

Silicon Photomultipliers for Scintillation Readout: Device Physics, Noise Mechanisms, and Temperature Stabilization

Raja Babu, Ashim Roy*

ORCID: 0000-0002-5622-4260

University Department of Physics, Bhupendra Narayan Mandal University, Madhepura, India - 852113

Corresponding author: Ashim Roy

DOI: <https://doi.org/10.51244/IJRSI.2026.1308000250>

Received: 03 September 2026; Accepted: 08 September 2026; Published: 18 September 2026

ABSTRACT

The silicon photomultiplier (SiPM) has displaced the vacuum photomultiplier tube (PMT) in most new scintillation-readout designs, combining single-photon sensitivity with a bias voltage of a few tens of volts, a compact solid-state package, and immunity to magnetic fields. This review treats three coupled aspects of SiPM behaviour that determine what a scintillation detector built around such a device can actually achieve: the physics of the Geiger-mode single-photon avalanche diode (SPAD) microcell and of the parallel microcell array, which fixes gain, photon detection efficiency (PDE), dynamic range and timing; the three correlated- and uncorrelated-noise mechanisms — dark counts, optical crosstalk and afterpulsing — that set an excess-noise floor independent of the coupled scintillator; and the temperature dependence that follows from the temperature coefficient of the breakdown voltage. Analytical relations for gain, PDE, saturation, dark count rate, excess noise factor and the bias-compensation law are stated explicitly, with symbols defined and internal numerical consistency checked. Published temperature-stabilization results are compared quantitatively across five distinct strategies, showing that photopeak or gain drift of 13–22 % per 10 °C can be reduced to 0.5–2 % by active bias correction. Remaining disagreements in the literature — chiefly the non-standardized reporting of correlated-noise metrics and the fact that bias compensation stabilizes overvoltage but not dark count rate or crosstalk — are identified as open problems. The synthesis is directed at the design of a prototype SiPM-coupled scintillation detector and complements companion reviews of scintillator materials and of system-level detector characterization prepared within the same research programme.

Keywords: silicon photomultiplier; single-photon avalanche diode; dark count rate; optical crosstalk; excess noise factor; breakdown-voltage temperature coefficient; bias compensation

1. INTRODUCTION

For most of the twentieth century the vacuum PMT — a photocathode followed by a dynode chain operated at several hundred to more than a thousand volts — was the default light sensor for scintillation detectors, offering high internal gain with single-photon sensitivity [1]. The SiPM, first demonstrated in its modern multi-pixel form by Buzhan and co-workers [2] and matured through two decades of device engineering [3],[4], is now the principal solid-state alternative. It reaches PMT-like gain (10^5 – 10^6) at a bias of only 24–60 V, in a mechanically robust package that is insensitive to magnetic fields [5],[6]. These properties drove the replacement of PMTs by SiPMs in commercial time-of-flight positron emission tomography (TOF-PET) scanners [7],[8] and their adoption in calorimetry, LIDAR, astroparticle instruments and portable radiation spectrometry [9].

Translating this potential into a working detector requires quantitative control of three linked issues. The first is device physics: how the SPAD microcell operates, and how the parallel array of microcells fixes gain, PDE, dynamic range and timing [1],[5]. The second is noise: thermally and tunnelling-generated dark counts, optical crosstalk between microcells, and afterpulsing from trapped carriers add excess fluctuations absent from an ideal photon counter [10],[11],[12], and the measurement conventions used to quantify them differ substantially

between laboratories [13]. The third is temperature: because the breakdown voltage itself is temperature dependent, essentially every parameter of interest drifts with ambient temperature, and an uncompensated detector can show photopeak shifts exceeding 10 % per 10 °C [14],[15],[16].

This paper reviews the three in turn. It states the governing relations explicitly rather than descriptively, checks their numerical consistency against reported measurements, and compares published temperature-stabilization strategies on a common quantitative basis. Original block diagrams of the SPAD-array equivalent circuit and of a complete energy/timing readout chain are presented; comparative plots of gain, PDE, dark count rate, excess noise factor and breakdown voltage are reconstructed from the stated equations and from coefficients reported in the cited literature, and are labelled as such; one figure reproduces a published experimental comparison. Content is classified throughout as measured (reported experimental values), reconstructed (computed here from a stated equation and literature coefficients), or indicative (representative ranges compiled from multiple sources and datasheets, to be re-verified against the datasheet of any specific device).

2. From PMT to SiPM: Comparative Device Landscape

Buzhan, Dolgoshein and co-workers reported the first modern multi-pixel SiPM: roughly 10^3 photon microcounters per mm^2 on a common silicon substrate, each operated in a limited Geiger mode, together delivering single-photoelectron gain of order 10^6 and a PDE comparable to a vacuum PMT [2]. Renker, and later Renker and Lorenz, catalogued the trade-offs between linear-mode avalanche photodiodes (APDs), individually operated SPADs, and the parallel SPAD arrays that define the SiPM [3],[4]; Acerbi and Gundacker and Piemonte and Gola later consolidated the device-level understanding of modern commercial technologies [5],[6].

Table 1 summarizes the comparison across parameters relevant to scintillation readout. The decisive entries for a portable instrument are the bias voltage, the magnetic-field immunity and the mechanical form factor; the decisive penalty is the last row, since the correlated noise of a SiPM has no counterpart in a PMT and cannot be removed by cooling alone.

Table 1: Comparison of PMTs, linear-mode APDs and SiPMs across parameters relevant to scintillation readout. Values are indicative ranges compiled from Refs. [1]–[6]; single-photon timing figures refer to small-area devices at high overvoltage.

Parameter	PMT	Linear-mode APD	SiPM
Typical bias voltage	500 – 1500 V	100 – 500 V	24 – 60 V
Internal gain	$10^5 - 10^7$	10 – 200	$10^5 - 10^6$
Single-photon sensitivity	Yes	No (external amplification required)	Yes
Magnetic-field immunity	No	Yes	Yes
Single-photon time resolution (FWHM)	~100 – 300 ps	Not applicable	$\lesssim 100$ ps to a few 100 ps
Physical form	Bulky, fragile vacuum envelope	Compact, solid-state	Very compact, solid-state
Uncorrelated dark noise	Low (thermionic)	Low	Moderate to high (DCR)
Correlated excess noise	Negligible	Negligible	Optical crosstalk + afterpulsing

Microcell pitch is the single most consequential selection parameter, because it simultaneously fixes microcell capacitance (hence gain), fill factor (hence PDE), and the per-unit-area density of dark and correlated noise. Table 2 gives representative parameters by pitch class. The gain column is calculated from Eq. (1) using the quoted microcell capacitances, so that Table 2 and Figure 2 are mutually consistent; the remaining columns are indicative ranges compiled from Refs. [1],[5],[6],[17] and commercial datasheets.

Table 2: Representative operating parameters for commercially available analog SiPMs, grouped by microcell pitch.

