Passive quenching, signal shapes, and space charge effects in SPADs and SiPMs

P. Windischhofer\textsuperscript{a}, W. Riegler\textsuperscript{b}

\textsuperscript{a}University of Oxford \hspace{1cm} \textsuperscript{b}CERN

Abstract

In this report we study the dynamics of passive quenching in a single-photon avalanche diode. Our discussion is based on a microscopic description of the electron-hole avalanche coupled to the equivalent circuit of the device, consisting of the quench resistor and the junction capacitance. Analytic expressions for the resulting signal shape are derived from this model for simple electric field configurations, and efficient numerical prescriptions are given for realistic device geometries. Space charge effects arising from the avalanche are included using simulations. They are shown to distort the signal shape, but alter neither its basic characteristics nor the underlying quenching mechanism.

1. Introduction

Single-photon avalanche diodes (SPADs) and silicon photomultipliers (SiPMs) are Geiger-mode avalanche detectors with direct single-particle sensitivity to photons and charged particles \[1, 2, 3\]. In these devices, the small amount of charge deposited by the primary particle triggers a self-sustaining electron-hole avalanche which leads to the creation of a detectable electric signal. The avalanche develops in a thin region with very high electric fields, created in a p-n diode which is initially reverse-biased above its breakdown voltage.

The resulting exponential growth of the avalanche must be stopped (“quenched”) at a suitable moment so that all charge carriers are removed from the junction and the detector can register the next event. In passively quenched devices, the diode is decoupled from the bias supply through a series resistance. The charge produced by the avalanche then reduces the electric field in the junction and suppresses the avalanche growth. Quenching occurs when the voltage across the diode drops below the breakdown limit. Once all charges have drifted out of the junction, the original bias voltage is restored.

The resulting voltage and current waveforms form the output signal. The signal properties are ultimately determined by the complicated interplay of impact ionisation, the quenching circuit, and the diode capacitance. The operating conditions are typically chosen such that the achieved signal charge is around $10^4$–$10^5$ electrons, i.e. very large avalanches must be considered.

Existing analyses of the signal formation in passively quenched SPADs either model the avalanche discharge and quenching purely in terms of an equivalent electrical circuit \[4, 5, 6\], or use a coarse effective description of the avalanche development, e.g. in terms of rate equations \[7, 8\]. In this report, we construct a simple analytic quenching model which is directly based on the microscopic dynamics of electron-hole avalanches. It naturally connects to our previous treatment of avalanche statistics given in Ref. \[9\]. As a result, our model describes the detector signal in terms of the underlying material parameters and the device geometry. It can therefore guide the evaluation of specific detector designs and allows direct comparisons to be made with measurements.

This paper is structured as follows. Section 2 briefly reviews the avalanche formation and summarises properties of large electron-hole avalanches used to formulate the quenching model. Analytic predictions
for the signal shape and the spatial distributions of charge carriers in the high-field region are derived in Section 3. The impact of the space charge field of the avalanche on the quenching process is then studied in Section 4.

2. Properties of large electron-hole avalanches

The development of the electron-hole avalanche is determined by the transport of charge carriers of both polarities through the multiplication region and their interactions with the material. For sufficiently strong electric fields, both electrons and holes can cause impact ionisation and produce additional electron-hole pairs. The probability for a charge carrier to undergo a multiplication reaction may be parameterised in terms of the material parameters and, in general, the past history of the carrier. The drifting carriers in the avalanche induce a current $I_{\text{ind}}$ on the metallised device contacts, which may be computed with the Ramo-Shockley theorem [10, 11].

Impact ionisation is a stochastic process. The avalanche development is thus subject to fluctuations, which are particularly pronounced at early times when the total number of participating charge carriers is still small. For avalanches containing many charge carriers, these fluctuations average out and the further evolution appears deterministic. (The continued evolution of an avalanche containing $N$ charges will generate relative fluctuations with a magnitude proportional to $1/\sqrt{N}$, as discussed in Section 3.5 of Ref. [9].) It is useful to introduce a current scale $I_{\text{det}}$ to separate these two regimes, chosen such that avalanche fluctuations are important whenever $I_{\text{ind}} \ll I_{\text{det}}$ and a deterministic description of the avalanche is appropriate when $I_{\text{ind}} \gtrsim I_{\text{det}}$. The current $I_{\text{det}}$ therefore defines the level of fluctuations that are deemed “important” in a particular situation. It implements a purely semantic definition, and expressions for physical observables do not depend explicitly on $I_{\text{det}}$.

Our model of the quenching process is based on the following two general properties of large electron-hole avalanches ($I_{\text{ind}} \gtrsim I_{\text{det}}$) developing in a time-independent electric field, illustrated in Fig. 1:

(i) The shape of the induced current $I_{\text{ind}}$ is identical for each avalanche event. It is directly proportional to the average current $\langle I_{\text{ind}} \rangle$, where the average is taken over all avalanche evolutions starting from the same initial conditions,

$$I_{\text{ind}}(t) = k \langle I_{\text{ind}}(t) \rangle .$$

The proportionality factor $k$ is a random variable. The resulting fluctuations around the average evolution determine the time resolution, but are irrelevant for the study of the quenching process.

Figure 1: (Colour online.) The grey lines represent typical simulated avalanche events, each starting from a single primary electron. For large avalanches, the evolution of $I_{\text{ind}}$ in each event becomes proportional to the average current $\langle I_{\text{ind}} \rangle$, parameterised by $I_0$ and $S_1$, as indicated by the red line.
Figure 2: (a) Electrons and holes drift in opposite directions with velocities \(v_e\) and \(v_h\) in the electric field \(E(x)\) of the p-n junction, inducing a current \(I_{\text{ind}}\). (b) Both charge carrier species can undergo impact ionisation and produce additional electron-hole pairs, parameterised by the impact ionisation coefficients \(\alpha\) and \(\beta\). (c) Equivalent circuit used for the analysis of the quenching process, consisting of the diode capacitance \(C_d\) and the quench resistor \(R_q\). The current source represents the induced current \(I_{\text{ind}}\).

