Copyright 2007 American Institute of Physics. This article may be downloaded for personal use only. Any other use requires prior permission of the author and the American Institute of Physics. The following article appeared in APPLIED PHYSICS LETTERS 90, 101110 (2007) and may be found at http://dx.doi.org/10.1063/1.2696926.

Constriction-limited detection efficiency of superconducting nanowire single-photon detectors

Andrew J. Kerman[a)]
Lincoln Laboratory, Massachusetts Institute of Technology, Lexington, Massachusetts, 02420

Eric A. Dauler, Joel K. W. Yang, Kristine M. Rosfjord, Vikas Anant, and Karl K. Berggren
Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Massachusetts, 02139

Gregory N. Gol'tsman and Boris M. Voronov
Moscow State Pedagogical University, Moscow 119345, Russia

(Received 22 November 2006; accepted 20 January 2007; published online 5 March 2007)

We investigate the source of the large variations in the observed detection efficiencies of superconducting nanowire single-photon detectors between many nominally identical devices. Through both electrical and optical measurements, we infer that these variations arise from "constrictions:" highly localized regions of the nanowires where the effective cross-sectional area for superconducting current is reduced. These constrictions limit the bias-current density to well below its critical value over the remainder of the wire, and thus prevent the detection efficiency from reaching the high values that occur in these devices when they are biased near the critical current density. © *2007 American Institute of Physics*. [DOI: 10.1063/1.2696926]

Superconducting nanowire single-photon detectors[1–4] (SNSPDs) provide a unique combination of high infrared detection efficiency [up to 57% at 1550 nm (Ref. 2) has been demonstrated] and high speed (<30 ps timing resolution[3,5] and few-nanosecond reset times[4]). Some applications for these devices already being pursued include high data-rate interplanetary optical communications,[6] spectroscopy of ultrafast quantum phenomena in biological and solid-state physics,[7,8] quantum key distribution,[9] and noninvasive, high-speed digital circuit testing.[10]

In many of these applications, large arrays of SNSPDs would be important.[5] For example, existing SNSPDs have small active areas, making optical coupling difficult and inefficient.[7,11] Their small size also limits the number of optical modes they can collect, which is critical in applications where photons are distributed over many modes, such as laser communication through the atmosphere and fluorescence detection. Furthermore, the maximum count rate for an individual SNSPD decreases as its active area is increased, due to its kinetic inductance,[4] forcing a trade-off between active area and high count rates. Detector arrays could solve these problems, giving larger active areas while *increasing* the maximum count rate by distributing the flux over many smaller (and therefore faster) pixels. Large arrays could also provide spatial and photon-number resolution. Although few-pixel detectors have been demonstrated,[5,9,11] fabrication and readout methods scalable to large arrays have not yet been discussed.

A first step towards producing large arrays of SNSPDs is to understand (and reduce) the large observed variation of detection efficiencies (DE) for nominally identical devices,[2] which would set an unacceptable limit on the yield of efficient arrays of any technologically interesting size. In this letter, we demonstrate that these DE variations can be understood in terms of "constrictions:" localized regions where the nanowire cross section is effectively reduced.

[a)]Electronic mail: ajkerman@ll.mit.edu

The electrical operation of these detectors has been discussed previously,[1–4] so we only summarize it here. The nanowires are biased with a current I_{bias} slightly below the critical value I_C. An incident photon produces a resistive "hotspot" which in turn disrupts the superconductivity across the wire, resulting in a series resistance which expands in size due to Joule heating.[4] The series resistance quickly becomes $\geqslant 50\ \Omega$, and the current is diverted out of the device and into the $50\ \Omega$ transmission line connected across it, resulting in a propagating voltage pulse on the line. The device can then return to the superconducting state, and the current through it recovers with the time constant $L_k/50\ \Omega$, where L_k is the kinetic inductance.[4]

The nanowires used in this work were patterned at the shared scanning-electron-beam-lithography facility in the MIT Research Laboratory of Electronics using a process described in Ref. 2 on ultrathin ($\sim$5 nm) NbN films grown at Moscow State Pedagogical University.[12] The devices had critical temperatures $T_C \sim 9-10$ K, and critical current densities $J_C \sim 5\times 10^{10}\ \text{A/m}^2$ at T=1.8 K. The testing was performed at MIT Lincoln Laboratory, using the procedures and apparatus discussed in Refs. 2 and 4. Briefly, the devices were cooled inside a cryogenic probing station, and electrical contact was established using a cooled $50\ \Omega$ microwave probe attached to a micromanipulator. We counted electrical pulses from the detectors using room temperature low-noise amplifiers and a gated pulse counter. To optically probe the devices, we used a 1550 nm modelocked fiber laser that was attenuated and sent into the probing station via an optical fiber. The devices were illuminated from the back (through the sapphire substrate) using a lens attached to the end of the fiber which was mounted to an automated micromanipulator. The focal spot had a measured e^{-2} radius of $\sim$25 μm.

