Electrothermal feedback in superconducting nanowire single-photon detectors

The MIT Faculty has made this article openly available. Please share how this access benefits you. Your story matters.

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>As Published</td>
<td><a href="http://dx.doi.org/10.1103/PhysRevB.79.100509">http://dx.doi.org/10.1103/PhysRevB.79.100509</a></td>
</tr>
<tr>
<td>Publisher</td>
<td>American Physical Society</td>
</tr>
<tr>
<td>Version</td>
<td>Final published version</td>
</tr>
<tr>
<td>Accessed</td>
<td>Mon Dec 17 17:03:43 EST 2018</td>
</tr>
<tr>
<td>Citable Link</td>
<td><a href="http://hdl.handle.net/1721.1/51752">http://hdl.handle.net/1721.1/51752</a></td>
</tr>
<tr>
<td>Terms of Use</td>
<td>Article is made available in accordance with the publisher's policy and may be subject to US copyright law. Please refer to the publisher's site for terms of use.</td>
</tr>
<tr>
<td>Detailed Terms</td>
<td></td>
</tr>
</tbody>
</table>
Electrothermal feedback in superconducting nanowire single-photon detectors

Andrew J. Kerman, Joel K. W. Yang, Richard J. Molnar, Eric A. Dauler, and Karl K. Berggren

1Lincoln Laboratory, Massachusetts Institute of Technology, Lexington, Massachusetts 02420, USA
2Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

DOI: 10.1103/PhysRevB.79.100509

Superconducting nanowire single-photon detectors (SNSPDs) combine high speed, high detection efficiency (DE) over a wide range of wavelengths, and low dark counts.1–4 Of particular importance is their high single-photon timing resolution of ~30 ps,4 which permits extremely high data rates in photon-counting communications applications.5,6 Full use of this electrical bandwidth is limited, however, by the fact that the maximum count rates of these devices are much smaller (a few hundred MHz for 10 μm² active area and decreasing as the area is increased), limited by their large kinetic inductance and the input impedance of the readout circuit.2,5 To increase the count rate, therefore, one must either reduce the kinetic inductance (by using a smaller active area or different materials or substrates) or increase the load impedance.2 However, either of these approaches causes the wire to “latch” into a stable resistive state where it no longer detects photons.8 This effect arises when negative electrothermal feedback, which in normal operation allows the device to reset itself, is made fast enough that it becomes stable. We present experiments which probe the stability of this feedback, and we develop a model which quantitatively explains our observations.

The operation of an SNSPD is illustrated in Fig. 1(a). A nanowire (typically ~100 nm wide and 5nm thick) is biased with a dc current Is near its critical current Ic. The nanowire has kinetic inductance L and is read out using a load impedance RL (typically a 50Ω transmission line). When a photon is absorbed, a short (<100 nm long) normal domain is nucleated, giving the wire a resistance Rs(t). This results in Joule heating which causes the normal domain (and consequently, Rs) to expand in time exponentially. The expansion is counteracted by negative electrothermal feedback from the load RL, which forms a current divider with Rs, and diverts a current Ic into the load (so that the current in the nanowire is reduced to Ic = Is − Ic), reducing the heating. However, in a correctly functioning device, this feedback is unstable: the inductive time constant is long enough so that before Ic becomes appreciable, Joule heating has already increased Rs, so that Rs >> RL. The current Ic then drops nearly to zero, turning off the heating and allowing the nanowire to quickly cool down and return to the superconducting state, after which Ic recovers with a time constant τc = L/Rs.2 If one attempts to shorten τc too much, the negative feedback becomes fast enough to counterbalance the Joule heating before it runs away, resulting in a stable resistive domain, known as a self-heating hotspot.9,10