Pitch class	Typical bias	Microcell capacitance (C_d)	Gain at $V_{ov}=3-5V$	Peak PDE	DCR at 25 °C per mm^2	Crosstalk probability
Small ($\approx 20-25 \mu m$)	24–38 V	≈ 40 fF	$0.75 - 1.25 \times 10^6$	35–45 %	50–150 kHz	5–10 %
Medium ($\approx 40-50 \mu m$)	38–55 V	≈ 90 fF	$1.7 - 2.8 \times 10^6$	40–50 %	100–400 kHz	10–20 %
Large ($\approx 75 \mu m$)	45–60 V	≈ 160 fF	$3.0 - 5.0 \times 10^6$	45–55 %	300–900 kHz	20–30 %

The trend across Table 2 is monotonic and unavoidable: gain, PDE, DCR and crosstalk all increase together with pitch and with overvoltage. There is therefore no pitch that is “best” in the abstract; the optimum depends on whether the application is photon-starved (favouring large cells and high PDE), rate- or dynamic-range-limited (favouring small cells and a large cell count), or resolution-limited (favouring low correlated noise).

3. SPAD and SiPM Device Physics

3.1 Geiger-mode avalanche and the microcell equivalent circuit

Each microcell is a p–n junction reverse-biased above its breakdown voltage V_{bd} , so that a single photo-generated carrier entering the high-field depletion region can trigger a self-sustaining avalanche by impact ionization [5],[6]. The avalanche is quenched — passively, by a series resistor R_q of order hundreds of kilo-ohms, or actively by dedicated circuitry — which drops the voltage across the junction below V_{bd} ; the cell then recharges through R_q with recovery time constant $\tau_{rec} \approx R_q C_d$, where C_d is the microcell junction capacitance [5],[17]. Figure 1 shows the resulting equivalent circuit.

Because every fired cell discharges through nominally identical C_d and R_q , each avalanche delivers a highly uniform unit charge, and the parallel connection of all cells to a common node makes the summed current proportional to the number of cells fired within the integration window — but only in the regime treated in Section 3.4. The uniformity of the unit charge, not the array size, is what produces the resolved single-photoelectron spectrum that makes SiPM self-calibration possible [13].

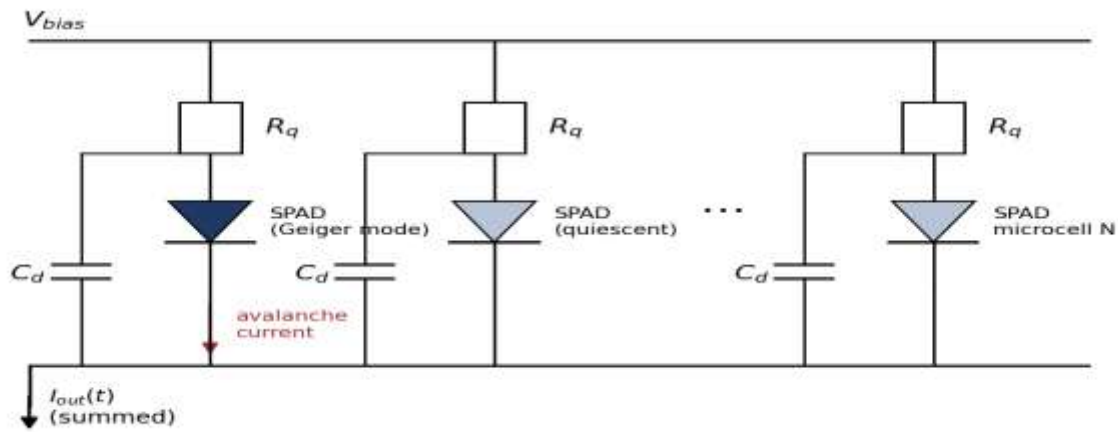


Figure 1: (Original schematic.) Equivalent circuit of an analog SiPM: an array of SPAD microcells, each in series with a quench resistor R_q and with junction capacitance C_d , biased above breakdown and connected in parallel to a common summing node.

3.2 Gain and overvoltage

The gain — elementary charges released per avalanche — is, to good approximation,

$$G = \frac{C_d V_{ov}}{e} \quad (1)$$

where $V_{ov} = V_{bias} - V_{bd}$ is the overvoltage, e the elementary charge, and C_d the microcell capacitance. Gain is thus linear in both overvoltage and cell capacitance and, for a given fabrication technology, increases with cell area. Figure 2 evaluates Eq. (1) for the three capacitances of Table 2.

Equation (1) is also the origin of the temperature sensitivity discussed in Section 5: differentiating with $V_{bd}(T)$ from Eq. (9) gives the fractional gain drift

$$\frac{1}{G} \frac{dG}{dT} = - \frac{\beta}{V_{ov}} \quad (2)$$

with $\beta \equiv dV_{bd}/dT$. For $\beta = 40 \text{ mV}/^\circ\text{C}$ and $V_{ov} = 3 \text{ V}$ this is $-1.3 \text{ }^\circ\text{C}$, i.e. $-13 \text{ }^\circ\text{C}$ from gain alone — the same order as the measured photopeak shift discussed in Section 5.2, which additionally includes the (weaker) temperature dependence of PDE.

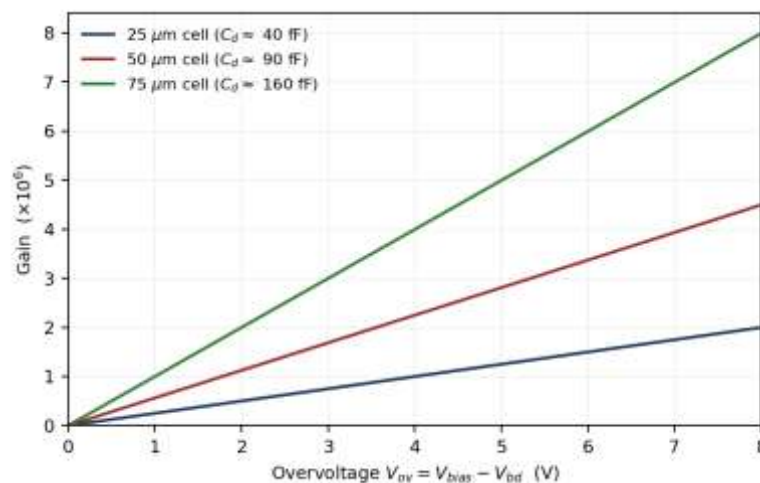


Figure 2: (Reconstructed.) SiPM gain versus overvoltage, computed from Eq. (1) for $C_d = 40, 90$ and 160 fF , representing $\approx 25 \text{ } \mu\text{m}$, $\approx 50 \text{ } \mu\text{m}$ and $\approx 75 \text{ } \mu\text{m}$ microcell pitch. Capacitances compiled from Refs. [5],[6],[17].