Figure 1 illustrates the DE variations observed on a single chip of 132 devices of the same geometry. These devices are the same ones reported in Ref. 2 before the optical cavities were added (maximum DE after the addition of cavities was 57%). In panel (a) we show a histogram of the measured DEs at $I_{\text{bias}}=0.975 I_C^{\text{obs}}$ (where I_C^{obs} is the observed

 © 2007 American Institute of Physics
Downloaded 26 Mar 2007 to 18.62.15.130. Redistribution subject to AIP license or copyright, see http://apl.aip.org/apl/copyright.jsp

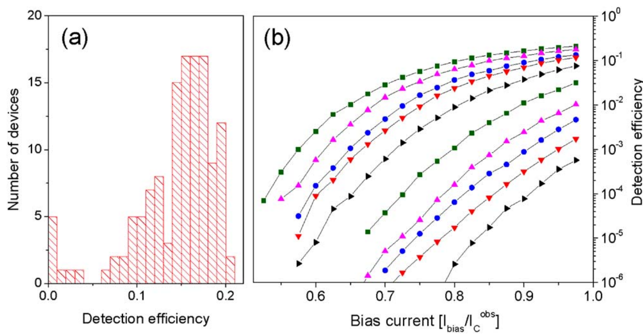

FIG. 1. (Color online) Variations in SNSPD detection efficiency. (a) Histogram of the DEs measured for 132 devices 3×3.3 μm in size, and composed of a 50 μm long, 90 nm wide nanowire in a meander pattern with 45% fill factor. The measurements were made at $T=1.8$ K and $I_{bias}=0.975I_C^{obs}$. (b) Measurements of DE vs I_{bias}/I_C^{obs} for a selection of these devices.

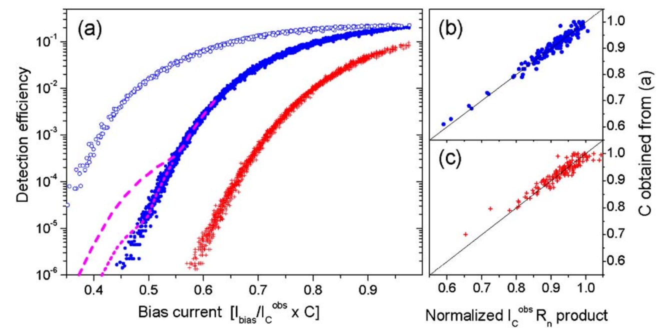

FIG. 2. (Color online) Constriction values extracted using DE vs I_{bias} data. (a) Universal DE curves for devices with 90 nm wide wires at $T=1.8$ K (●) and $T=4.2$ K (+). Data from 170 devices distributed over two separate fabrication runs are shown superposed. By rescaling the I_C^{obs} of each device such that all data lie on a single curve as shown, the constriction C can be obtained. Also shown (○) are data for 15 devices having 54 nm wide wires (and 36% fill factor) at $T=1.8$ K, indicating that narrower wires exhibit a different universal curve shape. These data provide evidence for the localized nature of the constrictions, since any appreciably long section of the wire having a smaller cross section should significantly alter the shape of the curves, making it impossible to superpose them as shown. This is illustrated by the broken lines, which show an estimate for a 90 nm wide wire having a 54 nm wide constriction with a length of 0.5 μm (dashed line) or 50 nm (dotted line). These estimates are obtained by adding together the universal curves for the two wire widths, in the ratios (0.5 μm/50 μm) and (0.05 μm/50 μm). [(b) and (c)] C values obtained from the data in (a) [for (b) $T=1.8$ K and (c) $T=4.2$ K] vs those obtained using the $I_C^{obs}R_n$ product. The C values in both cases are normalized absolutely using $L_k(I_{bias})$, as described below (see Fig. 3). The solid lines are straight lines through the origin with slope 1; no fitting was used.

critical current of each device), and (b) shows some representative data of the observed DE versus I_{bias}/I_C^{obs}. Note that the shape of these curves varies significantly. As we show below, these data can be explained with the hypothesis that some devices have constrictions: regions where the (superconducting) cross-sectional area A_{cs} of the wire is reduced by a factor we label C. This effectively reduces I_C^{obs} by that same factor ($I_C^{obs}=J_CA_{cs}C=I_CC$) and prevents the current density J everywhere but near the constriction from approaching J_C (and hence prevents the wire from having a high DE except locally near the constriction).

