Micrometre-scale refrigerators

Juha T Muhonen\textsuperscript{1,2}, Matthias Meschke\textsuperscript{1} and Jukka P Pekola\textsuperscript{1}

\textsuperscript{1} Low Temperature Laboratory, Aalto University, School of Science, PO Box 13500, 00076 Aalto, Finland
\textsuperscript{2} Department of Physics, University of Warwick, Coventry CV4 7AL, UK

E-mail: jukka.pekola@aalto.fi

Received 7 September 2011, in final form 17 November 2011
Published 9 March 2012
Online at stacks.iop.org/RoPP/75/046501

Abstract

A superconductor with a gap in the density of states or a quantum dot with discrete energy levels is a central building block in realizing an electronic on-chip cooler. They can work as energy filters, allowing only hot quasiparticles to tunnel out from the electrode to be cooled. This principle has been employed experimentally since the early 1990s in investigations and demonstrations of micrometre-scale coolers at sub-kelvin temperatures. In this paper, we review the basic experimental conditions in realizing the coolers and the main practical issues that are known to limit their performance. We give an update of experiments performed on cryogenic micrometre-scale coolers in the past five years.

(Some figures may appear in colour only in the online journal)

This article was invited by S Washburn.

1. Introduction

Electron transport in micro- and nano-structures has attracted lots of attention over the past several decades. Until recently, less concern has been paid to the associated energy currents and generation of heat. However, heat currents and dissipation often limit the performance of an electronic device in particular at cryogenic temperatures. Cooling a device to lower temperatures generally improves its characteristics in terms of increased sensitivity and decreased noise. Despite the fast progress in liquid cryogen free cooling techniques, refrigeration to cryogenic temperatures remains expensive and the proper infrastructure for cryogenic work is found typically only in specialized laboratories. Therefore, it is of interest to explore cooling techniques that operate directly on a chip, even though they may be an option only in special applications (such as direct detector cooling with limited cooling power). They could provide an alternative solution as the final stage of a refrigerator which is both economic and easy to use.

In this review, we will concentrate on low-temperature electronic on-chip coolers. The basic principle of operation is shown in figure 1. An energy filter allows only high energy electrons to be removed from an electron system, leading to cooling of the system. A possible energy filter is the superconducting gap. Experimental activity using on-chip cooling by NIS (N = normal conductor, I = insulator,
S = superconductor) junctions has a history of less than 20 years. We focus on progress over the past five years as the earlier achievements have been extensively covered in another review [1]. To our knowledge, only a handful of laboratories are actively experimenting on electronic coolers at sub-kelvin temperatures at the current time. This has made our task of covering most of the published activity on the topic hopefully successful. The aim in this review is to report mainly on experiments and the associated phenomena with less emphasis on the theoretical background. We do discuss various energy relaxation mechanisms in a cooler at length since their role has turned out to be centrally important in trying to achieve the optimum performance of an electronic cooler. For instance, the quasiparticle relaxation in a superconductor at the backside of the cooler is a major, still unresolved problem limiting the performance of an NIS refrigerator. This is a particularly important issue when low temperatures and enhanced cooling powers are to be achieved. After discussing heat transport and dissipation, we review recent conceptual and technological advances in terms of cooling principles, materials and physical realizations. Although most of the reported work deals with superconductor-based solutions, we want to note here that a relatively recent experiment using quantum dots as energy filters in semiconducting 2D electron gas (2DEG) structures [2] has spurred research in new potential realizations of practical on-chip coolers.

2. Temperatures and energy relaxation

2.1. Temperatures of a micrometre-scale conductor

It is not trivial to define the temperature of a micrometre-scale conductor at sub-kelvin temperatures. First, one does not have just one system but an ensemble of subsystems, which each have a characteristic internal energy and are coupled to each other very non-linearly. The most relevant subsystems concerning the micrometre-scale refrigerators discussed here are the electron system, the phonon system of the conductor and the phonon system of the substrate. Secondly, in order for temperature to be a well-defined concept, the relaxation rates inside each subsystem must be faster than the couplings between them. Then the systems will follow thermal distributions (Fermi–Dirac for electrons, Bose–Einstein for phonons) and temperatures can be related to them. The situation where this is true but the effective temperatures of different subsystems are not equal is generally called quasi-equilibrium. The situation where energy is exchanged with the system faster than it can relax and hence no temperature can be defined for it is called non-equilibrium. Full equilibrium would be the situation where all the subsystems are at the same temperature.

In what follows, we will be mainly concerned with the electron system, as the micrometre-scale refrigerators discussed in this review cool the electron system directly. Figure 2 gives a simplified thermal model of a conductor, S or N. The basic picture of cooling, described in the following chapters, only holds if the electron system stays in quasi-equilibrium. For this to be true, the electron–electron (e–e) and the injection rate (or the photon exchange rate \( \gamma_{\nu} \)) must be faster than the injection rate of quasiparticles and photons. Then the distribution can be described by a Fermi–Dirac one and we can ascribe a well-defined temperature to the subsystem. This is the prevailing situation, and it has turned out to be difficult to overcome this in tunnel-coupled systems, particularly in the N electrodes; see, however, [3]. In a system where electrons can be injected directly without a tunnel barrier, non-equilibrium distributions have been observed in several experiments (e.g. the seminal paper by Pothier et al [4]). On the other hand, the ratio between the e–p collision rate \( \gamma_{\nu} \) and the injection rate (or the photon exchange rate \( \gamma_{\nu} \)) determines whether the electron subsystem has the same temperature as the phonon bath or not. Low \( \gamma_{\nu} \) means that the electronic system can have a temperature different from that of the bath which makes direct electronic cooling possible. This happens in particular at low

Figure 1. The basic principle of direct electronic cooling. An energy filter (grey wall) allows only high energy electrons (red circles) to be removed from the electron system. This ejection leads to sharpening of the electron distribution, i.e. cooling. The corresponding Fermi distribution is shown on the left. The density of states of the conductor is to a good approximation constant on the narrow energy range of interest, as the thermal energy is small compared with the Fermi energy.

![Figure 1](https://example.com/figure1.png)

Figure 2. Simplified thermal model of a conductor at temperature \( T_e \) (S or N). External power \( P \) is exchanged with the system. Electron–electron (e–e) interaction drives the electron subsystem towards a quasi-equilibrium distribution and electron–phonon interaction couples it to the phonon bath (e–p). It is also coupled electromagnetically to the environment (at temperature \( T_{env} \)) which can be spatially well separated from the cooled volume.

![Figure 2](https://example.com/figure2.png)
temperatures since the relaxation gets increasingly slow at low temperatures (e.g. for normal metals \( \gamma_{e-p} \propto T^{-3} \)). If the coupling to the phonon system is suppressed, the photonic coupling to the environment \( \gamma_{e-v} \) can become the dominant relaxation mechanism as will be discussed later in the review.

The phonon systems in the conductor and in the substrate are also in principle two separate systems which can have differing temperatures. As a result of lattice mismatch between the two materials, the phonons can be scattered at the interface leading to thermal resistance, known as Kapitza resistance. However, at the low temperatures considered here, the dominant wavelength of the thermal phonons is of the order of several micrometres which is much larger than the thickness of a typical metallic or semiconducting film. Hence the interface (and difference between the two materials) should be quite transparent to these phonons. In addition, as the electron–phonon coupling (discussed below) decreases rapidly at low temperatures, this thermal resistance will be the dominant thermal bottleneck. (See, however, [5].) We will neglect the Kapitza resistance throughout.

2.2. Relaxation mechanisms at low temperatures

2.2.1. Electrons in metals and semiconductors. In most cases encountered experimentally, the electron–phonon coupling is the dominant inelastic scattering mechanism for the electron system at not very low temperatures. The electron–phonon relaxation in ordinary metals has been discussed and measured in various experiments over the past several decades. With quasi-equilibrium conditions, a straightforward first order perturbation theory calculation (assuming scalar deformation-potential coupling and 3D electron and phonon systems) yields an energy exchange rate [6]

\[
P_{e-p}^{n}(T_e, T_p) = \Sigma V(T_e^5 - T_p^5).
\]

Here \( \Sigma \) is the material parameter, known for most ordinary metals, \( V \) is the volume of the conductor, \( T_e \) and \( T_p \) are the temperatures of the electron and phonon system, respectively. Substantial deviations from this law, which is obeyed experimentally astonishingly well irrespective of the particular normal metal material or geometry, are expected in restricted dimensions, for superconductors and for semiconductors. We discuss these issues below. For most metals, one has \( \Sigma \sim 10^{3} \text{WK}^{-5} \text{m}^{-3} \).

In semiconductors, the situation is different from the normal metal case essentially because of two facts: the small amount of momentum-space that is occupied in the normal state essentially because of two facts: the small amount of momentum-space that is occupied in the normal state and the presence of the band gap. In essence, the coupling between phonons and the electron system can then be described with a deformation potential constant that describes how the minimum of the conduction band moves in response to the stresses caused by phonons. Hence, variables of the electron system (mainly the momentum distribution) can be neglected. We delay more detailed discussion about this issue to the section about Schottky coolers.

So far, we have implicitly assumed that the power is distributed uniformly on the conductor or that the conductor has a high enough thermal conductivity so that no temperature gradients exist inside it. In practice this is often not the case. Especially considering micrometre-scaled coolers, the prevailing situation is such that one has a pointlike cooling/heating source on one end of a conducting wire or a plate. To make an accurate model in such situations, it becomes compulsory to also consider the thermal conductivity inside the electron system. If one makes the assumptions outlined above (so that a position-dependent temperature can be defined), then the thermal conductivity in normal metals at low temperatures follows textbook models of the electron gas. The heat current density is related to the temperature gradient as \( \dot{Q} = -\kappa_a \nabla T \), where \( \kappa_a \) is the thermal conductivity. This can be related to the electrical conductivity \( \sigma \) via the Wiedemann–Franz law \( \kappa_a = L \sigma T \), where \( L \) is the Lorentz number. With these assumptions, a steady-state diffusion equation can be written for a differential volume element

\[
\nabla \cdot [-\kappa_a(T_e, x) \nabla T_e(x)] = \Sigma(T_e(x)^5 - T_p(x)^5) + \dot{P}_{\text{ext}}(x),
\]

where we have used \( \dot{P}_{\text{ext}} \) as the power density from all possible external heating sources. Solving this equation self-consistently and with proper boundary conditions will yield the temperature profile of the conductor.

2.2.2. Quasiparticle excitations in superconductors. In a superconductor, we are interested in the system of quasiparticle excitations. The Cooper-pair condensate carries no entropy and has no explicit role in the thermal properties discussed here. The typically dominant relaxation mechanisms are analogous to the normal metal case: the quasiparticle heat conductivity along the wire and quasiparticle-phonon relaxation (which is determined predominantly by the recombination of quasiparticles into Cooper pairs). The most obvious differences to the normal metal case are (i) the exponentially small amount of quasiparticles at temperatures \( T \ll T_c \) and (ii) the fact that the quasiparticles need to absorb or emit energy larger than the superconducting gap \( \Delta \). Combining these effects leads to exponentially suppressed heat conductivity and relaxation at low temperatures; \( \kappa, \Sigma \propto \exp(-\Delta/(k_B T)) \).

