring (experimental)	58.3% @ 500 nm (VEX = 3 V)	15 cps @ VEX = 1 V	71 ps @ 670 nm	Negligible	100 ns	–
[86]	180 nm CMOS, P-I-N SPAD with buried-N + DNW (experimental)	~40% @ 500–600 nm (VEX = 6 V)	256 cps @ 10 V (fresh)	–	–	–	23 °C
[77]	InGaAs/InP SAGCM, double-zinc diffusion SPAD (experimental)	~24–27% @ 1550 nm (VEX = 5 V)	~7 kcps (Δr = 7 μm, 225 K)	157 ps (Δr = 10 μm, 225 K)	Strong dependence on diffusion profile	–	225 K

5.8. Optimization in SPAD Arrays

Scaling SPAD technology from single-pixel detectors to large arrays introduces a new set of challenges that demand careful architectural and circuit-level optimization. One critical concern is optical and electrical crosstalk between adjacent pixels, which can generate false counts and degrade spatial resolution. Approaches such as deep trench isolation (DTI), guard rings, and optimized layout geometries have been adopted to mitigate these effects. However, these methods often reduce the active area per pixel, bringing up another fundamental issue; the trade-off between the fill factor and PDE. Microlens arrays are frequently employed to compensate for this loss by focusing incident photons onto the photosensitive region, but their effectiveness can vary depending on the angular distribution of incoming light [122]. Rech et al. [142] demonstrated that trench isolation between pixels can significantly reduce direct-path optical crosstalk in SPAD arrays, although residual indirect paths (e.g., substrate reflections) may still contribute to spurious counts. Morimoto and Charbon [143] demonstrated that implementing guard-ring sharing in SPAD arrays can significantly enhance pixel density without degrading crosstalk performance. Their experimental comparison between conventional well-sharing and the proposed shared guard-ring structures revealed comparable levels of optical and electrical crosstalk. This suggests that guard-ring sharing is an effective layout strategy to scale SPAD arrays while maintaining isolation integrity between adjacent pixels. To compen-

sate for the low native fill factor (~5%) in CMOS SPAD arrays, Buchholz et al. [144] implemented microlens arrays on the SwissSPAD chip. This approach significantly enhanced photon coupling efficiency and effectively increased the usable fill factor, enabling detection of weak fluorescent signals that were previously undetectable.

Another design consideration in SPAD arrays involves the quenching and readout architecture. While using a per-pixel quenching circuit enables fast and independent recovery, it increases pixel complexity, area, and power consumption [145]. For time-to-digital converters (TDCs), per-pixel TDCs offer maximal parallelism and minimal timing skew, but suffer from area and power scalability, whereas column-parallel TDC architectures trade off latency and resolution for higher pixel densities [146]. Ensuring precise synchronization and minimizing timing skew across large arrays is crucial, especially for time-resolved imaging and LiDAR applications, and may require global clock distribution networks or calibration routines.

For megapixel SPAD arrays operated in gigahertz gating mode, thermal management becomes a critical system-level limitation. Repeated high-speed charging and discharging of the SPAD junction capacitance, quenching circuits, gating transistors, time-to-digital converters, and peripheral readout electronics can generate substantial dynamic power dissipation across the chip. The heat is not uniformly distributed, leading to local hot spots and temperature gradients across the array. Since DCR, afterpulsing probability, breakdown voltage, and timing jitter are temperature-dependent, nonuniform heating can degrade pixel-to-pixel uniformity and reduce the reliability of photon counting measurements. Therefore, large-format SPAD arrays require careful design of low-power gating circuits, duty-cycle control, thermal floor planning, heat spreading layers, backside heat extraction, and calibration schemes that compensate for temperature-induced variations in breakdown voltage and noise.

As the array size increases, the challenges in scalability and readout congestion become more pronounced, particularly under high-count-rate conditions. Bottlenecks in data transmission and buffering must be addressed through techniques such as hierarchical readout, sparsity-based compression, or event-driven schemes [147]. Conventional SPADs are inherently binary detectors; however, there is growing interest in enabling multi-bit photon-number resolution (PNR) either via multi-pixel fusion, time multiplexing, or ultra-fast active quenching that limits avalanche growth [148]. These enhancements open up possibilities for more advanced imaging modalities such as HDR (high-dynamic range) reconstruction and intensity-resolved photon counting.

6. SPADs in Quantum Communication

Photons play a fundamental role in the development of quantum technologies. Their dual particle–wave behavior, along with quantum properties such as superposition and entanglement, make them ideal carriers for transmitting quantum information [149]. The ability of high-PDE and low-DCR SPADs (optimized in Section 5.5) to capture these single photons directly enables the successful implementation of quantum communication protocols. Quantum key distribution (QKD) is a cryptographic approach whose security is based on the fundamental principles of quantum mechanics. In this process, the sender (Alice) prepares and transmits qubits in distinct quantum states and the receiver (Bob) measures them. Any attempt to overhear the message will produce detectable disturbances due to the quantum non-simulation theorem and the impossibility of distinguishing non-orthogonal states without ambiguity [150–152].

Protocols used within a quantum communication system are all QKD, e.g., BB84, B92, E91, DPS, six-state, etc. [153], and depend on the precise performance of the detector. Parameters such as timing jitter (Sections 2.3.4 and 5.4), afterpulsing (Sections 2.3.3 and 5.3), and dead time (Section 2.3.5) directly influence the secure key rate and timing accuracy in

QKD. The fundamental distinction between classical and quantum communications lies in the detection method. Classical systems use energy-integrating detectors (EIDs) that measure the total optical energy [154–156], while quantum systems employ photon counting detectors (PCDs), such as SPADs and SNSPDs, which record single photons as discrete events [157–159]. This approach, which derives directly from the operating principle of the Geiger mode (Section 2.1), improves the SNR [160,161]. The following discussion shows how material selection (Section 4.2), junction engineering (Section 3.4), and quenching circuits (Section 2.2) converge to meet the requirements of QKD, transforming SPADs from laboratory components into key elements of quantum communication infrastructures.

6.1. SPADs and SNSPDs in QKD

In QKD systems, SPDs are of notable importance; hence, engineering their parameters, especially by the use of specific materials, is crucial in order to obtain better performance and reliability. For instance, the ternary alloy InGaAs/InP SPDs at 1.55 μm have been proposed as a suitable choice for QKD. These structures have good performance in terms of detection efficiency, DCR, temporal interference, and afterpulsing, and their behavior has been reported to be similar to that of the best commercial examples. Studies show that the activation energy of the traps responsible for afterpulsing is about 250 meV, where the InP layer plays a major role in this phenomenon. Improving the design of these layers can improve the efficiency of SPADs in QKD applications [162]. An experimental implementation of the BB84 protocol using an attenuated diode laser at 1553 nm and an InGaAs/InP SPAD detector was also demonstrated. In this scheme, the polarization of the photons was tuned by Alice in four linear modes, and after compensating for polarization changes in the fiber and passing through the variable Faraday rotator (VFR) and analyzer, the detection was performed by the SPAD in the gated mode. Experimental results showed that the refined key was obtained with a quantum bit error rate (QBER) of about 1.3%, which is much lower than the 11% limit for secure QKD. Studies have shown the implementation of the BB84 protocol with a single SPAD and measured QBER as a function of SPAD gate window width [163].

The results show that increasing the gate voltage amplitude (V_{pp}) significantly reduces the afterpulsing and improves the PDE because in this case the effective gate width is reduced and the number of carriers involved in the avalanche process decreases. On the other hand, the trap lifetime falls with the operating temperature, and as a result the probability of afterpulsing decreases. The higher intrinsic absorption coefficient of InGaAs structures allows for a significant increase in PDE, which is a critical performance metric for quantum communication and sensing applications. According to the reported results, at temperatures around 300 K, the PDE value is close to 60%, while the DCR and afterpulsing probability values remain within acceptable ranges. These findings show that by optimizing the operating conditions, the performance of SPAD detectors can be improved beyond that of commercial products, which is very important for increasing the efficiency of QKD systems [69].

The latest studies show that comprehensive models for analyzing various sources of QBER in practical QKD systems (Figure 16), especially the DPS protocol, can significantly optimize system performance. The accuracy of these models has been verified by high-clock-rate experiments, which revealed that secure key distribution up to a 265 km fiber length with a low QBER and a suitable key rate is possible by setting optimal parameters. These models can be extended to other protocols based on coherent laser pulses and phase coding, and by optimizing the operating conditions of single-photon detectors the system efficiency and the length of the key distribution path can be increased, so that distances of more than 380 km can be achieved. These advances significantly improve the prospects

for the application of QKD systems in urban and intercity quantum networks (Figure 17) [164]. Recent experimental studies have confirmed the ability of these models to predict and optimize the performance of real QKD systems. For example, in a field implementation on a 100 km underwater optical channel with 20 dB loss, the use of an advanced fast-gated single-photon detector based on SiP technology resulted in a secure key rate of 1.5 kbps, a 14-fold improvement over conventional commercial detectors (such as ID221). The superior performance of this new detector, which operates at higher operating temperatures and without the need for complex cooling systems, is due to its active quenching circuit, which significantly reduces the effect of afterpulsing and allows gating at repetition rates of up to 150 MHz [165].

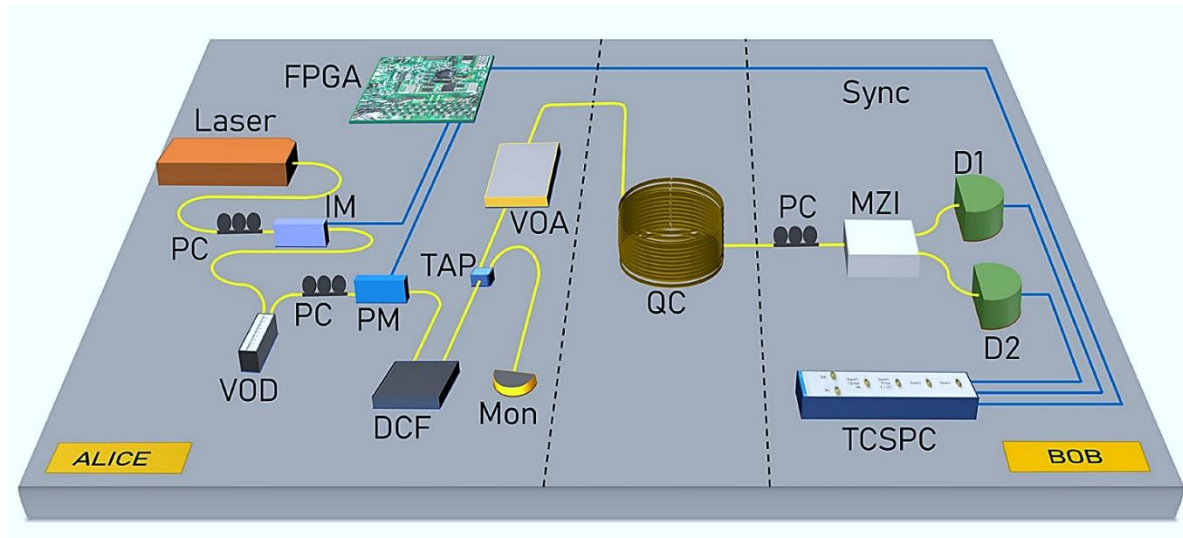


Figure 16. The 2.5 GHz clock DPS-QKD experimental setup; FPGA is used for synchronization and telecom-grade fiber is used as a quantum channel [164].

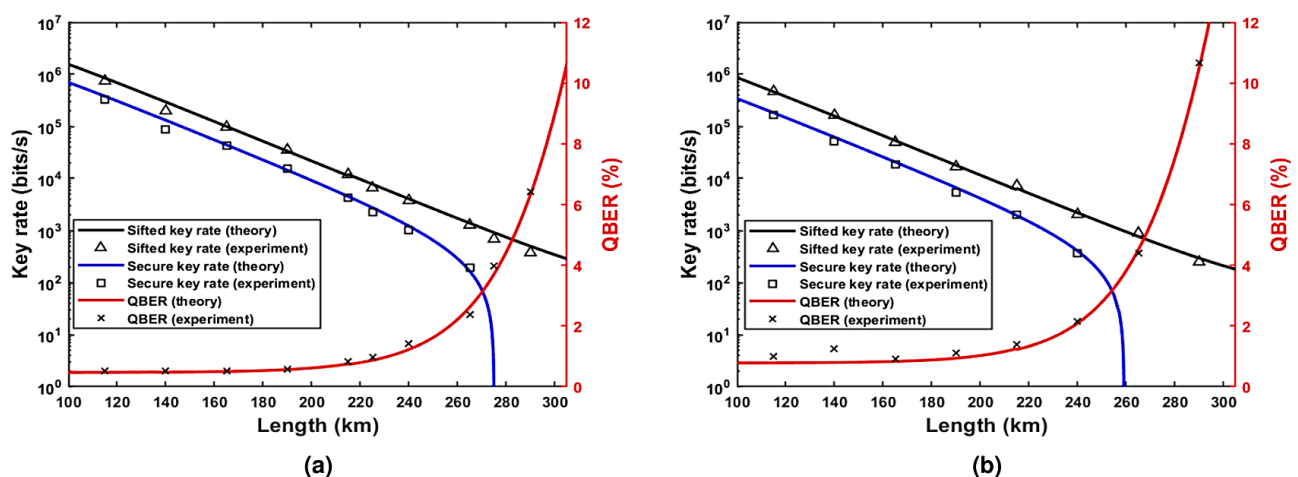


Figure 17. Performance of DPS-QKD with two choices of source linewidth: (a) 10 kHz linewidth and (b) 4 MHz linewidth [164].

The BB84 protocol has been implemented with SPAD detectors for free-space optical (FSO) systems, where qubits are encoded in linear polarization bases and photons are transmitted to the receiver through a quantum channel. The bases are randomly selected, and the photons are measured by SPADs. The high transmission rate and SPAD dead time directly affect the SNR and secure key rate (SKR). Increasing the transmission rate improves the SNR, but an excessive transmission rate reduces it, while longer gate times or

higher noise reduce the SKR. These results indicate that the practical design of FSO QKD systems requires a trade-off between transmission rate, SPAD dead time, and background noise level to optimize performance [40].

In addition to SPADs, the use of SNSPDs has also attracted much attention in recent years. These detectors offer better performance than SPADs due to their very high detection efficiency, negligible dark count rate, and short dead time, while they do not suffer from limitations such as afterpulsing. With the help of SNSPDs, it is possible to distribute secure keys over distances of more than 300 km in optical fiber and over 100 km in FSO systems. Furthermore, due to their high temporal resolution, these detectors are more robust to background noise and are considered an ideal choice for quantum networks over long distances and in noisy environments [166].

A comparison of the performance of two types of single-photon detectors, SPADs (e.g., InGaAs SPAD) and SNSPDs, in integrated QKD systems, has shown that although both are feasible for practical implementation, SNSPDs clearly have the advantage. Experiments have revealed that, using SNSPDs, secure key exchange is possible for up to 202 km of standard fiber with a stable secure key rate of about 9 kb/s and a QBER close to 0.9% (Figure 18). In contrast, InGaAs SPADs, although simpler to cool and less expensive, offer lower key rates and quality than SNSPDs due to lower efficiency (about 20%), higher afterpulsing, and larger timing jitter. This comparison suggests that SNSPDs are the optimal choice for long-distance quantum communications, while SPADs remain attractive for industrial applications due to their ease of operation. Tables 3 and 4 show a comparison between key parameters obtained for SNSPDs and SPADs, respectively [167].