3.3 Photon detection efficiency and spectral response

PDE is the product of three independent factors,

$$\text{PDE}(\lambda, V_{\text{ov}}) = \varepsilon_{\text{geo}} \cdot \text{QE}(\lambda) \cdot P_{\text{trig}}(\lambda, V_{\text{ov}}) \quad (3)$$

where ε_{geo} is the geometric fill factor (the photosensitive fraction of the cell area, reduced by inter-cell trenches, quench resistors and metal routing), $\text{QE}(\lambda)$ the quantum efficiency of silicon at wavelength λ , and P_{trig} the probability that an absorbed photon's carrier initiates a self-sustaining avalanche [1],[5]. Only P_{trig} depends on bias; it rises steeply with V_{ov} and saturates, so PDE saturates at high overvoltage while gain, DCR and crosstalk continue to grow. This asymmetry is the reason an optimum operating overvoltage exists at all, and it should be located experimentally for the chosen device rather than assumed.

P_{trig} depends on wavelength through the absorption depth, so the spectral response reflects the junction depth profile. Standard blue-green-optimized SiPMs peak near 450–480 nm; near-ultraviolet-optimized (NUV-HD) devices, developed for fast UV-emitting scintillators, shift the peak toward ≈ 380 –420 nm at the cost of reduced response beyond ≈ 550 nm [1],[19]. Figure 3 illustrates this contrast. Device choice must therefore be made jointly with the scintillator emission spectrum treated in the companion materials review, using the overlap integral of Eq. (3) with the emission spectrum rather than the peak PDE value quoted on a datasheet.

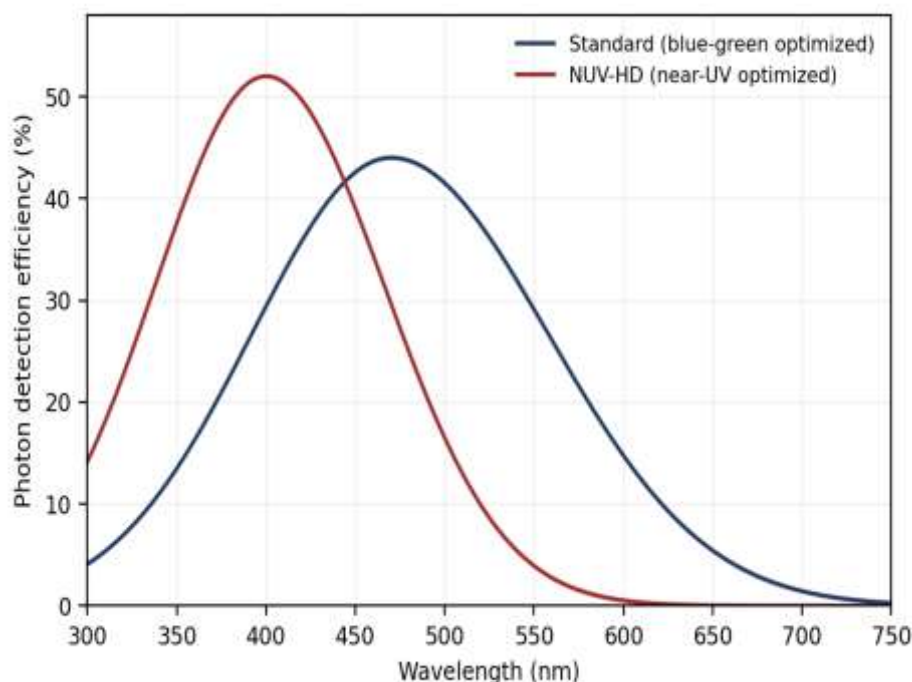


Figure 3: (Reconstructed, qualitative.) Representative PDE spectra for a standard blue-green-optimized SiPM and for a near-ultraviolet-optimized (NUV-HD) device, illustrating the spectral trade-off described in Refs. [1],[19]. Curve shapes are schematic; absolute values are indicative.

3.4 Dynamic range and saturation

A SiPM has a finite number of cells N_{cell} , and a cell that has already fired within the scintillation pulse cannot contribute again until it recovers. If the scintillation decay time is short compared with τ_{rec} , the number of fired cells is

$$N_{\text{fired}} = N_{\text{cell}} \left[1 - \exp\left(-\frac{N_{\text{ph}} \cdot \text{PDE}}{N_{\text{cell}}}\right) \right] \quad (4)$$

with N_{ph} the number of scintillation photons incident on the sensor. Equation (4) is the dominant source of energy non-linearity in SiPM-based spectroscopy and is frequently the limiting factor at high energy deposit. As a

worked example, a $3 \times 3 \text{ mm}^2$ device of $50 \text{ }\mu\text{m}$ pitch ($N_{\text{cell}} \approx 3600$) receiving 1000 detected photons fires ≈ 873 cells — 13 % below the linear expectation. Small-pitch devices, which have four to nine times more cells for the same area, are correspondingly more linear, which partially offsets their lower PDE and is the principal counter-argument to the “larger cells are better” reading of Table 2. When the scintillator decay time is comparable to or longer than τ_{rec} , partial recovery restores cells during the pulse and Eq. (4) becomes a lower bound; a scintillator-specific correction is then required [1],[5].

3.5 Readout architecture and timing

A practical SiPM couples the microcell array to front-end electronics that split the summed current into an energy channel — a charge-integrating or shaping amplifier feeding an ADC — and a timing channel — a constant-fraction or leading-edge discriminator feeding a time-to-digital converter. The split exploits the fact that the leading edge of the summed pulse carries the timing information while the integrated charge carries the energy information [1],[18]. Figure 4 shows this chain, including the temperature-sensing bias-control loop of Section 5.

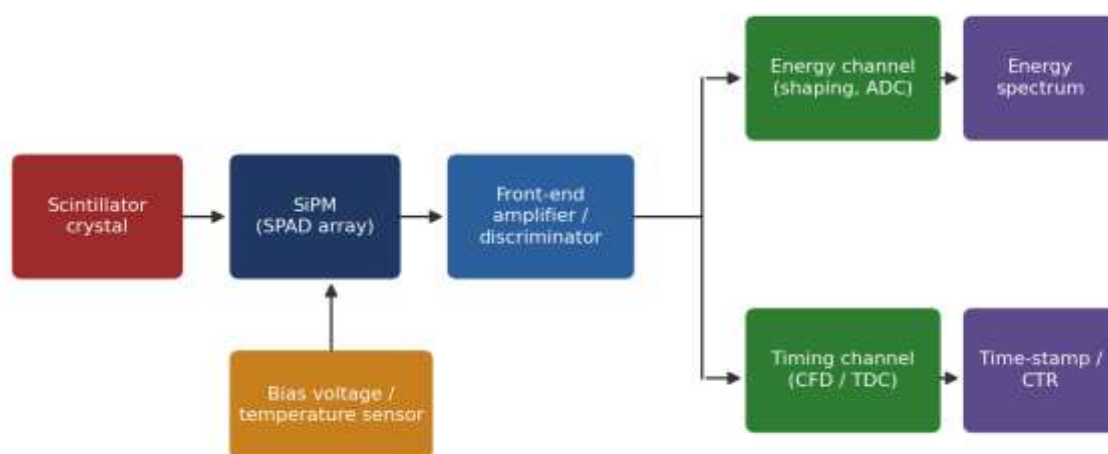


Figure 4: (Original schematic.) Signal chain from scintillation event to energy and timing outputs, showing the parallel energy and timing channels and the temperature-compensated bias loop that stabilizes overvoltage.