Ideally, a superconductor at a temperature well below \( T_c \) has a negligible number of quasiparticle excitations. Quantitatively, the Bardeen–Cooper–Schrieffer (BCS) theory predicts the quasiparticle density \( n_{qp} \) in thermal equilibrium to be

\[
n_{qp} = 2N_F\int_{\Delta}^{\infty} dE \frac{E}{\sqrt{E^2 - \Delta^2}} f(E)
\]

\[
\approx N_F \sqrt{2\pi k_B T} \Delta e^{-\Delta/(k_B T)},
\]

(3)

where \( N_F \) is the density of states (DOS) at the Fermi level (in the normal state), \( f \) is the Fermi distribution function. The last step applies \( \Delta/(k_B T) \gg 1 \). The factor of 2 comes from the fact that we should also integrate over the negative energies (holelike quasiparticles). For illustration, one can put the parameters of aluminium (Al) to equation (3) \( (\Delta/k_B = 2.4 \text{K}) \) at a temperature of \( T = 100 \text{mK} \); in this case, ideally, the quasiparticle density is phenomenally low \( n_{qp} \sim 10^{-5} (\mu \text{m})^{-3} \). However, invariably experiments
have shown quasiparticle densities above what is predicted by equation (3) at the lowest temperatures (an example is shown in figure 3). These excess quasiparticles can be explained in two ways: they are either (i) created by external pair-breaking sources or (ii) there are sub-gap quasiparticle states not present in an ideal BCS superconductor. According to our present understanding, the first option is the predominant one and the main external source is high-frequency noise radiated from the environment. This is discussed in more detail in the next subsection. However, often the experimental observations can and have been interpreted by adopting the second interpretation.

The amount and dynamics of the excess (often referred to as non-equilibrium) quasiparticles has been the subject of intense research lately, as they are a primary source of errors in almost all superconducting electronics. In addition to the anomaly at low temperatures, they are also produced when operating any superconducting device if there are dissipative elements in the circuit. Lately, these non-equilibrium quasiparticles have been considered in relation to qubits [8–10], radiation detectors [7, 11], single-electron turnstiles [12] and NIS tunnel junctions [13]. All of these have confirmed the existence of excess quasiparticles at the lowest temperatures as well as the assumed dependence equation (3) of the thermal quasiparticles at higher temperatures.

Quasiparticle recombination is a process where two quasiparticles of opposite momenta (\( k \) and \(-k \) where \( k \approx \Delta \)) recombine to form a Cooper pair and emit a phonon with energy equal to \( 2\Delta \). The recombination rate was studied several decades ago [14] but the associated heat flux from quasiparticles to the phonon system has been experimentally determined only very recently [15]. At the limit where \( T_p \ll T_{qp} \ll \Delta/k_B \), the heat flux can be calculated analytically from the quasiclassical theory to yield

\[
P_{qp-p}^s \approx \frac{0.98 e^{\gamma(T)\kappa_n}}{k^2} P_{p-p}(T_{qp}, T_p),
\]

where \( T_{qp} \) is the quasiparticle temperature. However, the authors of [15] found experimentally that the heat flux was larger than what was expected theoretically. The possible additional relaxation channels remain unclear at the moment (see figure 4 for the experimental setup and results).

Now, working under the quasi-equilibrium assumption, we can also write the heat diffusion equation in the superconducting case. The reduction in thermal conductivity at the superconducting state was calculated theoretically soon after the BCS theory appeared [16]. Assuming that the thermal conductivity is limited by impurities, in the superconducting state it can be written as \( \kappa_s = \gamma(T)\kappa_n \), where the suppression ratio \( \gamma(T) \) is given by

\[
\gamma(T) \equiv \frac{3}{2\pi^2} \int_{\Delta/\kappa_B T}^{\infty} \frac{x^2}{\cos^2(x/2)} \, dx \approx \frac{6}{\pi^2} \left( \frac{\Delta}{k_B T} \right)^2 e^{-\Delta/\kappa_B T},
\]

where the approximation shown on the right again applies for \( k_B T \ll \Delta \). Note that we assume everywhere that the superconducting gap \( \Delta \) has a temperature dependence given by BCS theory. With these equations, a diffusion equation for the superconductor can be constructed just by inserting equation (4) and equation (5) into equation (2)

\[
\nabla \cdot \left[ -\kappa_s(x, T_{qp}(x)) \nabla T_{qp}(x) \right] = P_{qp-p}^s(x, T_{qp}, T_p) + P_{ext}(x),
\]

where again \( P_{ext} \) is the power density from external sources and \( P_{qp-p}^s \) is \( P_{qp-p}/\gamma \).

The thermal conductivity of the superconductor can, however, be significantly modified if a normal metal is brought into contact with it [17]. A transparent normal metal–superconductor contact will modify both the properties of the normal metal (proximity effect) and the superconductor (inverse proximity effect) close to the interface. In a superconductor, this will lead to an effectively decreased superconducting gap as well as non-zero DOS inside the gap in the vicinity of the interface. This means that at short distances, quasiparticles with energy below the gap can also be transported to the superconductor (and are not Andreev reflected) which enhances the thermal conductivity. The length scale of this effect is roughly the superconducting coherence length \( \xi \). In [18], the thermal conductivity of inverse proximized superconducting wire was studied in detail. A superconducting wire with length \( L \) of the order of \( \xi \) was placed in between two normal metal wires and a temperature difference was applied over it. The results of the thermal conductivity were in agreement with the theory predictions from quasi古典 theory. The longest wire (\( L = 4.2 \mu m \)) behaved almost as an ideal BCS superconductor, i.e. according to equation (5), whereas the shortest wire (\( L = 0.425 \mu m \)) showed many orders of magnitude larger thermal conductance. These results are presented in figure 5.

2.3. Coupling of the electronic system to the electromagnetic environment

Of more recent interest is the coupling of the electronic system via radiation [19–21]. This can manifest itself as heating/cooling of a conductor due to the presence of another resistive conductor at higher/lower temperature coupled to the one being monitored. The hot environment can also lead to
Figure 4. Quasiparticle relaxation in Al, presented in [15]. (a) The experimental sample. Quasiparticle relaxation is probed in the thick central island, using two SIS junctions to inject the quasiparticles and two other junctions to probe the distribution function. The sample is fully aluminum. (b) Measured qp–p relaxation as a function of the electronic temperature. The dashed line shows the expected normal state relaxation and the solid line is the result from full quasiclassical theory. Data from three different samples all lie between the normal state and the theory predictions.

Figure 5. Thermal conductance of a short superconducting (aluminum) wire between two normal metal (copper) leads, presented in [18]. The thermal conductivity is enhanced by the inverse proximity effect. (a) SEM image of the sample and sketch of the measurement setup. Two Cu islands are connected via a short superconducting Al wire with transparent NS interfaces. Four S electrodes (top of the image) are connected to each of the two N islands through tunnel barriers for electronic thermometry and temperature control. Inset: sketch of the side profile of the NSN structure, consisting of an S wire connected via overlap junctions to two N reservoirs. (b) Thermal conductivity of the S wire, normalized to the normal state conductivity (left) and to the BCS theory prediction (right). These data were extracted from fits to the measurement results. The four samples correspond to different lengths of the superconducting wire: 4.2, 1.1, 0.875 and 0.425 µm. The longest wire behaves as a BCS superconductor, whereas the heat conductivity of the shortest wire is many orders of magnitude larger. Two other samples fall in between these extreme cases.
Photon-assisted tunnelling. These are both naturally well-known concepts but there has been some recent interest in these phenomena on the quantum level in mesoscopic structures, since they govern the ultimate heat-balance of the system. Generally, radiative heat flux from a resistor \( R_1 \) at (electronic) temperature \( T_1 \) to resistor \( R_2 \) at (electronic) temperature \( T_2 \) via a transmission line with a total series impedance \( Z_t(\omega) \) (see figure 6(b)) is described by [22]\

\[
P_v = \int_0^\infty \frac{d\omega}{2\pi} \frac{4R_1R_2\omega}{|Z_t(\omega)|^2} \left( \frac{1}{\gamma_{\nu_0}/k_BT_2 - 1} - \frac{1}{\gamma_{\nu_0}/k_BT_1 - 1} \right).
\]

The Bose–Einstein distributions \( (e^{\nu_0/k_BT} - 1)^{-1} \) of the two resistors at the corresponding temperatures and the matching between them \( r_0 \equiv 4R_1R_2/|Z_t(\omega)|^2 \) determine the total heat flux by the electromagnetic noise. Energy exchange with the environment via this photonic heat exchange can overcome the e–p coupling below the crossover temperature \( T_{cT} \sim r_0\nu_0/k_B(30\Sigma\Omega)^{1/3} \) [23]. For typical mesoscopic structures made of normal metals with a volume on the order of \( \Omega \sim 10^{-20} \text{m}^3 \) and electron–phonon coupling strength \( \Sigma \sim 10^8 \text{W} \text{m}^{-3} \text{K}^{-4} \), one obtains moderately low values of \( T_{cT} \sim 150 \text{mK} \) (see figure 6(a)). These parameters are within the range of experimental values for micrometre-scale refrigerators and have to be considered when describing the device operation. A strong coupling to a hot environment will degrade the cooler performance noticeably towards low temperatures. On the other hand, the total coupling strength can be minimized via the coupling \( r_0 \) between the environment and the device as demonstrated experimentally in [24, 25]. We discuss the latter cooling experiment in more detail in section 3.5. The experimental setup in [24] (see figure 6(b)) examines the influence of radiative heat exchange as the coupling \( r_0 \) between the two coupled resistors is varied using SQUIDs. Their Josephson inductance \( L_j \sim h/(2eI_C) \) is influenced by the penetrating external magnetic flux (\( \Phi \)) through the SQUIDs as \( I_C \propto |\cos(\Phi/\Phi_0)|. \) \( L_j \) is minimized at integer values of the flux quantum \( \Phi_0 \), thereby maximizing the Josephson inductance. This consequently minimizes the coupling \( r_0 \) between both resistors causing the measured temperature of \( R_2 \) to peak (see figure 6(c)).

It has recently been shown [26] that high-frequency noise radiated from higher temperatures \( T_{env} \) by some environment creates a leakage current with approximately linear bias voltage dependence in an NIS junction at low bath temperatures \( T \), if \( k_BT_{env} \gg \Delta \). This leakage current is due to photon-assisted tunnelling events (see section 3.4), where high energy photons are absorbed during the tunnelling event and hence can facilitate lower energy electrons from N to tunnel above the gap to S. For a resistive environment with high cut-off frequency, this leakage current is exactly equivalent [26] to the one which one gets by assuming instead of the pure BCS DOS \( n_S(E) = |E|/\sqrt{E^2 - \Delta^2} \), the so-called Dynes DOS with a lifetime broadening \( \gamma \)

\[
n_S(E) = \frac{|E + i\gamma\Delta|}{\sqrt{(E + i\gamma\Delta)^2 - \Delta^2}}.
\]

Experimentally \( \gamma \) can be extracted as the ratio of the measured zero bias conductance of the NIS junction and the asymptotic conductance at large voltages \( (R_T)^{-1} \).