If all the nanowires were identical save for constrictions, we would expect that if the data of Fig. 1(b) were plotted versus the absolute current I_{bias} (rather than I_{bias}/I_C^{obs}) it would all lie on a single, universal curve, with the data for more constricted devices simply not extending to as high currents. This turns out to be approximately true, but variations across the chip either in film thickness or in the nanowire width obscure this feature of the data. To exhibit it more clearly, we present our results as shown in Fig. 2. In panel (a), the DE data for each of 170 devices with 90 nm wide wires (across two chips fabricated in separate runs) are shown superposed (filled circles indicate $T=1.8$ K and crosses $T=4.2$ K). All of the data for each temperature can be made to lie on a single universal curve by scaling I_C^{obs} for each device by an adjustable factor (which is just $1/C$: $I_C=I_C^{obs}/C$). The very fact that data from this many devices can be so well superposed already indicates a universal shape. However, we can now take the C extracted for each device and cross-check it. Based on our previous discussion, if all wires were identical save for constrictions, we would expect $C=I_C^{obs}/I_C$, i.e., C to be exactly proportional to I_C^{obs}. Due to the variations across a chip discussed above, this is only partially true. However, we can normalize out these variations using a very simple method: instead of comparing C directly to I_C^{obs}, we instead compare it to the product $I_C^{obs}R_n=(C\times J_CA_{cs})(\rho_nl/A_{cs})=C\times J_C\rho_nl$, where A_{cs} and R_n are the cross-sectional area and room-temperature resistance of each nanowire, J_C and ρ_n are the critical current density and room-temperature resistivity of NbN, respectively, and l is the total wire length. This product depends on the wire geometry only through l (which is fixed lithographically and does not vary appreciably between devices) and not on each wire's individual A_{cs}. Figures 2(b) and 2(c) show a comparison between the C values extracted from the data in Fig. 2(a) and $I_C^{obs}R_n$. The data lie on a straight line through the origin, indicating that these two independent measures of C are self-consistent.

Our data can also be used to infer that the constrictions are very short in length. The open circles in Fig. 2(a) are data for 15 devices with 54 nm wide wires, and clearly show a different shape. The broken lines are predictions, based on the data for 90 and 54 nm wide wires, for a device having 90 nm wide wire, except at a single constriction 54 nm wide ($C\sim0.6$), and either 0.5 μm (dashed line) or 50 nm (dotted line) long. These curves have a different shape from the data for 90 nm wide wires, which should be distinguishable if it were present, and should prevent the data from being superposed onto a single curve. The absence of this in our data indicates that the constricted regions are likely much shorter than $\sim$0.5 μm.

So far, we have, only measured C in a relative sense; that is, we cannot tell if our best devices, in fact, have $C=1$. To address this, we can exploit the fact that kinetic inductance increases locally as $J\rightarrow J_C$ (due to the depletion of Cooper pair density and the consequent local increase in pair velocity necessary to maintain the current). The total kinetic inductance of the wire therefore provides a way to determine if J is indeed near J_C over the whole wire or only at one localized place. We measured the inductance of our nanowires by observing the phase of a reflected microwave signal versus frequency and then fitting these data with a suitable electrical model. A bias tee was inserted into the signal path to superpose I_{bias} with the microwaves. The phase contributions from the coaxial cable, bias tee, and microwave probe were removed by probing an *in situ* calibration standard (GGB Industries CS-5). The microwave power used corresponded to a peak current amplitude of $\leqslant0.5$ μA, and the critical current measured in the presence of the microwaves was within 10% of that measured in their absence ($\sim$20 μA for typical devices at $T=1.8$ K).

