Nguyen, H.Q.; Meschke, Matthias; Courtois, H; Pekola, J.P.

Sub-50-mK Electronic Cooling with Large-Area Superconducting Tunnel Junctions

Published in:
Physical Review Applied

DOI:
10.1103/PhysRevApplied.2.054001

Published: 04/11/2014

Please cite the original version:
I. INTRODUCTION

When a normal-metal–insulator–superconductor (N-I-S) junction is biased near the superconducting gap, hot electrons tunnel into the superconductor and the normal metal is cooled down [1–6]. Based on this effect, micron-size electronic coolers attached to a dielectric platform [7] can cool an external object at temperatures well below 300 mK, a regime that traditionally belongs to macroscopic cryostats, such as adiabatic demagnetization or $^3$He-$^4$He dilution refrigerators. In the best demonstration to date [8], a superconducting transition edge sensor has been cooled on a platform attached to N-I-S junctions from 300 to 200 mK. Clearly, such a platform is of great interest as an alternative method to bring ultrasensitive low-temperature detectors [9], including those at the frontier of astronomical studies of dark matter [10], neutrinos [11], or the cosmic microwave background [12], into their proper working temperature. It allows those devices to work from a bath temperature higher than their own critical temperatures. Moreover, for space-borne applications [13], N-I-S coolers could reduce the payload noticeably. The ability to reach the sub-50-mK regime with N-I-S coolers is crucial in order to open up new possibilities to cool qubits [14] and nanomechanical resonators [15] to their quantum ground states or to evacuate heat in electron-pumping devices [16]. This fact extends to the next generation of ultralow-temperature detectors, like superconducting quantum interference proximity transistor sensors [17] that are even more sensitive once operated in the sub-50-mK temperature regime.

To cool well, the normal metal needs to be isolated from the environment, and, at the same time, the hot superconducting leads need to be thermalized with the surrounding bath. In the normal metal, electrons interact with the lattice phonons, with a coupling strength decaying quickly towards low temperatures as $T^3$. Under some conditions, the lattice phonons can decouple from the thermal bath of the substrate phonons [18–20]. In the superconducting leads, hot quasiparticles at an energy just above the energy gap are generated [3,4,21]. It is a challenge to thermalize them, as quasiparticle relaxation rates decrease exponentially with lowering the temperature. Typically, these hot particles can be evacuated to a quasiparticle trap, which is a layer of normal metal in close contact with the superconductor [22–24].

When assuming that the electronic populations in both the normal metal and the superconductor can be described by Fermi distributions at respective temperatures $T_N$ and $T_S$, the cooling power of a N-I-S junction at its optimum cooling bias $eV = \Delta - 0.66k_BT_N$ is given by [3]

$$\dot{Q}_{NIS} \approx \frac{\Delta^2}{e^2R_T} \left[ 0.59 \left(\frac{k_BT_N}{\Delta}\right)^{3/2} - \sqrt{\frac{2\pi k_BT_S}{\Delta}} e^{-\Delta/k_BT_S} \right]. \tag{1}$$

Here, $R_T$ is the tunnel resistance, $\Delta$ is the superconducting gap, $k_B$ is the Boltzmann constant, and $e$ is the electron charge. If the superconducting lead is not properly thermalized so that $T_S > T_{\text{bath}}$, the term $\exp(-\Delta/k_BT_S)$ approaches the superconducting transition temperature $T_c$, the term $e^{-\Delta/k_BT_S}$ becomes significant and $\dot{Q}_{NIS}$ diminishes. In many cases, one assumes no overheating, and at low temperature one can neglect the second term; thus, $\dot{Q}_{NIS} \propto T_N^{-3/2}$. In the low-temperature regime $T_N \ll T_c$, the efficiency of the cooler is then given by [3]
\[
\dot{Q}_{NIS} = \frac{\dot{Q}_{NIS}}{TV} = 0.7 \frac{T_N}{T_c}.
\]

It amounts to about 20% near \( T_N = 350 \text{ mK} \) for aluminum, which is the standard choice of a superconductor. In general, the most significant opposing heat current to \( \dot{Q}_{NIS} \) comes from the electron-phonon interaction in the normal metal. The most accepted form for a metal is written
\[
\dot{Q}_{e-ph} = \Sigma V(T_e^5 - T_{ph}^5),
\]
where \( \Sigma = 2 \times 10^9 \text{ W K}^{-5} \text{ m}^{-3} \) for Cu, \( V \) is the volume of the normal island, and \( T_e \) and \( T_{ph} \) are the electron and phonon temperature, respectively.