The photon-assisted tunnelling events lead to \( \gamma = 2\pi R_Kk_BT_{env}/(R_K\Delta) \), where \( R \) is the effective resistance of the environment and \( R_K = h/e^2 \) the resistance quantum. In effect this linear behaviour amounts to viewing the sub-gap current of the NIS junction as if it were that of a fully normal (NIN) junction with tunnel resistance \( R_T/\gamma \). Then the tunnelling rates at zero bias have a value \( \Gamma_0 = (k_B T)/(e^2 R_T) \gamma \) and the power input is then roughly \( \dot{Q} \approx 2\gamma\beta\Delta \), where the factor 2 appears because of tunnelling in two directions, each creating one quasiparticle. The radiated noise can then, at least partially, also explain the low-temperature anomaly of extra quasiparticles mentioned above. Experimental confirmation for this hypothesis has been added recently [12, 27, 28], as decreases in the amount of excessive quasiparticles at low temperatures have been seen to depend on the filtering and shielding of experimental setups.
3. Cooling with NIS junctions

3.1. Basics of NIS cooling and thermometry

The basics of cooling by NIS junctions have been discussed in several works, and many experiments have confirmed the predicted overall behaviour of this system (see [1] and references therein). The phenomenon is based on the gap in the DOS of a superconductor acting as an energy filter for the electrons. If proper bias voltage ($eV$ just below the gap energy) is applied over the junction, only the most energetic electrons can tunnel out from the normal metal (see figure 7). This leads to a decrease in the average energy of electrons in the N conductor, i.e. cooling. The opposite process, quasiparticle tunnelling from the superconductor to the normal metal, has been viewed as tunnelling of hole-like quasiparticles from N to S. The cooling power, i.e. the amount of heat extracted from a normal conductor per unit time in an NIS junction biased at voltage $V$, is simply given by

$$P_{NIS} = \frac{1}{e^2 R_T} \int dE(E - eV)n_S(E)[f_N(E - eV) - f_S(E)],$$

where $R_T$ is the normal state resistance of the junction, $n_S(E)$ is the DOS in the superconductor (normalized by that of the corresponding normal metal) and $f_S(E), f_N(E)$ are the energy distributions of electrons in S and N, respectively. It should be noted that as equation (9) is symmetric in voltage, putting two NIS junctions in series (forming an SINIS structure) doubles the cooling power of the normal metal. This is a different behaviour as compared with traditional Peltier cooling elements. Typically one assumes the basic BCS DOS $n_S(E) = |E|/\sqrt{E^2 - \Delta^2}$. However, it is usually convenient, and sometimes also justified, to assume lifetime type broadening of the BCS DOS so that it follows the so-called Dynes form with the parameter $\gamma$ describing the sub-gap leakage (see section 2.3). The ideal behaviour of the cooler is achieved when the DOS is of pure BCS type, since in this case the electrons extracted from the normal metal can tunnel only to the states above the gap energy $\Delta$, leading to ideal energy filtering. Furthermore, to capture the ideal performance of the NIS cooler as a refrigerator in a traditional sense, one assumes that the occupations of quasiparticles in S and N follow the Fermi–Dirac distribution, i.e.

$$f_S(E)/f_N(E) = 1/(1 + e^{(E - E_F)/k_BT_S})$$

where $T_S$ and $T_N$ are the (electronic) temperatures of the two conductors.

Consider an idealized cooler with pure BCS DOS and with well-defined temperatures $T_S$ and $T_N$, which may both differ from the bath temperature $T_0$. In this case, for temperatures well below the critical temperature of the superconductor $k_BT \ll k_BT_C$, one obtains analytical expressions for the optimal cooling power [29] (see also [30]). The cooling power maximizes at bias voltages $V = (\Delta - 0.66k_BT)/e$ where it reaches

$$P_{opt} \simeq \frac{\Delta^2}{e^2 R_T} [0.59 \left( \frac{k_BT_S}{\Delta} \right)^{3/2} - \frac{2\pi k_BT_S}{\Delta} e^{-\Delta/k_BT_S}].$$

(10)

At the optimal bias point, the current through the cooler junction is (see below)

$$I_{V_{opt}} \simeq 0.48 \frac{\Delta}{eR_T} \sqrt{\frac{k_BT_S}{\Delta}}.$$

(11)

An important figure of merit of the cooler is its coefficient of performance (‘efficiency’) $\eta$, which we define as the cooling power at the optimum point divided by the power consumed in the voltage source, i.e.

$$\eta = \frac{P_{opt}}{I(V_{opt})V} \simeq 0.7 \frac{T}{T_C},$$

(12)

where the last approximation applies again at $T \ll T_C$. Some of the characteristics of an ideal cooler have been shown in figure 7.

The current–voltage ($I-V$) characteristics of NIS junctions are strongly non-linear and dependent on temperature, a fact that also enables the use of these junctions as thermometers. Consider the same kind of junction as...
should be noted that this insensitivity to the quasiparticle normal metal and are insensitive to the temperature of the characteristics depend only on the temperature of the SNS structure \[31–35\] where the N island can also be temperatures is to measure proximity induced supercurrent and demonstrated alternative to probe the lowest electronic through the junction (4). From the latter symmetrized form, it is clear that the \( I \sim V \) characteristics depend only on the temperature of the normal metal and are insensitive to the temperature of the superconductor (assuming constant superconducting gap). It should be noted that this insensitivity to the quasiparticle temperature of the superconductor does not apply to the cooling effect; no symmetrized form excluding \( f_S \) can be derived from equation (9).

As NIS thermometers are simple, easy to use and measure only the temperature of the normal metal, they are almost invariably used for thermometry in conjunction with NIS coolers. They do, however, have some drawbacks. Most notably, there is an inevitable power dissipation from the operation of the thermometer. This power can be made very small (\( \sim \text{MW} \)) but it can still have a notable influence on the temperature at the lowest temperatures that the coolers are operated (\( \sim 50 \text{ mK} \)). In addition, NIS thermometers tend to lose sensitivity at the lowest temperatures, as the \( I \sim V \) characteristics are more influenced by the leakage currents through the junction (\( \gamma \) parameter above). A proposed and demonstrated alternative to probe the lowest electronic temperatures is to measure proximity induced supercurrent in an SNS structure [31–35] where the N island can also be connected to NIS cooler junctions. In this type of structure, the measurement consists of sweeping current through the SNS system and measuring the critical current \( I_c \) where the structure switches to the resistive state. The switching current depends strongly on the temperature of the N island and the measurement becomes increasingly sensitive towards low temperatures as the supercurrent increases. Also, ideally the power dissipation is exactly zero before the switching happens, meaning that the thermometry has no self-heating effect. The approach also has some disadvantages, mainly that it is quite complicated and time consuming as in practice the measurement consists of making a switching histogram at each temperature point. In addition, as the exact dependence of the switching current from the temperature is a strong function of structure parameters, the extrapolation to lowest temperatures (where no calibration exists) is not straightforward.

Below we will give a review of the experimental and theoretical advances done with NIS cooling in recent years. For the earlier history, see, for example, \[1\].

### 3.2. Limitations on NIS cooler performance

Generally the performance of NIS coolers has been below what would be expected from equation (9), especially at temperatures below \( \sim 150 \text{ mK} \). Many possible reasons for this have been identified, here we review two of the most often raised effects: the excess density of quasiparticles in the superconducting leads and two-electron tunnelling processes.

#### 3.2.1. Effects of quasiparticle population on refrigeration.

As mentioned above the efficiency of an NIS cooler (the ratio of the cooling power over the input power) is roughly \( 0.7 / T_c \). At \( 0.3 \text{ K} \) this corresponds to \( 15\% \) (assuming \( T_C \) of 1.3 K, common to thin Al films). Put another way, the power dissipated into the superconducting leads (even in this ideal case and at rather high temperature) is always an order of magnitude larger than the cooling power. In any practical cooler, this can be a significant power. As both the quasiparticle–phonon (qp–p) relaxation rate and the diffusion of quasiparticles are additionally exponentially suppressed in the superconductor (as described in section 2.2.2), the dissipated power can create a high density of non-equilibrium quasiparticles on the superconducting side of the cooler, i.e. heat it up. This can have severe effects on the cooling power of NIS junctions.

Assuming that non-equilibrium quasiparticles can be described with an effective temperature, the reduction of cooling power due to this overheating can be understood from equation (9). In figure 8(a), the cooling power of an NIS junction is presented as a function of the superconductor temperature at a constant normal metal temperature. The effect of rising \( T_N \) is two-fold: it changes the distribution function \( f_S \) (meaning there are more quasiparticles above the gap that can tunnel into the normal metal) and it reduces the \( \Delta \) parameter in the DOS. Unlike for the electric current, for the heat flow the distribution function of the superconductor plays a major role and the rising \( T_N \) can destroy the cooling effect, even assuming a constant superconducting gap. The same effect is also visible in equation (10), which has been derived assuming a constant superconductor gap. Note that throughout the discussion we assume that the quasiparticle distribution still follows a Fermi distribution and is not in a true non-equilibrium.

In many instances, this overheating of the S has been modelled simply as a backflow parameter of heat, where a constant portion of the whole input power \( IV \) is assumed to ‘flow back’ to the normal metal and induce a parasitic heating power \( P_{\text{sc}} = \beta IV \). Typically \( \beta \) lies between 1% and 10%. Although this kind of model has had some success in fitting the experimental data, it does not really address the mechanisms behind the backflow.

Recently, there has been considerable interest in modelling this effect more precisely, based on the diffusion equations presented in section 2. Assuming a diffusion of quasiparticles away from the junction area as well as their relaxation (recombination), one can self-consistently calculate the effective temperature profile of the superconductor and hence the cooling power of the junction. In figure 8(b), we present a calculation of the cooling power using the effective temperature model presented in section 2.

To make some simple estimates with a diffusion model, let us consider a 1D temperature profile and make the assumptions \( T_p \ll T_{qp} \ll \Delta / k_B \). Then we can write the diffusion equation
for a superconductor in an analytically solvable form [25] (we neglect the $P_{\text{ext}}$ term and the prefactor of the order of unity in equation (4))

$$\frac{\partial}{\partial x} \left[ \frac{6}{\pi^2} \left( \frac{\Delta}{k_B T} \right)^2 e^{-\Delta/(k_B T)} \frac{\partial T}{\partial x} \right] = e^{-\Delta/(k_B T)} \sum \left( T^5 - T_p^5 \right),$$  \hspace{1cm} (14)

where we have written $T = T_{\text{qp}}$ for clarity. Linearizing equation (14) for small temperature differences $\delta T(x) = T(x) - T_p$, one obtains a simple expression for the temperature profile in a uniform 1D wire. For a wire extending to positive $x$, we then have $\delta T(x) = \delta T(0)e^{-x/\ell}$, where $\delta T(0)$ is determined by the heat input at the end of the wire and the relaxation length is given by

$$\ell = \frac{\Delta}{\pi k_B} \sqrt{\frac{5 D_e \sigma}{2 \tau_p}} T_p^{-5/2}.$$  \hspace{1cm} (15)

Putting the parameters of aluminum in equation (15), we find that $\ell \approx (50 \mu m K^{3/2})/T^{5/2}$. This means that at typical sub-kelvin temperatures, the quasiparticle distribution relaxes over millimetre distances. The magnitude of the temperature rise can be obtained in the same linearized approximation by employing the boundary condition $P = -k_A A(d\delta T(x))/(dx)|_{x=0}$, where $A$ is the cross-sectional area of the wire. Inverting this for the temperature rise for a given heat input, we find

$$\delta T(0) = \frac{I P}{k_A} = \frac{\pi k_B}{\Delta \sqrt{30 \Sigma}} e^{\Delta/(k_B T)} T^{-3/2} \frac{P}{A}.$$  \hspace{1cm} (16)

Inserting numbers for a $A = 100 \text{ nm} \times 100 \text{ nm}$ wire at $T = 200 \text{ mK}$ yields $\delta T(0) \approx 20P/A \simeq (2 \times 10^{15} \text{ KW}^{-1})P$. This means that in order to keep $\delta T(0) < T$, one needs to have $P \ll 10^{-16} \text{ W}$, i.e. a very small power input indeed. Assuming each quasiparticle brings energy $\Delta$ to the superconductor, this implies a tunnelling rate of $\Gamma = P/\Delta \ll 3 \times 10^6 \text{ s}^{-1}$.

This corresponds to bias current of only $0.5 \text{ pA}$. This example demonstrates that a bare superconducting wire is driven out of equilibrium even with a very small current injection.

It is instructive to compare this approach (using effective temperature), with the results one gets using just the quasiparticle density. In relation to micrometre-scaled coolers, this approach was first used in [36] and has been expanded in [37, 38]. In this case the diffusion equation in 1D is

$$D_{n} \frac{\partial^2 n(x)}{\partial x^2} = \Gamma_{\text{qp-p}} + \Gamma_{\text{ext}},$$  \hspace{1cm} (17)

where $\Gamma_{\text{qp-p}} + \Gamma_{\text{ext}}$ are now the relaxation (scattering) rates to phonons and the external environment, respectively, and can be converted to power by multiplying with the energy exchanged in each scattering event. $D_n$ can be related to the normal state diffusion constant $D_n$ by $D_n = \sqrt{1 - (\Delta/E)^2} D_o$, where $E$ is the energy of the quasiparticle and $D_o$ is related to normal state heat conductivity through $k_o = \Sigma D_o N_p e^2 T$. In this way, one does not need to make the assumption of quasi-equilibrium but one now needs to consider explicitly the energies of the quasiparticles. In practice one has to make one out of two assumptions. One can either assume that the quasiparticles follow a thermal distribution and, in fact, it is straightforward to show that in that case the left-hand side (lhs) of equation (17) is exactly equivalent to the lhs of equation (14) (the connection between $n$ and $T$ is from equation (3)). The other option (adapted in [36, 37]) is to replace $E$ in equation (17) with the average energy of quasiparticles in the sample $\langle E \rangle$. This makes the $D_n$ independent of the $x$ coordinate and essentially is an approximation for small temperature differences where the spatial dependence of the diffusion constant can be neglected.