In a standard treatment of these hotspots,9 solutions to a one-dimensional heat equation are found in which a normal-superconducting (NS) boundary propagates at constant velocity vNS for fixed device current Is.3,11 This results in a solution of the form

\[ v_{NS} = \frac{\alpha(I_{switch}/I_s)^2 - 2}{\sqrt{\alpha(I_{switch}/I_s)^2 - 1}} = \frac{1}{\gamma (I_d - I_c^2)} \]

where \( v_0 = \sqrt{A_cKh/c} \) is a characteristic velocity (A_c is the wire’s cross-sectional area, K is its thermal conductivity, and c and h are the heat capacity and heat transfer coefficient to the substrate, per unit length, respectively), Ic is the critical current, and \( \alpha = \rho c K h / h (T_c - T_0) \) is known as the Stekly parameter, which characterizes the ratio of Joule heating to...
conduction cooling in the normal state$^9$ $\rho_n$ is the normal resistance of the wire per unit length and $T_0$ and $T_c$ are the substrate and critical temperatures. Equation (1) is valid when $T_c < T_0$, and the approximate equality holds when $|I_c-I_0| \ll I_{ss}$ with $\gamma=(T_c-T_0)/(c/\rho_0)^{1/2}/\kappa A_0$ and $I_{ss} = 2h(T_c-T_0)/\rho_0$. The physical meaning of Eq. (1) is clear: the NS boundary is stationary only if the local power density $(\approx I_d^2)$ is equal to a fixed value; if it is greater, the hotspot will expand ($v_{NS}>0$), if less it will contract ($v_{NS}<0$).

We can use Eq. (1) to describe the electrothermal circuit in Fig. 1(a), by combining it with the circuit equation $I_d R_L + L dI_d/dt = I_{th} (I_c - I_0)$ (where $dR_L/dt=2\rho_0 v_{NS}$). To determine when the device will latch, we analyze the stability of the resulting second-order nonlinear system for small deviations from its steady-state solution $[I_d-I_{th} R_L = -R_s (I_c - I_0) + I_{th} I_{th} / I_{th} = 1]$ to obtain a damping coefficient $\zeta = I_{th} / \tau_{th} = \tau_m / \tau_e$, where $\tau_m = R_L/2\rho_0 v_{th}$ is a thermal time constant. This can be re-expressed in terms of $R_{tot} = R_L + R_s$, thus: $\zeta = 1 / 4 \tau_m / \tau_{tot}$ (where $\tau_{tot} = L / R_{tot}$ and $\tau_{th} = R_s / 2\rho_0 v_{th}$), which clearly shows that the stability is determined by a ratio of electrical and thermal time constants.

In normal device operation, where the damping $\zeta$ is small, the feedback cannot stabilize the hotspot during the initial photoresponse, as described above. However, as $I_0$ is increased, $\zeta$ increases, making the hotspot more stable (this occurs because $R_s \approx I_0$ and larger $R_s$ gives a shorter inductive time constant $\tau_{ind}$). Eventually, at a bias current $I_d = I_{th}$, the device latches. For a correctly functioning device, $I_{th} > I_c$, so that latching does not affect its operation. However, if $\tau_e$ is decreased, $I_{th}$ decreases, and eventually it becomes less than $I_c$. This prevents the device from being biased near $I_c$, resulting in a drastic reduction in performance.$^{12}$

Devices used in this work were fabricated from ~5-nm-thick NbN films, deposited on R-plane substrate sapphire substrates in a UHV dc magnetron sputtering system (base pressure $<10^{-10}$ mbar). Film deposition was performed at a wafer temperature of ~800 °C and a pressure of ~10⁻⁸ mbar.$^{13}$ Aligned photolithography and liftoff were used to pattern ~100-nm-thick Ti films for on-chip resistors$^8$ and Ti:Au contact pads. Patterning of the NbN was then performed with e-beam lithography.$^3$ Devices were tested in a cryogenic probing station at 2 K as described in Refs. 2 and 3.

Figures 1(c)–1(f) show data for a set of (3 $\mu$m $\times$ 3.3 $\mu$m area) devices having various resistors $R_s$ in series with the 50Ω readout line.$^8$ [Fig. 1(b)], so that $R_s=50\Omega+R_c$. For $R_s=0$, these devices had similar performance to those in Ref. 3. Panels (c) and (d) show averaged pulse shapes for devices with $R_s=0,250\Omega$, respectively. Clearly, the reset time can be reduced; however, this comes at a price. Panels (e) and (f) show, for devices with different $R_s$, the current $I_{switch} = \min(I_c, I_{th})$ above which each device no longer detects photons and the measured DE at $I_c=0.975I_{switch}$. The data show that as $R_s$ is increased, $I_{switch}$ decreases far below $I_c$ (due to reduction in $I_{th}$), resulting in a significantly reduced DE.$^{14}$