Downloaded 26 Mar 2007 to 18.62.15.130. Redistribution subject to AIP license or copyright, see http://apl.aip.org/apl/copyright.jsp

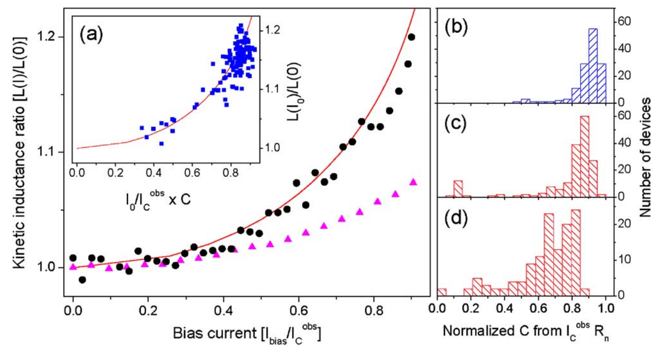

FIG. 3. (Color online) Absolute measurement of C using I_{bias} dependence of kinetic inductance. (a) The kinetic inductance of our nanowires should increase with I_{bias} (solid line). A detector with the highest DE on the chip (22%) behaves as expected (filled circles), with no free parameters. However, a detector with much lower DE (0.1%) does not (filled triangles). This is due to a constriction, which prevents J from approaching J_C except near one localized point. Inset: the inductance ratio $\mathcal{R}_L \equiv L(I_0 \approx 0.9 I_C^{\text{obs}})/L(0)$ measured for each device (filled squares) plotted vs $I_0/(I_C^{\text{obs}}/C) = I_0/I_C$, where C is obtained from the $I_C^{\text{obs}} R_n$ product. These data agree with the prediction (solid line), indicating that $\mathcal{R}_L$ and $I_C^{\text{obs}} R_n$ give mutually consistent results for the constriction. [(b)–(d)] Distribution of C values obtained using purely electrical measurements, for (b) the same set of devices shown in Fig. 1 and [(c) and (d)] for 310 additional devices on a separate chip, with active areas of both (c) 3×3.3 μm and (d) 10×10 μm. These data were obtained from $I_C^{\text{obs}} R_n$, with their normalization set by single measurements of $L_k(I_{\text{bias}})$ like that shown in (a). The devices from (b) and (c) are nominally identical, though they were fabricated on different NbN films. The difference between (c) and (d) indicates that larger devices have a higher probability of constriction.

In Fig. 3(a), we show the measured inductance versus I_{bias} for two devices: one which has nearly the highest DE observed on this chip (22%, filled circles), and the other having one of the lowest (0.1% filled triangles). Also shown is the prediction for $L_k(I_{\text{bias}})$ from Ginsburg-Landau theory with no free parameters (solid line). The data for the high-DE device show good agreement with this prediction, indicating that this device is indeed unconstricted. However, for the low-DE device the inductance increases less than predicted, due to a constriction: J is only near J_C at one localized place whereas everywhere else J is lower, producing a smaller total increase in inductance. The factor by which I_C^{obs} must be rescaled for a given device so that the $L_k(I_{\text{bias}})$ data match the prediction constitutes an absolute measurement of C. One such measurement for a single representative device, from among a large set of nominally identical devices, then allows us to correctly normalize the C values obtained using either of the previous methods described above (see Fig. 2) for all other devices in that set.

We can also verify that the observed $L_k(I_{\text{bias}})$ and $I_C^{\text{obs}} R_n$ product give mutually consistent results. To check this, for each device we measured the inductance ratio $\mathcal{R}_L \equiv L(I_0)/L(0)$, where $I_0 \approx 0.9 I_C^{\text{obs}}$. Using the C obtained from the normalized $I_C^{\text{obs}} R_n$ product, we also obtain $I_0/I_C = I_0 C / I_C^{\text{obs}}$ for each device. The inset in Fig. 3(a) shows $\mathcal{R}_L$ vs I_0/I_C (filled squares), and the data are in reasonable agreement with the Ginsburg-Landau prediction (solid line).

In addition to providing evidence for constriction, measurement of $L_k(I_{\text{bias}})$ and $I_C^{\text{obs}} R_n$ provide a powerful diagnostic tool, since they constitute a purely electrical measurement of C, which can then be used to predict the DE [Figs. 2(b) and 2(c)]. Since optical testing of many detectors is significantly more difficult and time consuming than electrical testing, this is potentially an important screening technique. An example of data obtained this way is shown in Figs. 3(b)–3(d). Panel (b) shows data for the set of devices from Fig. 1(a); panels (c) and (d) show results for a sample of 310 additional devices on another chip, all having 90 nm wide wires on a 200 nm pitch, but with active areas of both (c) 3×3.3 μm and (d) 10×10 μm. The devices from (b) and (c), which are nominally identical, exhibit very similar C distributions (though some yield fluctuation, which we commonly observe, is evident). Panels (c) and (d), however, clearly show that the larger area devices have a lower average C, which is qualitatively consistent with some fixed area density of constrictions.