From figure (8b) and equations (15) and (16), it is clear that for the design of efficient NIS coolers, it is
critical to also consider the quasiparticle thermalization of the superconducting leads. Fortunately, the situation is usually not as bad as the solid lines in figure 8(b) would seem to suggest, as the effects of so-called quasiparticle traps were neglected. These are usually normal metal (or lower gap superconductor) films which are in contact with the superconductor and act as heat sinks where the quasiparticles can be absorbed. The effect is based on the fact that (as described in section 2.2.2) the normal metal has exponentially stronger electronic heat diffusion and electron–phonon coupling than the superconductor and hence the excess heat is quickly absorbed to the bath. However, a normal metal island very close to the junction can severely decrease the performance of the cooler as the energy gap of the superconductor is smeared due to the proximity effect. Hence, optimizing the distance of the traps is difficult [39, 40].

A safe way of introducing a trap with moderate improvement in quasiparticle relaxation rate comes for free in junctions fabricated by shadow angle deposition (if some care is taken in designing the leads in the vicinity of the junctions). This fabrication procedure first produces the superconducting (e.g. aluminum) lead, which is subsequently oxidized, and thereafter a metal layer (e.g. copper) is deposited at another angle, forming the NIS junction. Such an overlap structure can be made in the same process outside the junctions to partially cover the superconducting leads by the normal metal via the oxide barrier. The mechanism of quasiparticle thermalization in this structure is via tunnelling of hot quasiparticles into the normal metal. The dashed lines in figure 8(b) show the same calculation as the solid lines but now also including a quasiparticle trap. The improvement is considerable, even though the trap is assumed to be behind a relatively thick (1 kΩ µm²) oxide barrier.

By similar arguments used in obtaining equation (15), we can obtain a thermal relaxation length with the trap,

$$\ell = \left(\frac{2\sqrt{2}\delta\rho_{\sigma}}{\sqrt{\pi}}\right)^{1/2} \left(\frac{k_B T}{\Delta}\right)^{1/4}. \quad (18)$$

Here \(d\) is the thickness of the superconducting lead and \(\rho_{\sigma}\) is the specific resistivity of the trap barrier. For a relatively resistive barrier \(\rho_{\sigma} = 1 k\Omega \mu m^2\), with \(d = 30 nm\) and \(T = 200 mK\), we obtain \(\ell \approx 20 \mu m\), which is about two orders of magnitude shorter than in a bare superconducting wire.

In [37], the diffusion model using quasiparticle density was extended to include the relaxation due to quasiparticle traps and the results seem to agree with experiments [41]. A recent paper [38] considers both the quasiparticle density in the superconductor, the temperature profile of the normal metal trap and the possible 'athermal phonons', i.e. very hot phonons emitted by quasiparticle recombination.

### 3.2.2. Influence of Andreev current on refrigeration

All the equations given above for the current and heatflow through an NIS interface have been derived assuming single-electron tunnelling events through the barrier. This approximation is generally valid for NIS junctions as proven by the good agreement between experiments and equation (13). However, when going to very high transparencies of the junction (in order to maximize the cooling power), a second order process, called Andreev current, starts to be a significant transport mechanism. The Andreev current is essentially a process where a Cooper pair in the superconductor is transported into two quasiparticle excitations in the normal metal or vice versa. It can dominate over the single-particle current at voltages \(V < \Delta/e\). Parameters are derived from the fit to the experimental result. (c) is a reprinted figure with permission from [47]. Copyright 2008 by the American Physical Society.
equals \( P_{AR} = I_{AR} V \), where \( I_{AR} \) is the electrical current due to the Andreev process.

The magnitude of the Andreev current depends on several parameters of the tunnel junction and its electrodes. For small junctions at the ballistic limit (meaning that the dimensions of the junctions are smaller than the mean free path of electrons in the normal metal), it is proportional to voltage such that \( I_{AR} \approx \frac{16N\hbar}{e} R_K V \), where \( N \) is the number of conduction channels, and \( R_K = \frac{\hbar}{e^2} \) is the quantum resistance [43]. This ballistic description gives typically very small values for Andreev current with transparent junctions common in NIS junctions. However, for larger diffusive junctions typical of NIS coolers, the Andreev current is not given by this simple expression. This is basically because disorder in the metals leads to quasiparticle confinement near the interface and they can experience multiple reflections before escaping the junction area. This can lead to orders of magnitude higher values of the Andreev current because of constructive interference between the consequent tunnelling amplitudes. Note that although the same confinement is also present in the single particle case, it gives no enhancement to current as there is no interference between the tunnelling amplitudes. The diffusive case can be analysed theoretically [44-46] and the results seem to agree with experiments [47, 48].

Andreev current can have a significant influence on the cooling power and hence on the temperature of the N electrode at low bias voltages \( |eV| \ll \Delta \) (see figures 9(b) and (c)). In practice, the N island of an SINIS cooler heats up at these intermediate voltages. Yet the contribution of Andreev current becomes less significant at voltages close to the optimal one at \( \sim 2\Delta /e \), and, luckily enough, the achieved temperature reduction is generally almost unaffected by the Andreev effect.

### 3.3. Coulomb-blockaded NIS cooler

In the basic description of a tunnel junction refrigerator, the Coulomb effects are neglected. However, as the size of the cooled normal metal island and the area of the tunnel junctions is decreased, the energy cost of extracting or putting electrons in the normal metal can become considerable. Although this is not a desirable effect for a basic cooler, it does bring up the possibility of controlling the heatflow with a gate, i.e. a heat transistor in analogue to a single-electron transistor (SET) for charge transport.

The energy cost of adding or extracting one electron from a normal metal island, connected to metallic leads through tunnel barriers, is determined by the charging energy \( E_C = e^2 / (2C_T) \), where \( e \) is the electron charge and \( C_T \) the total capacitance of the island. This total capacitance includes both the stray capacitance of the island (determined by the size of the island) and the capacitance of the tunnel junctions (determined by the size of the junctions). In the regime where \( E_C \gg k_B T \), there can be no electron tunnelling events to or from the island, unless the bias voltage over the island overcomes this charging energy cost. This is the effect of Coulomb blockade.

When combined with superconducting leads, a Coulomb-blockaded SINIS device is created. One then has two energy barriers to overcome in order to be able to push current through the device: the superconducting gap and the charging energy. There is, however, one crucial difference between these two: the charging energy level of the island can be adjusted with an external gate. In this way another control parameter is added to the SINIS device and one can now tune the electron current and, hence, the heatflow at a certain bias point by adjusting the gate voltage. Theoretically, the maximum on/off ratio of heatflow at the optimum cooling point reachable this way can be written at the limit \( E_C \gg k_B T \) as [50]

\[
\frac{P_{open}}{P_{closed}} \approx 0.45 \frac{k_B T}{E_C} e^{\frac{E_C}{k_B T}}, \tag{19}
\]

which, for example, for \( E_C / k_B T = 10 \) gives a ratio of 990. Note, however, that even at the gate open position, the maximum cooling power of an SINIS structure is only one half of the corresponding case where Coulomb blockade plays no role, as heatflow through one of the two junctions is effectively blocked at all times.

The heat transistor effect was experimentally demonstrated in [50], where an on/off ratio of over 3 was demonstrated (see figure 10). In the demonstration, Coulomb-blockaded SINIS structures were used for both thermometry and cooling. The correspondence between simulations based on Coulomb-blockaded single-electron tunnelling (orthodox theory) and

---

**Figure 10.** Heat transistor. (a) A schematic of the measurement setup. (b) SINIS thermometer readout (dots connected with lines) as a function of applied cooler voltage under the constant gate sweep condition. Gate open (orange) and gate closed (blue) lines are connected with lines. A cooling peak is present only for the gate open case. Note that the SINIS probe voltage depends both on the gate charge and the temperature, for details see [50]. (c) Theoretical (line) and experimental (dots) values for the on/off ratio \( P_{open} / P_{closed} \) for the cooler used in [50]. The charging energy of this particular structure was relatively low, leading to modest on/off ratios.
The idea was presented in [52] (before the current standard turnstile using only one island. This allows one to construct an accurate current blocked by the superconducting gap at the intermediate gate is because the parameters can be chosen so that the current is would be freely transferred, unlike in the normal case. This during one full pumping cycle between two stable charge states, one does not need to go through any point where current can be used for thermometry also in this regime, although this requires a careful modelling of the charge distribution on the island.

Another possibility arising in the charging energy dominated regime, again analogously with electric current, is the pumping of heat. In recent years, there has been a lot of interest in using superconductor–normal metal SETs as a current standard by pumping single electrons through it, thus connecting frequency to current by the relation $I = e f [51]$, where $f$ is the pumping frequency. The hybrid SET has a definite advantage as compared with the normal metal SET: during one full pumping cycle between two stable charge states, one does not need to go through any point where current would be freely transferred, unlike in the normal case. This is because the parameters can be chosen so that the current is blocked by the superconducting gap at the intermediate gate voltages. This allows one to construct an accurate current turnstile using only one island.

The same principle can also be applied to heatflow. The idea was presented in [52] (before the current standard proposition) and experimentally demonstrated in [53]. If a sinusoidally varying gate signal with proper amplitude is applied to a Coulomb-blockaded SINIS structure, a sequential tunnelling of single electrons with certain energy can be achieved. Each tunnelling will take (on average) an energy $k_B T$ from the island, producing a total cooling power of $k_B T f$. A schematic of the process is presented in figure 11. Although the energy power achievable this way is below the constant bias case, this kind of cooler has the obvious advantage that no bias voltage is in principle needed over the island and there is no need for a galvanic connection between the leads and the island. Also, it is an example of cyclical conversion of work into cooling, illustrating basic thermodynamics at the nanoscale.

3.4. Brownian refrigerator

Above we have discussed how an NIS junction can be used for cooling by either dc biasing or by ac driving (in the Coulomb-blockaded case) through a capacitive coupling. However, a third option exists. If one can couple the NIS junction to a hot environment, the voltage fluctuations of the environment can ‘drive’ the cooler. This leads to a somewhat counterintuitive phenomenon (which also at first sight seems to defy the laws of thermodynamics) where a coupling to a hot environment cools the normal metal.

This kind of Brownian refrigerator was proposed in [54] (see also [55]) and further extended in [56]. There the authors considered a case where one couples an NIS junction to a hot resistor with superconducting leads ideally providing only photonic energy exchange between the resistor and the NIS structure. The hot photons emitted by the resistor can then ‘kick’ the electrons to overcome the energy barrier needed for them to tunnel to the superconductor (no voltage is applied over the junction). If the temperature of the resistor is properly set, then preferentially only the high energy electrons are removed from the normal metal and this again leads to cooling.

Quantitatively, one way to model this effect is the so-called $P(E)$ theory. Environmentally assisted tunnelling (or photon-assisted tunnelling), where photons are exchanged with the electromagnetic environment during the tunnelling event, has been studied extensively in relation to tunnel junctions [57]. In the case considered here (un-biased NIS junction), the heatflow from the normal metal can be shown to follow

$$P_n = \frac{2}{e^2 R_T} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} dE d' n_S(E') E f_S(E) [1 - f_S(E')] P(E - E'),$$

where $P(E)$ is essentially the probability density of the environment to emit a photon with energy $E$. If the
Figure 12. The Brownian refrigerator presented in [54, 56]. (a) Basic circuit diagram. A hot resistor at temperature $T_R$ is connected to an NIS junction with junction resistance $R_T$, and provides energy allowing it to work as a cooler. (b) Energy diagram of photon-assisted tunnelling. Energy $E' - E$ is provided by the environment (the hot resistor) allowing an electron to tunnel to the superconductor from the normal metal. (c) Calculated cooling power of the Brownian refrigerator $P_n$ as a function of the resistor (environment) temperature, calculated with different values of the charging energy $E_C$. Solid lines correspond to junction resistance $R_T = 0.5 R_K$ and dashed lines to $R_T = 10 R_K$.