(ii) The average current \(\langle I_{\text{ind}}(t) \rangle\) evolves exponentially as

\[
\langle I_{\text{ind}}(t) \rangle = I_0 \exp(S_1 t),
\]

and is parameterised by the dimensionful quantities \(I_0\) and \(S_1\). The current \(I_0\) depends on the number and the spatial distribution of the primary charges initiating the avalanche. The parameter \(S_1\) represents the asymptotic net growth rate of the avalanche, i.e. it takes into account the generation of charges in the high-field region and their outflow across its boundaries. It depends on the device geometry and the reverse bias voltage \(V_d\) applied across the junction, i.e. \(S_1 = S_1(V_d)\). In the following, the breakdown voltage \(V_{\text{br}}\) is defined such that \(S_1(V_{\text{br}}) = 0\). This implies that \(S_1(V_d) > 0\) for \(V_d > V_{\text{br}}\), i.e. exponentially growing avalanches are possible if the device is operated “above breakdown”. For a bias voltage below the breakdown limit, \(V_d < V_{\text{br}}\), impact ionisation cannot sustain the discharge. The avalanche current then decays exponentially, i.e. \(S_1(V_d) < 0\).

For a given device and electric field configuration, the parameters \(S_1\) and \(I_0\) are always accessible through Monte Carlo (MC) simulations of the avalanche development (such as the one in Fig. 2). Efficient numerical prescriptions for quasi-one-dimensional situations are referenced below. The above properties arise as explicit predictions of the avalanche model studied analytically in a one-dimensional geometry in Ref. 9. It is briefly reviewed in the following. We use this model to illustrate our results in Sections 3 and 4. However, our description of the quenching process is more general and applies to all scenarios in which the avalanche obeys the properties (i)–(ii) above, including realistic three-dimensional device geometries.

One-dimensional memoryless avalanche model. The diameter \(D\) of a typical SPAD pixel is significantly larger than the longitudinal thickness \(d\) of the high-field region, typically \(D/d \gtrsim 10\). Close to the centre of a pixel, this results in a quasi-one-dimensional field geometry through which electrons and holes drift in opposite directions with velocities \(v_e\) and \(v_h\), respectively. The drift direction is taken to be along the \(x\)-axis (Fig. 2a).

The one-dimensional model of Ref. 9 describes the avalanche development in this geometry. In this model, impact ionisation is parameterised by the impact ionisation coefficients \(\alpha\) and \(\beta\), which represent the probability per unit length for an electron or hole to undergo a multiplication reaction (Fig. 2b). This probability is assumed to depend only on the local electric field strength \(E(x)\), i.e. \(\alpha(x) = \alpha(E(x))\) and
\( \beta(x) = \beta(\mathbf{E}(x)) \). Any dependence on the history of the charge carriers is neglected, i.e. the avalanche is assumed to be a memoryless stochastic process.

For general field profiles \( \mathbf{E}(x) \), the growth parameter \( S_1 \) may be obtained from the numerical solution of the differential equations describing the evolution of the average charge content of the avalanche, Eqs. 30 and 31 in Ref. [9]. These equations also determine the spatial carrier densities in the junction. As shown in Appendix B of Ref. [3], the breakdown condition \( S_1 = 0 \) may be expressed in terms of the impact ionisation coefficients as

\[
\int_0^d dx \alpha(x) \exp \left[ - \int_0^x dx' \left( \alpha(x') - \beta(x') \right) \right] = 1. \tag{3}
\]

It is thus equivalent to the breakdown condition of Ref. [12], which evaluates the point at which the gain for a constant current injected into the junction diverges.

For constant fields, \( S_1 \) may be expressed as the product of the effective carrier drift velocity \( v^* = 2v_e v_h/(v_e + v_h) \) and the growth factor \( \gamma_1 \), i.e. \( S_1 = \gamma_1 v^* \). The latter is defined in Eq. 35 in Ref. [9]. The drift velocities for electrons and holes, \( v_e \) and \( v_h \), are themselves dependent on the local electric field. The spatial carrier densities for electrons and holes, \( n_e(x,t) \) and \( n_h(x,t) \), are proportional to the functions \( f_{\lambda_1}^e(x,t) \) and \( f_{\lambda_1}^h(x,t) \), respectively. Their analytic expressions are given in Eqs. 36 and 37 of Ref. [9]. The numerical parameter \( \lambda_1 \) is related to the time constant \( S_1 \) and defined in Eq. 38 of Ref. [9].

The proportionality factor \( k \) in Eq. (4) is approximately distributed according to a gamma distribution with shape parameter \( A \) and scale parameter \( A^{-1} \) (cf. Section 3.6 of Ref. [9]),

\[
p(k)dk = \frac{k^{A-1}e^{-k/A}}{\Gamma(A)A^{-A}}dk, \tag{4}
\]

where \( \Gamma(z) \) is the gamma function, and the avalanche parameter \( A \) is defined as

\[
A = \frac{\alpha v_e N^0_e + \beta v_h N^0_h}{\alpha v_e + \beta v_h} \tag{5}
\]

for position-independent electric fields. In Eq. (5), \( N^0_e \) and \( N^0_h \) label the numbers of primary electrons and holes, respectively. For general field profiles \( \mathbf{E}(x) \), the avalanche parameter may be computed as shown in Eq. 79 of Ref. [9].

### 3. Passive quenching in the absence of space charge effects

To describe the quenching process, the avalanche evolution summarised in Section 2 is coupled to the electrical equivalent circuit of a single SPAD (Fig. 2c). In this circuit, the p-n junction is represented by its capacitance \( C_d \), and the current \( I^{\text{ind}} \) acts as a parallel current source. (Note that the dynamic resistance \( R_d \) of the diode is automatically included through the dependence of the carrier drift velocities \( v_e \) and \( v_h \) on the electric field and is thus not present as an explicit circuit component.) In the absence of free charge carriers in the junction, \( I^{\text{ind}} = 0 \) and \( V_d = V_{\text{supply}} \). The excess voltage \( V_{\text{ex}} \) is defined as \( V_{\text{ex}} = V_{\text{supply}} - V_{\text{br}} > 0 \), where the breakdown voltage \( V_{\text{br}} \) is determined by Eq. (3).