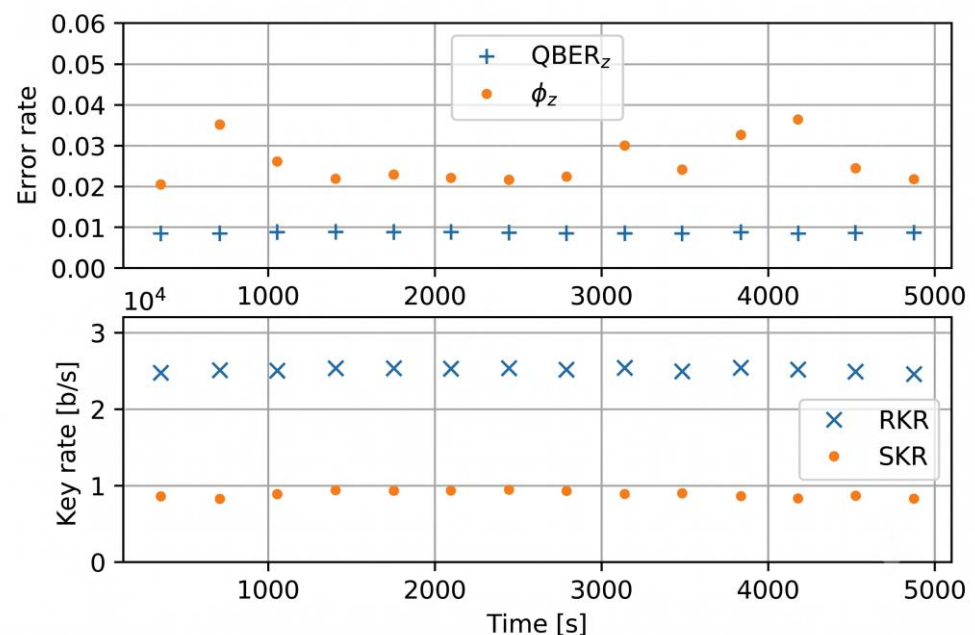


Figure 18. QBER, Φ (phase error rate), and secret key rate (SKR) during several secret key exchanges over 80 min using SNSPDs at a distance of 202 km single-mode fiber [167].

Table 3. Parameters and results of secret key exchange when using SNSPDs [167].

Length [km]	Attenuation [dB]	Block Time [s]	RKR † [kb/s]	QBER [%]	ϕ [%]	SKR † [kb/s]
-	30	37	216	0.9	1.0	91.0 *
-	36	124	66	0.8	1.1	28.3
-	38	168	42	0.8	1.4	17.2
-	40	306	27	0.8	2.1	10.6
202.0	39.5	351	25	0.9	2.2	9.4

251.7	42.7	720	12	0.5	2.2	4.9
-------	------	-----	----	-----	-----	-----

† RKR = raw key rate. * The value is estimated from raw data.

Table 4. Parameters and results of secret key exchange when using InGaAs SPADs [167].

Length [km]	Attenuation [dB]	Dead Time [μs]	Temperature [K]	Block Time [s]	RKR [kb/s]	QBER [%]	Φ [%]	SKR [kb/s]
-	30	20	188	453	18.0	3.6	2.1	2.9
-	35	32	183	858	9.6	3.1	4.5	1.3
-	40	20	188	1590	4.0	4.4	6.0	0.2
151.5	29.7	40	188	716	11.0	3.3	2.7	1.3
151.6	30.2	19	183	360	22.8	3.2	2.1	7.2

In QKD systems at the 1550 nm telecommunication wavelength, single-photon detectors play an important role in determining the secure key rate, QBER, and communication distance. InGaAs/InP APD detectors are widely used in QKD systems, but their performance is affected by factors such as dark count, afterpulsing, and dead time. Increasing dark and afterpulsing counts raises the QBER and reduces the secure key rate; therefore, methods such as fast gating and sinusoidal gating have been developed to mitigate these effects. Studies on the DPS-QKD protocol have shown that improving the performance of detectors can increase the secure key rate and enable quantum communications over longer distances. In addition to physical limitations, practical flaws in detectors can be a source of security attacks. In the DPS-QKD protocol, attacks such as photon-number-splitting (PNS), beam-splitting, and intercept-resend attacks have been investigated. Although DPS-QKD is more resistant to PNS attacks due to encoding information in the phase difference of consecutive pulses, the performance of the detector still has a direct impact on the ultimate security of the system. Studies have also shown that frequency-conversion and Si-APD detectors can provide better performance than conventional InGaAs/InP detectors due to their higher detection efficiency, lower DCR, and shorter dead time. Overall, the development of high-efficiency, low-noise single-photon detectors is one of the key factors for increasing the security and efficiency of quantum encryption and key distribution systems at telecommunication wavelengths [168,169].

SPADs are also used in biophotonics and quantum imaging applications. These single-photon detectors are used in TCSPC to reconstruct fast pulses and study phenomena such as fluorescence decay [170]. Combining them with artificial intelligence algorithms and recurrent neural networks has enabled direct extraction of fluorescence lifetimes with high accuracy and in real time. For example, the 32 × 32 Piccolo sensor has enabled imaging at a rate of about 10 fps (frames per second) [171]. The development of CMOS-SPAD arrays with a time resolution of 119 ps and an efficiency of 35% has also demonstrated the achievement of fast and near-real-time fluorescence imaging of biological processes, such as neural dynamics [172].

6.2. Security and Reliability Considerations in Quantum Communication

When it comes to communication, several factors are important, e.g., noise, which is crucial in the detection of original signals, photon transmission and receiving rates, and security. In quantum communications, noise is any disturbance or interference in the quantum state of a signal that can arise from photon collisions with the environment, electromagnetic disturbances, detector errors, or effects of eavesdropping. This noise reduces the fidelity of qubit transmission and increases the QBER, which if it exceeds a certain threshold, makes secure communication impossible. Estimating and isolating the error lines caused by noise and spying attacks using error correction and information reduction

protocols is essential to maintain the security of the final key. In practice, the balance between transmission distance, signal attenuation, and noise tolerance is a fundamental challenge in implementing global quantum networks [173]. In QKD protocols such as BB84, analysis by model checking showed that to obtain security, one must lower the probability of eavesdropping, which depends on the number of transmitted photons, the noise level of the channel, and the power of the eavesdropper [174]. A technology employed for security is the quantum random number generator (QRNG), classified as a type of true random number generator (TRNG). QRNGs fall into three categories: practical (trusting), self-testing, and semi-self-testing [175,176].

The operation of QRNGs is simple in general and it can be considered according to the sequence shown in Figure 19a. There are lots of configurations for the implementation of a QRNG. Figure 19b–e show a few simple ones with their measurement results. Figure 19b,c show an entanglement-based QRNG using a Hadamard gate. While all circuit inputs start in identical states with full probability, the results are two distinct strings, each occurring with a roughly equal likelihood—akin to a fair coin.

This is proven by the Shannon entropy value, which is approximately equal to 1, and, of course, this is the maximum possible entropy for this situation. Another configuration is shown in Figure 19d,e. The phase randomness, with the phase imposed by R_z gates, will change randomly at each time of execution of the circuit; thus, in the outcome measurement, we observe random changes in the probability of each outcome. In this case, the Shannon entropy could not give enough information about the randomness. For 16 unique outcomes, the maximum entropy is 4 bits. However, because the probability of each outcome changes randomly with every execution, the Shannon entropy also fluctuates unpredictably. Other configurations may follow different algorithms, but all QRNG systems, in their general form, adhere to the same fundamental procedure, as shown in Figure 19.

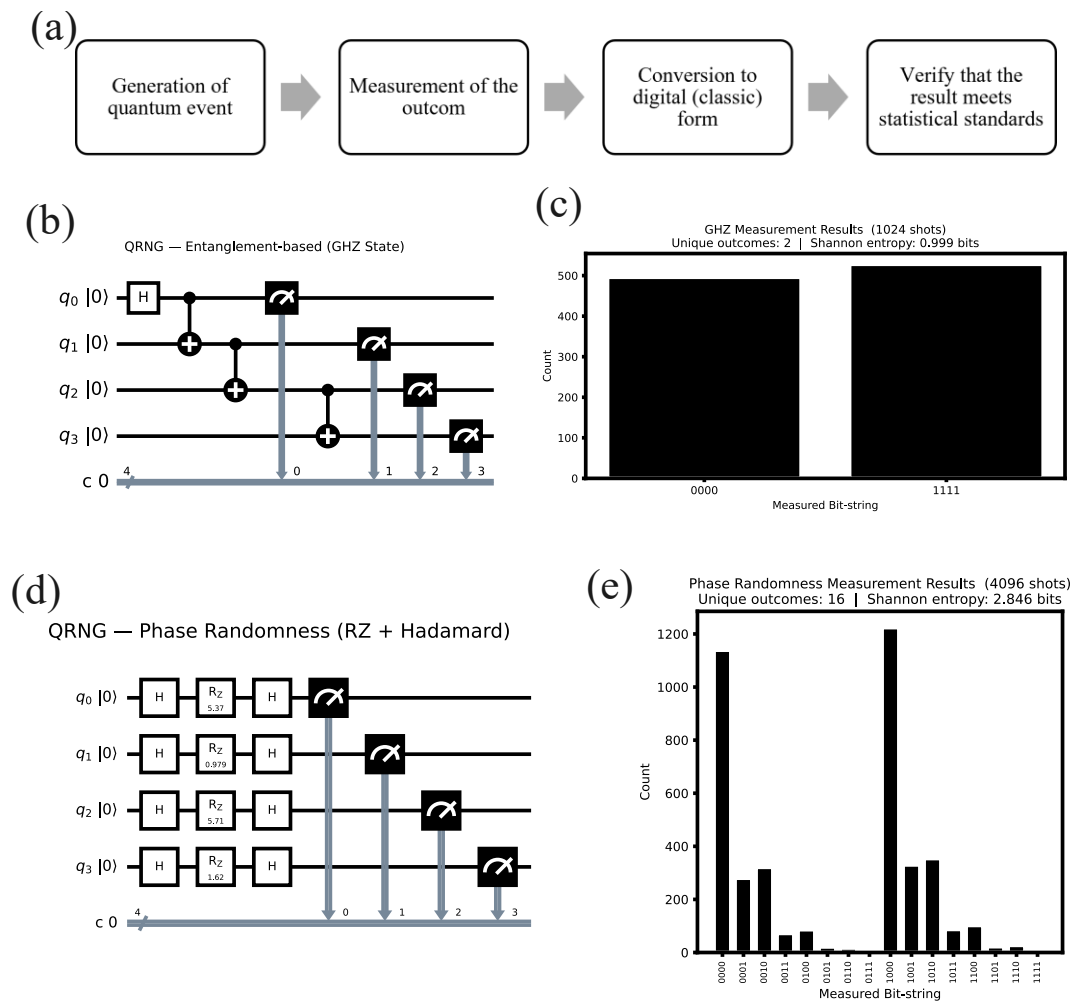
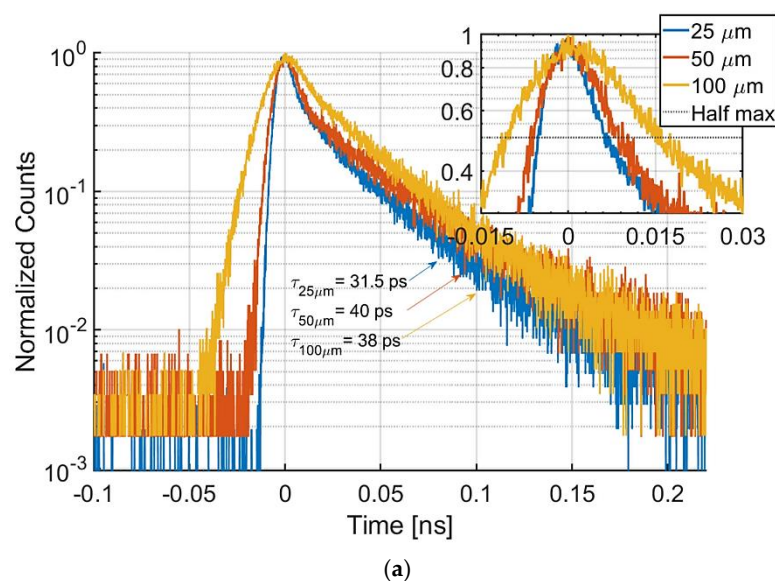


Figure 19. (a) Steps of operation for QRNGs. Entanglement-based (GHZ) (b) quantum circuit and (c) outcome measurements (conversion to digital form) for 1024 shots; the 2 unique outcomes have approximately equal probability, creating a fair-coin situation (Shannon entropy defines the least necessary number of bits to describe the uncertainty of the outcome). (d) Phase randomness quantum circuit and (e) outcome measurement for 4096 shots; there are 16 unique outcomes and it must be considered that each time the circuit is run, the probabilities of outcomes will change due to random changes in the phase, imposed on R_z gates, and guarantee the randomness within the outcome measurements.

The important role of QRNGs in security has led to hardware designs for practical usages. Among different implementations, SPAD-based QRNGs, which fall under the optical trusted QRNG category, are reported for applications such as LiDAR interference suppression, quantum imaging, quantum communication systems, and encryption [176–179]. Multidimensional QRNGs can be made using all-silicon and CMOS SPAD arrays [180,181]. One of the limits is the decrease in the bit rate with increasing photon flux, which is directly related to SPAD dead time. Researchers have shown that the combination of SPADs with different digital devices, such as FPGAs and flip-flops, can influence the threshold of the bit rate drop imposed by photon flux with different SPAD dead times. For instance, a SPAD-based QRNG system, designed by combining a SPAD with an FPGA and using a selection algorithm, typically operates with linear complexity and is faster than performing a full sort. Using a high-rank algorithm on an FPGA, it is possible to deal with the rapid drop in the bit rate caused by a higher photon flux. Furthermore, increasing the rank will increase the slope of the bit rate versus the SPAD count rate [182].

The device geometry, particularly the SPAD diameter, is another parameter that can influence SPAD action. Researchers working on SPAD pixel systems have shown that timing jitter, dead time, and DCR are all related to SPAD diameter and excess bias voltage (Figure 20). Afterpulsing, on the other hand, is important when one aims to minimize the dead time [183]. Therefore, given the relation between bit rate and SPAD dead time, all parameters that impact the dead time will impact the achievable bit rate as well. This identifies the dead time as a key performance-limiting factor in both classical and quantum communication systems.