Figure 13. Remote cooling, presented in [25]. (a) An SEM picture of the sample. The right inset shows a schematic of the operating principle. Two normal metal islands are connected in a long superconducting loop. Both islands’ electronic temperature is probed with NIS thermometers and the other island is also cooled with NIS junctions. Coupling through the transmission line couples the two temperatures at low bath temperatures. (b) Experimental results. Lower lines are the temperature of the directly cooled island and the upper lines are the temperature of the other island, located at 50 $\mu$m distance.

The calculated heat flow out of the normal metal as a function of the environment temperature is presented in figure 12(c). The cooling power is maximized when the temperature of the resistor is around $T_N/\Delta_1/E_C$. It should also be noted that when the resistor is cooler than the normal metal, there is a net heatflow from the superconductor to the normal metal, leading to cooling of the superconductor. This is a regime which is unattainable in a voltage biased NIS junction.

The Brownian refrigerator is still to be experimentally demonstrated. In actual experiments, there are some complications not included in the ideal treatment above. First, in practice the hot resistor must be heated with electric current, which then has to be prevented from flowing through the NIS junction. Second, the finite frequency impedance of the circuit dictates that the heated resistor has to be quite close to the cooler junction, making heatflow through the phonon system a considerable concern. And third, if the resistance of the normal metal island is not totally negligible compared with the resistance of the cooling junction, there will be a parasitic heating power from direct photonic heating from the hot resistor. Nevertheless, taking all these effects into account, it was concluded in [56] that the effect should still be experimentally detectable with realistic parameters.

3.5. Remote cooling

Interestingly, the concept of radiative heat exchange introduced in section 2.3 allows the spatial separation of the micrometre-scale cooler and the cooled object itself when both objects are coupled together in a matched circuit [25]. The scheme in figure 13 depicts the device concept: the actively cooled metal acts as a cold environment for the device, the latter is cooled via a transmission line. Thermometry of both the cooler island and the remotely connected island is done with standard SINIS thermometry. Above about 300 mK, heat is mainly transported between both islands by quasiparticles as the quasiparticle population is not frozen out yet. Superconducting aluminum employed in the experiment as transmission lines provides sufficient thermal isolation, from heat transported by quasiparticles, only at the lowest temperatures. Nevertheless, towards lower temperatures, when in addition the electron–phonon coupling diminishes, the photonic heat transport becomes dominant and couples both islands effectively. This
results in a temperature drop of the remotely cooled island reaching about 60% of the directly cooled island. Moreover, a complete galvanic isolation of the cooler and the cooled device would be achievable if the structures were coupled capacitively or inductively to realize cooling at a distance.

3.6. Effect of magnetic field on NIS cooling

When a magnetic field is applied over a superconductor, it will create surface currents that will cancel the field completely within a penetration depth from the surface, this is the well-known Meissner effect. In a type I superconductor, such as bulk Al, there exists a single well-defined critical field above which the superconductivity is totally suppressed and below which the Meissner effect prevents magnetic field from entering the bulk of the superconductor. The situation changes, however, when the dimensions of the metal film become comparable to the penetration depth. In this thin wire form all superconductors display type II behaviour, having two critical fields. At the lower critical field $B_{c1}$, magnetic vortices start to penetrate the material, creating areas where the superconducting energy gap is locally suppressed but the overall superconducting behaviour is retained. The superconductivity is suppressed only at the higher critical field $B_{c2}$. In addition, the lower critical field is not primarily determined by the material but by the geometry of the wire. It has a universal characteristic value $B_{c1} \sim \Phi_0 / W^2$, where $\Phi_0 = h / 2e$ is the flux quantum and $W$ is the width of the wire (assuming a wire with thickness $\ll W$ and a magnetic field perpendicular to the wire) [58].

It has been shown that the creation of magnetic vortices leads to faster quasiparticle relaxation [59]. The vortices act as quasiparticle traps: the areas with locally suppressed gap also have correspondingly higher e–p coupling and can absorb the ‘hot’ quasiparticles. Combined with the feature that the vortices will first be introduced into the widest parts of the superconductor, this allows one to design the superconducting leads of an NIS cooler so that the thermalization of the lead will be optimized in a small magnetic field. This effect was recently reported in [60], where a very significant improvement of NIS cooler performance was seen in small magnetic fields, where the lead geometry was designed so that it was narrower at the junction area than elsewhere (see figure 14). In the opposite case, vortices are first created at the junction area and will deteriorate the cooler performance.

3.7. Cooling phonons with NIS coolers

A concern regarding NIS coolers in practical applications has been the fact that they directly cool the electron system. In many cases, it would be desirable that the cooler structure itself were electronically isolated from the sample to be cooled. In order to achieve this, one has to somehow thermally couple the sample to the cooled normal metal volume without electrically coupling it. In practice, this means coupling through a phonon system. However, in order to cool a phonon system with an NIS cooler, one has to make the coupling from the phonon system to the environment smaller than the electron–phonon coupling in the normal metal. As the e–p heat current decreases as $T^5$ at low temperatures, this is a very challenging condition.

The most straightforward way to achieve this isolation from the environment is to have the phonon system as a micromachined membrane, on top of which the samples to be cooled are fabricated. This membrane can then be cooled with so-called cold fingers, normal metal leads extending from an NIS cooler to the membrane. The junctions need to be located on the bulk, in order not to let the heat dissipate to the superconductor to couple to the cooled volume.

This kind of membrane cooler would be of considerable interest in many applications of superconducting electronics, ranging from quantum information technology to space borne radiation detectors. In principle, all of the community utilizing aluminum as a superconductor are facing a technological challenge in providing temperatures below 0.1 K where the superconducting properties of Al are optimized. Current solutions, mainly adiabatic demagnetization refrigerators and dilution refrigerators, are complicated to use and, more importantly for space applications, heavy. It would be enormously advantageous to replace these refrigerators with a simple 3He refrigerator or, even better, a pumped 4He bath, combined with an NIS membrane cooler. The first applications to benefit would be the ones where the fabrication onto a membrane is straightforward. This group especially includes radiation detectors which are already often fabricated on a membrane.

Membrane cooling was first demonstrated by Luukanen et al [61] with a small membrane volume coupled to the bath through four few hundred micrometre long and $\sim 5 \mu m$ wide bridges. A considerable temperature decrease was achieved (from 200 to 100 mK), although the actual cooling power was modest ($\sim$ pW). However, actual application demonstrations have been done recently by Ullom’s group at NIST. They first demonstrated the cooling of a macroscopic size Ge cube [62] and then an aluminum transition-edge detector, designed for

![Figure 14. Magnetic field effect to NIS cooling, presented in [60], showing the maximum cooling $\Delta T$, starting from a bath temperature of 285 mK, as a function of applied perpendicular magnetic field. The cooling effect is seen to be enhanced by over a factor of 1.5 by applying a small magnetic field. This is explained as magnetic vortices acting as quasiparticle traps and preventing the overheating of the superconducting lead. In larger fields vortices are created at the junction area, hence degrading the cooling effect.](image-url)
x-ray detection [63]. In the latter experiment, an effective temperature reduction from 300 to 190 mK was achieved in the noise properties of the detector, presenting a significant technological advance (see figure 15). The authors tested that inducing a 22 pW heating power to the membrane reduced the cooling by 7 mK, which would suggest an effective total cooling power of the order of some hundreds of pW.

Another geometry where the needed isolation of the phonon system from the environment can be achieved is in the form of a nanosized beam. The integration of other samples in the beam geometry is generally not very convenient but in these cases the beam itself can be the sample. In recent years, there has been much interest in the cooling down of local mechanical objects. This has so far been achieved only in piezoelectric modes [64]. The problem here is that in order to demonstrate the quantization, these modes need to be very weakly coupled to the phonon bath of the bulk substrate (i.e. the Q-value of the resonator needs to be high). This condition makes the cooling mediated by the bulk phonon bath more difficult as there is inevitably some dissipated power generated by the measurement of the vibrations. Making the beam out of normal metal connected to NIS junctions would circumvent this problem as the local modes would then be directly cooled (through the electrons).

This scenario was considered theoretically in [65] (see figure 16(a)). It had been experimentally verified earlier that e–p coupling can depend on the dimensionality of the phonon system [66]. In [65], the authors showed that assuming a 1D phonon population and doing the conventional calculation [6] for electron–phonon coupling yields a $T^3$ power law for the heat flow

$$P_{e-p}^{1D} = \frac{\pi \xi(3)}{6 \xi(5)} \left( \frac{\hbar c}{k_B} \right)^2 \Sigma L(T_c^3 - T_{ph}^3),$$

where $\xi(3)/\xi(5) \simeq 1.16$, $c_1$ is the speed of sound of the longitudinal modes, $\Sigma$ is the same electron–phonon coupling constant as in the 3D case and $L$ is the length of the beam. This coupling is between the electrons and the longitudinal phonon modes, since in the first order perturbation theory the electrons do not couple to the transversal modes, i.e. the flexural modes of the beam. With this coupling, cooling the longitudinal modes of the resonator below the phonon bath temperature with an NIS cooler should be possible as long as the $Q$-factor of the resonator is above 100, an easy requirement to meet with mechanical resonators. The fabrication techniques needed for this kind of beam cooler were demonstrated in [67] (see figure 16(b)), but the coupling between the electron system and the local mechanical modes remains an open question. Lately, many other proposals on cooling down the mechanical modes of a metallic resonator have also been reported [68–71]. These do not rely on NIS junctions.

In [72], a hybrid solution between the beam geometry and membrane geometry was fabricated (see figure 16(c)). Here the beam was made of silicon nitride and was connected to the bath only through narrow bridges. The thermal conductivity model is hence the same as in the earlier demonstrations of membrane cooling (including the cold fingers), but now the cooled phonon system is in the form of a beam. In this way, the authors were able to also cool the presumably 1D phonons of the beam but the experiment was not directly sensitive to any dimensionality or localization effects of the wire phonons. Nevertheless, the authors saw a power law of 2.8 at the lowest temperatures and attributed this to the 1D–2D phonon scattering at the bridge–bulk interface.

4. SIS/ coolers

In all the NIS coolers presented in the previous section, the superconducting material used was aluminum (Al). As Al naturally forms a very high quality oxide layer when it is exposed to oxygen, this makes the fabrication of high quality tunnel barriers easy and has made Al by far the most popular superconducting material to use in the fabrication of tunnel barriers. However, when considering coolers, using some other materials with higher $T_C$ would have the obvious advantage of moving the optimum cooling temperature higher and providing higher cooling powers. This would be a very important step for technological applications. One might even envision a cascade cooler cooling from 4 K to below 100 mK utilizing different superconducting materials. So far,
However, most attempts have been hindered by fabrication difficulties. The main problem lies in achieving low leakage junctions with other insulating barrier materials than aluminum oxide. One of the most common superconducting material is niobium (Nb), which has the advantage of having the highest $T_c$ ($\sim 9$ K) among metallic elements and is also relatively easy to sputter deposit. However, achieving tunnel junctions with sufficiently low leakage currents has proved problematic with Nb and has so far prevented all demonstrations of significant cooling. Some progress has been recently made, however, using aluminum nitride instead of aluminum oxide [73].