The time-dependent voltage \( V_d(t) \) determines the electric field \( \mathbf{E}(x,t) = \mathbf{E}(x,V_d(t)) \) which is relevant for the development of the avalanche. To ensure quenching, the decoupling resistor \( R_q \) must be chosen such that the recharging time constant \( R_q C_d \) is much longer than the time scale \( 1/S_1(V_{\text{supply}}) \) on which the avalanche develops. The initial phase of the quenching process thus remains unchanged as \( R_q \to \infty \).

In this dynamical situation, the instantaneous growth rate of the avalanche, \( S_1(t) \), becomes itself time-dependent. Provided that \( V_d(t) \) changes on time scales that are long compared to the internal time constants of the avalanche, the changing growth rate may be approximated as \( S_1(t) \approx S_1(V_d(t)) \), where
$S_1(V_d)$ is the function introduced in Section 2 for static situations. This “adiabatic approximation” is shown below to be applicable to practically relevant situations. For $R_q \to \infty$ and $I^{\text{ind}} \gtrsim I_{\text{det}}$, the combined deterministic system of avalanche and equivalent circuit is then described by the equations

\begin{align}
\frac{dI^{\text{ind}}}{dt} &= S_1(V_d) I^{\text{ind}}, \\
\frac{dV_d}{dt} &= -\frac{I^{\text{ind}}}{C_d}.
\end{align}

More complicated equivalent circuits that include additional parasitic elements or finite values of the quench resistance $R_q$ are readily accommodated using the same steps. Eq. 6, which describes the evolution of the avalanche, must then be solved together with the corresponding circuit equations. For the situation discussed here, Eqs. 6–7 may be solved numerically for general $S_1(V_d)$. This is similar to the rate model constructed in Ref. [7], but in contrast to the calculations performed there automatically includes boundary effects.

In many practically relevant situations, $V_{\text{ex}}/V_{\text{br}} \ll 1$, i.e. the voltage $V_d$ remains close to $V_{\text{br}}$ during the full evolution of the system. In this case, the function $S_1(V_d)$ in Eq. 6 may be replaced by its linear expansion around $V_{\text{br}}$. With the breakdown condition $S_1(V_{\text{br}}) = 0$, this gives

\begin{equation}
S_1(V_d) \approx \left. \frac{dS_1}{dV_d} \right|_{V_{\text{br}}} (V_d - V_{\text{br}}) =: K_{\text{br}} \cdot (V_d - V_{\text{br}}).
\end{equation}

The parameter $K_{\text{br}}$ may be easily obtained through finite differences provided that $S_1(V_d)$ is known around $V_d \approx V_{\text{br}}$. An analytic expression for $K_{\text{br}}$ for a memoryless avalanche and position-independent fields is given in Eq. 26 below.

If the primary charge deposit occurs at $t = 0$, Eqs. 4–7 must be solved for $t \geq t_{\text{det}}$, where $t_{\text{det}}$ is the time at which the avalanche current first exceeds $I_{\text{det}}$. The time $t_{\text{det}}$ is a random variable whose statistics are further discussed below. Defining $V_d(t_{\text{det}}) = V_{\text{det}}$, the relevant initial condition is $I^{\text{ind}}(t_{\text{det}}) = I_{\text{det}}$, leading to the following analytic solutions,

\begin{align}
I^{\text{ind}}(t) &= \frac{C_d V_{\text{eff}}}{\tau_q} \left[ 1 - \tanh^2 \left( \frac{t - t_{\text{det}} - \Delta t}{\tau_q} \right) \right], \\
V_d(t) &= V_{\text{det}} - V_{\text{eff}} \left[ \tanh \left( \frac{t - t_{\text{det}} - \Delta t}{\tau_q} \right) + \tanh \left( \frac{\Delta t}{\tau_q} \right) \right],
\end{align}

with

\begin{equation}
V_{\text{eff}} = \sqrt{1 + \frac{2I_{\text{det}}}{C_d K_{\text{br}} (V_{\text{det}} - V_{\text{br}})^2}},
\end{equation}

the quenching time constant $\tau_q$,

\begin{equation}
\tau_q = \frac{2}{K_{\text{br}} V_{\text{eff}}},
\end{equation}

and the time $\Delta t$,

\begin{equation}
\Delta t = \tau_q \text{artanh} \left( \frac{V_{\text{det}} - V_{\text{br}}}{V_{\text{eff}}} \right).
\end{equation}

The quantity $\tau_q$ parameterises the time scale on which $V_d$ changes as the avalanche is quenched. For Eqs. 9–10 to be a valid description of the quenching process, i.e. for the adiabatic approximation to be applicable, $\tau_q$ must be larger than the transit time $d/v^*$. The latter approximates the time scale on which the spatial charge distribution in the junction, and therefore also the instantaneous growth rate $S_1(t)$, react to changes of $V_d$. 

5
The general expressions in Eqs. 9–10 can be further simplified. For practical devices, the diode capacitance $C_d$ typically amounts to a few femtofarads. With the elementary charge $e_0$, around $C_dV_{ex}/e_0 \sim 10^4$ charge carriers need to be produced by the avalanche in the high-field region before any significant discharge of $C_d$ occurs, i.e. any avalanche fluctuations remaining during the quenching process amount to a few percent at most. Working to this accuracy, $V_d(t_{det}) = V_{det} \approx V_{supply} \text{ and } V_{det} - V_{br} \approx V_{ex}$; furthermore, $I_{det}$ can be selected so that $I_{det} \ll C_dK_{br}(V_{det} - V_{br})^2$. In this limit, Eq. 11 simplifies to $V_{eff} \approx V_{ex}$ and the quenching time constant is directly given in terms of the excess voltage,

$$\tau_q \approx \frac{2}{K_{br}V_{ex}},$$

and the time $\Delta t$ defined in Eq. 13 becomes

$$\Delta t \approx \frac{\tau_q}{2} \log \frac{2C_dK_{br}V_{ex}^2}{I_{det}} \approx \frac{1}{S_1(V_{supply})} \log \frac{2C_dK_{br}V_{ex}^2}{I_{det}}.$$ (15)