In addition, the angle of incidence and the optical power impact SPAD operation, as they affect the bit error rate (BER)—the ratio of erroneous bits to the total number of bits received over a given time interval [184,185]. It is possible to reduce the BER effectively by increasing the optical power in SPAD operating modes, such as clock-driven (CD), free-running (FR), and time-gate (TG) modes, when no digital modulation is applied [185]; however, if a modulation scheme like on–off keying (OOK) is used with linear equalization (LEQ) or non-linear equalization (NEQ), there exists an optimal optical power [186]. The scenario is different when quadrature amplitude modulation (QAM) is employed. Reports showed that in DC-biased optical (DCO) orthogonal frequency-division multiplexing (OFDM), an increase in optical irradiance effectively decreases BER, especially when fewer constellation points are used, say, DCO-4QAM-7dB [187]. Another type of digital modulation that has been studied is pulse amplitude modulation (PAM). The study by Almer et al. showed that PAM can resemble QAM in behavior, yet it offers improved efficiency under certain conditions when comparing 4-QAM to 2-PAM [188]. On the other hand, in continuous-variable QKD (CV-QKD) systems, Gaussian modulation usually provides the best performance in terms of key rate and security. However, to reduce hardware complexity, discrete modulation, such as QPSK, is one of the most widely used methods in CV-QKD because it provides a good balance between data rate, system complexity, and security performance [189,190]. In conclusion, all factors, such as noise, photon rate, and security, are related and the limits imposed on each are dependent on the communication approach used, i.e., several concepts like digital modulation, SPAD geometry, and combination with different digital devices are of importance in order to decrease BER, which is inversely proportional to SNR. In other words, a higher SNR leads to a lower BER and error vector magnitude (EVM) [191–193].



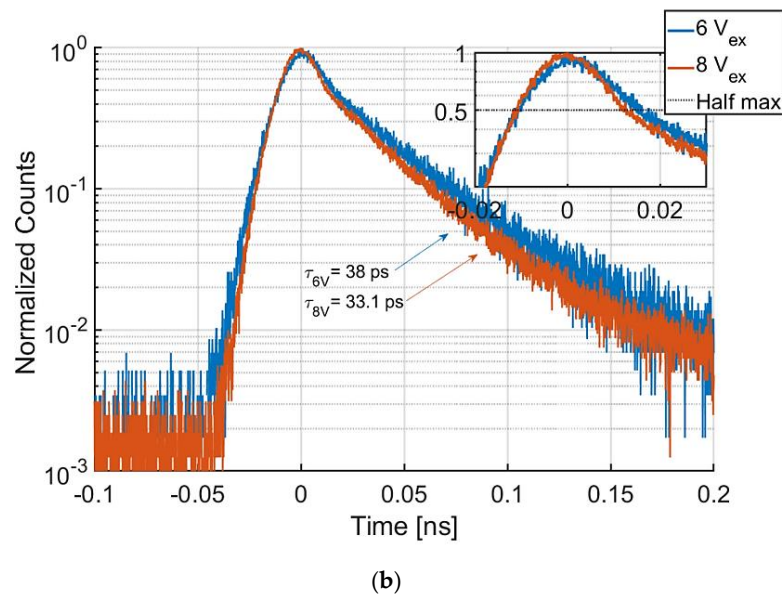


Figure 20. Jitter measurement for different (a) diameters and (b) excess bias voltages [183].

7. Emerging Trends and Future Directions

Despite significant advances in the design, fabrication, and optimization of SPADs, the performance frontiers of these detectors are still expanding. The present review confirms that major challenges persist as high R&D priorities, including noise reduction, quantum yield enhancement, array uniformity improvement, and efficient integration with readout integrated circuits. At the same time, the increasing need for single-photon detection in advanced photonic systems, including quantum communications, time-lapse imaging, and optical sensing, has driven the research trajectory in this area toward the integration of novel technologies and the use of hybrid platforms. In this regard, emerging approaches are focused on several key axes: the use of hybrid material substrates, including silicon and III–V materials or 2D materials (such as graphene and MoS₂), to enhance optoelectronic properties; the integration of photonics and electronics on a chip to achieve integrated and high-speed detectors; and intelligent control of the quenching process based on data-driven optimization and machine learning algorithms. In addition, adapting SPAD architectures and performance to the requirements of quantum communication systems, including high detection rates, precise timing, and high quantum fidelity, will determine future development paths. These trends outline the evolutionary perspective for the design of future generations of SPADs and establish the role of this technology as a key component in future photonic and quantum infrastructures. In this context, SNSPDs have become one of the best options for long-distance and high-rate QKD due to their very high detection efficiency, very low dark count rate, low time jitter, and detection capability at the telecommunication wavelength of 1550 nm. The use of SNSPDs allows for increased transmission distance, reduced QBER, and improved system security. On the other hand, CMOS technology plays a key role in the integration, cost reduction, scalability, and implementation of readout and signal processing circuits in QKD systems. CMOS SPAD arrays also enable the fabrication of compact, low-power, and integrated photonics-compatible single-photon detectors, which are of great importance for quantum networks and portable QKD systems [194,195].

Last but not least, it is worth briefly mentioning PNR techniques. PNR techniques are able to measure the exact number of incoming photons and thus increase the accuracy and security of advanced QKD protocols. However, conventional CMOS-SPAD arrays are more suitable for scalability in telecommunication networks because of their lower cost,

easier integration, and higher compatibility with existing photonic circuits. In contrast, many PNR detectors require complex refrigeration systems and higher implementation costs, which limit their widespread use. Therefore, PNRs offer better performance, but SPAD arrays are currently considered a more practical option for large-scale deployment in quantum telecommunication networks [38].

7.1. Hybrid Material Platforms (Si + III-V, 2D, Combination of Graphene, and MoS₂ on CMOS)

The evolution of SPAD technology in recent years has moved towards the hybrid integration of different material platforms, aiming to exploit the unique advantages of each platform. Hybrid III-V/silicon platforms allow the combination of the superior optical properties of III-V materials, such as direct band gaps and strong second-order non-linearity, with the processing capabilities and compatibility of CMOS silicon. This hybrid integration approach offers significant advantages compared to monolithic growth, which faces major challenges, such as lattice mismatch and the generation of high-density dislocations ($\sim 10^{10} \text{ cm}^{-2}$). Direct molecular bonding and fusion bonding techniques, performed at relatively low temperatures (300 K), allow preservation of the quality of the active materials and the creation of ohmic junctions with high conductivity [196–200].

On the other hand, two-dimensional materials, including graphene and transition metal dichalcogenides such as MoS₂, have attracted considerable attention as novel platforms for single-photon photodetectors. Graphene, with its extremely high mobility and broad absorption spectrum, when combined with MoS₂, which has a tunable energy gap, produces hybrid structures with superior performance. Recent studies have shown that graphene/MoS₂ detectors can have a responsivity up to 10^6 times better than MoS₂ detectors alone, which is due to the ultrashort transit time of charge carriers in the graphene layer and the relatively long carrier lifetime in the MoS₂ layer. This mechanism results in high photoconductivity gains that are comparable to those of traditional avalanche detectors [201–204].

Successful integration of these novel platforms with CMOS systems requires solving several technical challenges, including precise control of optical properties, mode distribution, and nonlinear responses. Recent advances in wafer-scale 2D material growth and the development of optimized transfer and coupling techniques have paved the way for the mass production of single-photon detectors based on this technology. Furthermore, recent research on hybrid integration of InGaAs/InP SPADs with silicon photonic chips, as shown in Figure 21 for instance, has shown that this approach can provide a practical solution for achieving scalable quantum information systems. These advances, along with continuous improvements in lithography and material processing techniques, point to a promising future for next-generation SPADs with improved performance and mass production capability [205–207].

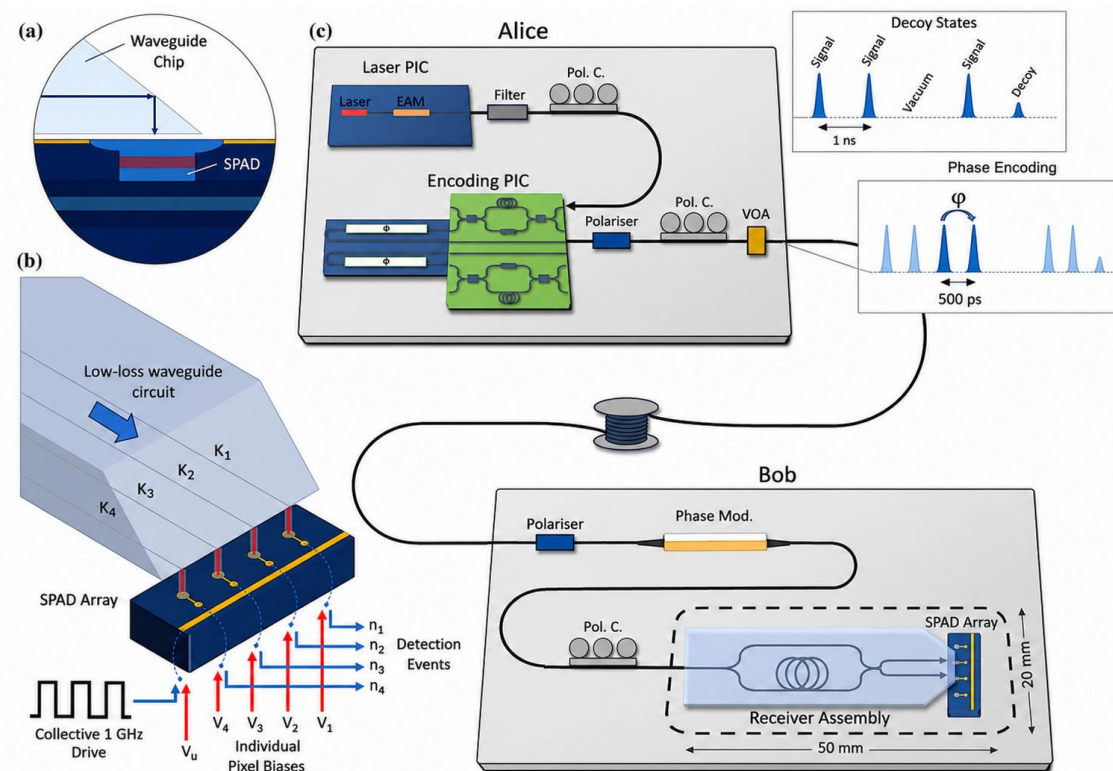


Figure 21. (a) Quasi-plane coupling: a diagonal cut in the silica chip reflects waveguide light into the SPAD. (b) A 1 GHz-gated SPAD array is coupled to the waveguide chip. Per-pixel DC bias compensates for variations in waveguide loss (κ_i) and detector gain (η_i). (c) On-chip BB84 QKD setup: Alice generates 1 GHz decoy-state time-bin qubits. Bob uses a SPAD array and an asymmetric Mach-Zehnder interferometer (AMZI). Polarization controllers, a VOA, and a phase modulator complete the system [205].

7.2. On-Chip Photonic–Electronic Co-Integration

The simultaneous integration of photonic and electronic circuits on a single integrated chip is one of the most important developments in next-generation SPAD technology, enabling fully integrated and high-performance systems. This monolithic integration method allows the fabrication of single-photon detectors based on Si and Si₃N₄ waveguides on a silicon substrate using standard CMOS processes. Recent studies have shown that monolithic SiN-on-Si detectors can maintain external quantum efficiencies of more than 60% over a wide spectral range from 400 to 640 nm, while achieving optoelectronic bandwidths of up to 9 GHz. In addition, the 3D integration technology using the 28 nm FD-SOI (fully depleted silicon on insulator) process (a very thin Si film implementing a transistor channel on an ultra-thin layer of insulator called a buried oxide positioned on top of a Si substrate) allows the placement of processing electronics under the sensitive layer of the SPAD, providing optimal fill factor and real-time data processing capabilities [208–211]. The monolithic integration of photonic and electronic components, as exemplified by Figure 22, offers significant benefits, such as reduced system complexity, improved signal integrity, and lower power consumption. However, this approach presents several technical challenges. Efficient coupling of light from transparent optical waveguides to light-absorbing regions in Si is one of the main challenges, requiring precise design of waveguide geometry and cladding materials, as can be seen in Figure 23, which reveals that including an intrinsic region impacts the performance. The need for precise phase calibration in coherent systems and management of propagation losses in integrated circuits are other challenges. However, recent advances in smart pixel-level integration,

which allows complex calculations such as histogramming and photon correlation analysis to be performed directly at the detection site, point to a promising future for this technology. These capabilities are especially important in quantum applications that require simultaneous detection of photons and fast data processing [212–216].

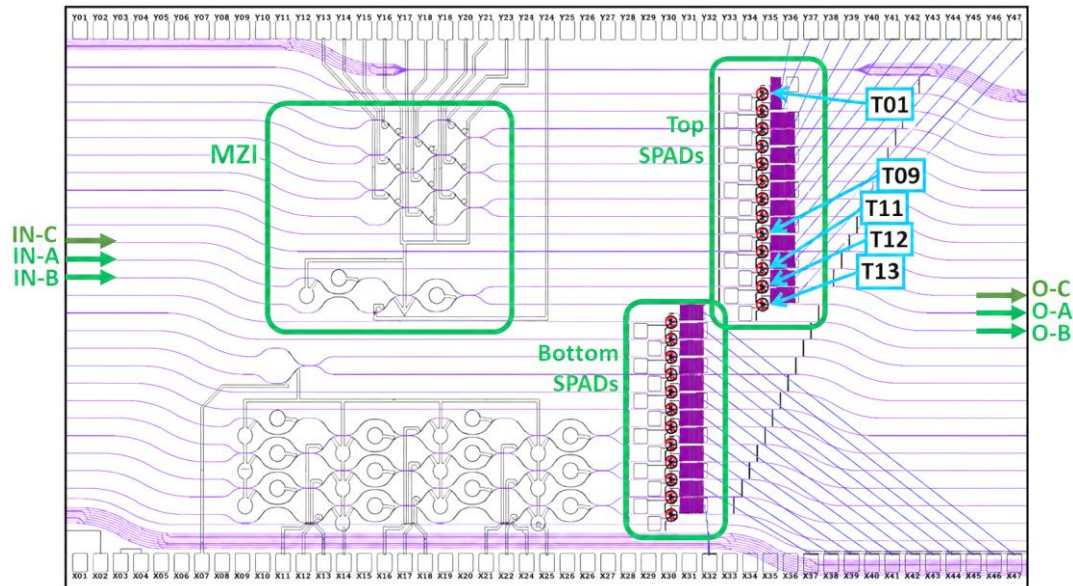


Figure 22. Monolithically integrated photonic chip. The layout includes waveguides, a Mach–Zehnder interferometer (MZI), and a SPAD detector fabricated on a single substrate. Light enters the waveguides from the left [211].