Time resolution is governed by photon statistics at the leading edge. For a scintillator of rise time τ_r and decay time τ_d read out with N_{pe} detected photoelectrons, the single-detector time jitter scales as

$$\sigma_t \propto \sqrt{\frac{\tau_r \tau_d}{N_{\text{pe}}}} \quad (5)$$

to which the sensor adds its intrinsic single-photon time resolution (SPTR), set by avalanche build-up and lateral propagation statistics and by the front-end signal-to-noise ratio [18],[20]. Equation (5) makes explicit why timing improves with overvoltage — through both PDE in N_{pe} and SPTR — and why timing and noise optimization pull in the same direction, in contrast to energy resolution, which is degraded by the correlated noise that high overvoltage also produces. Experimental studies of coincidence time resolution across scintillator families, including the contribution of prompt Cherenkov emission, quantify how close modern devices approach the limit implied by Eq. (5) [21], and analytical treatments show that crystal geometry and light-transport delays impose a further floor independent of the photosensor [18].

4. Noise Mechanisms in SiPMs

4.1 Dark count rate

A dark count is an avalanche triggered by a thermally generated or field-assisted tunnelling carrier rather than by an absorbed photon; pulse by pulse it is indistinguishable from a genuine single-photon signal [5],[13]. The generation-dominated contribution follows an Arrhenius form,

$$\text{DCR}(T) \propto T^{3/2} \cdot \exp\left(-\frac{E_a}{k_B T}\right) \quad (6)$$

with activation energy $E_a \approx E_g/2 \approx 0.56$ eV for Shockley–Read–Hall generation through mid-gap defects, k_B the Boltzmann constant and T the absolute temperature. Evaluating $d(\ln \text{DCR})/dT$ from Eq. (6) at $T = 300$ K gives 0.077 K^{-1} , i.e. a factor-of-two change every 9.0°C — which is precisely the “halving per $8\text{--}9^\circ \text{C}$ ” rule of thumb quoted in the review and characterization literature [1],[15]. Equation (6) and that rule are therefore not independent statements, and the agreement is a consistency check rather than corroboration. The tunnelling contribution, significant at high field and at low temperature, has a much weaker temperature dependence, which is why the halving rule breaks down below roughly -40°C [15]. DCR also grows with overvoltage, since the field raises both the generation rate and P_{trig} .

Figure 5 reconstructs the temperature dependence over the range relevant to a field-deployed instrument. In a scintillation event lasting a few tens of nanoseconds, the dark contribution from a single small-area channel is modest, but it scales with the summed area of a multi-channel array and with the integration window, and it sets the practical low-energy threshold. Cooling, area minimization and careful threshold-setting are the available remedies.

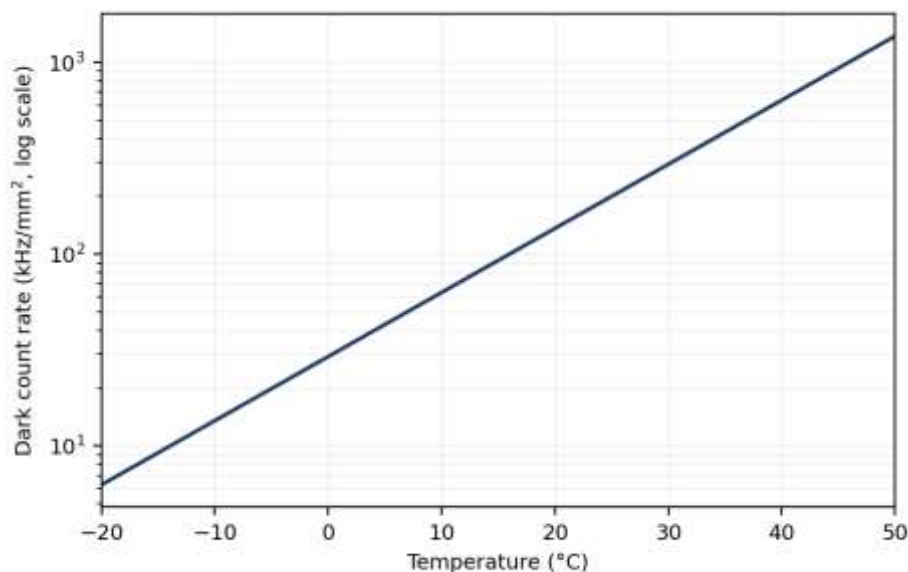


Figure 5: (Reconstructed.) Dark count rate versus temperature, computed from the factor-of-two-per- 9°C scaling implied by Eq. (6) and normalized to a representative room-temperature value of 200 kHz/mm^2 at 25°C . Behaviour consistent with Refs. [1],[15].

4.2 Optical crosstalk

Hot-carrier luminescence during an avalanche emits a small number of secondary photons — of order 10^{-5} photons per carrier crossing the junction, hence a few tens of photons per avalanche at $G \approx 10^6$ [1],[10]. A fraction of these reach a neighbouring cell and trigger it, either within the same die (internal crosstalk) or, after reflection from the package or from an adjacent coupled device, across devices (external crosstalk) [10],[11],[22]. Crosstalk therefore breaks the proportionality between detected photons and fired cells in a way

that no threshold setting can recover, because the secondary avalanche is prompt and indistinguishable from a coincident photon.

Vinogradov modelled the cascade as a branching process, giving a Borel distribution for the total number of cells fired per primary and closed-form expressions for the resulting excess noise factor [10]. Writing λ for the mean number of direct secondaries per avalanche, the Borel distribution has mean $1/(1 - \lambda)$ and variance $\lambda/(1 - \lambda)^3$, so that

$$\text{ENF} = 1 + \frac{\text{Var}}{\text{Mean}^2} = \frac{1}{1 - \lambda}, \quad \lambda = -\ln(1 - P_{\text{XT}}) \quad (7)$$

where P_{XT} is the experimentally reported crosstalk probability, i.e. the probability that a primary avalanche produces at least one direct secondary. Equation (7) gives $\text{ENF} = 1.054, 1.118, 1.287$ and 1.554 for $P_{\text{XT}} = 5, 10, 20$ and 30% respectively; Figure 6 plots it continuously. Gallego, Rosado, Blanco and Arqueros developed a related model validated against measured single- and multi-photoelectron spectra [11]. Because the excess noise enters the energy resolution as

$$\left(\frac{\Delta E}{E}\right)_{\text{FWHM}}^2 = R_{\text{int}}^2 + (2.355)^2 \cdot \frac{\text{ENF}}{N_{\text{pe}}} \quad (8)$$

with R_{int} the scintillator-intrinsic resolution and N_{pe} the number of detected photoelectrons, a device with 30% crosstalk requires $\approx 55\%$ more detected photons than a noiseless counter to reach the same statistical term. Suppressing crosstalk is thus quantitatively as valuable as raising PDE, and deep inter-cell trenches — the principal mitigation in modern devices — buy energy resolution at a modest fill-factor cost.