Recently, we have developed a technique [25] to fabricate large-area \( S-I-N-I-S \) coolers targeted at optimizing both \( \dot{Q}_{e-ph} \) and \( \dot{Q}_{NIS} \). First, the cooled normal metal is suspended on top of the superconducting electrodes and, thus, quite decoupled from the substrate phonons. Second, hot quasiparticles in the leads are efficiently thermalized with a normal-metal drain coupled to the superconductor through a transparent tunnel barrier [26].

In this paper, we show that these two advanced features, combined with an optimized tunnel junction transparency, improve the performance of a \( S-I-N-I-S \) cooler significantly. At intermediate temperatures where electron-phonon coupling is substantial, phonons in the suspended normal metal are cooled. At low temperatures, where the cooler is almost free from electron-phonon interaction, we tune the overheating in the superconducting leads by adjusting the transparency of the cooling junctions tunnel barrier. The most efficient \( S-I-N-I-S \) cooler reaches about 30 mK, which is about 3% of \( \Delta/k_B \). We discuss the heat transport and the benefit of having a quasiparticle drain coupled to the superconductor, as well as the possible reasons for the saturation at the lowest temperature.

**II. FABRICATION AND MEASUREMENT METHODS**

By following Ref. [25], the devices are fabricated by using photolithography and metal wet etch; see Fig. 1(a) for a schematic view and the inset in Fig. 1(b) for a top view of the samples. In this work, all coolers have a suspended normal metal bridging two 200-nm-thick Al superconducting electrodes sitting on top of a 200-nm-thick AlMn quasiparticle drain [26]. The Al and AlMn layers are separated by a thin AlO\(_{2}\) layer, oxidized in a mixture of Ar:O\(_2\), ratio 10:1 at pressure \( 2 \times 10^{-2} \text{ mbar} \) for 2 min; see Table I for more sample parameters. All samples except C2 (see below) have in addition (to the drain) a quasiparticle trap of Cu next to the junction. Coolers are measured with the standard four-probe technique in a dilution cryostat. The electronic temperature \( T_N \) on the normal metal is

![FIG. 1.](image)

**TABLE I.** Parameters of the measured \( S-I-N-I-S \) coolers. \( P_{O_2}/I_{O_2} \) refers to the oxidation pressure and time used for producing the tunnel barrier of the cooler. \( l \times w \) is the length \& width of a rectangular \( N-I-S \) junction. We write the area for A5 in \( \mu \text{m}^2 \), as it has an interdigitated shape. \( d_{\text{Cu}} \) is the thickness of the normal-metal island. \( 2R_T \) is the tunnel resistance of the two \( N-I-S \) cooling junctions in series. All samples gap \( 2\Delta = 375 \mu \text{eV} \). “Figure” indicates the figure where the data on this sample are shown.

<table>
<thead>
<tr>
<th>Samples</th>
<th>( P_{O_2}/I_{O_2} ) (mbar/min)</th>
<th>( l \times w ) (( \mu \text{m} \times \mu \text{m} ))</th>
<th>( d_{\text{Cu}} ) (nm)</th>
<th>( 2R_T ) (( \Omega ))</th>
<th>Figure</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>1.3/5</td>
<td>70 \times 4</td>
<td>100</td>
<td>1.7</td>
<td>1, 2, 3</td>
</tr>
<tr>
<td>A2</td>
<td>1.3/5</td>
<td>70 \times 10</td>
<td>100</td>
<td>0.75</td>
<td>1</td>
</tr>
<tr>
<td>A3</td>
<td>1.3/5</td>
<td>70 \times 2</td>
<td>100</td>
<td>3.4</td>
<td>1</td>
</tr>
<tr>
<td>A4</td>
<td>1.3/5</td>
<td>70 \times 4</td>
<td>20</td>
<td>1.2</td>
<td>1</td>
</tr>
<tr>
<td>A5</td>
<td>1.3/5</td>
<td>950</td>
<td>100</td>
<td>0.8</td>
<td>1</td>
</tr>
<tr>
<td>B</td>
<td>9/5</td>
<td>70 \times 4</td>
<td>60</td>
<td>3.7</td>
<td>2</td>
</tr>
<tr>
<td>C1</td>
<td>13/5</td>
<td>70 \times 4</td>
<td>60</td>
<td>4.8</td>
<td>2, 3</td>
</tr>
<tr>
<td>C2</td>
<td>13/5</td>
<td>70 \times 4</td>
<td>60</td>
<td>4.6</td>
<td>2</td>
</tr>
<tr>
<td>D</td>
<td>50/120</td>
<td>70 \times 4</td>
<td>60</td>
<td>10.5</td>
<td>2, 3</td>
</tr>
</tbody>
</table>
measured by using a pair of smaller N-I-S junctions [3], where the voltage drop under a constant current (about 1 nA) is calibrated to the cryostat temperature. Below the lowest cryostat temperature of 50 mK, the temperature $T_N$ is extracted based on an extrapolation of BCS theory using data from the higher-temperature regime. The measurement noise-related uncertainty is less than 300 μK; see Fig. 2. Extracting the temperature from fitting the current-voltage characteristic to isothermal theory curves [20,22] gives identical results, including those below the lowest temperature of the cryostat where the difference is smaller than 1 mK between the two methods.