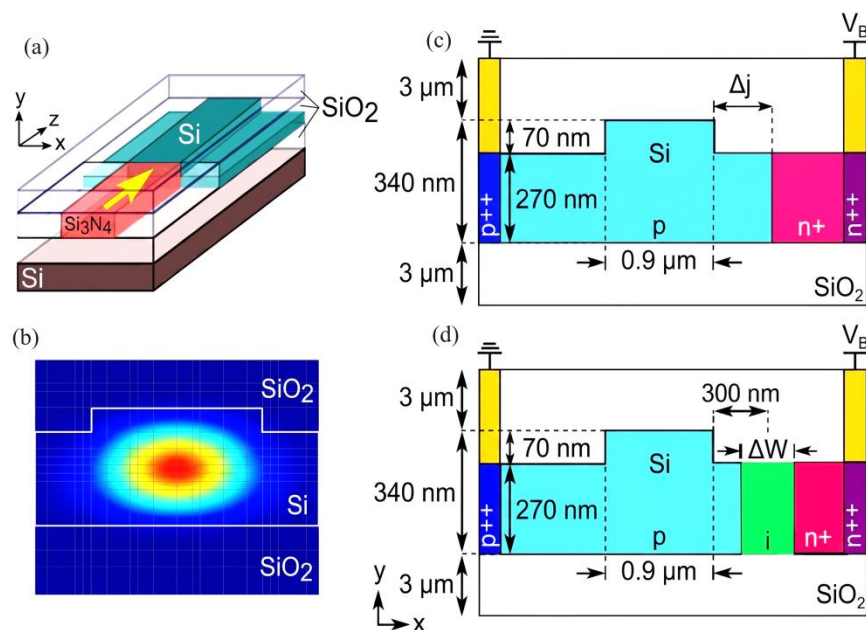


Figure 23. (a) SPAD structure with an end-fire coupled Si rib to Si_3N_4 input waveguide. (b) Optical mode profile at 640 nm for the fundamental quasi-TE mode. (c) p–n⁺ doping configuration showing the junction offset, Δj , from the waveguide edge. (d) p–i–n⁺ configuration with an intrinsic region of width, ΔW . The cross section of the waveguide remains constant along the length of the waveguide. Images are not drawn to scale [210].

7.3. Intelligent Quenching Control (AI-Driven Optimization and Adaptive Electronics)

Recent advances in SPADs have specifically focused on the design of active and adaptive quenching circuits, which aim to reduce dead time, reduce afterpulsing, and increase detector recovery speed, as shown in Figure 24. In this context, the use of adaptive resistive switch circuits that dynamically switch between different ohmic states based on voltage allows for high-speed avalanche quenching and rapid recovery without reducing detector sensitivity. In this approach, which is schematically shown in Figure 25, the dynamic resistance is changed in such a way that its value is very high during the avalanche (fast quenching) and gets reduced after the avalanche is over, allowing for faster recharge. Inspired by concepts of artificial intelligence and adaptive control, these structures can automatically optimize detector behavior. The design of active quenching circuits based on machine learning algorithms or adaptive control has been introduced in recent years as an innovative way to improve the performance of SPADs. These architectures allow for automatic adjustment of the quenching parameters by monitoring the detector response in real time and analyzing statistical data on parameters such as dead time, afterpulsing, and timing jitter. Studies indicate that, especially in high-speed applications such as QKD systems and time-of-flight imaging, the use of this type of intelligent control can lead to stable time response, reduced dead time, and overall improved detection efficiency. These intelligent approaches can enable continuous self-optimization of detectors for quantum and bioimaging applications in the future by connecting to processing chips [25,113].

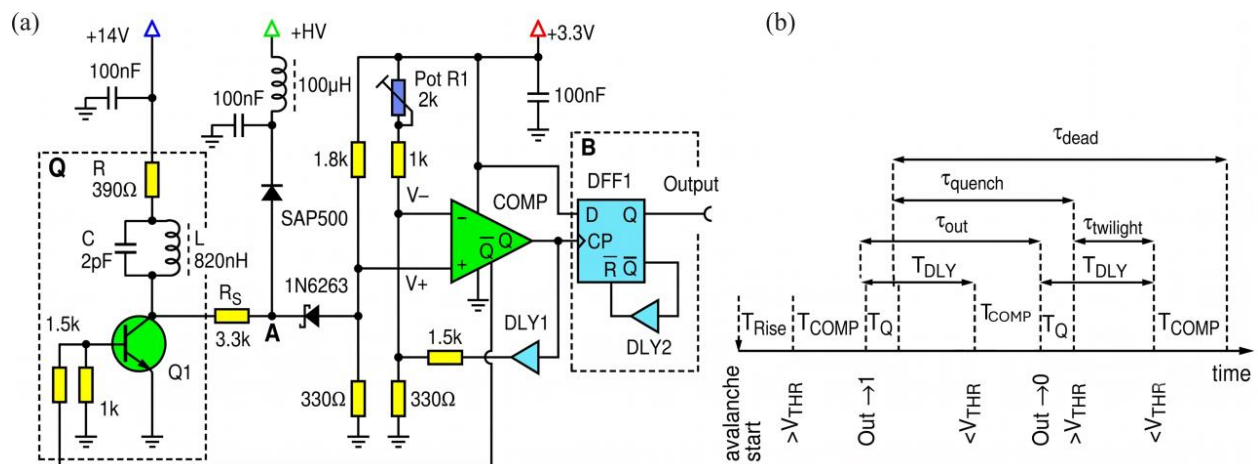


Figure 24. (a) Schematic diagram of an improved avalanche quenching circuit using the AD8611 fast comparator from analog devices. (b) Timing diagram of the photon detection cycle in this avalanche quenching circuit [113].

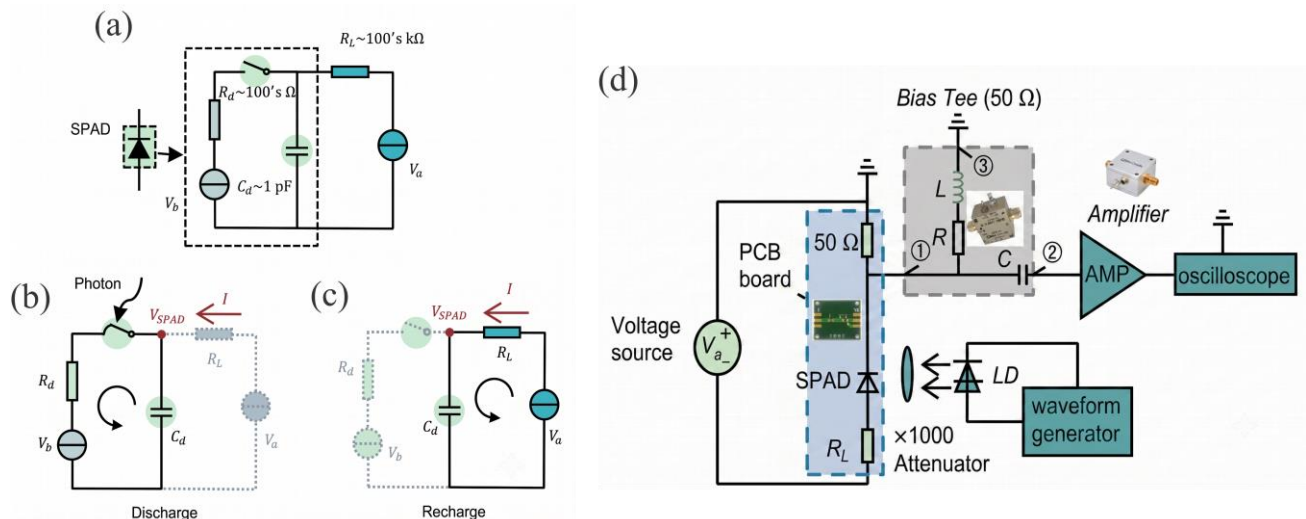


Figure 25. Schematic and experimental setup for a passively quenched SPAD. (a) Equivalent circuit model; the SPAD (hatched) is represented as a combination of a capacitor, resistor, photon-triggered switch, and voltage source in series with a quenching resistor. (b) Discharge process during photon-triggered avalanche. (c) Recharge process after quenching. (d) Experimental single-photon detection setup. A laser diode provides the weak optical signal, while the SPAD is reverse-biased [25].

7.4. Future Paths to Adapt the Needs of Quantum Communications

One of the most important challenges facing SPADs for quantum communication applications is the performance limitations of QKD over long distances, mainly due to the increase in QBER due to detector dark counts. Recent research has shown that the hybrid integration of InGaAs/InP-based SPAD arrays with low-loss silica waveguide chips can provide a practical solution for compact and scalable QKD systems. For example, SPAD arrays with gigahertz gating operation have achieved secure key rates of 2 Mbit/s over short distances and 15 Kbit/s over 100 km of optical fiber in BB84 protocol experiments. In addition to lowering the electronics costs, employing architectures that drive detectors collectively through a common cathode while individually optimizing each pixel's bias voltage also enables implicit synchronization of the gates across detectors [205,217]. For the future, progress towards increasing the quantum efficiency, reducing dark counts, and improving the temporal resolution of SPADs is essential to meet the needs of large-scale quantum communication networks. The development of single-photon detectors with high quantum efficiency at the telecommunication wavelength (1550 nm) and room-temperature operation is a major research priority that can overcome the limitations of current cryogenically cooled systems. Furthermore, the integration of capabilities such as photon-number detection, dead-time reduction, and smart signal processing at the chip level can lead to increased secure key rates and the development of regional accessible quantum networks. Ultimately, the hybrid integration approach of SPADs with integrated photonic circuits is a promising path to achieve deployable, cost-effective, and scalable QKD systems that can be exploited in existing telecommunications infrastructures [33,205,218–220].

8. Conclusions

This review has surveyed the state of single-photon avalanche diode (SPAD) technology across its full scientific and engineering landscape, spanning fundamental device physics, material platforms, fabrication strategies, and system-level performance optimization. By integrating perspectives from mature silicon-based detectors to advanced III–V heterostructures, Ge-on-Si hybrid platforms, and emerging materials, this work high-

lights how single-photon detection has evolved into a multidisciplinary field where progress is driven by the convergence of materials science, nanoscale device engineering, and integrated photonics. A central motivation of this review has been to clarify how these diverse aspects jointly influence sensitivity, noise characteristics, timing resolution, and scalability, which are performance metrics that underpin modern quantum technologies, imaging modalities, and photonic information systems.

A recurring theme across the survey literature is the dominant role of material selection and interface quality in shaping SPAD operation. Silicon remains unrivaled in its crystalline perfection, manufacturability, and CMOS co-integration, enabling low-noise detectors in the visible spectrum and large-scale imaging arrays. III–V architectures, particularly InGaAs/InP SACM and SAGCM designs, extend single-photon sensitivity into the telecom wavelengths, though challenges such as tunneling, afterpulsing, and cooling remain inherent to their narrow bandgaps and complex heterointerfaces. Hybrid Ge-on-Si structures now represent a compelling intermediate path: recent advances in mesa geometries, defect-filtered epitaxy, and sidewall field engineering have significantly narrowed the performance gap with III–V counterparts, while preserving compatibility with silicon photonics and CMOS readout. Complementarily, superconducting detectors define the performance ceiling in timing precision and dark-count suppression, illustrating the range of physical mechanisms available for single-photon detection.

Fabrication innovations have been equally decisive. Progress in epitaxial growth (MOCVD and MBE), multi-step diffusion, optimized charge-control layers, advanced passivation, and low-temperature bonding have collectively improved field uniformity, suppressed interface-mediated tunneling, and enhanced long-term stability. Across all platforms, defect management emerges as a governing constraint: trap states, threading dislocations, sidewall damage, and oxide-interface defects directly influence DCR, afterpulsing dynamics, and pixel-to-pixel reproducibility. Correspondingly, high-temperature annealing, epitaxial filtering, interface gettering, and conformal dielectric encapsulation have become foundational elements of modern SPAD process flows. On the performance front, this review has outlined the physical origins and mitigation pathways for dark counts, afterpulsing, timing jitter, and PDE. Active and gated quenching schemes, thin-junction and back-illuminated device geometries, microlens-assisted photon collection, photonic crystal absorbers, and field-engineered SACM layers have enabled substantial improvements in efficiency and timing without compromising noise. At the array level, scaling introduces new constraints related to optical crosstalk, fill factor, quenching architecture, and readout bandwidth. Recent work on deep-trench isolation, guard-ring sharing, and compact TDC topologies highlights how pixel-level physics and system-level architecture are increasingly interdependent. The relevance of these developments to quantum communication is growing rapidly. High-speed QKD experiments, long-distance fiber links, underwater channels, and free-space optical demonstrations show that SPAD performance directly determines achievable secure key rates and communication distances. Although SNSPDs currently set the benchmark for deep quantum applications, advances in room-temperature NIR SPADs, hybrid III–V/Si integration, and high-speed gating circuits suggest a path toward compact and deployable quantum-ready detectors.

Looking ahead, the emergence of heterogeneous integration (III–V/Si-, 2D/Si-, and SiN-based photonics), monolithic photonic–electronic co-design, and intelligent quenching architectures indicates a transition from material-limited SPADs to function-tailored platforms. These technologies promise detectors that combine high PDE, ultralow noise, and sub-nanosecond timing with CMOS-scale manufacturability. As quantum communication, bio-photonics, autonomous sensing, and neuromorphic photonic systems mature,

SPADs are poised to become a central enabling technology. Continued progress in interface engineering, materials synthesis, and circuit co-integration will be essential for realizing the next generation of scalable, high-performance single-photon detectors.

Author Contributions: M.A. contributed to the investigation and review of materials, fabrication strategies and material platforms, performance optimization, and writing of the Introduction section. S.T.S. contributed to the review of operating principles and the literature related to device engineering and architectures. P.N. contributed to the review of SPADs in quantum communication, as well as emerging trends and future directions. M.G. secured funding, supervised the work, guided the research group, and contributed to manuscript writing and editing. All authors have read and agreed to the published version of the manuscript.

Funding: The Iran National Science Foundation (grant no. 4001694), the National Natural Science Foundation of China (grant no. 52161145404), and the CAS-Iranian Vice Presidency for Science and Technology (project nos. 4001124 and 211134KYSB20210011) are acknowledged for funding this research in the framework of international projects.

Data Availability Statement: No new data were produced or examined for this study. This review draws solely on previously published works, all of which are referenced in the manuscript.