A simple solution to this problem is to use the Al as a normal metal on top of which one can grow the oxide layer and then deposit a superconducting layer. This can be done and has been done by suppressing the superconductivity of Al with magnetic impurities, specifically manganese (Mn) [74]. In [74] the actual superconductor was, however, still Al. The approach had the advantage that the superconducting layer could be made arbitrarily thick (which is not possible when a superconductor is deposited as the first layer due to fabrication technicalities). The thicker Al layer makes the quasiparticle overheating effects discussed in section 3.2.1 less harmful. This kind of Al–AlMn cooler was used in the membrane cooling demonstrations presented in section 3.7 [62, 63]. More recently, a trap layer was also introduced on top of the superconducting layer [38] in AlMn coolers. This means that the superconductor is then sandwiched between two normal metal films, and this improves the cooling performance.

Another solution is to use a different $T_c$ (and hence different energy gap) superconductor on top of an Al layer. This kind of SIS’ (superconductor–insulator–superconductor with a different gap) cooler [75] was first demonstrated with Ti/Al junctions [76]. The cooling power of an SIS’ junction is, in analogy to NIS cooling,

$$P_{SIS} = \frac{1}{e^2 R_T} \int dE (E - eV) n_S(E) n_S(E - eV)$$

$$\times \left[ f_S(E - eV) - f_S(E) \right].$$

Equation (22) is very similar to equation (9) but has a few important qualitative differences: (i) cooling power is maximized when the voltage is $\Delta_S - \Delta_S/e$ and (ii) at the optimum cooling point, the DOS diverges on both sides of the tunnel junction giving rise to nominally infinite cooling power. In figure 17, we plot the normalized $P_{SIS}$ as a function of $V$ for different values of $\Delta_S$.

The SINIS analogy was later found to be too simplistic [77] for the case of an SIS’IS structure. When applying a current over a superconducting tunnel junction, the voltage will not in general develop gradually but rather there is an abrupt change from the supercurrent state to a quasiparticle current state when the current exceeds the critical current of the junction. This is commonly known as switching of a Josephson junction and will lead voltage $\Delta/e$ to abruptly appear over the junction. As two tunnel junctions will never be exactly the same in a double tunnel junction structure, one junction will tend to switch with lower current and develop a voltage over it. When more current is then applied over the structure, the voltage across the first junction will increase until the other junction also switches. This takes the voltage $\sim \Delta/e$ over it, causing the voltage over the first junction to decrease. When the first switching happens, the voltage reaches the difference of the two gaps $V \approx (\Delta_S - \Delta_S)/e$. The other junction is then driven to the cooling regime and a drop in $S'$ quasiparticle temperature can be seen (see figure 18 from [77]). With increasing bias, the sharp cooling peak will turn into heating until the other junction also switches, at which point a second cooling peak is seen at $V \approx 2(\Delta_S - \Delta_S)/e$.

In [77] the authors used the critical current of two additional SIS’ probe junctions attached to the cooled Ti island as thermometers. With the SIS’IS cooler biased at the optimum point, they could reach a critical current on the probe junctions at 0.5 K that exceeded the critical current they could see at the lowest cryostat temperature $(\sim 50$ mK) without biasing the cooler. This suggests that the $S'$ island quasiparticle population was cooled from 0.5 K to below what it could be cooled through the phonon system.

A fully Al SIS’ cooler was presented in [78], where the difference in energy gaps was achieved using Al layers with different thicknesses. The energy gap of aluminum increases...
Figure 17. SIS' cooler. (a) The energy diagram of an SIS' structure. Basic principles are identical to figure 7. The qualitative differences are the different optimal cooling voltage and the divergent DOS when the two gap edges are aligned. (b) Bias dependence of the cooling power of an SIS' junction. The smooth bottom line is the calculation for the corresponding SIN junction. The different lines are for different values of the energy gap of the second superconductor $\Delta_{S'}$. From left to right, $\Delta_{S'}$ is 0.6, 0.5, 0.4, 0.3, 0.2, 0.1 and 0.05 times $\Delta_{S}$. The cooling power diverges at $eV = \Delta_{S} - \Delta_{S'}$.

Figure 18. SIS'IS cooling. (a) The experimental structure. The Ti island is connected to four Al leads through aluminum oxide tunnel barriers. Two of the junctions are used for cooling and the other two for thermometry. (b) Current–voltage characteristics of the thermometer junctions measured with different biasing of the cooling junctions. The critical current of the junctions is seen to first decrease, then increase, decrease and increase again, before the supercurrent finally quenches at high biases. (c) The thermometer critical current seen at different bath temperatures as a function of the biasing of the cooler junctions. An increase in the maximum supercurrent corresponds to cooling of the quasiparticle system of the Ti island. A double peak structure is seen (see text). (d) The maximum thermometer supercurrent as a function of the bath temperature at zero cooler bias, optimal bias and one point in between. Reprinted figure with permission from [77]. Copyright 2008 by the American Physical Society.

when the thickness of the metal layer is below a few tens of nanometres. In [78] Al layers of 30 and 10 nm were used, giving energy gaps of 209 $\mu$eV and 250 $\mu$eV, respectively. The SIS' structure was then used to cool down one of the electrodes of a single Cooper pair transistor (SCPT). An SCPT is essentially an SET working with Cooper pairs instead of single electrons. SCPTs are, however, prone to quasiparticle poisoning [79] causing the devices to have 1$e$ periodicity in their gate voltage characteristics in addition to the expected 2$e$. By biasing the SIS' structure which then cooled one of the leads of an SCPT, the author was able to show that the probability of having unpaired electrons on the island went down by a factor of two.

None of the aforementioned coolers addressed the cooling from higher temperatures that was given as a motivation at the beginning of the section. Lately significant progress in this direction has been made in [80] where the authors succeeded in making a vanadium (V, $T_{c} \sim 5$ K)–aluminum SIS' cooler (see figure 19). In order to make a good quality junction it was necessary to first cover the oxide layer with a small amount of Al before depositing V. The Al/V bilayer had a $T_{c}$ of 4 K and the authors were able to achieve a significant temperature reduction as deduced from the critical current of two additional probe junctions: they cooled the quasiparticle system of Al from 1 K to 400 mK.

5. Schottky barrier coolers

The basic principles of NIS cooling also apply if the normal metal island is replaced by a heavily doped semiconductor (see figure 20(a)). The superconductor–semiconductor (S–Sm) cooler presents some benefits compared with normal metals: (i) the electron–phonon coupling strength is generally weaker in semiconductors than in normal metals (at 100 mK, Si, depending on the doping level, has roughly 1–2 orders of magnitude smaller e–p coupling than Cu) and (ii) the Schottky barrier can play the role of the tunnelling barrier and hence no oxide layer is needed between the superconductor and the
Figure 19. Vanadium cooler. (a) The cooler structure. An Al island is connected to four V leads, two for thermometry and two for cooling. Above the aluminum oxide a thin layer of Al is deposited before the V deposition in order to protect the oxide layer. (b) Thermometer current–voltage characteristics at different values of the cooler biasing. From left to right, the curves refer to biasing points shown by circles in (c). (c) The supercurrent of the thermometer junctions as a function of the cooler biasing. A clear cooling peak can be seen at $V \sim 1.2\,\text{mV}$. In the middle curve ($T_{\text{bath}} = 1\,\text{K}$), the increase in critical current corresponds to quasiparticle temperature of 0.4 K at the Al island. Reprinted with permission from [80]. Copyright 2011, American Institute of Physics.

Figure 20. The energy diagram of an S–Sm–S structure. Basic principles are identical to those in figure 7. The insulating layers are replaced by a Schottky barrier at the Al–Si interface. The energy gap of the semiconductor plays no role, as the island is degenerately doped.

This makes fabrication of especially large area junctions more straightforward than with the standard shadow evaporation techniques. In addition, both the Schottky barrier resistance and the electron–phonon coupling can be tuned by varying the doping level of the semiconducting island. The most obvious drawbacks are that even highly doped semiconductors have a higher resistivity than metals and hence more parasitic Joule heating is generated. Furthermore, relatively large sub-gap currents are typically observed leading to non-ideal cooler performance.

The cooling effect in S–Sm structures was first presented in [81] and extended in [82, 83]. In [81], a cooling power of roughly 0.5 pW was achieved with two $5 \times 18 \,\mu\text{m}^2$ junctions having total $R_T$ of 800 $\Omega$. This led to a 30% drop in temperature from 175 mK because of the small e–p coupling. The doping level of the n$^+$ silicon was $4 \times 10^{20}\,\text{cm}^{-3}$. In [82, 83], the work was extended to multiple n$^+$ doping levels of the semiconducting island. It was found that, in agreement with the theory, the contact resistance $R_T$ of the Al–Si interface scaled as $\exp(N^{-1/2})$ where $N$ is the doping level. For the cooling effect, this is partly compensated by the increase in the e–p coupling due to higher doping. However, the latter effect was found to be only linearly proportional to doping and hence larger doping should lead to increase in cooling power. Yet the larger cooling effect was seen only at higher temperatures (above $\sim 300\,\text{mK}$) and increasing doping to above $1 \times 10^{20}\,\text{cm}^{-3}$ made the cooling effect smaller. This was attributed to large ohmic leakage currents through the barrier at lower transparencies, i.e. effectively the $\gamma$ parameter in equation (8). The $\gamma$ generally found in Al–Si junctions has been $10^{-2}$–$10^{-1}$, which is a few orders of magnitude worse than in Al–Al$_2$O$_3$–Cu junctions. In [84], also niobium–silicon junctions were studied. The basic IV's of SINIS structures could also be reproduced with Nb, but no cooling was observed. This was again due to large sub-gap leakage currents. The contact resistance between Nb and Si was found to be much smaller than with Al and Si, in accordance with expectations as the Schottky barrier height is smaller in this case.

In [81–83], e–p coupling in Si was modelled with the $T^5$ power law as in the normal metal case. However, later it was confirmed to follow a higher $T^6$ law [85]. Theoretically, this power law was expected for (single-valley) semiconductors in two dimensions at the diffusive limit [86] but the theoretical prefactor was several orders of magnitude smaller than measured. In [87, 88], the fact that multiple conduction band valleys exist in Si was included in the theoretical analysis. Phonons can lift the degeneracy between the different valleys so that the valley degree of freedom starts to play a role in the low temperature e–p coupling. This was shown to lead to the $T^6$ power law and, because this channel is unscreened at low temperatures, have a prefactor of the correct order of magnitude with experiments. The prefactor and temperature dependence were also experimentally confirmed in [87].
Recently [89, 90], it has been tested how removing this relaxation channel will modify the e–p relaxation in Si. Similarly as phonons lift the degeneracy between the different valleys, one can also lift the degeneracy ‘permanently’ by inducing strain to the silicon layer. If the strain induced energy splitting is larger than the Fermi energy (as measured from the bottom of the conduction band), the in-plane valleys will depopulate and hence the screened, single-valley case prevails. Theoretically, the e–p coupling should decrease by several orders of magnitude [88]. The effect was tested experimentally in [89] (see figure 21). The authors found that the e–p coupling was indeed lower in the strained sample as compared with an unstrained control sample and to previous experiments, although the reduction factor was only between one and two orders of magnitude. Nevertheless, the heat flow from the phonon system decreases significantly and it can be useful for cooler applications, as was demonstrated in [90] where enhanced electron cooling in strained silicon sample was seen.

6. Quantum dot refrigerator

Refrigeration using a semiconducting quantum dot, instead of a metal hybrid, had been proposed by Edwards et al [91, 92]. The principle of such a quantum dot refrigerator (QDR) is shown in figure 22. The proposed device consisted of a central island which is separated from the leads by two quantum dots, A and B. The dots have an energy level separation of $\Delta$ and the dot energies, $E_A$ and $E_B$, can be adjusted by separate gate voltages. It is assumed that the dots are tunnel coupled both to the leads and to the island but no coupling exists directly between the leads and the island. If one applies a small dc voltage across the leads so that the energy separation between the chemical potentials $\mu_L$ and $\mu_R$ is smaller than $2\Delta$, the chemical potential of the island $\mu_0$ will lie midway between $\mu_L$ and $\mu_R$. There will then be exactly one or zero energy levels between $\mu_0$ and the chemical potentials of either lead. By positioning the energy level $E_A$ ($E_B$) so that it is slightly above (below) $\mu_0$, an energy $E_B - E_A$ is removed from the central area as an electron travels from one lead to another. For this process to be energetically possible, the separations between the energy levels of the quantum dots and the island must be of the order of $k_0T$.