III. EXPERIMENTAL RESULTS

Figure 1(b) demonstrates that, compared to the cooling power, electron-phonon interactions are negligible in our devices. Here, all coolers are made of the same AlMn/AlOx/Al/AlOx/Cu multilayer from a single wafer, but with a varying junction geometry. Taking A1 as the reference sample, the A2 junction has a double junction size [27] and A3 half of that in A1. A4 has the same junction size as A1, but its Cu thickness is 20% of the value in A1. The junction in sample A5 has an interdigitated shape, where quasiparticle traps surround the N-I-S junctions. As $Q_{NIS} \propto R_T^3$ and $Q_{e-ph} \propto V$, varying the junction size and the normal-metal volume can give information about the heat balance. Nonetheless, it is hard to see a real trend in the data, even near 300 mK, where $Q_{e-ph} \sim \Sigma T_{ph}^5$ is expected to be significant [28]. We conclude that the present electronic coolers are almost free of electron-phonon interaction so that $Q_{e-ph}$ is not responsible for the saturation of the normal-metal temperature $T_{N,\text{min}}$ in the low-temperature end. As a consequence, adjusting the normal-metal geometry does not improve its low-temperature performance.

Our main practical achievement is presented in Fig. 2: Starting from a 150-mK bath temperature, the most powerful coolers reach a 30-mK electronic temperature, a fivefold reduction of temperature. Samples A1, B, C1, and D are made by using an identical recipe and differ only in the cooler barrier resistance $2R_T$ with respective values 1.7, 3.7, 4.8, and 10.5 Ω. A smaller $R_T$ leads to a larger cooling power, which is beneficial. Nevertheless, it also leads to a stronger quasiparticle injection, which overheats the superconducting leads and degrades cooling at low temperature. Adjusting the tunnel resistance $R_T$ is therefore essential for optimizing electronic cooling. Within our sample set, sample A1 with a large cooling power works best near 300 mK but saturates at 94 mK. Sample B reaches a lower temperature of 60 mK thanks to a smaller dissipation by the injection current. Sample D has the lowest cooling power at high temperature but cools down to 32 mK. Sample C1 is a compromise between B and D and performs well over a wide temperature range. We conclude that the higher cooling power that is desirable at high temperatures compromises the performance of the device at the low-temperature end, due to the backflow of heat from the overheated leads.

In order to investigate the limits of electronic cooling in the samples, we have improved sample C1 in a number of ways, so that it is afterwards called C2. First, C2 is equipped with a pair of direct quasiparticle traps, which locate 1 μm away from the junction. The superconducting leads are then affected by the direct contact with the normal metal and are thus expected to conduct heat better [23,24]. This modification would improve the performance of C2 if the quasiparticle drain were limiting the performance in C1. Second, we bond sample C2 with Au wires, which have a much higher thermal conductance as compared to the usual superconducting Al wires employed in other samples. Finally, we measure sample C2 in a rf-tight double-shielded sample stage [29]. These improvements work toward eliminating extra heating from quasiparticles and phonons as well as the radiation from the high-temperature parts of the cryostat. Despite all of the effort, sample C2 performs only slightly better than C1 in the whole temperature range covered. This result implies that the drain constitutes by itself an efficient quasiparticle trap that crucially helps an electronic cooler to reach 30 mK.