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

References

- McIntyre, R.J. Theory of Microplasma Instability in Silicon. *J. Appl. Phys.* **1961**, *32*, 983–995.
- Haitz, R.H. Model for the Electrical Behavior of a Microplasma. *J. Appl. Phys.* **1964**, *35*, 1370–1376.
- Renker, D. Geiger-mode avalanche photodiodes, history, properties and problems. *Nucl. Instrum. Methods Phys. Res. Sect. A* **2006**, *567*, 48–56.
- Lacaita, A.; Cova, S.; Samori, C.; Ghioni, M. Performance optimization of active quenching circuits for picosecond timing with single photon avalanche diodes. *Rev. Sci. Instrum.* **1995**, *66*, 4289–4295.
- Carrara, L.; Niclass, C.; Scheidegger, N.; Shea, H.; Charbon, E. A gamma, x-ray and high energy proton radiation-tolerant CIS for space applications. In *Proceedings of the 2009 IEEE International Solid-State Circuits Conference (ISSCC)*; IEEE: Piscataway, NJ, USA, 2009; pp. 40–41.
- Rochas, A.; Gani, M.; Furrer, B.; Besse, P.A.; Popovic, R.S.; Ribordy, G.; Gisin, N. Single photon detector fabricated in a complementary metal–oxide–semiconductor high-voltage technology. *Rev. Sci. Instrum.* **2003**, *74*, 3263–3270.
- Rochas, A.; Gosch, M.; Serov, A.; Besse, P.A.; Popovic, R.S.; Lasser, T.; Rigler, R. First fully integrated 2-D array of single-photon detectors in standard CMOS technology. *IEEE Photonics Technol. Lett.* **2003**, *15*, 963–965.
- Bondarenko, G.; Buzhan, P.; Dolgoshein, B.; Golovin, V.; Guschin, E.; Ilyin, A.; Kaplin, V.; Karakash, A.; Klanner, R.; Pokachalov, V.; et al. Limited Geiger-mode microcell silicon photodiode: New results. *Nucl. Instrum. Methods Phys. Res. Sect. A* **2000**, *442*, 187–192.
- Morimoto, K.; Ardelean, A.; Wu, M.-L.; Ulku, A.C.; Antolovic, I.M.; Bruschini, C.; Charbon, E. Megapixel time-gated SPAD image sensor for 2D and 3D imaging applications. *Optica* **2020**, *7*, 346–354.
- Eisaman, M.D.; Fan, J.; Migdall, A.; Polyakov, S.V. Invited Review Article: Single-photon sources and detectors. *Rev. Sci. Instrum.* **2011**, *82*, 071101.
- Yu, Y.; Lv, H.; Tariq, K.; Liu, D.; Sheng, X.; Feng, C. Study of the performance of photomultiplier tubes at high variable counting rates. *Nucl. Instrum. Methods Phys. Res. Sect. A* **2021**, *1008*, 165433.
- Lolli, L.; Taralli, E.; Portesi, C.; Monticone, E.; Rajteri, M. High intrinsic energy resolution photon number resolving detectors. *Appl. Phys. Lett.* **2013**, *103*, 41107.
- Baek, B.; McKay, K.S.; Stevens, M.J.; Kim, J.; Hogue, H.H.; Nam, S.W. Single-Photon Detection Timing Jitter in a Visible Light Photon Counter. *IEEE J. Quantum Electron.* **2010**, *46*, 991–995.
- Fukasawa, A.; Hotta, Y.; Ishizu, T.; Negi, Y.; Nakano, G.; Ichikawa, S.; Nagasawa, T.; Egawa, Y.; Kageyama, A.; Adachi, I.; et al. Development of a new 2-inch hybrid photo-detector using MPPC. *Nucl. Instrum. Methods Phys. Res. Sect. A* **2018**, *912*, 290–293.
- Izhnin, I.I.; Lozovoy, K.A.; Kokhanenko, P.A.; Khomyakova, K.I.; Douhan, R.M.H.; Dirko, V.V.; Voitsekhovskii, A.V.; Fitsych, O.I.; Akimenko, N.Y. Single-photon avalanche diode detectors based on group IV materials. *Appl. Nanosci.* **2022**, *12*, 253–263.

16. Rawat, A.; Ahamed, A.; Bartolo-Perez, C.; Mayet, A.S.; McPhillips, L.N.; Islam, M.S. Design and Fabrication of High-Efficiency, Low-Power, and Low-Leakage Si-Avalanche Photodiodes for Low-Light Sensing. *ACS Photonics* **2023**, *10*, 1416–1423.
17. Sze, S.M. *Semiconductor Devices: Physics and Technology*; John Wiley & Sons: Hoboken, NJ, USA, 2008.
18. Huntington, A.; Compton, M.; Coykendall, S.; Soli, G.; Williams, G.M. Linear-mode single-photon-sensitive avalanche photodiodes for GHz-rate near-infrared quantum communications. In *Proceedings of the MILCOM 2008—2008 IEEE Military Communications Conference*; IEEE: Piscataway, NJ, USA, 2008; pp. 1–6.
19. Qian, X.; Jiang, W.; Elsharabasy, A.; Deen, M.J. Modeling for single-photon avalanche diodes: State-of-the-art and research challenges. *Sensors* **2023**, *23*, 3412.
20. Yu, Y.; Wang, C.; Shi, H.; Yang, B.; Liu, X. A Review of Quenching Circuit Design Based on Geiger-Mode APD. In *Proceedings of the 2018 IEEE International Conference on Mechatronics and Automation (ICMA)*; IEEE: Piscataway, NJ, USA, 2018; pp. 28–33.
21. Cusini, I.; Berretta, D.; Conca, E.; Incoronato, A.; Madonini, F.; Maurina, A.A.; Nonne, C.; Riccardo, S.; Villa, F. Historical perspectives, state of art and research trends of single photon avalanche diodes and their applications (part 1: Single pixels). *Front. Phys.* **2022**, *10*, 906675.
22. Hayat, M.M.; Itzler, M.A.; Ramirez, D.A.; Rees, G.J. Model for passive quenching of SPADs. In *Proceedings of the SPIE—The International Society for Optical Engineering*; SPIE: Bellingham, WA, USA, 2010; Volume 7608, pp. 709–716.
23. Gallivanoni, A.; Rech, I.; Ghioni, M. Progress in quenching circuits for single photon avalanche diodes. *IEEE Trans. Nucl. Sci.* **2010**, *57*, 3815–3826.
24. Hall, D.; Liu, Y.; Lo, Y. Single photon avalanche detectors: Prospects of new quenching and gain mechanisms. *Nanophotonics* **2015**, *4*, 397–412.
25. Zheng, J.; Xue, X.; Ji, C.; Yuan, Y.; Sun, K.; Rosenmann, D.; Wang, L.; Wu, J.; Campbell, J.C.; Guha, S. Dynamic-quenching of a single-photon avalanche photodetector using an adaptive resistive switch. *Nat. Commun.* **2022**, *13*, 1517.
26. Liu, J.; Li, Y.; Ding, L.; Wang, Y.; Zhang, T.; Wang, Q.; Fang, J. Fast active-quenching circuit for free-running InGaAs (P)/InP single-photon avalanche diodes. *IEEE J. Quantum Electron.* **2016**, *52*, 4000306.
27. Thew, R.T.; Stucki, D.; Gautier, J.-D.; Zbinden, H.; Rochas, A. Free-running InGaAs/InP avalanche photodiode with active quenching for single photon counting at telecom wavelengths. *Appl. Phys. Lett.* **2007**, *91*, 201114.
28. Acerbi, F.; Della Frera, A.; Tosi, A.; Zappa, F. Fast active quenching circuit for reducing avalanche charge and afterpulsing in InGaAs/InP single-photon avalanche diode. *IEEE J. Quantum Electron.* **2013**, *49*, 563–569.
29. Ceccarelli, F.; Acconcia, G.; Gulinatti, A.; Ghioni, M.; Rech, I. Fully integrated active quenching circuit driving custom-technology SPADs with 6.2-ns dead time. *IEEE Photonics Technol. Lett.* **2018**, *31*, 102–105.
30. Lakeh, M.D.; Kammerer, J.-B.; Schell, J.-B.; Issartel, D.; Gao, S.; Cathelin, A.; Dartigues, A.; Calmon, F.; Uhring, W. Integration of an ultra-fast active quenching circuit with a monolithic 3D SPAD pixel in a 28 nm FD-SOI CMOS technology. *Sensors Actuators A Phys.* **2023**, *363*, 114744.
31. Lakeh, M.D.; Kammerer, J.-B.; Aguénounon, E.; Issartel, D.; Schell, J.-B.; Rink, S.; Cathelin, A.; Calmon, F.; Uhring, W. An ultra-fast active quenching active reset circuit with 50% SPAD afterpulsing reduction in a 28 nm FD-SOI CMOS technology using body biasing technique. *Sensors* **2021**, *21*, 4014.
32. Savuskan, V.; Brouk, I.; Javitt, M.; Nemirovsky, Y. An estimation of single photon avalanche diode (SPAD) photon detection efficiency (PDE) nonuniformity. *IEEE Sens. J.* **2013**, *13*, 1637–1640.
33. Panglosse, A.; Martin-Gonthier, P.; Marcelot, O.; Virmontois, C.; Saint-Pé, O.; Magnan, P. Dark count rate modeling in single-photon avalanche diodes. *IEEE Trans. Circuits Syst. I Regul. Pap.* **2020**, *67*, 1507–1515.
34. Ceccarelli, F.; Acconcia, G.; Gulinatti, A.; Ghioni, M.; Rech, I.; Osellame, R. Recent advances and future perspectives of single-photon avalanche diodes for quantum photonics applications. *Adv. Quantum Technol.* **2021**, *4*, 2000102.
35. Haitz, R.H. Mechanisms contributing to the noise pulse rate of avalanche diodes. *J. Appl. Phys.* **1965**, *36*, 3123–3131.
36. Bronzi, D. Front-Side and Back-Side Illuminated Spad Arrays for 2D Imaging and 3D Ranging. Doctoral Dissertation, Politecnico di Milano, Milan, Italy, 2014.
37. Liu, L.; Wang, X. Germanium/Silicon Single Photon Avalanche Diodes for SWIR Applications: A Review. *IEEE Sens. J.* **2025**, *25*, 23499–23513.
38. Hadfield, R.H. Single-photon detectors for optical quantum information applications. *Nat. Photonics* **2009**, *3*, 696–705.
39. Qian, X. Modeling and Design of Single Photon Avalanche Diodes for Lidar Applications. Doctoral Dissertation, University of Electronic Science and Technology of China, Chengdu, China, 2023.
40. Jiang, S.; Safari, M. High-speed free-space QKD in the presence of SPAD dead time. In *Proceedings of the 2022 IEEE International Conference on Communications Workshops (ICC Workshops)*; IEEE: Piscataway, NJ, USA, 2022; pp. 457–462.

41. Hoque, S.; Abdullah, M.A.; Alayed, M.; Elamien, M.; Deen, M.J. Silicon-based Single Photon Avalanche Diode Technologies: Structures, Performance, and Challenges. *IEEE Electron Devices Rev.* **2025**, *2*, 69–97.
42. Piron, F.; Morrison, D.; Yuce, M.R.; Redouté, J.-M. A Review of Single-Photon Avalanche Diode Time-of-Flight Imaging Sensor Arrays. *IEEE Sens. J.* **2021**, *21*, 12654–12666.
43. Baek, S.-H.; Yang, S.-C.; Park, C.-Y.; Park, C.-W.; Cho, S.-B.; Ryu, S.-W. Room temperature quantum key distribution characteristics of low-noise InGaAs/InP single-photon avalanche diode. *J. Korean Phys. Soc.* **2021**, *78*, 634–641.
44. Calandri, N.; Sanzaro, M.; Motta, L.; Savoia, C.; Tosi, A. Optical crosstalk in InGaAs/InP SPAD array: Analysis and reduction with FIB-etched trenches. *IEEE Photonics Technol. Lett.* **2016**, *28*, 1767–1770.
45. Tang, Y.; Wang, R.; Yang, X.; He, T.; Liu, Y.; Zhao, M. High crosstalk suppression in InGaAs/InP single-photon avalanche diode arrays by carrier extraction structure. *Nat. Commun.* **2024**, *15*, 593.
46. Lubin, G.; Tenne, R.; Antolovic, I.M.; Charbon, E.; Bruschini, C.; Oron, D. Quantum correlation measurement with single photon avalanche diode arrays. *Opt. Express* **2019**, *27*, 32863–32882.
47. Piacentini, F.; Avella, A.; Levi, M.P.; Gramegna, M.; Brida, G.; Degiovanni, I.P.; Cohen, E.; Lussana, R.; Villa, F.; Tosi, A. Measuring incompatible observables by exploiting sequential weak values. *Phys. Rev. Lett.* **2016**, *117*, 170402.
48. Tisa, S.; Villa, F.; Giudice, A.; Simmerle, G.; Zappa, F. High-speed quantum random number generation using CMOS photon counting detectors. *IEEE J. Sel. Top. Quantum Electron.* **2014**, *21*, 23–29.
49. Perenzoni, M.; Pancheri, L.; Stoppa, D. Compact SPAD-based pixel architectures for time-resolved image sensors. *Sensors* **2016**, *16*, 745.
50. Shin, D.; Park, B.; Chae, Y.; Yun, I. The effect of a deep virtual guard ring on the device characteristics of silicon single photon avalanche diodes. *IEEE Trans. Electron Devices* **2019**, *66*, 2986–2991.
51. Accarino, C.; Al-Rawhani, M.; Shah, Y.D.; Maneuski, D.; Mitra, S.; Buttar, C.; Cumming, D.R.S. Low noise and high photodetection probability SPAD in 180 nm standard CMOS technology. In *Proceedings of the 2018 IEEE International Symposium on Circuits and Systems (ISCAS)*; IEEE: Piscataway, NJ, USA, 2018; pp. 1–4.
52. Xu, H.; Pancheri, L.; Betta, G.-F.D.; Stoppa, D. Design and characterization of a p+/n-well SPAD array in 150nm CMOS process. *Opt. Express* **2017**, *25*, 12765–12778.
53. Karami, M.A.; Gersbach, M.; Yoon, H.-J.; Charbon, E. A new single-photon avalanche diode in 90 nm standard CMOS technology. *Opt. Express* **2010**, *18*, 22158–22166.
54. Liu, Y.; Fan, R.; Zhao, Y.; Hu, J.; Ma, R.; Zhu, Z. Junction-Optimized SPAD with 50.6% Peak PDP and 0.64 cps/ μm^2 DCR at 2 V Excess Bias Voltage in 130 nm CMOS. *IEEE Electron Device Lett.* **2024**, *45*, 308–311.
55. Oh, J.; Jeong, H.; Lee, M.S.; Kwon, I. Characterization of small single photon avalanche diode fabricated using standard 180 nm CMOS process for digital SiPM. *Nucl. Eng. Technol.* **2024**, *56*, 3076–3083.
56. Niclass, C.; Soga, M. A miniature actively recharged single-photon detector free of afterpulsing effects with 6ns dead time in a 0.18 μm CMOS technology. In *Proceedings of the 2010 IEEE International Electron Devices Meeting (IEDM)*; IEEE: Piscataway, NJ, USA, 2010; pp. 13–14.
57. Karami, M.A.K.; Amiri-Sani, A.A.-S.; Ghormishi, M.H.G. Tunneling in submicron CMOS single-photon avalanche diodes. *Chinese Opt. Lett.* **2014**, *12*, 2501.
58. Stoppa, D.; Mosconi, D.; Pancheri, L.; Gonzo, L. Single-photon avalanche diode CMOS sensor for time-resolved fluorescence measurements. *IEEE Sens. J.* **2009**, *9*, 1084–1090.
59. Lee, M.-J.; Sun, P.; Gandraud, G.; Bruschini, C.; Charbon, E. First near-ultraviolet-and blue-enhanced backside-illuminated single-photon avalanche diode based on standard SOI CMOS technology. *IEEE J. Sel. Top. Quantum Electron.* **2019**, *25*, 3800206.
60. Dautet, H.; Deschamps, P.; Dion, B.; MacGregor, A.D.; MacSween, D.; McIntyre, R.J.; Trottier, C.; Webb, P.P. Photon counting techniques with silicon avalanche photodiodes. *Appl. Opt.* **1993**, *32*, 3894–3900.
61. Rech, I.; Ingargiola, A.; Spinelli, R.; Labanca, I.; Marangoni, S.; Ghioni, M.; Cova, S. A new approach to optical crosstalk modeling in single-photon avalanche diodes. *IEEE Photonics Technol. Lett.* **2008**, *20*, 330–332.
62. Chen, Q.; Wu, S.; Zhang, L.; Fan, W.; Tan, C.S. Simulation of high-efficiency resonant-cavity-enhanced GeSn single-photon avalanche photodiodes for sensing and optical quantum applications. *IEEE Sens. J.* **2021**, *21*, 14789–14798.
63. Jegannathan, G.; Van den Dries, T.; Kuijk, M. Current-assisted SPAD with improved pn junction and enhanced NIR performance. *Sensors* **2020**, *20*, 7105.
64. Park, E.; Ha, W.-Y.; Park, H.-S.; Eom, D.; Choi, H.-S.; Ahn, D.-H.; Choi, W.-Y.; Lee, M.-J. A back-illuminated SPAD fabricated with 40 nm CMOS image sensor technology achieving near 40% PDP at 940 nm. *IEEE J. Sel. Top. Quantum Electron.* **2023**, *30*, 3800207.