The concept was experimentally tested and demonstrated recently by Prance et al [2] in a 2DEG, which is also the system Edwards et al originally proposed. 2DEGs have a very weak coupling to the acoustic phonons of the lattice which both makes cooling them by conventional methods hard and facilitates significant changes to the electron temperature with even modest cooling powers by direct electronic cooling. In the experiment of [2], a $6 \mu m^2$ central area of a 2DEG was cooled from 280 mK down to 187 mK under optimized bias conditions of the device, see figure 23. The data are consistent with a thermal model (quasi-equilibrium) down to about 120 mK bath temperature. Below that the cooling becomes ineffective and the data cannot be fit to a simple model where conductance is parametrized by temperature. The authors ascribe this to poor electron–electron relaxation at low temperatures leading to a non-equilibrium energy distribution.

More recently [93], local thermometry of the 2DEG reservoir was demonstrated with a similar structure. The temperature of the reservoir was deduced from the thermally broadened conductance of a quantum dot, allowing the measurement of the electron–phonon coupling constant of the 2DEG. An alternative cooler mechanism for 2DEG has been proposed in [94].
7. Perspectives

It has been 45 years since sub-100 mK temperatures were first opened up to researchers by the advent of dilution refrigerators. Although this technique has steadily matured and proven to be a very important workhorse for low-temperature research, it remains complicated and is getting increasingly expensive due to the unstable price of $^3$He. A significant demand exists for alternative techniques that would not require users to be experts in low-temperature techniques and, ideally, would not use cryoliquids. The micrometre-scale coolers presented in this review have the potential to provide these advantages in the future, although a significant amount of research is still needed. This goal could be accomplished, for example, by combining already available commercial pulsed cryocoolers with different superconducting materials.

Important aspects of the development process of micrometre-scale coolers, which have been mostly not mentioned in this review, are the fabrication techniques. The overwhelming majority of superconducting coolers demonstrated so far have been fabricated with e-beam lithography (EBL) combined with multiple angle evaporation. Although a very useful technique in laboratory settings, this can hardly be considered an industrial process and does place severe limitations on the junction sizes (and hence the cooling powers) that can be produced. Larger Al–Al$_2$O$_3$–Cu junctions were demonstrated ten years ago with mechanical masks combined with multiple angle evaporation [95]. Another possibility raised early on is to use a degenerately doped semiconductor instead of normal metal (see section 5).

However, more recently, demonstrations have been made with photolithography. A prime example is the AlMn based coolers presented in section 4 and used in the membrane cooling demonstrations (see section 3.7). Another very recent example is presented in figure 24. NIS coolers can now achieve effective cooling power of hundreds of picowatts, which is a considerable increase from the sub-picowatt range of the original demonstrations, but there is still a lot of room for improvement.

![Fabrication process for large area junctions presented in [96]. (a) Process flow: Al–Al$_2$O$_3$–Cu structure is evaporated on a substrate and covered with a resist. Dots are patterned to the resist (with photolithography) and etched through the copper with an ion-beam-etcher. Aluminum is wet etched through the holes, creating a suspended Cu structure in the middle. Finally, in a second lithography (EBL) step, the cooler junctions are defined by cutting through the Cu at a distance of a few micrometres from the holes. With this technique very large area junctions can be achieved and the thickness of the Al layer is not limited (facilitating easier removal of hot quasiparticles). (b) Optical microscope image of a cooler structure. The two junctions at the top of the picture are for thermometry. The large cooling junctions are defined by the cut in copper (showing as light vertical stripes) and the cut in Al in the middle. (c) View of the middle normal metal part, showing that it is freely suspended.](image)

It has, however, become increasingly clear that the generalization of superconducting coolers to high cooling powers is not as straightforward as one might think. The high density of non-equilibrium quasiparticles created means that thermalization of the ‘backside’ of the cooler and local phonon heating become significant concerns which degrade the efficiency of the cooler at high cooling powers. These issues must be addressed in designing any high power cooler.
The solutions presented so far rely on quasiparticle traps (see section 3.2.1) and separating the cooled phonon system from the phonons in contact with the cooler (perforated membranes as in section 3.7). As was mentioned in section 3.6, it was also recently found out that small magnetic fields can help in this respect.

On the other hand, in some applications, the high cooling power is not important but rather the goal is to reach as low as possible electronic temperatures. In this regime, significant progress has also been made lately, as the significance of the coupling of an electronic conductor to its electromagnetic environment has become clear in this context (see section 2.3). This has also helped to increase the understanding of the ultimate limitations of the quality of NIS junctions.

Recently, several new kinds of micrometre-scale coolers have been proposed. Partly this has been driven by the explosion of interest in cooling of the local mechanical modes of nanomechanical resonators. This is good news for the field and shows that for micrometre-scaled coolers, there remains a lot of research to do, both for finding totally new directions as well as in improving the known ones.

Acknowledgments

We wish to thank T Aref for proofreading the manuscript and V F Maisi, J T Peltonen, M Prunnila and A V Timofeev for providing material for the article. We acknowledge financial support by the EPSRC grant EP/F040784/1 and the European Community’s FP7 Programme under Grant Agreement No 228464 (MICROKELVIN, Capacities Specific Programme).

References

[1] Giazotto F, Heikkilä T T, Luukanen A, Savin A M and Pekola J P 2006 Opportunities for mesoscopics in thermometry and refrigeration: physics and applications Rev. Mod. Phys. 78 217–74
[2] Prance J R, Smith C G, Griffiths J P, Chorley S J, Anderson D, Jones G A C, Farrer I and Ritchie D A 2009 Electronic refrigeration of a two-dimensional electron gas Phys. Rev. Lett. 102 146602
[3] Pekola J P, Heikkilä T T, Savin A M, Flyktman J T, Giazotto F and Heikkilä K J 2004 Limitations in cooling electrons using normal-metal–superconductor tunnel junctions Phys. Rev. Lett. 92 056804
[4] Pothier H, Guérin S, Birge N O, Esteve D and Devoret M H 1997 Energy distribution function of quasiparticles in mesoscopic wires Phys. Rev. Lett. 79 3490–3
[5] Rajauria S, Luo P S, Fournier T, Heikkilä F W J, Courtois H and Pannetier B 2007 Electron and phonon cooling in a superconductor–normal-metal–superconductor tunnel junction Phys. Rev. Lett. 99 047004
[6] Wellstood F C, Urbina C and Clarke J 1994 Hot-electron effects in metals Phys. Rev. B 49 5942–55
[7] de Visser P J, Baselmans J J A, Diener P, Yates S J C, Endo A and Klapwijk T M 2011 Number fluctuations of sparse quasiparticles in a superconductor Phys. Rev. Lett. 106 167004
[8] Palmer B S, Sanchez C A, Naik A, Manheimer M A, Schneiderman J F, Echternach P M and Wellstood F C 2007 Steady-state thermodynamics of nonequilibrium quasiparticles in a Cooper-pair box Phys. Rev. B 76 054501
[9] Shaw M D, Lutchyn R M, Delsing P and Echternach P M 2008 Kinetics of nonequilibrium quasiparticle tunnelling in superconducting charge qubits Phys. Rev. B 78 024503
[10] Martins J M, Ansmann M and Aumentado J 2009 Energy decay in superconducting Josephson-junction qubits from nonequilibrium quasiparticle excitations Phys. Rev. Lett. 103 097002
[11] Barends R, Baselmans J J A, Yates S J C, Gao J R, Hovenier J N and Klapwijk T M 2008 Quasiparticle relaxation in optimally excited high-q superconducting resonators Phys. Rev. Lett. 100 257002
[12] Saira O P, Kempttinen A, Maisi V F and Pekola J P 2011 Is aluminum a perfect superconductor? arXiv:1106.1326v2
[13] Arutyunov K Yu, Auranева H-P and Vasenko A S 2011 Spatially resolved measurement of nonequilibrium quasiparticle relaxation in superconducting AI Phys. Rev. B 83 104509
[14] Kaplan S B, Chi C C, Langenberg D N, Chang J J, Jafarey S and Scalapino D J 1976 Quasiparticle and phonon lifetimes in superconductors Phys. Rev. B 14 4854–73
[15] Timofeev A V, García C P, Kopnin N B, Savin A M, Meschke M, Giazotto F and Pekola J P 2009 Reconfiguration-limited energy relaxation in a Bardeen–Cooper–Schrieffer superconductor Phys. Rev. Lett. 102 017003
[16] Bardeen J, Rickayzen G and Tewordt L 1959 Theory of the thermal conductivity of superconductors Phys. Rev. 113 982–94
[17] Virtanen P and Heikkilä T T 2007 Thermoelectric effects in superconducting proximity structures Appl. Phys. A 89 625–37
[18] Pekola J T, Virtanen P, Meschke M, Koski J V, Heikkilä T T and Pekola J P 2010 Thermal conductance by the inverse proximity effect in a superconductor Phys. Rev. Lett. 105 097004
[19] Ojanen T and Heikkilä T T 2007 Photon heat transport in low-dimensional nanostructures Phys. Rev. B 76 073414
[20] Ojanen T and Jauho A-P 2008 Mesoscopic photon heat transistor Phys. Rev. Lett. 100 155902
[21] Pascal L M A, Courtois H and Heikkilä F W J 2011 Circuit approach to photonic heat transport Phys. Rev. B 83 125113
[22] Pendry J B 1983 Quantum limits to the flow of information and entropy J. Phys. A: Math. Gen. 16 2161
[23] Schmidt D R, Schoelkopf R J and Cleland A N 2004 Electron and phonon cooling in a superconductor–normal-metal–superconductor tunnel junction Nature 444 187–90
[24] Timofeev A V, Helle M, Meschke M, Möttönen M and Pekola J P 2009 Electronic refrigeration at the quantum limit Phys. Rev. Lett. 102 200801
[25] Pekola J P, Maisi V F, Kafanov S, Chekurov N, Kempttinen A, Pashkin Yu A, Saira O-P, Möttönen M and Tsai J S 2010 Environment-assisted tunnelling as an origin of the Dynes density of states Phys. Rev. Lett. 105 026803
[26] Barends R et al 2011 Minimizing quasiparticle generation from stray infrared light in superconducting quantum circuits Appl. Phys. Lett. 99 113507
[27] Paik H et al 2011 Observation of high coherence in Josephson junction qubits measured in a three-dimensional circuit QED architecture Phys. Rev. Lett. 107 240501
[28] Angell D V and Pekola J P 2001 Noise in refrigerating tunnel junctions and in microbolometers J. Low Temp. Phys. 123 197–218
[29] Müller H-O and Chao K A 1997 Electron refrigeration in the tunnelling approach J. Appl. Phys. 82 453–6
[30] Dubois P, Courtois H, Pannetier B, Wilhelm F K, Zaikin A D and Schön G 2001 Josephson critical current in a long mesoscopic S-N-S junction Phys. Rev. B 63 064502
[32] Heikkilä T T, Särkkä J and Wilhelm F K 2002 Supercurrent-carrying density of states in diffusive mesoscopic Josephson weak links Phys. Rev. B 66 184513

[33] Jiang Z, Lim H, Chandrasekhar V and Eom J 2003 Local thermometry technique based on proximity-coupled superconductor/normal-metal/superconductor devices Appl. Phys. Lett. 83 2190–2

[34] Courtois H, Meschke M, Peltonen J T and Pekola J P 2008 Origin of hysteresis in a proximity Josephson junction Phys. Rev. Lett. 101 067002