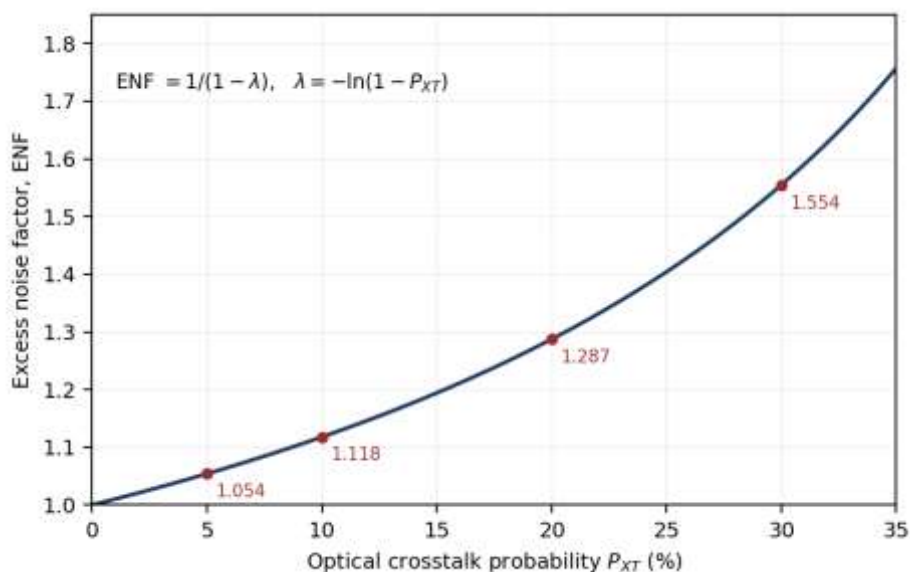


Figure 6: (Calculated.) Excess noise factor versus optical crosstalk probability, computed from Eq. (7) with $\lambda = -\ln(1 - P_{\text{XT}})$, following the branching-process formalism of Vinogradov [10] and Gallego et al. [11]. Marked points are the tabulated check values quoted in the text.

Reported crosstalk probabilities for modern devices span a few percent to $20\text{--}30\%$ at high overvoltage, depending strongly on pitch and trench design [5],[6],[10]. Crosstalk is also temperature dependent: measurements between $-60\text{ }^{\circ}\text{C}$ and $-20\text{ }^{\circ}\text{C}$ show the probability falling with temperature down to about $-40\text{ }^{\circ}\text{C}$ and then saturating, indicating that the low-temperature limit is set by photon emission and transport rather than by triggering statistics [15]. External crosstalk, less studied than the internal component, has been measured at liquid-nitrogen temperature using two identical facing devices, yielding a photon-emission probability per avalanche that is linear in overvoltage [22]. Any SiPM operated away from room temperature therefore requires re-characterization of P_{XT} rather than extrapolation.

4.3 Afterpulsing

Afterpulsing occurs when a carrier generated in a primary avalanche is trapped at a defect level and released after tens of nanoseconds to microseconds, re-triggering the same cell once it has recharged above breakdown [1],[12]. Because the delay is comparable to τ_{rec} , the afterpulse amplitude is a function of delay, and pulse-shape and inter-arrival-time analysis is required to separate afterpulses from genuine dark counts and from prompt crosstalk. Nagy and co-workers characterized afterpulsing and delayed crosstalk jointly on a commercial device, showing that both grow with overvoltage and that the two cannot be distinguished by amplitude alone [12]. In current-generation devices, improved silicon purity has reduced afterpulsing below the crosstalk probability, but it remains non-negligible at high overvoltage and enters Eq. (8) alongside crosstalk.

Table 3: The three principal SiPM noise mechanisms. Magnitudes are indicative ranges for modern commercial devices, compiled from Refs. [1],[5],[10],[11],[12],[13],[15],[22].

Mechanism	Physical origin	Correlated?	Typical magnitude	Dominant dependence
Dark count rate	Thermal (SRH) and field-assisted tunnelling carrier generation	No	$10^2 - 10^3$ kHz/mm ² at 20–25 °C	Exponential in T (Eq. 6); rises with V_{ov} via generation rate and P_{trig}
Optical crosstalk	Hot-carrier luminescence triggering neighbouring cells	Yes (prompt)	~5–30 % probability	Strongly increasing in V_{ov} ; decreases with T down to ≈ -40 °C, then saturates
Afterpulsing	Delayed release of carriers trapped at defect levels	Yes (delayed)	Typically below P_{XT}	Increasing in V_{ov} ; set by trap density and τ_{rec}

4.4 Why reported values disagree

A recurring difficulty in comparing published noise figures is that the measurement conventions are not standardized [13]. DCR depends on the discriminator threshold used (0.5 p.e. thresholds include crosstalk-induced counts that a 1.5 p.e. threshold excludes); crosstalk probability is variously defined from the ratio of counting rates above 1.5 and 0.5 p.e., from pulse-amplitude spectra, or from the fitted branching parameter λ , and these differ by the factor implicit in Eq. (7); afterpulse probability depends on the inter-arrival window and on the assumed recovery model. Cross-study comparisons at the 10 % level are therefore generally not meaningful unless the convention is stated. For a prototype programme, the practical consequence is that the three noise metrics must be measured in-house on the selected device, at the intended overvoltage and temperature, using a documented convention.

5. Temperature Dependence and Stabilization

5.1 Breakdown-voltage temperature coefficient

Breakdown voltage rises approximately linearly with temperature because increased phonon scattering shortens the mean free path of carriers accelerating in the junction field, so a higher field — hence higher bias — is needed to sustain avalanche multiplication [1],[15]. Over the range of interest,

$$V_{\text{bd}}(T) = V_{\text{bd}}(T_0) + \beta(T - T_0) \quad (9)$$

Reported coefficients β range from roughly 20–24 mV/°C for small-cell, NUV-optimized devices to 40–55 mV/°C for larger-cell devices, as established by multi-temperature current–voltage and gain characterization spanning -20 °C to $+50$ °C [15],[16]. Figure 7 reconstructs this family. Because gain and P_{trig} depend on V_{ov}

and not on V_{bias} , a fixed bias means a falling effective overvoltage as temperature rises, with the fractional gain drift given by Eq. (2).

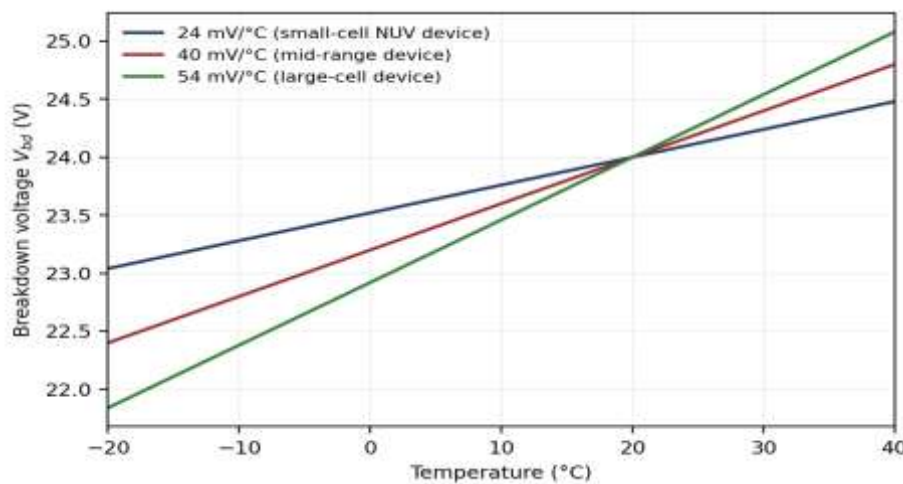


Figure 7: (Reconstructed.) Breakdown voltage versus temperature computed from Eq. (9) for $\beta = 24, 40$ and 54 mV/°C, representing small-cell NUV, mid-range and large-cell devices, referenced to a common $V_{bd} = 24.0$ V at 20 °C. Coefficient ranges from Refs. [15],[16].