With this, $\Delta t/\tau_q \gg 1$ and Eqs. 9–10 simplify to

$$I_{ind}(t) = \frac{C_dV_{ex}}{\tau_q} \left[1 - \tanh^2 \left(\frac{t - t_{det} - \Delta t}{\tau_q}\right)\right],$$ (16)

$$V_d(t) = V_{supply} - V_{ex} \left[1 + \tanh \left(\frac{t - t_{det} - \Delta t}{\tau_q}\right)\right].$$ (17)

and the adiabatic approximation is valid provided that $K_{br}V_{ex} \lesssim 2e^*/d$. The current signal $I_{ind}$ attains a peak value of $I_{peak}^{ind} = C_dV_{ex}/\tau_q$ at time $t_{max} = t_{det} + \Delta t$. If $t = 0$ is instead regarded as the position of the peak, the signal shape reads

$$I_{peak}^{ind}(t) = C_dV_{ex} \left[1 - \tanh^2 \left(\frac{t}{\tau_q}\right)\right],$$ (18)

$$V_{d,peak}(t) = V_{supply} - V_{ex} \left[1 + \tanh \left(\frac{t}{\tau_q}\right)\right].$$ (19)

Its evolution is symmetric around this peak with a FWHM of

$$\text{FWHM} (I^{ind}) = 2\tau_q \arctan \frac{1}{\sqrt{2}} \approx 1.76 \tau_q.$$ (20)

The observable voltage step $\Delta V_d = V_{supply} - \lim_{t \rightarrow \infty} V_d(t)$ is

$$\Delta V_d = 2V_{ex},$$ (21)

and $V_d$ falls from $V_{supply} - 0.1\Delta V_d$ to $V_{supply} - 0.9\Delta V_d$ within about $2.2 \tau_q$. The total signal charge is

$$Q = \int_{-\infty}^{\infty} dt I^{ind}(t) = 2C_dV_{ex}.$$ (22)
appropriate, the collapse of the avalanche exactly mirrors its initial growth; the evolution of \( V_d(t) \) is then symmetric around \( V_{br} \) and so \( \Delta V_d = G V_{ex} \) with \( G = 2 \). Note that this conclusion deviates from other commonly-used quenching models, which take the avalanche discharge to stop already at \( V_d(t) \approx V_{br} \) and thus lead to \( G \approx 1 \) \cite{4,5,6}. These differences deserve a detailed comparison and discussion, which we carry out at the end of this Section.

The timing of the detector signal also carries important information. The average time of maximum current,

\[
\langle t_{\text{max}} \rangle = \langle t_{\text{det}} \rangle + \Delta t,
\]

is experimentally accessible in a setup where the primary charge deposit occurs at a well-defined time, e.g. where charges are deposited by means of a laser pulse or a charged-particle beam. It contains information about the early evolution of the avalanche (through \( \langle t_{\text{det}} \rangle \)), as well as the quenching process itself (through \( \Delta t \)). Eqs. 1 and 2 result in the following expression for \( \langle t_{\text{det}} \rangle \),

\[
\langle t_{\text{det}} \rangle = \frac{1}{S_1(V_{\text{supply}})} \left( \log \frac{I_{\text{det}}}{I_0} - \langle \log k \rangle \right),
\]

which, together with Eq. 15 gives for \( \langle t_{\text{max}} \rangle \)

\[
\langle t_{\text{max}} \rangle = \frac{1}{S_1(V_{\text{supply}})} \left( \log \frac{2C_dK_{br}V_{ex}^2}{I_0} - \langle \log k \rangle \right).
\]

The logarithmic expectation \( \langle \log k \rangle \) depends on the magnitude of the avalanche fluctuations; an analytic expression valid for the memoryless avalanche model is derived in Eq. 27 below.

**Quenching for one-dimensional memoryless avalanche.** We illustrate these results for the case of a simplified SPAD, for which we approximate the field distribution in the multiplication region with a position-independent, one-dimensional electric field. The field is thus taken to be \( \mathbf{E}(x, V_d) = -x V_d/d \), where \( \mathbf{x} \) is the unit vector along \( x \) (cf. Fig. 2a). In this case, the breakdown voltage is related to the breakdown field

![Figure 3: (a) Dependence of the growth parameter \( \gamma_1 \) on the electric field and the thickness \( d \), for the case of silicon and the impact ionisation coefficients of Ref. 13. The limit of an unbounded multiplication region \( (d \to \infty) \) is also shown. (b) The parameter \( \gamma_1 \) is expressed in terms of the voltage \( V_d \) close to the breakdown point.](image)
$E_{br}$ as $V_{br} = E_{br}d$. The effective drift velocity $v^*$ typically saturates in the high-field region, i.e. becomes independent of $V_d$. Using the memoryless avalanche model from Section 2, $S_1(V_d) = \gamma_1(V_d)v^*$ and

$$K_{br} = \frac{v^* d\gamma_1}{dE_{br}} = \left(\frac{\alpha'_{br} + \beta'_{br}}{2d}\right) + \frac{(\sinh 2\kappa_{br} - 2\kappa_{br}) (\alpha'_{br} \beta_{br} + \alpha_{br} \beta'_{br}) v^*}{2\kappa_{br} (\cosh 2\kappa_{br} - 1) - d (\sinh 2\kappa_{br} - 2\kappa_{br}) (\alpha_{br} + \beta_{br})},$$

(26)

where $\alpha_{br} = \alpha(E_{br})$, $\beta_{br} = \beta(E_{br})$, $\alpha'_{br} = d\alpha/dE|_{E_{br}}$, $\beta'_{br} = d\beta/dE|_{E_{br}}$, and $\kappa_{br} = d(\alpha_{br} - \beta_{br})/2$. In the adiabatic approximation, predictions for the charge carrier densities of electrons and holes are proportional to the functions $f_A^1(V_d)$ and $f_A^h(V_d)$. The linear expansion in Eq. 3 describes $\gamma_1(V_d)$ to excellent precision for typical values of $V_{ex}$. This is visualised in Fig. 3 for silicon, using the parameterisations of the impact ionisation coefficients from Ref. 13.