65. Ha, W.-Y.; Park, E.; Eom, D.; Park, H.-S.; Chong, D.; Tan, S.S.; Tng, M.; Quek, E.; Bruschini, C.; Charbon, E.; et al. Single-photon avalanche diode fabricated in standard 55 nm bipolar-CMOS-DMOS technology with sub-20 V breakdown voltage. *Opt. Express* **2023**, *31*, 13798–13805.
66. Ha, W.-Y.; Park, E.; Eom, D.; Park, H.-S.; Gramuglia, F.; Keshavarzian, P.; Kizilkan, E.; Bruschini, C.; Chong, D.; Tan, S.S. SPAD developed in 55 nm bipolar-CMOS-DMOS technology achieving near 90% peak PDP. *IEEE J. Sel. Top. Quantum Electron.* **2023**, *30*, 1–10.
67. Itzler, M.A.; Ben-Michael, R.; Hsu, C.-F.; Slomkowski, K.; Tosi, A.; Cova, S.; Zappa, F.; Ispasoiu, R. Single photon avalanche diodes (SPADs) for 1.5 μm photon counting applications. *J. Mod. Opt.* **2007**, *54*, 283–304.
68. Takai, I.; Matsubara, H.; Soga, M.; Ohta, M.; Ogawa, M.; Yamashita, T. Single-photon avalanche diode with enhanced NIR-sensitivity for automotive LIDAR systems. *Sensors* **2016**, *16*, 459.
69. Fang, Y.-Q.; Chen, W.; Ao, T.-H.; Liu, C.; Wang, L.; Gao, X.-J.; Zhang, J.; Pan, J.-W. InGaAs/InP single-photon detectors with 60% detection efficiency at 1550 nm. *Rev. Sci. Instrum.* **2020**, *91*, 083102.
70. Issartel, D.; Gao, S.; Pittet, P.; Cellier, R.; Golanski, D.; Cathelin, A.; Calmon, F. Architecture optimization of SPAD integrated in 28 nm FD-SOI CMOS technology to reduce the DCR. *Solid. State. Electron.* **2022**, *191*, 108297.
71. Wanitzek, M.; Schulze, J.; Oehme, M. Ge-on-Si single-photon avalanche diode using a double mesa structure. *Opt. Lett.* **2024**, *49*, 6345–6348.
72. Tian, Y.; Li, Q.; Ding, W.; Wu, D.; Lin, Z.; Feng, X.; Zhang, H.; Yu, X.; Zhao, Y. High speed and high sensitivity InGaAs/InAlAs single photon avalanche diodes for photon counting communication. *J. Light. Technol.* **2022**, *40*, 5245–5253.
73. Farrell, A.C.; Meng, X.; Ren, D.; Kim, H.; Senanayake, P.; Hsieh, N.Y.; Rong, Z.; Chang, T.-Y.; Azizur-Rahman, K.M.; Huffaker, D.L. InGaAs–GaAs nanowire avalanche photodiodes toward single-photon detection in free-running mode. *Nano Lett.* **2018**, *19*, 582–590.
74. Petticrew, J.D. Design and Device Fabrication of Silicon Single Photon Avalanche Diodes. Doctoral Dissertation, The University of Sheffield, Sheffield, UK, 2020.
75. Zou, K.; Hu, X. Fabrication Development of High-Performance Fractal Superconducting Nanowire Single-Photon Detectors. *IEEE J. Sel. Top. Quantum Electron.* **2024**, *31*, 3801410.
76. Xie, S.; Liu, J.; Zhang, F. An accurate circuit model for the statistical behavior of InP/InGaAs SPAD. *Electronics* **2020**, *9*, 2059.
77. Telesca, F.; Signorelli, F.; Tosi, A. Double zinc diffusion optimization for charge persistence reduction in InGaAs/InP SPADs. *IEEE J. Sel. Top. Quantum Electron.* **2023**, *30*, 3801207.
78. Warburton, R.E.; Intermite, G.; Myronov, M.; Allred, P.; Leadley, D.R.; Gallacher, K.; Paul, D.J.; Pilgrim, N.J.; Lever, L.J.M.; Ikonik, Z. Ge-on-Si single-photon avalanche diode detectors: Design, modeling, fabrication, and characterization at wavelengths 1310 and 1550 nm. *IEEE Trans. Electron Devices* **2013**, *60*, 3807–3813.
79. Vines, P.; Kuzmenko, K.; Kirdoda, J.; Dumas, D.C.S.; Mirza, M.M.; Millar, R.W.; Paul, D.J.; Buller, G.S. High performance planar germanium-on-silicon single-photon avalanche diode detectors. *Nat. Commun.* **2019**, *10*, 1086.
80. Na, N.; Lu, Y.-C.; Liu, Y.-H.; Chen, P.-W.; Lai, Y.-C.; Lin, Y.-R.; Lin, C.-C.; Shia, T.; Cheng, C.-H.; Chen, S.-L. Room temperature operation of germanium–silicon single-photon avalanche diode. *Nature* **2024**, *627*, 295–300.
81. Wu, J.Y.; Chen, C.-E.; Wu, C.-H.; Lin, G.-R. CMOS Design of Ge-on-Si Single-Photon Avalanche Diode with Ultralow Noise and Jitter. *IEEE J. Sel. Top. Quantum Electron.* **2025**, *31*, 3801307.
82. Palubiak, D.P.; Deen, M.J. CMOS SPADs: Design issues and research challenges for detectors, circuits, and arrays. *IEEE J. Sel. Top. Quantum Electron.* **2014**, *20*, 409–426.
83. Wu, M.-L.; Fenoglio, C.A.; Ripicini, E.; Kizilkan, E.; Gramuglia, F.; Morimoto, K.; Paolozzi, L.; Charbon, E.; Bruschini, C.; Keshavarzian, P. CMOS SPADs for High Radiation Environments. In *Proceedings of the 2022 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC)*; IEEE: Piscataway, NJ, USA, 2022.
84. Wu, J.-Y.; Lu, Y.-W.; Liu, M.-H.; Chang, T.-N.; Liu, C.-H. An Analysis of Temperature-Dependent Timing Jitter Factors in the Structural Design of Complementary Metal-Oxide-Semiconductor Single-Photon Avalanche Detectors. *Sensors* **2025**, *25*, 391.
85. Sicre, M.; Roy, D.; Calmon, F. Hot-Carrier Degradation modeling of DCR drift in SPADs. In *Proceedings of the ESSDERC 2023—IEEE 53rd European Solid-State Device Research Conference*; IEEE: Piscataway, NJ, USA, 2023; pp. 61–64.
86. Xun, M.; Li, Y.; Feng, J.; He, C.; Liu, M.; Guo, Q. Effect of proton irradiation on complementary metal oxide semiconductor (CMOS) single-photon avalanche diodes. *Electronics* **2024**, *13*, 224.
87. Jiang, W.; Scott, R.; Deen, M.J. Improved noise performance of CMOS poly gate single-photon avalanche diodes. *IEEE Photonics J.* **2021**, *14*, 6802208.

88. Jouni, A.; Malherbe, V.; Mamdy, B.; Thery, T.; Sicre, M.; Soussan, D.; Lorquet, V.; De Paoli, S.; Belloir, J.-M.; Lalluca, V. Study of proton-induced defects in 40-nm CMOS SPADs. *IEEE Trans. Nucl. Sci.* **2023**, *70*, 1680–1686.
89. Coughlan, C.; Mirza, M.M.A.; Kirdoda, J.; Dumas, D.; Smith, C.; McCarthy, C.; Mowbray, H.; Watson, S.; Millar, R.; Paul, D.J. Variation of sidewall passivation on sub- μm selectively grown Ge-on-Si devices towards single photon avalanche detectors. In *Proceedings of the Optical Components and Materials XX (SPIE)*; SPIE: Bellingham, WA, USA, 2023; pp. 114–120.
90. Villa, F.; Bronzi, D.; Zou, Y.; Scarcella, C.; Boso, G.; Tisa, S.; Tosi, A.; Zappa, F.; Durini, D.; Weyers, S. CMOS SPADs with up to 500 μm diameter and 55% detection efficiency at 420 nm. *J. Mod. Opt.* **2014**, *61*, 102–115.
91. Roy, K.; Ahmed, T.; Dubey, H.; Sai, T.P.; Kashid, R.; Maliakal, S.; Hsieh, K.; Shamim, S.; Ghosh, A. Number-Resolved Single-Photon Detection with Ultralow Noise van der Waals Hybrid. *Adv. Mater.* **2018**, *30*, 1704412.
92. Ren, X.; Fan, Y.; Shi, Y.; Yuan, Z.; Huang, Y.; Zhang, W. Hybrid integration of single-photon avalanche diode array and silicon photonic chip. *Acta Opt. Sin.* **2024**, *1*, 104001.
93. Lei, Q.; Li, L.; Lu, W.; Tao, J.; Ling, R.; Zhang, L.; Chen, X.; Wu, S.; Li, S.; Li, S. Preparation and research progress of GaN-based avalanche photodetectors. *Microstructures* **2025**, *5*, 2025079.
94. Yu, C.; Li, T.; Zhao, X.-S.; Lu, H.; Zhang, R.; Xu, F.; Zhang, J.; Pan, J.-W. Free-running 4H-SiC single-photon detector with ultralow afterpulse probability at 266 nm. *Rev. Sci. Instrum.* **2023**, *94*, 033101.
95. Dou, L.; Yang, Y.; You, J.; Hong, Z.; Chang, W.-H.; Li, G.; Yang, Y. Solution-processed hybrid perovskite photodetectors with high detectivity. *Nat. Commun.* **2014**, *5*, 5404.
96. Sun, Z.; Fang, L.; Shi, F.; Shi, Z.; Wei, T.; Bi, W.; Wang, Y.; Yan, J. Dual-functional GaN photodiode for photodetection and neuromorphic computing with high specific detectivity. *Opt. Lett.* **2024**, *49*, 6821–6824.
97. Chen, J.; Du, H.; Qu, H.; Gao, H.; Gu, Y.; Zhu, Y.; Ye, W.; Zou, J.; Wang, H.; Zou, X. AlGaIn/GaN MOS-HEMT enabled optoelectronic artificial synaptic devices for neuromorphic computing. *APL Mach. Learn.* **2024**, *2*, 026113.
98. Jang, H.; Ju, S.; Lee, S.; Choi, J.; Byun, U.; Min, K.; Rasheed, M.; Kim, S. Recent Advances in Optoelectronic Synaptic Devices for Neuromorphic Computing. *Biomimetics* **2025**, *10*, 584.
99. Cui, D.; Pei, M.; Lin, Z.; Zhang, H.; Kang, M.; Wang, Y.; Gao, X.; Su, J.; Miao, J.; Li, Y.; et al. Versatile optoelectronic memristor based on wide-bandgap Ga_2O_3 for artificial synapses and neuromorphic computing. *Light Sci. Appl.* **2025**, *14*, 161.
100. Sun, H.; Ye, B.; Ge, M.; Gong, B.; Qian, L.; Parkhomenko, I.N.; Komarov, F.F.; Liu, Y.; Yang, G. Artificial optoelectronic synapses based on Ga_2O_3 metal–semiconductor–metal solar-blind ultraviolet photodetectors with asymmetric electrodes for neuromorphic computing. *Responsive Mater.* **2025**, *3*, e20240038.
101. Vasilopoulou, M.; bin Mohd Yusoff, A.R.; Chai, Y.; Kourtis, M.-A.; Matsushima, T.; Gasparini, N.; Du, R.; Gao, F.; Nazeeruddin, M.K.; Anthopoulos, T.D.; et al. Neuromorphic computing based on halide perovskites. *Nat. Electron.* **2023**, *6*, 949–962.
102. Vats, G.; Hodges, B.; Ferguson, A.J.; Wheeler, L.M.; Blackburn, J.L. Optical Memory, Switching, and Neuromorphic Functionality in Metal Halide Perovskite Materials and Devices. *Adv. Mater.* **2023**, *35*, 2205459.
103. Zhou, X.; Hu, F.; Hou, Q.; Hu, J.; Wang, Y.; Chen, X. All-photonic artificial synapses based on photochromic perovskites for noncontact neuromorphic visual perception. *Commun. Mater.* **2024**, *5*, 116.
104. Sun, H.; Wang, H.; Dong, S.; Dai, S.; Li, X.; Zhang, X.; Deng, L.; Liu, K.; Liu, F.; Tan, H.; et al. Optoelectronic synapses based on a triple cation perovskite and Al/MoO₃ interface for neuromorphic information processing. *Nanoscale Adv.* **2024**, *6*, 559–569.
105. Liu, Z.; Cheng, P.; Kang, R.; Zhou, J.; Wang, X.; Zhao, X.; Zhao, J.; Zuo, Z. All-Inorganic CsPbBr₃ Perovskite Planar-Type Memristors as Optoelectronic Synapses. *ACS Appl. Mater. Interfaces* **2024**, *16*, 51065–51079.
106. Lee, S.-U.; Kim, S.-Y.; Lee, J.-H.; Baek, J.H.; Lee, J.-W.; Jang, H.W.; Park, N.-G. Artificial Synapse Based on a δ -FAPbI₃/Atomic-Layer-Deposited SnO₂ Bilayer Memristor. *Nano Lett.* **2024**, *24*, 4869–4876.
107. Chen, C.-Y.; Lin, H.-C.; Jan, P.-E.; Chen, H.-M.; Chuang, Y.-T.; Li, C.-F.; Huang, Y.-C.; Lin, H.-W. Dual-Mode Photonic Synapse Based on a Lead-Free 2D Ruddlesden–Popper Perovskite for Neuromorphic Vision. *ACS Appl. Mater. Interfaces* **2025**, *17*, 48547–48554.
108. Wu, M.-L.; Ripicini, E.; Kizilkan, E.; Gramuglia, F.; Keshavarzian, P.; Fenoglio, C.A.; Morimoto, K.; Charbon, E. Radiation hardness study of single-photon avalanche diode for space and high energy physics applications. *Sensors* **2022**, *22*, 2919.
109. Finazzi, L.; Signorelli, F.; Bagatin, M.; Gerardin, S.; Zahidy, M.; Vallone, G.; Jennewein, T.; Villoresi, P.; Tosi, A. Radiation hardness properties and DCR reduction via laser annealing of InGaAs/InP SPADs for space applications. *Opt. Express* **2025**, *33*, 47876–47889.
110. Pellegrini, S.; Warburton, R.E.; Tan, L.J.J.; Ng, J.S.; Krysa, A.B.; Groom, K.; David, J.P.R.; Cova, S.; Robertson, M.J.; Buller, G.S. Design and performance of an InGaAs-InP single-photon avalanche diode detector. *IEEE J. Quantum Electron.* **2006**, *42*, 397–403.