5.2 Impact on scintillation-detector performance

Since pulse amplitude for a fixed energy deposit scales as $G \times PDE$, an uncompensated detector shows a directly measurable photopeak drift. Shim, Park and Lee quantified this for a single-channel SiPM coupled to an LGSO crystal, reporting a 511 keV photopeak shift of 13.27 % per 10 °C between 10 °C and 30 °C without compensation, reduced to 1.52 % per 10 °C by a temperature-sensor-driven bias-correction circuit that tracks $V_{bd}(T)$, and showed the technique is transferable to devices of different breakdown voltage by retuning the sensing element's coefficient [14]. The uncompensated figure agrees to within the uncertainty of β with the -13 % per 10 °C predicted by Eq. (2) for $\beta = 40$ mV/°C at $V_{ov} = 3$ V, which supports the interpretation that gain drift, not PDE drift, dominates. Figure 8 reproduces the comparison.

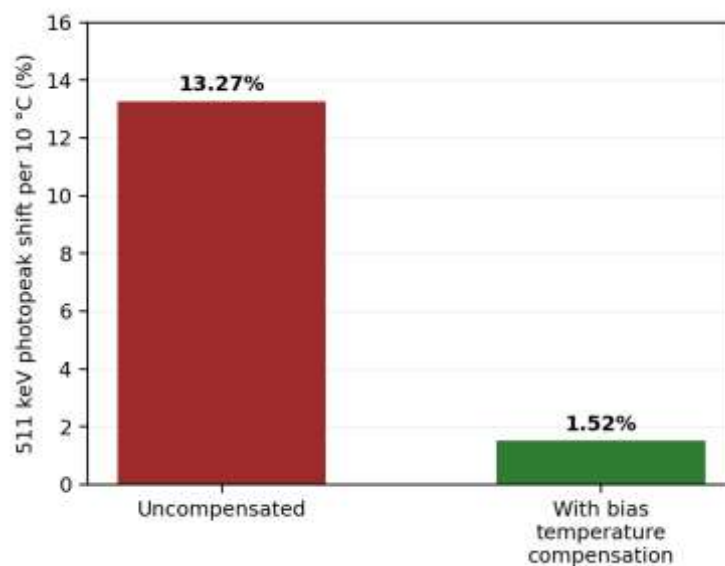


Figure 8: (Reproduced experimental data.) Effect of bias-voltage temperature compensation on 511 keV photopeak stability; single-channel SiPM coupled to an LGSO crystal over 10–30 °C. Data from Shim, Park and Lee, Ref. [14].

5.3 Compensation strategies compared

The compensation law follows directly from Eq. (9): to hold overvoltage at a set value V_{ov}^{set} ,

$$V_{bias}(T) = V_{bd}(T_0) + \beta(T - T_0) + V_{ov}^{set} \quad (10)$$

Implementations differ in how β is obtained and how the correction is applied. A temperature sensor — a diode chain, thermistor or dedicated IC — placed in thermal contact with the sensor drives an analog or DAC-controlled bias supply [14],[26]. Alternatively, a light-tight reference SiPM can serve as its own thermometer: the mean amplitude of its dark pulses is held constant by a negative feedback loop that adjusts the bias, and the same correction is applied to the light-sensitive channels [23],[24]. This second approach removes the need to know β independently and automatically tracks device-to-device variation, at the cost of a second sensor and of a loop bandwidth limited by the dark-pulse averaging time. A third route abandons electrical compensation and stabilizes the sensor's physical temperature by thermoelectric or fluid cooling [27], and a fourth exploits the intrinsically digital output of dSiPM architectures (Section 6) [28],[29].

Table 4 compares reported performance. The uncompensated baselines are consistent across three independent groups at 13–22 % per 10 °C, and all four active-bias implementations converge on residual drift of 0.5–2 %, an improvement of roughly one order of magnitude. On present evidence there is no strong performance discriminator between temperature-sensor-driven and reference-SiPM feedback schemes; the choice should be made on complexity, calibration effort and channel count.

Table 4: Quantitative comparison of temperature-stabilization strategies. All entries are measured values as reported by the cited authors; the stabilized quantity differs between studies (photopeak position, gain, or dark-pulse amplitude) and the temperature ranges differ, so entries indicate achievable performance rather than being directly inter-comparable.

Strategy	Stabilized quantity	Temperature range	Uncompensated drift	Compensated drift	Ref.
None (fixed bias, baseline)	511 keV photopeak, LGSO	10 – 30 °C	13.27 % per 10 °C	—	[14]
Temperature-sensor-driven analog bias correction	511 keV photopeak, LGSO	10 – 30 °C	13.27 % per 10 °C	1.52 % per 10 °C	[14]
Reference (dark) SiPM in negative-feedback loop	Gain	20 – 30 °C	> 20 %	≈ 2 %	[23]
Reference SiPM feedback, refined loop	Dark-pulse amplitude	10 °C span	≈ 22 %	0.78 %	[24]
Adaptive supply, single linear dV_{bias}/dT shared by 4 devices	Gain (30 devices surveyed)	1 – 48 °C	—	< 0.5 % over 20–30 °C	[25]
DAC-based programmable bias with temperature look-up	Gain	–20 – +30 °C	—	within 2 %	[26]
Fluid cooling (no electrical compensation)	Gain / energy resolution at fixed T	1 h at stabilized T	—	0.2 % gain, 0.7 % energy resolution	[27]

Two limitations of bias compensation deserve emphasis, since they are often left implicit. First, Eq. (10) holds V_{ov} constant and therefore stabilizes gain and, to a good approximation, PDE — but it does not stabilize DCR, which continues to follow Eq. (6), nor crosstalk, which retains its own temperature dependence [15]. A compensated detector will therefore still show a temperature-dependent trigger rate, low-energy threshold and, through Eq. (8), a residual variation in energy resolution. Second, β is itself weakly bias- and device-dependent, so a single coefficient applied across a large array leaves a residual spread; the $< 0.5\%$ result of Ref. [25] across four simultaneously stabilized devices indicates the magnitude of what is achievable with one shared coefficient. For a portable instrument, active bias compensation remains the preferred route because it avoids the power, mass and condensation problems of cooling [27], but where DCR itself limits performance, cooling and compensation are complementary rather than alternative.