With the distribution of $k$ in Eq. 4, $(\log k) = \psi(A) - \log(A)$ with the digamma function $\psi(z)$, and the expression for $\langle t_{max}\rangle$ in Eq. 25 becomes

$$\langle t_{max}\rangle = \frac{1}{S_1(V_{supply})} \left( \frac{2C_dK_{br}V_{ex}^2}{I_0} + \log(\psi(A) - \log(A)) \right),$$

(27)

and $I_0$ may be computed for a given initial charge deposit as shown in the paragraph following Eq. 46 in Ref. 9.

The analytic results derived above are compared to a MC implementation of the memoryless avalanche model, coupled to the quenching circuit of Fig. 2c. The avalanche simulation is performed as described in the Appendix, using the impact ionisation coefficients for silicon from Ref. 13 and saturated drift velocities $v_e = v_h = v^* = 0.1 \mu m/ps$ [14]. The circuit equations are integrated using the forward Euler method. The induced current $I_{ind}$ is computed assuming the weighting field of a parallel-plate geometry with an electrode spacing of $d$ for the multiplication region. A dedicated conversion layer, common in SPADs optimised for infrared photon detection, is not included. The full simulation is referred to as “memoryless avalanche MC + circuit” (MAMC + circuit). For the comparisons shown here, the thickness $d$ of the high-field region is chosen as $d = 0.5 \mu m$ and its diameter is $D = 10 \mu m$. Assuming a parallel field geometry and using $\epsilon_r = 11.7$ [15], the junction capacitance is $C_d \approx 16 \mu F$. A quench resistor with $R_q = 200 \Omega$ is used. The breakdown voltage for this geometry is $V_{br} \approx 20.34 V$ and with Eq. 26 the model parameter $K_{br}$ evaluates to $K_{br} \approx 0.105 (V \ ps)^{-1}$. The quenching time constant is $\tau_q \approx 9.5 ps$ for $V_{ex} = 2 V$ and $\tau_q \approx 6.3 ps$ for $V_{ex} = 3 V$. The transit time $d/v^*$ is 5 ps, i.e. the adiabatic approximation is valid for excess voltages not significantly above 3 V.

Fig. 4 shows the induced current $I_{ind}$ for several avalanche events and different excess voltages, obtained with the MAMC + circuit simulation model. Fluctuations in the early development of the avalanche lead to a stochastic displacement of the deterministic current waveform from Eq. 16. The simulated time of maximal current $\langle t_{max}\rangle$ is 55.9 ps (40.8 ps) for an excess voltage $V_{ex}$ of 2 V (3 V). These values are well reproduced by Eq. 27 which evaluates to $\langle t_{max}\rangle \approx 56.1 ps$ (40.9 ps) for $V_{ex} = 2 V$ (3 V).

Fig. 5 compares the shapes of the approximate solutions in Eqs. 16–17 to the result from the MAMC + circuit simulation model around the time $t_{max}$, demonstrating very good agreement between the analytic result and the simulation. Small deviations arise as $V_d$ crosses the breakdown voltage, resulting in a slight underestimation of $\Delta V$ for large excess voltages. This is caused by additional dynamic effects originating from changes in the shapes of the spatial carrier densities of electrons $n_e(x,t)$ and holes $n_h(x,t)$ that are not included in the adiabatic model. The temporal evolution of these densities is compared in Fig. 6 (For the simulation, the average densities $\langle n_e \rangle$ and $\langle n_h \rangle$ are shown to ensure a meaningful comparison with the adiabatic model even at early times where the number of charge carriers is small.) Charge carriers of both polarities are present throughout the entire multiplication region. The highest densities are attained at the corresponding downstream ends, i.e. at $x = d$ for electrons and at $x = 0$ for holes. The shape of the distributions, and in particular the curvature of $n_h(x)$, depends on $V_d$. 

8
Discussion and comparison with other quenching models. The relation $\Delta V_d = G V_{ex}$ with $G = 2$ is derived in Eq. \ref{eq:21} as an exact analytic result that holds in situations where the linear expansion in Eq. \ref{eq:8} is appropriate.

This conclusion differs from predictions made with simpler, widely-used, quenching models \cite{4, 5, 6}. In these models, the avalanche is taken to directly transition the device from its initial inactive configuration ($V_d = V_{\text{supply}}$ and $I_{\text{ind}} = 0$) into a quasi-time-independent, active state where $V_d = V_{br}$ and a persistent current $I_p = V_{ex}/R_q$ flows through the quench resistor. If $R_q$ is chosen sufficiently large, the current $I_p$ is
small enough so that avalanche fluctuations may spontaneously quench the discharge after a brief period of time (the probability per unit time for this to happen is exponentially suppressed in $I_p$ [6]). This leads to a voltage step of $\Delta V_d = V_{ex}$, i.e. $G = 1$. The central proposition is that this behaviour can be modelled by the equivalent circuit shown in Fig. 7, where the diode resistance $R_d$ is explicitly included and the avalanche is represented by a switch $S$ whose state marks the presence (closed) or absence (open) of an avalanche discharge.

The microscopic analysis conducted here indeed confirms the existence of this active equilibrium state: Eq. 6 implies $dI^{\text{ind}}/dt = 0$ at $V_d = V_{br}$, which reveals $I^{\text{ind}} = I(R_d) = I_p$ as a valid configuration of the circuit in Fig. 2a. This shows that the simple model in Fig. 7 can adequately describe the possible active (switch closed) and inactive (switch open) equilibrium configurations of the device.

There is, however, no a-priori reason to expect this passive circuit to correctly model the fast transition between these two states. In particular, the avalanche current $I^{\text{ind}}$ in Fig. 7 decays exponentially as $V_d$ approaches $V_{br}$, and $V_d \geq V_{br}$ throughout the entire quenching process. This is apparently inconsistent with the definition of $V_{br}$ as the voltage above which each existing carrier produces a diverging number of secondary charges [6]. To the best of our knowledge, the existing literature does not address (or resolve) this contradiction for switch-based models.

The microscopic treatment reveals that it is not the absolute value of the avalanche current (or, equivalently, the total number of charges in the avalanche) that vanishes close to the breakdown point, but rather its growth rate $S_1$, thus giving rise to a different transient behaviour and a proportionality
factor $G = 2$. A similar factor $G \approx 2$ has also been obtained in Refs. [7, 8] through the numerical solution of a system of rate equations describing the avalanche, and in Ref. [16] through a microscopic MC simulation similar to the one presented here. Refs. [7, 8] furthermore report measurements of $\Delta V_d$ and find $G > 1$.