To investigate the latched state, we fabricated devices designed to probe the stability of self-heating hotspots as a function of $I_0$, $L$, and $R_s$. Each device consisted of three sections in series, as shown in Fig. 2(a): a 3-$\mu$m-long 100-nm-wide nanowire where the hotspot was nucleated, a wider (200 nm) meandered section acting as an inductance, and a series of nine contact pads interspersed with Ti-film resistors. Also shown are the two electrical probes, which result in the circuit of Fig. 2(b): a high-impedance ($R_s=20$ kΩ) three-point measurement of $R_{tot}$. We varied $R_s$ by moving the probes along the line of contact pads and $L$ by testing different devices (with different $L$). We tested 66 devices on three chips and selected from these only unconstricted$^{12}$ nanowires with nearly identical linewidths ($I_c=22–24$ $\mu$A) and with $R_s=200\Omega–1000\Omega$ and $L=6–600$ $\mu$H.

For each $L$ and $R_s$ we acquired a dc V-I curve like those shown in Figs. 2(c) and 2(d), sweeping $I_0$ downward starting from high values where the hotspot was stable.$^{15}$ These data can be converted to $I_d$ and $R_s$, as shown in Figs. 2(e) and 2(f) [for the data of Fig. 2(d)]. From data of this kind, $I_{th}$ can be identified by the sudden jumps in $I_d$ for $I_0>I_{th}$; $I_d$ is fixed (at $I_{th}$), independent of $I_0$ and $R_s$, as predicted by Eq. (1). For the largest values of $R_s$, $I_d$ never reaches $I_c$ (shown by a horizontal dashed line in Fig. 2(e)) because once it latches $I_d \rightarrow I_{th} < I_c$. As $R_s$ is decreased, $I_{th}$ increases as expected, until another feature appears when $I_{th}>I_c$. In this region ($I_c < I_{th} < I_{th}$) the nanowire can neither superconduct nor latch and instead undergoes relaxation oscillations.$^{9,11}$ as indicated in the figure, producing a periodic pulse train with a frequency that increases as $I_0$ is increased.$^{14}$ The average resistance (from the dc V-I curve) increases with this frequency, producing the observed continuous decrease in $I_d$ until $I_{th}$ is reached.

The data in Fig. 3 show the measured $I_{th}$ as $R_s$ and $L$ are varied, plotted in dimensionless form as $2\pi / \tau_{th} = (L/R_s^2)$ vs $(I_{th}/I_0^2)$, which can be thought of as defining the boundary between stable and unstable hotspots. Our simple model de-
scribed above predicts a line of slope 1 (indicated by the dashed line). The data approach this line, although only in the \( \tau_\epsilon \gg \tau_n \) limit. This is consistent with the assumption of constant (or slowly varying) \( I_d \) under which Eq. (1) was derived. As \( \tau_\epsilon/\tau_n \) is decreased, the data trend downward, away from this line, and \( I_{\text{latch}}/I_{\text{ss}} \) becomes almost independent of \( \tau_\epsilon/\tau_n \) (all data approach the same vertical asymptote); this implies a minimum \( I_{\text{latch}}/I_{\text{ss}} \), or equivalently, a minimum \( R_c/R_L \), below which the hotspot is always unstable. This is shown in the inset: the measured minimum stable \( R_c \) is always greater than \( \sim R_{L} \), over a range of \( L \) values from 6 to 217 nH (shown by solid symbols—solid lines are guides for the eyes).

This behavior can be explained in terms of a time scale \( \tau_\epsilon \) over which the temperature profile of the hotspot stabilizes into the quasisteady-state form which yields Eq. (1). For power-density variations faster than this, the NS boundaries do not have time to start moving, resulting instead in a temperature deviation \( \Delta T \). Since the NS boundary occurs at \( T \approx T_c \), where \( p_s \) is strongly temperature dependent (defined by \( dp/dT = \beta > 0 \)), this changes \( R_c \), giving a second parallel electrothermal feedback path which dominates for frequencies \( \omega \gg \tau_\epsilon^{-1} \). We can describe this by replacing Eq. (1) with

\[
\frac{d\Delta T}{dt} = \gamma p_s \frac{d^2 \Delta T}{dt^2} \frac{d^2}{dt^2} = -\frac{1}{2} \rho h \Delta T. \tag{3}
\]