6. Digital SiPMs and Emerging Architectures

Frach and co-workers introduced the digital SiPM, in which each microcell integrates active quenching, recharge and a one-bit memory in a standard CMOS process, with cell trigger signals combined on-chip through a balanced network feeding an integrated time-to-digital converter, so that photons are counted and time-stamped digitally rather than summed as an analog current [28]. Two consequences matter here. The digital output is insensitive to the analog gain drift of Eq. (2), removing the dominant temperature term; and individual noisy cells can be disabled through the on-chip memory, a targeted DCR-reduction route unavailable to analog devices. Neither eliminates temperature dependence entirely, since P_{trig} — and hence PDE through Eq. (3) — still depends on overvoltage and therefore on temperature. Schaart and co-workers reviewed the subsequent maturation of the technology and its biomedical applications, together with the trade-offs in timing architecture, dead time and power relative to analog arrays read out with discrete electronics [29].

For a research-scale prototype, analog SiPMs with actively compensated bias remain the more accessible and lower-cost option. Digital architectures are best regarded as a benchmark and an upgrade path where intrinsic temperature and noise robustness would outweigh higher unit cost, higher power and reduced front-end design flexibility.

7. Applications and Design Implications for the Prototype

SiPMs now dominate clinical TOF-PET, where magnetic-field immunity enables PET/MRI and where Eq. (5) and SPTR jointly set the achievable coincidence time resolution [7],[8],[21]. They are equally established in sampling calorimetry, astroparticle and Cherenkov instruments, LIDAR, muon scattering tomography and portable radiation spectrometry [9],[16]. Three implications follow for the detector proposed in the associated doctoral programme.

Microcell pitch must be selected against a stated criterion. Larger cells give higher gain at lower overvoltage and higher peak PDE, but higher DCR, higher crosstalk (Table 2) and, through Eq. (4), earlier saturation. If the prototype prioritizes energy resolution at moderate energy, a medium pitch with deep trenches is the defensible compromise; if it prioritizes dynamic range or high-rate operation, small pitch is preferable despite the PDE penalty. The selection should be made from the overlap of Eq. (3) with the measured scintillator emission spectrum, not from datasheet peak PDE.

All three noise metrics must be measured in-house before assembly. DCR, crosstalk and afterpulsing enter Eq. (8) independently of the coupled scintillator, and — for the reasons given in Section 4.4 — published values for nominally identical devices are not reliably comparable. The measurement should be performed at the intended overvoltage and across the intended temperature range.

Active bias compensation is a baseline requirement, not a refinement. Table 4 shows a consistent order-of-magnitude improvement across independent implementations. For a detector intended to operate outside a temperature-controlled laboratory, the compensation loop of Eq. (10) should be designed in from the start, with the residual DCR and crosstalk drift budgeted separately.

8. Limitations And Research Gaps

Several issues remain unresolved and bear directly on prototype design.

Non-standardized noise metrology. As set out in Section 4.4, DCR, crosstalk and afterpulsing are reported under differing thresholds, definitions and time windows [13]. A community convention — or at minimum, mandatory reporting of the convention used — would make the large body of characterization data far more useful than it currently is.

Separation of correlated-noise components. Delayed crosstalk and afterpulsing overlap in both amplitude and delay, and their experimental separation depends on the assumed recovery model [12]. Reported afterpulse probabilities are consequently model-dependent, and the resulting uncertainty propagates into the ENF used in Eq. (8).

External crosstalk. Internal crosstalk is well modelled [10],[11]; external crosstalk between adjacent devices in a packed array is not, and the available quantitative measurements are at cryogenic temperature [22]. For multi-channel scintillation detectors with reflective wrapping, this is a plausible and largely unquantified contribution.

Residual temperature dependence after compensation. No published study appears to characterize the full spectroscopic chain — photopeak position, energy resolution, threshold and trigger rate simultaneously — across a wide ambient range with compensation active. Table 4 records photopeak or gain stability only, leaving the resolution and threshold budgets unverified.

Temperature dependence of PDE. PDE varies more weakly with temperature than gain, but the coefficient is rarely quoted separately, so the extent to which Eq. (10) restores PDE as well as gain is not well established.

Scope exclusions. Radiation-induced DCR increase, relevant to accelerator and space environments, and long-term ageing are outside the scope of this review and are not addressed by any of the compensation strategies in Table 4.

9. Future Directions

Four lines of work follow. First, wide-range validation: measurement of photopeak position, energy resolution, low-energy threshold and count rate with compensation active over at least $-20\text{ }^{\circ}\text{C}$ to $+50\text{ }^{\circ}\text{C}$, closing the gap identified above. Second, joint sensor–scintillator optimization performed on the overlap integral of Eq. (3) with the emission spectrum and on the interplay of scintillator decay time with τ_{rec} in Eq. (4), rather than on separately optimized components. Third, evaluation of reference-cell and on-chip compensation against conventional sensor-driven schemes at realistic channel counts, where per-channel calibration effort — not single-channel performance — becomes the deciding factor. Fourth, adoption of a documented reporting convention for correlated-noise metrics, without which the accumulating characterization literature will remain difficult to aggregate.

10. CONCLUSIONS

The silicon photomultiplier delivers PMT-like gain and single-photon sensitivity in a compact, low-voltage, magnetically insensitive package, at the cost of three correlated and uncorrelated noise mechanisms and a pronounced temperature dependence inherited from the breakdown voltage. The device physics is captured by a small set of relations — Eqs. (1), (3) and (4) for gain, PDE and dynamic range; Eqs. (6) and (7) for the dark and correlated noise entering the excess noise factor; Eqs. (9) and (10) for the temperature dependence and its correction — whose internal consistency has been checked here against reported measurements, including the agreement between the derived $-1.3\text{ }^{\circ}\text{C}^{-1}$ gain coefficient and the measured $13.27\text{ }^{\circ}\text{C}^{-1}$ per $10\text{ }^{\circ}\text{C}$ photopeak drift.

The evidence assembled in Table 4 is unambiguous on the practical question: active bias compensation reduces temperature-driven drift by roughly an order of magnitude across four independent implementations and should be treated as a baseline requirement for any field-deployed SiPM scintillation detector. It does not, however, stabilize dark count rate or crosstalk, both of which retain their own temperature dependence and must be

budgeted separately. The remaining weaknesses in the literature are metrological rather than physical: the noise metrics that enter energy resolution are reported under inconsistent conventions, and the full spectroscopic chain has not been validated across a wide ambient range with compensation active. These device-physics foundations, together with the block diagrams and reconstructed relations presented here, provide the basis for selecting operating parameters and designing the compensation circuitry for the prototype scintillation detector proposed in the associated doctoral research, and should be read alongside the companion review of scintillator materials and the further review of system-level detector characterization.