Let us now compare the cooler performance to theoretical predictions. We first assume a good thermalization of the metal phonons to the bath temperature: $T_{ph} = T_{bath}$. Figure 3(a) shows the efficiency $Q_{e-ph}/IV$ calculated within this framework for samples A1, C1, and D at the optimum
cooling point. The measured quantity exceeds the prediction of Eq. (2) over a wide temperature range. In the inset, the comparison of the N-I-S cooling power Eq. (1) with the electron-phonon coupling power Eq. (3) confirms this conclusion for sample C1: The two curves cross near $T_N = 220 \text{ mK}$, above which $\dot{Q}_{e-ph} > \dot{Q}_{NIS}$. Here we assumed no overheating in the superconductor: $T_S = T_{bath}$ in Eq. (1), which gives an upper estimate for $\dot{Q}_{NIS}$. As the excess efficiency occurs at intermediate temperatures near $T_{bath} = 300 \text{ mK}$, it is best explained by assuming that not only the electrons but also the phonons of the normal island cool, i.e., $T_{ph} < T_{bath}$. Our samples are particularly suited for this cooling to take place, as the normal-metal island is suspended on the superconducting electrodes and thus decoupled from the substrate. In order to estimate the phonon cooling, we calculate the drop in phonon temperature $\Delta T_{ph} = T_{bath} - T_{ph}$ necessary to fulfill the heat balance Eq. (2); see Fig. 3(b). A phonon cooling of about 20 mK is obtained in samples A1 and C1. Again, the probable overestimation of $\dot{Q}_{NIS}$ makes the present phonon-temperature drop estimation a minimum value.

IV. ANALYSIS OF THE RESULTS AND DISCUSSION

To illustrate the benefit of having an efficient quasiparticle drain, we now analyze the heat transport by using a diffusion model sketched in Fig. 4(a). In the $y$ direction, the temperature change $\delta T_S(y)$ across the superconductor thickness $d_S$ is estimated to be $\frac{\sqrt{2\pi k_BT_S}}{d_S}$, where $k_S$ is the electronic thermal conductivity of the superconductor and $\dot{Q}_{in}$ is the input power from the cooler barrier of area $A$. The temperature drop is only $\delta T_S = 0.4 \text{ mK}$ under a 1-nW input power at 250 mK and, thus, negligible. Along the $x$ direction, we use a thermal conductivity $\kappa_D = \zeta_0/\rho_D T_D$ in the normal-metal drain, where $\zeta_0$ is the Lorenz number, $T_D$ the local temperature, and $\rho_D$ the resistivity. The density of the heat current is $\dot{q}_D = \Sigma_D (T_D^3 - T_{ph}^3)$, where $\Sigma_D$ is the electron-phonon coupling in the drain material. For a small temperature change, we obtain a heat relaxation length $\ell_D = \sqrt{\zeta_0 / \rho_D \Sigma_D T_{ph}^{-3/2}}$. For AlMn, this yields $\ell_D = 117 T_0^{-3/2} \mu\text{m K}^{-3/2}$ (120 $\mu\text{m}$ at 200 mK) [30]. In the superconducting leads, the heat conductivity is decreased by a factor of $6(\frac{\Delta}{k_BT_S})^2 \exp(-\Delta/k_BT_S)$, while the heat exchange is reduced by $\exp(-\Delta/k_BT_S)$ [31]. We obtain $\ell_S = \frac{\sqrt{2\pi k_BT_S}}{\frac{\rho_D}{2\pi^2} \sqrt{6\zeta_0/\rho_D \Sigma_S T_{ph}^{-3/2}}}$. For pure Al without a trap, $\ell_S = 50 T_0^{-5/2} \mu\text{m K}^{-5/2}$ (3 mm at 200 mK). The heat transport along the superconducting leads is thus much less efficient than in the drain, as expected.

When the superconductor couples to the drain, the heat flux through the barrier in between is given by [32]

$$\dot{q} = \frac{\sqrt{2\pi k_BT_S^3}}{e^2 d_S r_D} \left( \sqrt{T_S e^{-\Delta/k_BT_S}} - \sqrt{T_D e^{-\Delta/k_BT_D}} \right),$$

where $r_D$ is the junction resistance per area. The relaxation length in the $x$ direction is then written

$$\ell_{S,D} = \left( \frac{r_D d_S}{\rho_S} \sqrt{\frac{k_BT_D}{\Delta}} \right)^{1/2} = 2\sqrt{r_D \Omega_D^{1/2}} \mu\text{m}$$
at 300 mK. A small specific resistance \( r_D \) is thus needed so that the superconductor is locally efficiently thermalized to the drain. The similar cooling behavior of sample C2 with a direct quasiparticle trap and the original sample C1 leads to the conclusion that heat relaxation in C1 or C2 occurs at a distance from the junction smaller than the distance of the direct trap, i.e., \( \ell_{SD} \lesssim 5 \mu m \). This statement leads to the estimate \( r_D \lesssim 10 \Omega \mu m^2 \), which is consistent with expectations based on the fabrication recipe and with a previous evaluation [26].