4. Space charge effects

For sufficiently large avalanches, the space charge field $\mathbf{E}_{sc}(x, t)$ caused by the charge carriers in the avalanche may become comparable to the field $\mathbf{E}_{ext}(x, V_d(t))$ produced as a result of the externally applied voltage $V_d$. The electric field profile $\mathbf{E}$ that is relevant for the development of the avalanche is now the sum of both components, i.e. $\mathbf{E} = \mathbf{E}_{ext} + \mathbf{E}_{sc}$. A full description of the effects of space charge (SC) on the quenching process requires a three-dimensional simulation of the avalanche together with the computation of a self-consistent solution for the space charge field at every time step. In particular, $\mathbf{E}_{sc}$ depends on the spread of the avalanche in the direction transverse to the externally impressed field $\mathbf{E}_{ext}$.

The situation is studied here in a simplified way as illustrated in Fig. 8. The evolution of the avalanche continues to be simulated by the MAMC + circuit simulation model, with impact ionisation resulting from the drift of charge carriers along the $x$-axis. A parameterised model is used to describe the development of the avalanche in the transverse $yz$-plane, defining a radially symmetric charge profile $\rho(r, x, t)$. Different physical effects may be included in this parameterisation, e.g. charge carrier diffusion as well as drift due to nonvanishing radial electric field components. The prescribed charge distribution $\rho$ is then used to compute the space charge field $\mathbf{E}_{sc}$ along the $x$-axis, which, by symmetry, only has a nonvanishing $x$-component, $\mathbf{E}_{sc}(r = 0, x, t) = E_{sc}^x(x, t) \hat{x}$. This (one-dimensional) field distribution is used to determine the impact ionisation coefficients $\alpha(x, t)$ and $\beta(x, t)$. The finite transverse extent of the avalanche is thus taken into account when computing its space charge field, but it is neglected when evaluating the backaction of this field on the further evolution of the avalanche. This is analogous to the “1.5-dimensional” model of space charge effects in resistive plate chambers (RPCs) discussed in Ref. [17]. As before, the field $\mathbf{E}_{ext}$ is assumed to be position-independent, i.e. $\mathbf{E}_{ext} = -\hat{x} V_d/d$. The resulting simulation model is referred to as “MAMC + circuit + SC”.

For the results presented here, only transverse diffusion is considered for the construction of $\rho(r, x, t)$. This simplified model is likely to overestimate the true effect of space charge on the development of the avalanche. First, the space charge field $\mathbf{E}_{sc}$ itself is overestimated: the $x$-component of the field...
Figure 8: For the computation of the space charge field, the high-field region is taken to be delimited by two conducting plates of infinite transverse extent, located at \(x = 0\) and \(x = d\). Charge carriers are constrained to move along the \(x\)-axis. This charge distribution is extended into a three-dimensional charge density \(\rho(r, x_0, t)\), illustrated here in the plane \(x = x_0\). The field \(E_{\text{sc}}(r, x, t)\) is evaluated on the \(x\)-axis as the electric field generated by \(\rho(r, x, t)\). The external field \(E_{\text{ext}}(t)\) also contributes to the total electric field \(E(x, t)\) relevant for the development of the avalanche.

Originating from the chosen radially symmetric charge density \(\rho(r, x, t)\) attains a maximum at \(r = 0\) (cf. Eqs. 28 and 29 below). All drifting charge carriers are thus subject to the maximal field, while regions of lower field strengths at \(r > 0\) are not probed by the avalanche. Second, the transverse size of the avalanche is underestimated by only considering diffusion, but ignoring carrier drift. This leads to a narrower distribution of charges and also enhances the computed space charge field.

Transverse spread of the avalanche and space charge field. Charge carrier diffusion is modelled by prescribing a normal distribution for the transverse distribution of charge carriers, while the longitudinal distribution continues to be given by the densities \(n_e(x, t)\) and \(n_h(x, t)\). The densities \(\rho_e\) and \(\rho_h\) that are relevant for the computation of the space charge field are thus

\[
\rho_e(r, x, t) = -\frac{e_0 n_e(x, t)}{2\pi \sigma_e^2(t)} \exp \left[ -\frac{r^2}{2\sigma_e^2(t)} \right],
\]

(28)

\[
\rho_h(r, x, t) = \frac{e_0 n_h(x, t)}{2\pi \sigma_h^2(t)} \exp \left[ -\frac{r^2}{2\sigma_h^2(t)} \right].
\]

(29)

The transverse extent is governed by the transverse diffusion coefficients \(D_T\) for electrons and holes, i.e. \(\sigma_e(t) = \sqrt{2D_{T,e}t}\) and \(\sigma_h(t) = \sqrt{2D_{T,h}t}\). Values of \(D_{T,e} = 15 \text{ cm}^2/\text{s}\) [18] and \(D_{T,h} = 11 \text{ cm}^2/\text{s}\) [19] are used in the following. To the best of our knowledge, no measurements of the transverse diffusion coefficients are available in the literature for the very strong longitudinal fields \(|E_{\text{ext}}| > 20 \text{ V/\mu m}\) that occur above breakdown. The selected values correspond to a field of 5 \text{ V/\mu m} in silicon. The transverse diffusion coefficients depend only relatively mildly on the longitudinal electric field, and it is expected that these values give at least an approximate description of diffusion in the relevant regime. This is sufficient for the qualitative assessment of space charge effects attempted here.

The electrostatic Green’s function for the parallel-plate geometry of Fig. 8 derived in Refs. [20, 21] is used to compute the space charge field \(E_{\text{sc}}\).

Quenching dynamics with space charge. Fig. 9 shows the evolution of the carrier densities, the electric field, and the induced signal. The same simulation parameters and material properties as in Section 3 are used.
are used. At early times, space charge effects are of subleading importance, and the resulting charge carrier densities are described by the adiabatic model in good approximation. The separation of charge caused by $E_{\text{ext}}$ leads to a space charge field which reduces the electric field in the centre of the gain layer and enhances it close to its boundaries. (Space charge cannot alter the voltage $V_d$ measured across the junction contacts, i.e. $\int_{0}^{d} \text{d} s \cdot \mathbf{E}_{\text{sc}} = 0$.) This results in increased impact ionisation rates and carrier densities near the edges of the multiplication region. For the chosen material parameters and device geometry, this effect overcompensates the suppressed avalanche growth in regions with reduced field. The signal charge and the voltage step $\Delta V_d$ increase by about 50% compared to the situation without space charge effects, but the signal shape is not significantly altered.