REFERENCES

1. S. Gundacker and A. Heering, "The silicon photomultiplier: fundamentals and applications of a modern solid-state photon detector," *Phys. Med. Biol.*, vol. 65, no. 17, art. 17TR01, 2020.
2. P. Buzhan, B. Dolgoshein, L. Filatov, A. Ilyin, V. Kantzerov, V. Kaplin, A. Karakash, F. Kayumov, S. Klemm, E. Popova, and S. Smirnov, "Silicon photomultiplier and its possible applications," *Nucl. Instrum. Methods Phys. Res. A*, vol. 504, no. 1–3, pp. 48–52, 2003.
3. D. Renker, "Geiger-mode avalanche photodiodes, history, properties and problems," *Nucl. Instrum. Methods Phys. Res. A*, vol. 567, no. 1, pp. 48–56, 2006.
4. D. Renker and E. Lorenz, "Advances in solid state photon detectors," *J. Instrum.*, vol. 4, art. P04004, 2009.
5. F. Acerbi and S. Gundacker, "Understanding and simulating SiPMs," *Nucl. Instrum. Methods Phys. Res. A*, vol. 926, pp. 16–35, 2019.
6. C. Piemonte and A. Gola, "Overview on the main parameters and technology of modern silicon photomultipliers," *Nucl. Instrum. Methods Phys. Res. A*, vol. 926, pp. 2–15, 2019.
7. E. Roncali and S. R. Cherry, "Application of silicon photomultipliers to positron emission tomography," *Ann. Biomed. Eng.*, vol. 39, no. 4, pp. 1358–1377, 2011.
8. D. R. Schaart, "Physics and technology of time-of-flight PET detectors," *Phys. Med. Biol.*, vol. 66, no. 9, art. 09TR01, 2021.
9. Z. Sadygov, A. Sadigov, and S. Khorev, "Silicon photomultipliers: status and prospects," *Phys. Part. Nucl. Lett.*, vol. 17, no. 2, pp. 160–176, 2020.
10. S. Vinogradov, "Analytical models of probability distribution and excess noise factor of solid state photomultiplier signals with crosstalk," *Nucl. Instrum. Methods Phys. Res. A*, vol. 695, pp. 247–251, 2012.
11. L. Gallego, J. Rosado, F. Blanco, and F. Arqueros, "Modeling crosstalk in silicon photomultipliers," *J. Instrum.*, vol. 8, art. P05010, 2013.
12. F. Nagy, M. Mazzillo, L. Renna, G. Valvo, D. Sanfilippo, B. Carbone, A. Piana, G. Fallica, and J. Molnár, "Afterpulse and delayed crosstalk analysis on a STMicroelectronics silicon photomultiplier," *Nucl. Instrum. Methods Phys. Res. A*, vol. 759, pp. 44–49, 2014.
13. R. Klanner, "Characterisation of SiPMs," *Nucl. Instrum. Methods Phys. Res. A*, vol. 926, pp. 36–56, 2019.
14. H. S. Shim, H. Park, and J. S. Lee, "A temperature-dependent gain compensation technique for positron emission tomography detectors based on a silicon photomultiplier," *Phys. Med. Biol.*, vol. 66, no. 20, art. 205015, 2021.
15. C. Zhong et al., "Study on SiPM performance at low temperatures between $-60\text{ }^{\circ}\text{C}$ and $-20\text{ }^{\circ}\text{C}$," *J. Instrum.*, vol. 17, art. T11003, 2022.
16. B. Sun, H. Li, Q. Li, H. Liang, C. Liu, H. Wang, Z. Wu, S. Xiao, and W. Xu, "Characterization of silicon photomultipliers for their application in muon scattering tomography," *J. Instrum.*, vol. 20, no. 02, art. T02003, 2025.
17. P. Eckert, H.-C. Schultz-Coulon, W. Shen, R. Stamen, and A. Tadday, "Characterisation studies of silicon photomultipliers," *Nucl. Instrum. Methods Phys. Res. A*, vol. 620, no. 2–3, pp. 217–226, 2010.
18. J. W. Cates, R. Vinke, and C. S. Levin, "Analytical calculation of the lower bound on timing resolution for PET scintillation detectors comprising high-aspect-ratio crystal elements," *Phys. Med. Biol.*, vol. 60, no. 13, pp. 5141–5161, 2015.

19. F. Acerbi, G. Paternoster, M. Capasso, M. Marcante, A. Mazzi, V. Regazzoni, N. Zorzi, and A. Gola, "Silicon photomultipliers: technology optimizations for ultraviolet, visible and near-infrared range," *Instruments*, vol. 3, no. 1, art. 15, 2019.
20. M. V. Nemallapudi, S. Gundacker, P. Lecoq, and E. Auffray, "Single photon time resolution of state of the art SiPMs," *J. Instrum.*, vol. 11, art. P10016, 2016.
21. S. Gundacker, R. M. Turtos, N. Kratochwil, R. H. Pots, M. Paganoni, P. Lecoq, and E. Auffray, "Experimental time resolution limits of modern SiPMs and TOF-PET detectors exploring different scintillators and Cherenkov emission," *Phys. Med. Biol.*, vol. 65, no. 2, art. 025001, 2020.
22. G. Li, L. Wang, X. Sun, F. Liu, C. Guo, K. Zhao, L. Tian, Z. Yu, Z. Hou, C. Li, Y. Lei, B. Wang, and R. Zhou, "A novel measurement method for SiPM external crosstalk probability at low temperature," *IEEE Trans. Nucl. Sci.*, vol. 71, no. 6, pp. 1357–1364, 2024.
23. F. Licciulli, I. Indiveri, and C. Marzocca, "A novel technique for the stabilization of SiPM gain against temperature variations," *IEEE Trans. Nucl. Sci.*, vol. 60, no. 2, pp. 606–611, 2013.
24. F. Licciulli and C. Marzocca, "An active compensation system for the temperature dependence of SiPM gain," *IEEE Trans. Nucl. Sci.*, vol. 62, no. 1, pp. 228–235, 2015.
25. G. Eigen, A. Tret, J. Zalieckas, J. Cvach, J. Kvasnička, and I. Polák, "Gain stabilization of SiPMs and afterpulsing," *J. Phys. Conf. Ser.*, vol. 1162, art. 012013, 2019.
26. C. Zi, J. Chang, M. Yang et al., "A compensation circuit for the gain temperature drift of silicon photomultiplier tube," *Radiat. Detect. Technol. Methods*, vol. 7, no. 4, pp. 571–577, 2023.
27. A. V. Stolin, S. Majewski, and R. R. Raylman..., "Novel method of temperature stabilization for SiPM-based detectors," *IEEE Trans. Nucl. Sci.*, vol. 60, no. 5, pp. 3181–3187, 2013.
28. T. Frach, G. Prescher, C. Degenhardt, R. de Gruyter, A. Schmitz, and R. Ballizany, "The digital silicon photomultiplier — principle of operation and intrinsic detector performance," in *Proc. 2009 IEEE Nucl. Sci. Symp. Conf. Record (NSS/MIC)*, 2009, pp. 1959–1965.
29. D. R. Schaart, E. Charbon, T. Frach, and V. Schulz, "Advances in digital SiPMs and their application in biomedical imaging," *Nucl. Instrum. Methods Phys. Res. A*, vol. 809, pp. 31–52, 2016.