[35] Meschke M, Peltonen J, Courtios H and Pekola J 2009 Calorimetric readout of a superconducting proximity-effect thermometer J. Low Temp. Phys. 154 190–8

[36] Ullom J N, Fisher P A and Nahum M 1998 Energy-dependent quasiparticle group velocity in a superconductor Phys. Rev. B 58 8225

[37] Rajauria S, Courtios H and Pannetier B 2009 Quasiparticle-diffusion-based heating in superconductor tunnelling microcoolers Phys. Rev. B 80 144521

[38] O’Neill G C, Lowell P J, Underwood J M and Ullom J N 2011 Observations and modeling of large area normal-metal/insulator/superconductor refrigerator cooling from 300 mK to below 100 mK arXiv:1109.1273

[39] Pekola J P, Anghel D V, Suppula T I, Suoknuuti J K, Manninen A J and Manninen M 2000 Trapping of quasiparticles of a nonequilibrium superconductor Appl. Phys. Lett. 76 2782–4

[40] Court N A, Ferguson A J, Lutchyn R and Clark R G 2008 Quantitative study of quasiparticle traps using the single-Coooper-pair transistor Phys. Rev. B 77 100501

[41] Rajauria S, Pascal L M A, Gandit P, Hekking F W J, Pannetier B and Courtios H 2011 Efficient quasiparticle evacuation in superconducting devices arXiv:1106.4949

[42] Bardas A and Averin D 1995 Peltier effect in normal-metal–superconductor microcontacts Phys. Rev. B 52 12873

[43] Averin D V and Pekola J P 2008 Nonadiabatic charge pumping in a hybrid single-electron transistor Phys. Rev. Lett. 101 066801

[44] Hekking F W J and Nazarov Yu V 1993 Interference of two electrons entering a superconductor Phys. Rev. Lett. 71 1625–8

[45] Hekking F W J and Nazarov Yu V 1994 Subgap conductivity of a superconductor–normal-metal tunnel interface Phys. Rev. B 49 6847–52

[46] Vasenko A S, Bezguliy E V, Courtios H and Hekking F W J 2010 Electron cooling by diffusive normal metal—superconductor tunnel junctions Phys. Rev. B 81 094513

[47] Rajauria S, Gandit P, Fournier T, Hekking F W J, Pannetier B and Courtios H 2008 Andreev current-induced dissipation in a hybrid superconducting tunnel junction Phys. Rev. Lett. 100 207002

[48] Lowell P J, O’Neil G C, Underwood J M and Ullom J N 2011 Andreev reflections in micrometre-scale normal-insulator–superconductor tunnel junctions arXiv:1110.4839

[49] Rajauria S 2008 Electronic refrigeration using superconducting tunnel junctions PhD Thesis Université Joseph Fourier

[50] Saira O-P, Meschke M, Giazotto F, Savin A M, Möttönen M and Pekola J P 2007 Heat transistor: demonstration of gate-controlled electronic refrigeration Phys. Rev. Lett. 99 027203

[51] Pekola J P, Vartiainen J J, Möttönen M, Saira O-P, Meschke M and Averin D V 2008 Hybrid single-electron transistor as a source of quantized electric current Nature Phys. 4 120–4

[52] Pekola J P, Giazotto F and Saira O-P 2007 Radio-frequency single-electron refrigerator Phys. Rev. Lett. 98 037201

[53] Kafanov S, Kemppinen A, Pashkin Yu A, Meschke M, Tsai J S and Pekola J P 2009 Single-electronic radio-frequency refrigerator Phys. Rev. Lett. 103 120801

[54] Pekola J P and Hekking F W J 2007 Normal-metal–superconductor tunnel junction as a Brownian refrigerator Phys. Rev. Lett. 98 210604

[55] Van den Broeck C and Kawai R 2006 Brownian refrigerator Phys. Rev. Lett. 96 210601

[56] Peltonen J T, Helle M, Timofeev A V, Solinas P, Hekking F W J and Pekola J P 2011 Brownian refrigeration by hybrid tunnel junctions Phys. Rev. B 84 144505

[57] Grabert H and Devoret M H (ed) 1992 Single Charge Tunneling—Coulomb Blockade Phenomena in Nanostructures (NATO ASI Series B: Physics vol 294) (New York: Plenum)

[58] Stan G, Field S B and Martinis J M 2004 Critical field for complete vortex expulsion from normal superconducting strips Phys. Rev. Lett. 92 097003

[59] Ullom J N, Fisher P A and Nahum M 1998 Magnetic field dependence of quasiparticle losses in a superconductor Appl. Phys. Lett. 73 2494–6

[60] Peltonen J T, Muonenen J T, Meschke M, Kopnin N B and Pekola J P 2011 Magnetic-field-induced stabilization of non-equilibrium superconductivity arXiv:1108.1544

[61] Luukanen A, Leivo M M, Suoknuuti J K, Manninen A J and Pekola J P 2000 On-chip refrigeration by evaporation of hot electrons at sub-kelvin temperatures J. Low Temp. Phys. 120 281–90

[62] Clark A M, Miller N A, Williams A, Ruggiero S T, Hilton G C, Vale L R, Beall J A, Irwin K D and Ullom J N 2005 Cooling of bulk material by electron-tunnelling refrigerators Appl. Phys. Lett. 86 173508

[63] Miller N A, O’Neil G C, Beall J A, Hilton G C, Irwin K D, Schmidt D R, Vale L R and Ullom J N 2008 High resolution x-ray transition-edge sensor cooled by tunnel junction refrigerators Appl. Phys. Lett. 92 163501

[64] O’Connell A D et al 2010 Quantum ground state and single-photon control of a mechanical resonator Nature 464 697–703

[65] Hekking F W J, Niskanen A O and Pekola J P 2008 Electron–phonon coupling and longitudinal mechanical-mode cooling in a metallic nanowire Phys. Rev. B 77 033401

[66] Karvonen J T and Maasilta I J 2007 Influence of phonon dimensionality on electron energy relaxation Phys. Rev. Lett. 99 145503

[67] Muonenen J T, Niskanen A O, Meschke M, Pashkin Yu A, Tsai J S, Sainiemi L, Franssila S and Pekola J P 2009 Electronic cooling of a submicron-sized metallic beam Appl. Phys. Lett. 94 073101

[68] Sonne G, Peña Aza M E, Gorelik L Y, Shekhter R I and Jonson M 2010 Cooling of a suspended nanowire by an ac Josephson current flow Phys. Rev. Lett. 104 226802

[69] Sonne G and Gorelik L Y 2011 Ground-state cooling of a suspended nanowire through inelastic macroscopic quantum tunnelling in a current-biased Josephson junction Phys. Rev. Lett. 106 167205

[70] Santandrea F, Gorelik L Y, Shekhter R I and Jonson M 2011 Cooling of nanomechanical resonators by thermally activated single-electron transport Phys. Rev. Lett. 106 186803

[71] Chamon C, Mucciolo E R, Arrachea L and Capaz R B 2011 Heat pumping in nanomechanical systems Phys. Rev. Lett. 106 135504

[72] Koppinen P J and Maasilta I J 2009 Phonon cooling of nanomechanical beams with tunnel junctions Phys. Rev. Lett. 102 165502

[73] Zijlstra T, Lodewijk C F J, Vercruyssen N, Tichelaar F D, Loudkov D N and Klapwijk T M 2007 Epitaxial aluminum
nitride tunnel barriers grown by nitridation with a plasma source Appl. Phys. Lett. 91 233102

[74] Clark A M, Williams A, Ruggiero S T, van den Berg M L and Ullom J N 2004 Practical electron-tunnelling refrigerator Appl. Phys. Lett. 84 625–7

[75] Frank B and Krech W 1997 Electronic cooling in superconducting tunnel junctions Phys. Lett. A 235 281–4

[76] Manninen A J, Suoknahti J K, Leivo M M and Pekola J P 1999 Cooling of a superconductor by quasiparticle tunnelling Appl. Phys. Lett. 74 3020–2

[77] Tirelli S, Savin A M, García C P, Pekola J P, Beltram F and Giazotto F 2008 Manipulation and generation of supercurrent in out-of-equilibrium Josephson tunnel nanojunctions Phys. Rev. Lett 101 077004

[78] Ferguson A J 2008 Quasiparticle cooling of a single Cooper pair transistor Appl. Phys. Lett. 93 052501

[79] Tuominen M T, Hergenrother J M, Tighe T S and Tinkham M 1992 Experimental evidence for parity-based 2e periodicity in a superconducting single-electron tunnelling transistor Phys. Rev. Lett. 69 1997–2000

[80] Quaranta O, Spathis P, Beltram F and Giazotto F 2011 Cooling electrons from 1 to 0.4 K with V-based nanorefrigerators Appl. Phys. Lett. 98 032501

[81] Savin A M, Prunnila M, Kivinen P P, Pekola J P, Ahopelto J and Manninen A J 2001 Efficient electronic cooling in heavily doped silicon by quasiparticle tunnelling Appl. Phys. Lett. 79 1471–3

[82] Savin A, Prunnila M, Ahopelto J, Kivinen P, Törmä P and Pekola J 2003 Application of superconductor–semiconductor Schottky barrier for electron cooling Physica B 329–333 1481–4

[83] Savin A, Pekola J, Prunnila M, Ahopelto J and Kivinen P 2004 Electronic cooling and hot electron effects in heavily doped silicononinsulator film Phys. Scr. T114 57

[84] Buonomo B, Leoni R, Castellano M G, Mattioli F, Torrioli G, Di Gaspare L and Evangelisti F 2003 Electron thermometry and refrigeration with doped silicon and superconducting electrodes J. Appl. Phys. 94 7784

[85] Kivinen P, Savin A, Zgirski M, Törmä P, Pekola J, Prunnila M and Ahopelto J 2003 Electron phonon heat transport and electronic thermal conductivity in heavily doped silicon-on-insulator film J. Appl. Phys. 94 3201–5

[86] Sergeev A, Reizer M Yu and Mitin V 2005 Deformation electron–phonon coupling in disordered semiconductors and nanostructures Phys. Rev. Lett. 94 136602

[87] Prunnila M, Kivinen P, Savin A, Törmä P and Ahopelto J 2005 Intervalley-scattering-induced electron–phonon energy relaxation in many-valley semiconductors at low temperatures Phys. Rev. Lett. 95 206602

[88] Prunnila M 2007 Electron–acoustic-phonon energy-loss rate in multicomponent electron systems with symmetric and asymmetric coupling constants Phys. Rev. B 75 165322

[89] Muhonen J T et al 2011 Strain dependence of electron–phonon energy loss rate in many-valley semiconductors Appl. Phys. Lett. 98 182103

[90] Prest M J et al 2011 Strain enhanced electron cooling in a degenerately doped semiconductor arXiv:1111.0465

[91] Edwards H L, Niu Q and de Lozanne A L 1993 Quantum dot refrigerator Appl. Phys. Lett. 63 1815–17

[92] Edwards H L, Niu Q, Georgakis G A and de Lozanne A L 1995 Cryogenic cooling using tunnelling structures with sharp energy features Phys. Rev. B 52 5714–36

[93] Gasparinetti S, Deon F, Biasiol G, Sorba L, Beltram F and Giazotto F 2011 Probing the local temperature of a two-dimensional electron gas microdomain with a quantum dot: measurement of electron–phonon interaction Phys. Rev. B 83 201306

[94] Giazotto F, Taddei F, Governale M, Fazio R and Beltram F 2007 Landau cooling in metal–semiconductor nanostructures New J. Phys. 9 439

[95] Pekola J P, Manninen A J, Leivo M M, Arutyunov K, Suoknahti J K, Suppula T I and Collaudin B 2000 Microrefrigeration by quasiparticle tunnelling in NIS and SIS junctions Physica B 280 485–90

[96] Nguyen H Q, Pascal L M A, Peng Z H, Buisson O, Gilles B, Winkelmann C and Courtois H 2011 Etching suspended superconducting hybrid junctions from a multilayer arXiv:1111.3541