Here, \( l \) is the hotspot length, \( \rho(\Delta T) \) is the resistance per unit length [with \( \rho(0) = \rho_\infty \)], and \( R_n = \rho l \Delta T/l_t \). In Eq. (2), \( \tau_\epsilon \) is the characteristic time over which \( 2c \Delta T(\Delta T/dt) \) adapts to changes in power density: for slow time scales \( dt \gg \tau_\epsilon \), we have \( \tau_\epsilon (d^2/l^2 dt^2) \ll dt/\Delta T \) and Eq. (2) reduces to Eq. (1) with \( \Delta T = 0 \). For faster time scales, \( \tau_\epsilon (d^2/l^2 dt^2) \gg \Delta T \) becomes appreciable and acts as a source term for temperature deviations in Eq. (3). When \( dt \ll \tau_\epsilon \), \( \tau_\epsilon (d^2/l^2 dt^2) \gg \Delta T \) and Eqs. (2) and (3) can be combined to give \( c \cdot d\Delta T/ dt = \int_0^t \rho(\Delta T) dt - \int_0^t \rho_\infty \Delta T \). In this limit, if \( R_c \approx R_n \), the bias circuit including \( R_c \) begins to look like a current source, which then results in positive feedback: a current change produces a temperature and resistance change of the same sign. Therefore, the hotspot is always unstable when \( R_c \leq R_n \).

Expressing Eqs. (2) and (3) in dimensionless units (\( l = I_d/l_0, \ r = R_c/R_L, \ \lambda = \beta T_c/\rho_\infty, \ \text{and} \ \theta = T/T_c \)) and expanding to first order in small deviations (\( \delta T, \delta \theta, \delta \lambda, \delta \theta \)) from steady state, we obtain

\[
\delta \theta' = -(\delta T \delta + \delta l_0 \delta \lambda), \tag{4}
\]

\[
\delta \theta = \eta \delta l_0 - \eta \delta \lambda, \tag{5}
\]

\[
\frac{\tau_c}{\tau_\epsilon} \delta \lambda + \delta \lambda = 2 \eta \Delta T_{l_0} (\delta \theta + 2 \delta l_0 \eta^{-1} \delta \lambda), \tag{6}
\]

\[
\frac{\delta \theta'}{\delta \theta} = \frac{\Theta R_c \tau_\epsilon}{\eta \delta l_0 - \delta \lambda}. \tag{7}
\]

Here, the prime denotes differentiation with respect to \( t/\tau_\epsilon, \ l_0 = I_0/l_\infty, \ \Theta = (T_c-\theta)/T_c, \ \eta = \beta T_c/\rho_\infty \) characterizes the resistive transition slope, and \( \tau_\epsilon = c/l \) is a cooling time constant. When \( \tau_\epsilon \gg \tau_n, \ \tau_\epsilon \), the system reduces to \( \delta \theta' \sim \delta T \delta + \delta l_0 \delta \lambda, \delta \lambda = 0 \), which has damping coefficient \( \xi = \eta \delta l_0 (4 \tau_\epsilon/\tau_n)^{-1} \), as above. In the opposite limit, where \( \tau_\epsilon \ll \tau_n, \ \tau_\epsilon \), we obtain \( \delta l_0 \delta \lambda = (2 \eta \Theta \tau_c/\tau_\epsilon)(\delta l_0 - 2) = 0 \). In agreement with our argument above, the oscillation frequency becomes negative for \( R_c \approx R_c < R_{L} (l_0 < 2I_0) \).