111. Liu, D.; Li, M.; Xu, T.; Dong, J.; Fang, Y.; Xu, Y. Study of the influence of virtual guard ring width on the performance of SPAD detectors in 180 nm standard CMOS technology. *J. Semicond.* **2023**, *44*, 114102.
112. Jiang, X.; Itzler, M.A.; Ben-Michael, R.; Slomkowski, K. InGaAsP-InP avalanche photodiodes for single photon detection. *IEEE J. Sel. Top. Quantum Electron.* **2007**, *13*, 895–905.
113. Stipčević, M.; Christensen, B.G.; Kwiat, P.G.; Gauthier, D.J. Advanced active quenching circuit for ultra-fast quantum cryptography. *Opt. Express* **2017**, *25*, 21861–21876.
114. Liu, J.; Xu, Y.; Wang, Z.; Li, Y.; Gu, Y.; Liu, Z.; Zhao, X. Reducing afterpulsing in InGaAs (P) single-photon detectors with hybrid quenching. *Sensors* **2020**, *20*, 4384.
115. Lai, Y.; Shen, Z.; Chen, Y.; Wang, J.; Guo, J.; Wei, Z. The Impact of Afterpulsing Effects in Single-Photon Detectors on the Performance Metrics of Single-Photon Detection Systems. *Photonics* **2024**, *11*, 1074.
116. Restelli, A.; Bienfang, C.J.; Migdall, A.L. Time-domain measurements of afterpulsing in InGaAs/InP SPAD gated with sub-nanosecond pulses. *J. Mod. Opt.* **2012**, *59*, 1465–1471.
117. Nolet, F.; Gauthier, V.; Parent, S.; Vachon, F.; Roy, N.; St-Jean, N.; Charlebois, S.A.; Pratte, J.-F. Quenching circuit discriminator architecture impact on a sub-10 ps fwhm single-photon timing resolution spad. *Instruments* **2023**, *7*, 16.
118. Goll, B.; Saadi Nejad, M.; Schneider-Hornstein, K.; Zimmermann, H. Reducing Avalanche Build-Up Time by Integrating a Single-Photon Avalanche Diode with a BiCMOS Gating Circuit. *Sensors* **2024**, *24*, 7598.
119. Fujisaki, Y.; Tsugawa, H.; Sakai, K.; Humagai, H.; Nakamura, R.; Ogita, T.; Endo, S.; Iwase, T.; Takase, H.; Yokochi, K. A back-illuminated 6 μm SPAD depth sensor with PDE 36.5% at 940 nm via combination of dual diffraction structure and 2×2 on-chip lens. In *Proceedings of the 2023 IEEE Symposium on VLSI Technology and Circuits (VLSI Technology and Circuits)*; IEEE: Piscataway, NJ, USA, 2023; pp. 1–2.
120. Nolet, F.; Parent, S.; Roy, N.; Mercier, M.-O.; Charlebois, S.A.; Fontaine, R.; Pratte, J.-F. Quenching Circuit and SPAD Integrated in CMOS 65 nm with 7.8 ps FWHM Single Photon Timing Resolution. *Instruments* **2018**, *2*, 19.
121. Kohne Poushi, S.S.; Mahmoudi, H.; Hofbauer, M.; Steindl, B.; Zimmermann, H. Photon detection probability enhancement using an anti-reflection coating in CMOS-based SPADs. *Appl. Opt.* **2021**, *60*, 7815–7820.
122. Pavia, J.M.; Wolf, M.; Charbon, E. Measurement and modeling of microlenses fabricated on single-photon avalanche diode arrays for fill factor recovery. *Opt. Express* **2014**, *22*, 4202–4213.
123. McCarthy, C.; Smith, C.; Mowbray, H.; Paul, D.J.; Millar, R.W. Ge-on-Si single photon avalanche diode performance enhancement with photonic crystal nano-hole arrays. In *Proceedings of the Physics and Simulation of Optoelectronic Devices XXXII, San Francisco, CA, USA, 30 January–1 February 2024*; SPIE: Bellingham, WA, USA, 2024; Volume 12891, pp. 37–43.
124. Govdeli, A.; Straguzzi, J.N.; Yong, Z.; Lin, Y.; Luo, X.; Chua, H.; Lo, G.-Q.; Sacher, W.D.; Poon, J.K.S. Room-temperature waveguide-coupled silicon single-photon avalanche diodes. *npj Nanophotonics* **2024**, *1*, 2.
125. Yu, C.; Xu, Q.; Zhang, J. Recent advances in InGaAs/InP single-photon detectors. *Meas. Sci. Technol.* **2024**, *35*, 122003.
126. Zhang, J.; Itzler, M.A.; Zbinden, H.; Pan, J.-W. Advances in InGaAs/InP single-photon detector systems for quantum communication. *Light Sci. Appl.* **2015**, *4*, e286.
127. Gulinatti, A.; Ceccarelli, F.; Ghioni, M.; Rech, I. Custom silicon technology for SPAD-arrays with red-enhanced sensitivity and low timing jitter. *Opt. Express* **2021**, *29*, 4559–4581.
128. Giustolisi, G.; Mita, R.; Palumbo, G. Behavioral modeling of statistical phenomena of single-photon avalanche diodes. *Int. J. Circuit Theory Appl.* **2012**, *40*, 661–679.
129. Zhang, B.; Li, Z.; Zaghloul, M.E. A single-photon avalanche diode in CMOS 0.5 μm n-well process. In *Proceedings of the IEEE SENSORS*; IEEE: Piscataway, NJ, USA, 2012; pp. 1–4.
130. Chitnis, D.; Collins, S. A SPAD-based photon detecting system for optical communications. *J. Light. Technol.* **2014**, *32*, 2028–2034.
131. Vornicu, I.; Carmona-Galán, R.; Pérez-Verdú, B.; Rodríguez-Vázquez, Á. Compact CMOS active quenching/recharge circuit for SPAD arrays. *Int. J. Circuit Theory Appl.* **2016**, *44*, 917–928.
132. Zang, K.; Jiang, X.; Huo, Y.; Ding, X.; Morea, M.; Chen, X.; Lu, C.-Y.; Ma, J.; Zhou, M.; Xia, Z. Silicon single-photon avalanche diodes with nano-structured light trapping. *Nat. Commun.* **2017**, *8*, 628.
133. Hofbauer, M.; Steindl, B.; Zimmermann, H. Temperature Dependence of Dark Count Rate and After Pulsing of a Single-Photon Avalanche Diode with an Integrated Active Quenching Circuit in 0.35 μm CMOS. *J. Sensors* **2018**, *2018*, 9585931.
134. Xu, Y.; Lu, J.; Wu, Z. A compact high-speed active quenching and recharging circuit for SPAD detectors. *IEEE Photonics J.* **2020**, *12*, 6803208.
135. Pahlavanzadeh, H.; Karami, M.A. Low-power compact tunable quenching configuration for minimizing afterpulsing in single-photon avalanche diodes. *Int. J. Circuit Theory Appl.* **2020**, *48*, 1534–1543.

136. Goll, B.; Steindl, B.; Zimmermann, H. Avalanche Transients of Thick 0.35 μm CMOS Single-Photon Avalanche Diodes. *Micromachines* **2020**, *11*, 869.
137. Wu, J.Y.; Liu, C.-H. Design and characterization of n/p-well CMOS SPAD with low dark count rate and high photon detection efficiency. *IEEE Trans. Electron Devices* **2022**, *70*, 582–587.
138. Karaca, U.; Kizilkan, E.; Bruschini, C.; Charbon, E. A new double multiplication region method to design high sensitivity and wide spectrum SPADs in standard CMOS technologies. *Sci. Rep.* **2024**, *14*, 27118.
139. Smith, C.K.; McCarthy, C.; Millar, R.W. Simulation of photonic crystal enhanced Ge-on-Si single photon avalanche diodes. *Opt. Express* **2024**, *32*, 38805–38818.
140. Ma, J.; Zhou, M.; Yu, Z.; Jiang, X.; Huo, Y.; Zang, K.; Zhang, J.; Harris, J.S.; Jin, G.; Zhang, Q. Simulation of a high-efficiency and low-jitter nanostructured silicon single-photon avalanche diode. *Optica* **2015**, *2*, 974–979.
141. Ha, W.-Y.; Park, E.; Park, B.; Chae, Y.; Choi, W.-Y.; Lee, M.-J. Noise optimization of single-photon avalanche diodes fabricated in 110 nm CMOS image sensor technology. *Opt. Express* **2022**, *30*, 14958–14965.
142. Rech, I.; Ingargiola, A.; Spinelli, R.; Labanca, I.; Marangoni, S.; Ghioni, M.; Cova, S. Optical crosstalk in single photon avalanche diode arrays: A new complete model. *Opt. Express* **2008**, *16*, 8381–8394.
143. Morimoto, K.; Charbon, E. High fill-factor miniaturized SPAD arrays with a guard-ring-sharing technique. *Opt. Express* **2020**, *28*, 13068–13080.
144. Buchholz, J.; Krieger, J.; Bruschini, C.; Burri, S.; Ardelean, A.; Charbon, E.; Langowski, J. Widefield high frame rate single-photon SPAD imagers for SPIM-FCS. *Biophys. J.* **2018**, *114*, 2455–2464.
145. Tisa, S.; Guerrieri, F.; Zappa, F. Variable-load quenching circuit for single-photon avalanche diodes. *Opt. Express* **2008**, *16*, 2232–2244.
146. Zhang, C.; Lindner, S.; Antolovic, I.M.; Wolf, M.; Charbon, E. A CMOS SPAD imager with collision detection and 128 dynamically reallocating TDCs for single-photon counting and 3D time-of-flight imaging. *Sensors* **2018**, *18*, 4016.
147. Buttafava, M.; Villa, F.; Castello, M.; Tortarolo, G.; Conca, E.; Sanzaro, M.; Piazza, S.; Bianchini, P.; Diaspro, A.; Zappa, F. SPAD-based asynchronous-readout array detectors for image-scanning microscopy. *Optica* **2020**, *7*, 755–765.
148. Nehra, R.; Chang, C.-H.; Yu, Q.; Beling, A.; Pfister, O. Photon-number-resolving segmented detectors based on single-photon avalanche-photodiodes. *Opt. Express* **2020**, *28*, 3660–3675.
149. Krenn, M.; Malik, M.; Scheidl, T.; Ursin, R.; Zeilinger, A. Quantum communication with photons. *Opt. Our Time* **2016**, *18*, 455.
150. Scarani, V.; Bechmann-Pasquinucci, H.; Cerf, N.J.; Dušek, M.; Lütkenhaus, N.; Peev, M. The security of practical quantum key distribution. *Rev. Mod. Phys.* **2009**, *81*, 1301–1350.
151. Pirandola, S.; Andersen, U.L.; Banchi, L.; Berta, M.; Bunandar, D.; Colbeck, R.; Englund, D.; Gehring, T.; Lupo, C.; Ottaviani, C. Advances in quantum cryptography. *Adv. Opt. Photonics* **2020**, *12*, 1012–1236.
152. Djordjevic, I.B. Quantum Key Distribution (QKD) Fundamentals. In *Physical Security, Quantum Key Distribution, and Post-Quantum Cryptography*; Springer: Berlin/Heidelberg, Germany, 2025; pp. 305–362.
153. Sabani, M.; Savvas, I.; Poulakis, D.; Makris, G. Quantum key distribution: Basic protocols and threats. In *Proceedings of the 26th Pan-Hellenic Conference on Informatics*; ACM: New York, NY, USA, 2022; pp. 383–388.
154. Heismann, B.; Kreisler, B.; Fasbender, R. Photon counting CT versus energy-integrating CT: A comparative evaluation of advances in image resolution, noise, and dose efficiency. *Med. Phys.* **2025**, *52*, 1526–1535.
155. Kreisler, B. Photon counting Detectors: Concept, technical Challenges, and clinical outlook. *Eur. J. Radiol.* **2022**, *149*, 110229.
156. Nakamura, Y.; Higaki, T.; Kondo, S.; Kawashita, I.; Takahashi, I.; Awai, K. An introduction to photon-counting detector CT (PCD CT) for radiologists. *Jpn. J. Radiol.* **2023**, *41*, 266–282.
157. Couteau, C.; Barz, S.; Durt, T.; Gerrits, T.; Huwer, J.; Prevedel, R.; Rarity, J.; Shields, A.; Weihs, G. Applications of single photons to quantum communication and computing. *Nat. Rev. Phys.* **2023**, *5*, 326–338.
158. Liu, J.; Peng, Z.; Tan, C.; Yang, L.; Xu, R.; Wang, Z. Emerging single-photon detection technique for high-performance photo-detector. *Front. Phys.* **2024**, *19*, 62502.
159. Singh, R.; Bodile, R.M. A quick guide to quantum communication. *arXiv* **2024**, arXiv:2402.15707.
160. Tong, Q.; Lei, Y.; Shen, X.; Chen, C.; Chen, W.; Zhao, J.; Ren, L.; Cui, D.; Wang, L.; Cai, S. High-performance free-running InGaAs/InP single-photon detector for space laser communications. In *Proceedings of the Tenth Symposium on Novel Optoelectronic Detection Technology and Applications, Beijing, China, 1–3 November 2024*; SPIE: Bellingham, WA, USA, 2025; pp. 1171–1175.
161. Cova, S.; Ghioni, M.; Zappa, F.; Rech, I.; Gulinatti, A.; Maccagnani, P. Silicon single photon avalanche diodes: Situation and prospect. In *Proceedings of the Photon Counting Applications, Quantum Optics, and Quantum Cryptography, Prague, Czech Republic, 18–19 April 2007*; SPIE: Bellingham, WA, USA, 2007; pp. 17–28.