This is very different from the space charge effects in RPCs studied in detail in Ref. [17]. Also in the RPC geometry the field integral throughout the sensitive volume remains unchanged so there will be regions with reduced field and regions with increased field. Since in a gas detector there is however only electron multiplication and no ion multiplication, the electrons end up only in these low field regions and the gain is suppressed by many orders of magnitude.

Measurements of the voltage step $\Delta V_d$ in SPADs can thus serve as a direct probe of the importance of space charge effects for the avalanche development.
5. Conclusions

In a passively quenched SPAD, the diverging avalanche discharges the detector capacitance, which in turn lowers the electric field in the multiplication region and reduces the probability for impact ionisation. The dynamics of this process is here described quantitatively based on a microscopic model of the avalanche development, applicable as long as space charge effects are negligible. The main conclusions of this treatment are:

- The duration of the avalanche discharge scales with the quenching time constant \( \tau_q \) (cf. Eq. [14]). It is determined in terms of the excess voltage \( V_{ex} \) and the parameter \( K_{br} \) (cf. Eq. [8]). The latter may be extracted starting from the field profile \( E(x) \) around the breakdown point. It is known analytically for position-independent fields (cf. Eq. [26]).

- The avalanche attains its maximum size as the voltage \( V_d \) measured across the diode crosses the breakdown voltage. The total voltage step \( \Delta V_d \) caused by the discharge is therefore \( \Delta V_d = 2V_{ex} \). The total signal charge is \( Q = 2V_{ex}C_d \), with the junction capacitance \( C_d \) (cf. Eqs. [21] and [22]).

- Analytic expressions for the evolution of \( V_d(t) \) and the current \( I^{ind}(t) \) induced by the avalanche are available in Eqs. [16] and [17]. Analytic formulae for the spatial charge carrier densities also exist.

The space charge field produced by the charge carriers in the avalanche enhances the electric field (and therefore impact ionisation) close to the boundaries of the multiplication region and reduces it in its interior. Charge carriers are present throughout the entire junction, and the avalanche growth may, in principle, get enhanced or suppressed by the altered field configuration. A simplified MC simulation of a representative silicon device geometry shows a modest enhancement of the signal charge as a result of space charge. This is in stark contrast to gas detectors such as RPCs, where space charge considerably reduces the gas gain.

Our results in Eqs. [21] and [22] deviate from the prevalent description of passive quenching, in which the avalanche discharge is modelled as a bistable switch. We encourage further experimental work to investigate and clarify this discrepancy.

Appendix

To simulate the development of the avalanche for equal carrier drift velocities, \( v_e = v_h = v^* \), the avalanche region \( x \in [0, d] \) is partitioned into \( N \) equidistant bins of width \( \Delta x = d/N \). The units are chosen such that charges drift a distance of \( \Delta x \) during the simulation time step \( \Delta t = \Delta x/v^* \). The bin centre of bin \( j \) is located at coordinate \( x^j \) and the simulation time steps are indexed as \( t^i = i\Delta t \).

The electric field distribution \( E(x, t) = -E^x(x, t) \hat{x} \) and the resulting impact ionisation coefficients \( \alpha(x, t) \) and \( \beta(x, t) \) are discretised into \( E^j(t^i) = E^x(x^j, t^i) \), \( \alpha^j(t^i) = \alpha(x^j, t^i) \), and \( \beta^j(t^i) = \beta(x^j, t^i) \). The charge carrier densities \( n_e(x, t) \) and \( n_h(x, t) \) are represented by the number of electrons \( N^j_e(t^i) \) and holes \( N^j_h(t^i) \) in each bin. These carriers are assumed to be uniformly distributed within the bin.

In the absence of impact ionisation, the \( N^j_e \) holes present in bin \( j \) at time \( t^i \) drift into bin \( j + 1 \) at \( t^{i+1} \), and the \( N^j_h \) holes move into bin \( j - 1 \). Impact ionisation is simulated as illustrated in Fig. 10. During the time interval \( \Delta t \), the \( N^j_e \) electrons (\( N^j_h \) holes) in bin \( j \) produce \( N^j_{e \rightarrow e, ch} \) (\( N^j_{h \rightarrow e, ch} \) additional electron-hole pairs. These numbers are drawn from a Poisson distribution with rate parameters \( \alpha^j \Delta x \) and \( \beta^j \Delta x \), respectively,

\[
N^j_{e \rightarrow e, ch}(t^i) \sim \text{Po} \left( N^j_e(t^i); \alpha^j(t^i) \Delta x \right),
\]

\[
N^j_{h \rightarrow e, ch}(t^i) \sim \text{Po} \left( N^j_h(t^i); \beta^j(t^i) \Delta x \right).
\]
Figure 10: (Colour online.) Illustration of the update step $t^i \rightarrow t^{i+1}$ for the evolution of electrons in (a) and for the evolution of holes in (b). The number of electrons $N^j_e$ and holes $N^j_h$ contained in a bin $j$ are represented by coloured bars within the respective bin. A number of exemplary drift lines are shown to illustrate the movement of charge carriers. Impact ionisation events are indicated by black dots.