We characterize the stability of the system of Eqs. (4)–(7) using its “open loop” gain \( A_{\text{ol}} \); we assume a small oscillatory perturbation by replacing \( \delta \theta \) in Eq. (4) with \( \Delta T \) and responses \( \delta \theta, \delta \lambda, \delta \theta_e, \delta \theta_e \). Solving for \( A_{\text{ol}} = \delta \theta / \delta \theta_e \), we obtain

\[
A_{\text{ol}} = \frac{4 \tau_\epsilon}{l_0} \left( 1 + j \omega \tau_\epsilon^2 \right) (2 \eta \Theta \tau_\epsilon - 1) (1 + j \omega \tau_\epsilon^2) \left( 1 + j \omega \tau_\epsilon^2 \right). \tag{8}
\]

The stability of the system can then be quantified by the phase margin \( \pi + \arg[A_{\text{ol}}(\omega_0)] \), where \( \omega_0 \) is the unity gain \((|A_{\text{ol}}| = 1) \) frequency. In the extreme case, when the phase margin is zero \(( \arg[A_{\text{ol}}(\omega_0)] = -\pi) \), the feedback is positive. The solid lines in Fig. 3 show our best fit to the data. Note that although the stability is determined only by \( \tau_\epsilon/\tau_n \) and \( l_0 \) in the two extreme limits (not visible in the figure), in the
intermediate region of interest here this is not the case, so several curves are shown. Each solid curve segment corresponds to a single $L$, over the range of $R_L$ tested; the dotted lines continue these curves for a wider range of $R_L$. The data are grouped into three inductance ranges: 6–12, 15–60, and 120–600 mH, indicated by circles, squares, and triangles, respectively. We used fixed values of $\Theta=0.8$ and $\eta=6.5$, which are based on independent measurements, and fitted $\tau_h=1.9$ ns and $\tau_e=0.47$ ns to all data. Separate values of $\rho_e \nu_0$ were fitted to data from each of the three chips, differing at most by a factor of $\sim 2$. These fitted values were $\rho_e \nu_0 \sim 1 \times 10^{11} \Omega/s$; since $\rho_e \sim 3 \times 10^8 \Omega/m$, this gives $\nu_0 \sim 30$ m/s, which is a reasonable value.

A natural question to ask in light of this analysis is whether it suggests a method for speeding up these devices. The most obvious way would be to increase the heat transfer coefficient $h$, which increases both $I_{\text{th}}$ and $\nu_0$, moving the wire further into the unstable region and allowing its speed to be increased further without latching. However, at present it is unknown how much $h$ can be increased before the DE begins to suffer. At some point, the photon-generated hotspot will disappear too quickly for the wire to respond in the desired fashion. In any case, experiments like those described here will be a useful measurement tool in future work for understanding the impact of changes in the material and/or substrate on the thermal coupling and electrothermal feedback.

We acknowledge helpful discussions with Sae Woo Nam, Aaron Miller, and Enectaú Figueroa-Feliciano. This work is sponsored by the United States Air Force under Contract No. FA8721-05-C-0002.


10Since the negative feedback opposes the fast joule heating that produces the rising edge of the output pulse, it may also affect the timing jitter of this edge.

11This description is further simplified by the fact that near the NS boundary all material properties can be approximated by their values at $T_c$.


14For the circuit in Fig. 1(b), if $I_{\text{latch}}>I_c$, the detector oscillates [see Fig. 2(e) and associated discussion] for the time constant of the bias tee, after which it senses the larger $R_{\text{A}} \times 5$ kΩ and latches. The absence of this burst of pulses is a signature for $I_{\text{latch}}<I_c$.

15The results were almost identical when sweeping $I_0$ upward, since the dark counts of the device allow it to lock into the latched state if it is stable.

16From the observed $I_{\text{th}} = 5 \mu A$ and $\rho_e \sim 3 \times 10^9 \Omega m^{-1}$. Eq. (1) gives $h \sim 5 \times 10^{-3} W m^{-1} K^{-1}$; this gives $\alpha \sim 30$ ($T_c=10 K$, $T_0=2 K$, and $I_c=22 \mu A$).


18This equation is identical to that governing a transition-edge sensor under the action of negative electrothermal feedback; see, e.g., K. D. Irwin, G. C. Hilton, D. A. Wollman, and J. M. Martinis, J. Appl. Phys. 83, 3978 (1998).

19From the fitted $\tau_h=0.47$ ns and the $h$ inferred above (Ref. 16), we obtain $c=2.2 \times 10^{-12} J m^{-1} K^{-1}$. Reference 8 gives $c_{\text{el}}=1.2 \times 10^{-12}$ and $c_{\text{eh}}=4.9 \times 10^{-12} J m^{-1} K^{-1}$.