162. Buller, G.S.; Warburton, R.E.; Pellegrini, S.; Ng, J.S.; David, J.P.R.; Tan, L.J.J.; Krysa, A.B.; Cova, S. Single-photon avalanche diode detectors for quantum key distribution. *IET Optoelectron.* **2007**, *1*, 249–254.
163. Martelli, P.; Brunero, M.; Fasiello, A.; Rossi, F.; Tosi, A.; Martinelli, M. Single-SPAD implementation of quantum key distribution. In *Proceedings of the 2019 21st International Conference on Transparent Optical Networks (ICTON)*; IEEE: Piscataway, NJ, USA, 2019; pp. 1–3.
164. Pathak, N.K.; Chaudhary, S.; Sangeeta; Kanseri, B. Phase encoded quantum key distribution up to 380 km in standard telecom grade fiber enabled by baseline error optimization. *Sci. Rep.* **2023**, *13*, 15868.
165. Ribezzo, D.; Zahidy, M.; Lemmi, G.; Petitjean, A.; De Lazzari, C.; Vagniluca, I.; Conca, E.; Tosi, A.; Occhipinti, T.; Oxenløwe, L.K.; et al. Quantum key distribution over 100 km of underwater optical fiber assisted by a fast-gated single-photon detector. *Phys. Rev. Appl.* **2023**, *20*, 044052.
166. Miyajima, S.; Yabuno, M.; Miki, S.; Yamashita, T.; Terai, H. High-time-resolved 64-channel single-flux quantum-based address encoder integrated with a multi-pixel superconducting nanowire single-photon detector. *Opt. Express* **2018**, *26*, 29045–29054.
167. Sax, R.; Boaron, A.; Goso, G.; Atzeni, S.; Crespi, A.; Grünenfelder, F.; Rusca, D.; Al-Saadi, A.; Bronzi, D.; Kupijai, S.; et al. High-speed integrated QKD system. *Photonics Res.* **2023**, *11*, 1007–1014.
168. Sharma, V. Analysis of single photon detectors in differential phase shift quantum key distribution. *arXiv* **2023**, arXiv:2307.03593.
169. Sharma, V.; Bhardwaj, A. Analysis of Differential Phase Shift Quantum Key Distribution using single-photon detectors. In *Proceedings of the 2022 International Conference on Numerical Simulation of Optoelectronic Devices, Turin, Italy, 12–16 September 2022*; IEEE: Piscataway, NJ, USA, 2022; pp. 17–18.
170. Cusini, I.; Berretta, D.; Conca, E.; Incoronato, A.; Madonini, F.; Maurina, A.A.; Nonne, C.; Riccardo, S.; Villa, F. Historical perspectives, state of art and research trends of SPAD arrays and their applications (Part II: SPAD arrays). *Front. Phys.* **2022**, *10*, 906671.
171. Lin, Y.; Mos, P.; Ardelean, A.; Bruschini, C.; Charbon, E. Coupling a recurrent neural network to SPAD TCSPC systems for real-time fluorescence lifetime imaging. *Sci. Rep.* **2024**, *14*, 3286.
172. Gersbach, M.; Trimananda, R.; Maruyama, Y.; Fishburn, M.; Stoppa, D.; Richardson, J.; Walker, R.; Henderson, R.K.; Charbon, E. High frame-rate TCSPC-FLIM using a novel SPAD-based image sensor. *SPIE Detect. Imaging Devices* **2010**, *7780*, 357–369.
173. Shu, H.; Zhang, C.-Y.; Chen, Y.-Q.; Zheng, Z.-J.; Fei, S.-M. Quantum key distribution over noisy channels by the testing state method. *Int. J. Theor. Phys.* **2023**, *62*, 160.
174. Elboukhari, M.; Azizi, M.; Azizi, A. Analysis of the Security of BB84 by Model Checking. *arXiv* **2010**, arXiv:1005.4504.
175. Ma, X.; Yuan, X.; Cao, Z.; Qi, B.; Zhang, Z. Quantum random number generation. *npj Quantum Inf.* **2016**, *2*, 16021.
176. Mannalath, V.; Mishra, S.; Pathak, A. A comprehensive review of quantum random number generators: Concepts, classification and the origin of randomness. *arXiv* **2022**, arXiv:2203.00261.
177. Haase, J.F.; Grollius, S.; Grosse, S.; Buchner, A.; Ligges, M. A 32x24 pixel SPAD detector system for LiDAR and quantum imaging. In *Proceedings of the Photonic Instrumentation Engineering VIII*; SPIE: Bellingham, WA, USA, 2021; Volume 11693, pp. 83–93.
178. Wang, Y.; Wang, C.; Yang, X.; Zhao, L.; Hou, A.; Hu, Y.; Li, F. GpDiff-QRNG: An improved quantum random number generator for lidar interference suppression. *IEEE Sens. J.* **2024**, *19*, 30215–30226.
179. Wang, X.; Yang, X.; Pan, H.; Guo, Y.; Zuo, L.; Chen, Y. Optimized Quantum Random Number Generator for Single Photon LiDAR Interference Suppression. *IEEE Photonics Technol. Lett.* **2025**, *37*, 513–516.
180. Bisadi, Z. All-Silicon-Based Photonic Quantum Random Number Generators. Doctoral Dissertation, University of Trento, Trento, Italy, 2017.
181. Aldama, J.; Sarmiento, S.; Grande, I.H.L.; Signorini, S.; Vidarte, L.T.; Pruneri, V. Integrated QKD and QRNG photonic technologies. *J. Lightwave Technol.* **2022**, *40*, 7498–7517.
182. Tontini, A.; Gasparini, L.; Massari, N.; Passerone, R. SPAD-Based Quantum Random Number Generator with an N^{th} -Order Rank Algorithm on FPGA. *IEEE Trans. Circuits Syst. II Express Briefs* **2019**, *66*, 2067–2071.
183. Gramuglia, F.; Wu, M.-L.; Bruschini, C.; Lee, M.-J.; Charbon, E. A low-noise CMOS SPAD pixel with 12.1 ps SPTR and 3 ns dead time. *IEEE J. Sel. Top. Quantum Electron.* **2021**, *28*, 3800809.
184. Mahmoudi, H.; Zimmermann, H. Bit error performance of APD and SPAD receivers in optical wireless communication. *Electronics* **2021**, *10*, 2731.
185. Liu, J.; Jiang, W.; Kumar, S.; Deen, M.J. Time-controlled SPAD receivers in optical wireless communication system. *IEEE Photonics J.* **2023**, *15*, 7304113.

186. Huang, S.; Chen, C.; Bian, R.; Haas, H.; Safari, M. 5 Gbps optical wireless communication using commercial SPAD array receivers. *Opt. Lett.* **2022**, *47*, 2294–2297.
187. Li, Y.; Safari, M.; Henderson, R.; Haas, H. Optical OFDM with single-photon avalanche diode. *IEEE Photonics Technol. Lett.* **2015**, *27*, 943–946.
188. Almer, O.; Tsonev, D.; Dutton, N.A.W.; Al Abbas, T.; Videv, S.; Gneccchi, S.; Haas, H.; Henderson, R.K. A SPAD-based visible light communications receiver employing higher order modulation. In *Proceedings of the 2015 IEEE Global Communications Conference (GLOBECOM)*; IEEE: Piscataway, NJ, USA, 2015; pp. 1–6.
189. Bae, S.; Koh, S.-T. Optical Link Design for Quantum Key Distribution-Integrated Optical Access Networks. *Photonics* **2025**, *12*, 418.
190. Walton, A.; Ghesquière, A.; Varcoe, B.T.H. Quantum Key Distribution with Displaced Thermal States. *Entropy* **2024**, *26*, 488.
191. Freude, W.; Schmogrow, R.; Nebendahl, B.; Winter, M.; Josten, A.; Hillerkuss, D.; Koenig, S.; Meyer, J.; Dreschmann, M.; Huebner, M. Quality metrics for optical signals: Eye diagram, Q-factor, OSNR, EVM and BER. In *Proceedings of the 2012 14th International Conference on Transparent Optical Networks*; IEEE: Piscataway, NJ, USA, 2012; pp. 1–4.
192. Shafik, R.A.; Rahman, M.S.; Islam, A.H.M.R. On the extended relationships among EVM, BER and SNR as performance metrics. In *Proceedings of the 2006 International Conference on Electrical and Computer Engineering, Dhaka, Bangladesh, 19–21 December 2006*; IEEE: Piscataway, NJ, USA, 2006; pp. 408–411.
193. Meghdadi, V. *BER Calculation*; Université de Limoges: Limoges, France, 2008; pp. 1–9.
194. You, L. Superconducting nanowire single-photon detectors for quantum information. *Nanophotonics* **2020**, *9*, 2673–2692.
195. Gonzalez-Zalba, M.F.; de Franceschi, S.; Charbon, E.; Meunier, T.; Vinet, M.; Dzurak, A.S. Scaling silicon-based quantum computing using CMOS technology. *Nat. Electron.* **2021**, *4*, 872–884.
196. Schuhmann, J.; Lazzari, L.; Morassi, M.; Lemaître, A.; Sagnes, I.; Beaudoin, G.; Amanti, M.I.; Boeuf, F.; Raineri, F.; Baboux, F. Hybrid III-V/silicon quantum photonic device generating broadband entangled photon pairs. *PRX Quantum* **2024**, *5*, 040321.
197. Wen, P.; Tiwari, P.; Mauthe, S.; Schmid, H.; Sousa, M.; Scherrer, M.; Baumann, M.; Bitachon, B.I.; Leuthold, J.; Gotsmann, B. Waveguide coupled III-V photodiodes monolithically integrated on Si. *Nat. Commun.* **2022**, *13*, 909.
198. Yuan, Y.; Tossoun, B.; Huang, Z.; Zeng, X.; Kurczveil, G.; Fiorentino, M.; Liang, D.; Beausoleil, R.G. Avalanche photodiodes on silicon photonics. *J. Semicond.* **2022**, *43*, 021301.
199. Tanabe, K.; Watanabe, K.; Arakawa, Y. III-V/Si hybrid photonic devices by direct fusion bonding. *Sci. Rep.* **2012**, *2*, 349.
200. Tang, M.; Park, J.-S.; Wang, Z.; Chen, S.; Jurczak, P.; Seeds, A.; Liu, H. Integration of III-V lasers on Si for Si photonics. *Prog. Quantum Electron.* **2019**, *66*, 1–18.
201. Huang, Y.-H.; Chen, C.-Y.; Chiang, Y.-T.; Wu, C.-H.; Chang, S.-J.; Lin, S.-Y. Two-Dimensional Material Photodetectors: High Responsivities and Short Response Times of Graphene/Multilayer MoS₂ Heterostructures. *ACS Appl. Electron. Mater.* **2025**, *7*, 3947–3954.
202. Malik, M.A.; Iqbal, M.A.; Choi, J.R.; Pham, P.V. 2D materials for efficient photodetection: Overview, mechanisms, performance and UV-IR range applications. *Front. Chem.* **2022**, *10*, 905404.
203. Roy, K.; Padmanabhan, M.; Goswami, S.; Sai, T.P.; Ramalingam, G.; Raghavan, S.; Ghosh, A. Graphene–MoS₂ hybrid structures for multifunctional photoresponsive memory devices. *Nat. Nanotechnol.* **2013**, *8*, 826–830.
204. Liu, C.; Guo, J.; Yu, L.; Li, J.; Zhang, M.; Li, H.; Shi, Y.; Dai, D. Silicon/2D-material photodetectors: From near-infrared to mid-infrared. *Light Sci. Appl.* **2021**, *10*, 123.
205. Dolphin, J.A.; Scowen, R.O.E.; Wells, L.M.; Ellis, D.J.P.; Lowe, A.L.; Ramsay, B.; Davies, J.I.; Shields, A.J.; Paraiso, T.K.; Stevenson, R.M. Hybrid-integrated InGaAs/InP SPAD arrays for quantum communications. *Opt. Quantum* **2025**, *4*, 7–13.
206. Peyskens, F.; Chakraborty, C.; Muneeb, M.; Van Thourhout, D.; Englund, D. Integration of single photon emitters in 2D layered materials with a silicon nitride photonic chip. *Nat. Commun.* **2019**, *10*, 4435.
207. Pelgrin, V.; Yoon, H.H.; Cassan, E.; Sun, Z. Hybrid integration of 2D materials for on-chip nonlinear photonics. *Light Adv. Manuf.* **2023**, *4*, 311–333.
208. Acerbi, F.; Bernard, M.; Goll, B.; Gola, A.; Zimmermann, H.; Pucker, G.; Ghulinyan, M. Monolithically integrated SiON photonic circuit and silicon single-photon detectors for NIR-range operation. *J. Lightwave Technol.* **2024**, *42*, 2831–2841.
209. Lin, Y.; Yong, Z.; Luo, X.; Azadeh, S.S.; Mikkelsen, J.C.; Sharma, A.; Chen, H.; Mak, J.C.C.; Lo, P.G.-Q.; Sacher, W.D.; et al. Monolithically integrated, broadband, high-efficiency silicon nitride-on-silicon waveguide photodetectors in a visible-light integrated photonics platform. *Nat. Commun.* **2022**, *13*, 6362.
210. Yanikgonul, S.; Leong, V.; Ong, J.R.; Png, C.E.; Krivitsky, L. Simulation of silicon waveguide single-photon avalanche detectors for integrated quantum photonics. *IEEE J. Sel. Top. Quantum Electron.* **2019**, *26*, 6300108.

211. Acerbi, F.; Bernard, M.; Gola, A.; Pucker, G.; Ghulinyan, M. Silicon SPAD monolithically integrated with SiON-based photonic circuit. *Phys. Rev. Lett.* **2019**, *123*, 250503.
212. Incoronato, A.; Severini, F.; Villa, F.; Zappa, F. Multi-channel SPAD chip for silicon photonics with multi-photon coincidence detection. *IEEE J. Sel. Top. Quantum Electron.* **2021**, *28*, 3803607.
213. Trenti, A.; Achleitner, M.; Prawits, F.; Schrenk, B.; Conradi, H.; Kleinert, M.; Incoronato, A.; Zanetto, F.; Zappa, F.; Di Luch, I.; et al. On-Chip Quantum Communication Devices. *J. Lightwave Technol.* **2022**, *40*, 7485–7497.
214. Gupta, R.; Kumar, A.; Kumar, M.; Singh, R.; Gehlot, A.; Pandey, S.S.; Yadav, N.; Pandey, K.; Yadav, A.; Gupta, N. The integration of microelectronic and photonic circuits on a single silicon chip for high-speed and low-power optoelectronic technology. *Nano Mater. Sci.* **2025**, *7*, 305–315.
215. Ning, S.; Zhu, H.; Feng, C.; Gu, J.; Jiang, Z.; Ying, Z.; Midkiff, J.; Jain, S.; Hlaing, M.H.; Pan, D.Z.; et al. Photonic-electronic integrated circuits for high-performance computing and ai accelerators. *J. Lightwave Technol.* **2024**, *42*, 7834–7859.
216. Madonini, F.; Severini, F.; Zappa, F.; Villa, F. Single photon avalanche diode arrays for quantum imaging and microscopy. *Adv. Quantum Technol.* **2021**, *4*, 2100005.
217. Shu, H. Solve single photon detector problems. *Quantum* **2023**, *7*, 1187.
218. Incoronato, A.; Severini, F.; Madonini, F.; Villa, F.; Zappa, F. Linear SPAD array for quantum communication. In *Proceedings of the Quantum Optics and Photon Counting 2021*; SPIE: Bellingham, WA, USA, 2021; Volume 11771, p. 117710A.
219. Diamanti, E.; Lo, H.-K.; Qi, B.; Yuan, Z. Practical challenges in quantum key distribution. *npj Quantum Inf.* **2016**, *2*, 16025.
220. Manzalini, A. Quantum communications in future networks and services. *Quantum Rep.* **2020**, *2*, 221–232.

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.