Impact ionisation occurs uniformly along $\Delta x$. The $N^j_{e \rightarrow eh}$ electrons created in bin $j$ at $t^i$ drift into bin $j+1$ at $t^{i+1}$, while the $N^j_{h \rightarrow eh}$ holes spread across multiple bins: $N^j_{e \rightarrow eh}$ holes reach bin $j-1$, $N^j_{e \rightarrow eh}$ remain in bin $j$, and $N^j_{e \rightarrow eh}$ enter bin $j+1$. The same consideration holds for the $N^j_{h \rightarrow eh}$ electrons created from hole-initiated impact ionisation events, leading to $N^j_{h \rightarrow eh}$, $N^j_{h \rightarrow eh}$, and $N^j_{h \rightarrow eh}$ electrons in bins $j-1$, $j$, and $j+1$, respectively. These numbers are drawn from a multinomial distribution with event probabilities of $p_{j \rightarrow j-1} = 0.25$, $p_{j \rightarrow j} = 0.5$, and $p_{j \rightarrow j+1} = 0.25$, respectively,

\[
\begin{align*}
\{N^j_{e \rightarrow eh}^{-1}(t^i), N^j_{e \rightarrow eh}(t^i), N^j_{e \rightarrow eh}(t^{i+1})\} & \sim \text{Mult} \left( N^j_{e \rightarrow eh}(t^i); \{p_{j \rightarrow j-1}, p_{j \rightarrow j}, p_{j \rightarrow j+1}\} \right), \\
\{N^j_{h \rightarrow eh}^{-1}(t^i), N^j_{h \rightarrow eh}(t^i), N^j_{h \rightarrow eh}(t^{i+1})\} & \sim \text{Mult} \left( N^j_{h \rightarrow eh}(t^i); \{p_{j \rightarrow j-1}, p_{j \rightarrow j}, p_{j \rightarrow j+1}\} \right).
\end{align*}
\]

The updated charge carrier distributions are then computed as

\[
\begin{align*}
N^j_e(t^{i+1}) &= N^j_e(t^i) + N^j_{e \rightarrow eh}(t^i) + N^j_{e \rightarrow eh}(t^{i+1}) + N^j_{h \rightarrow eh}(t^i) + N^j_{h \rightarrow eh}(t^{i+1}), \\
N^j_h(t^{i+1}) &= N^j_h(t^i) + N^j_{h \rightarrow eh}(t^i) + N^j_{h \rightarrow eh}(t^{i+1}) + N^j_{e \rightarrow eh}(t^i) + N^j_{e \rightarrow eh}(t^{i+1}).
\end{align*}
\]

Assuming a uniform weighting field $E_w/V_w = 1/d$, the current $I^{\text{ind}}(t^i)$ induced by the drifting charges is computed as

\[
I^{\text{ind}}(t^i) = \frac{e_0 v^*}{d} \sum_j \left( N^j_e(t^i) + N^j_h(t^i) \right),
\]

where $e_0$ is the elementary charge.

In the continuum limit where $\alpha \cdot \Delta x \rightarrow 0$ and $\beta \cdot \Delta x \rightarrow 0$, this simulation model converges to the avalanche model described in Section 2. For the results presented in the main body, the avalanche region is divided into $N = 500$ bins.

References

[1] D. Renker, “Geiger-mode avalanche photodiodes, history, properties and problems”, Nucl. Instrum. Meth. A 567, 48–56 (2006)
[2] F. Acerbi, S. Gundacker, “Understanding and simulating SiPMs”, Nucl. Instrum. Meth. A 926, 16–35 (2019)

[3] W. Riegler, P. Windischhofer, “Time resolution and efficiency of SPADs and SiPMs for photons and charged particles”, Nucl. Instrum. Meth. A 1003, 165265 (2021)

[4] S. Seifert et al., “Simulation of Silicon Photomultiplier Signals”, IEEE Trans. Nucl. Sci. 56, 6 (2009)

[5] S. Cova, M. Ghioni, A. Lacaita, C. Samori, F. Zappa, “Avalanche photodiodes and quenching circuits for single-photon detection”, Appl. Opt. 35, 1956-1976 (1996)

[6] R. H. Haitz, “Model for the Electrical Behavior of a Microplasma”, J. Appl. Phys. 35, 1370 (1964)

[7] A. Inoue et al., “Modeling and Analysis of Capacitive Relaxation Quenching in a Single Photon Avalanche Diode (SPAD) Applied to a CMOS Image Sensor”, Sensors 20, 3007 (2020)

[8] A. Inoue et al., “Modeling and verification of capacitive quenching in a single photon avalanche diode”, Proc. SPIE 11721, Advanced Photon Counting Techniques XV, 117210O (2021)

[9] P. Windischhofer, W. Riegler, “The statistics of electron-hole avalanches”, Nucl. Instrum. Meth. A 1003, 165327 (2021)

[10] S. Ramo, “Currents induced in electron motion”, PROC IRE 27, 584 (1939)

[11] W. Shockley, “Currents to conductors induced by a moving point charge”, J. Appl. Phys. 9, 635 (1938)

[12] G. E. Stillman, C. M. Wolfe, “Avalanche Photodiodes”, Semiconductors and Semimetals, Chapter 5, Vol. 12, 291–393 (1977)

[13] R. Van Overstraeten, H. de Man, “Measurement of the ionization rates in diffused silicon p-n junctions”, Solid-St. Electron. 13, 583 (1969)

[14] C. Canali et al., “Electron and hole drift velocity measurements in silicon and their empirical relation to electric field and temperature”, IEEE Trans. Electron Dev. 22, 1045 (1975)

[15] W. C. Dunlap Jr., R. L. Watters, “Direct Measurement of the Dielectric Constants of Silicon and Germanium”, Phys. Rev. 92, 1396 (1953)

[16] M. M. Hayat, D. A. Ramirez, G. J. Rees, M. A. Itzler, “Modeling negative feedback in single-photon avalanche diodes”, Proc. SPIE 7681, Advanced Photon Counting Techniques IV, 76810W (2010)

[17] C. Lippmann, W. Riegler, “Space charge effects in Resistive Plate Chambers”, Nucl. Instrum. Meth. A 517, 54–76 (2004)

[18] D. J. Bartelink, G. Persky, “Diffusion of Electrons in Silicon Transverse to a High Electric Field”, Appl. Phys. Lett 16, 191 (1970)

[19] J. M. Hinckley, J. Singh, “Anisotropic high-field diffusion of holes in silicon”, Appl. Phys. Lett. 66, 2727 (1995)

[20] T. Heubrandtner et al., “Static electric fields in an infinite plane condenser with one or three homogeneous layers”, Nucl. Instrum. Meth. A 489, 439 (2002)

[21] W. Riegler, “Electric fields, weighting fields, signals and charge diffusion in detectors including resistive materials”, JINST 11, P11002 (2016)