As a final note, we remark on the origin of these constrictions. The most natural explanation would be lithographic patterning errors; e.g., a localized narrow section of the wire. However, we have performed extensive scanning electron microscopy of the devices that were measured to be severely constricted (e.g., $C \sim 0.5$) and no such errors were observed. This suggests that constrictions in our devices result either from thickness variations or material defects, which may have been present in the film before patterning, and may even be due to defects present in the substrate itself before film growth.

In conclusion, we have verified that the large variations in detection efficiency between nominally identical superconducting nanowire single-photon detectors are the result of localized constrictions which limit the device current. Further work is ongoing to pin down the exact source of these constrictions, with the hope of eventually fabricating large arrays of these detectors.

This work is sponsored by the United States Air Force under Contract No. FA8721-05-C-0002.

[1] G. Goltsman, O. Okunev, G. Chulkova, A. Lipatov, A. Dzardanov, K. Smirnov, A. Semenov, B. Voronov, C. Williams, and R. Sobolewski, IEEE Trans. Appl. Supercond. **11**, 574 (2001).

[2] K. M. Rosfjord, J. K. W. Yang, E. A. Dauler, A. J. Kerman, V. Anant, B. M. Voronov, G. N. Gol'tsman, and K. K. Berggren, Opt. Express **14**, 527 (2006).

[3] J. Zhang, W. Slysz, A. Verevkin, O. Okunev, G. Chulkova, A. Korneev, A. Lipatov, G. N. Gol'tsman, and R. Sobolewski, IEEE Trans. Appl. Supercond. **13**, 180 (2003).

[4] A. J. Kerman, E. A. Dauler, W. E. Keicher, J. K. W. Yang, K. K. Berggren, G. N. Gol'tsman, and B. M. Voronov, Appl. Phys. Lett. **88**, 111116 (2006).

[5] E. A. Dauler, B. S. Robinson, A. J. Kerman, J. K. W. Yang, K. M. Rosfjord, V. Anant, B. Voronov, G. Gol'tsman, and K. K. Berggren, IEEE Trans. Appl. Supercond. (to be published).

[6] B. S. Robinson, A. J. Kerman, E. A. Dauler, R. J. Barron, D. O. Caplan, M. L. Stevens, J. J. Carney, S. A. Hamilton, J. K. W. Yang, and K. K. Berggren, Opt. Lett. **31**, 444 (2006).

[7] R. H. Hadfield, M. J. Stevens, S. S. Gruber, A. J. Miller, R. E. Schwall, R. P. Mirin, and S. W. Nam, Opt. Express **13**, 10846 (2005).

[8] M. J. Stevens, R. H. Hadfield, R. E. Schwall, S. W. Nam, R. P. Mirin, and J. A. Gupta, Appl. Phys. Lett. **89**, 031109 (2006).

[9] M. A. Jaspan, J. L. Habif, R. H. Hadfield, and S. W. Nam, Appl. Phys. Lett. **89**, 031112 (2006).

[10] A. Korneev, A. Lipatov, O. Okunev, G. Chulkova, K. Smirnov, G. Gol'tsman, J. Zhang, W. Słysz, A. Verevkin, and R. Sobolewski, Microelectron. Eng. **69**, 274 (2003).

[11] W. Słysz, M. Węgrzecki, J. Bar, P. Grabiec, M. Górska, V. Zwiller, C. Latta, P. Bohi, I. Milostnaya, O. Minaeva, A. Antipov, O. Okunev, A. Korneev, K. Smirnov, B. Voronov, N. Kaurova, G. Goltsman, A. Pearlman, A. Cross, I. Komissarov, A. Verevkin, and R. Sobolewski, Appl. Phys. Lett. **88**, 261113 (2006).

[12] S. Cherednichenko, P. Yagoubov, K. Il'in, G. Gol'tsman, and E. Gershenzon, Proceedings of the Eight International Symposium On Space Terahertz Technology, Harvard University, Cambridge, MA, 1997, pp. 245–252.

Downloaded 26 Mar 2007 to 18.62.15.130. Redistribution subject to AIP license or copyright, see http://apl.aip.org/apl/copyright.jsp