Let us now estimate the different terms in the heat balance at the lowest temperature range. We consider sample C1 at its lowest bath temperature of 70 mK, assuming \( T_S = T_{ph} = T_{bath} \). Joule heating on the normal island, with a resistance of 0.02 \( \Omega \) and a typical 1-\( \mu \)A injection current, is only 20 fW and thus negligible. The electron-phonon coupling power in the Cu island \( \dot{Q}_{e-ph} = 0.17 \) pW is small in comparison to the cooling power \( \dot{Q}_{NIS} = 20 \) pW. The phenomenological Dynes factor \( \gamma = G_0/G_N \) is the ratio of the conductance at zero bias to its normal-state value. It captures the contribution of possible pin holes in the tunnel barrier, the inverse proximity effect in the superconductor, and the effect of the environment on electron tunneling. The fit of the current-voltage characteristics (data not shown) gives \( \gamma = 2.5 \times 10^{-4} \), which brings a related parasitic heating power \( \Delta^2 \gamma/R_T \approx 8 \) pW. As \( T_{N,min} \approx 2.5 \gamma^2/3T_c \) [4], the observed minimum temperature \( T_{N,min} \approx 30 \) mK could be related to a \( \gamma \) parameter at least 3 times larger than the measured value. Thus, the Dynes smearing of the superconductor density of states cannot account for the observed temperature saturation.

The discrepancy in the power balance necessarily comes from other sources that we do not have a direct way to probe. Besides potential candidates such as phonon heating [33] or near-field heat transport [34], we also suspect that the nonuniformity of the superconducting gap contributes to the saturation of \( T_{N,min} \). It is well known that the superconducting gap \( \Delta \) of a thin Al film can have different values depending on the fabrication details. The gap can be tuned by the grain size of the film [35,36], and different crystal orientations can have a different value up to 3% [37,38]. To model the performance of the cooler, we assume a Gaussian distribution of \( \Delta \) with a standard deviation \( \delta \Delta \) and calculate the maximum cooling power \( \dot{Q}_{NIS}^{opt} \) normalized to its value at \( \delta \Delta = 0 \). Figure 4(b) shows our result at different temperatures. With \( \delta \Delta/\Delta = 3% \) at 30 mK (\( T/T_c = 0.03 \)), the cooling power would be reduced to one-half of its nominal value.

V. CONCLUSIONS

In conclusion, the accumulation of hot quasiparticles in the superconducting leads limits the cooler performance, even with an efficient quasiparticle drain. We reduce this effect by tuning the tunnel barrier and demonstrate that such a cooler reaches a 32-mK electron temperature from a 150-mK bath temperature and performs outstandingly over a wide range of temperatures. At temperatures where electron-phonon interaction is strong, phonon cooling in the suspended normal-metal island boosts the performance of the cooler. This refrigerator has a significant power and can be easily integrated, so that it opens new possibilities to cool practical devices to the sub-50-mK regime.

ACKNOWLEDGMENTS

We have benefited from discussions with D.V. Averin, Y. A. Pushkin, D. S. Golubev, M. Prunnila, V. F. Maisi, I. Khaymovich, T. T. Heikkilä, V. J. Kauppila, and C. B. Winkelmann. We acknowledge the support of the European Community Research Infrastructures under the FP7 Capacities Specific Program, MICROKELVIN Project No. 228464, the Engineering and Physical Sciences Research Council (EPSRC) Grant No. EP/F040784/1, the Academy of Finland through its LTQ CoE grant (Project No. 250280 and Project No. 139172), and the OtaNano infrastructure. Samples are fabricated in the Micronova Nanofabrication Center of Aalto University.

[27] Because of the large junction size, A2 has 1-nW cooling power at 300 mK. Still, it performs equally well compared to A1 or A3.
[28] The crossing of data from sample A4 with those from other samples is vaguely inline with expectation, as A4 has the smallest metal volume V. Its higher $T_{N, min}$ is presumably due to Joule heating because of the thinner Cu layer.
[30] Because of the weak dependence $\ell_D \propto 1/\sqrt{\Sigma p\rho}$, the corresponding numbers for Cu and Au are similar. There is no advantage of replacing AlMn by another material.