THE UNIVERSITY OF
WARWICK

University of Warwick institutional repository: http://go.warwick.ac.uk/wrap

**A Thesis Submitted for the Degree of PhD at the University of Warwick**

http://go.warwick.ac.uk/wrap/50221

AUTHOR: **Juha Muhonen** DEGREE: **Ph.D.**

TITLE: **Cooling and heat transport in low dimensional phonon systems, superconductors and silicon**

DATE OF DEPOSIT:…………………………………

I agree that this thesis shall be available in accordance with the regulations governing the University of Warwick theses.
I agree that the summary of this thesis may be submitted for publication.
I agree that the thesis may be photocopied (single copies for study purposes only).

Theses with no restriction on photocopying will also be made available to the British Library for microfilming. The British Library may supply copies to individuals or libraries. Subject to a statement from them that the copy is supplied for non-publishing purposes. All copies supplied by the British Library will carry the following statement:

"Attention is drawn to the fact that the copyright of this thesis rests with its author. This copy of the thesis has been supplied on the condition that anyone who consults it is understood to recognise that its copyright rests with its author and that no quotation from the thesis and no information derived from it may be published without the author's written consent."

AUTHOR'S SIGNATURE:……………………………………………………………….

_____
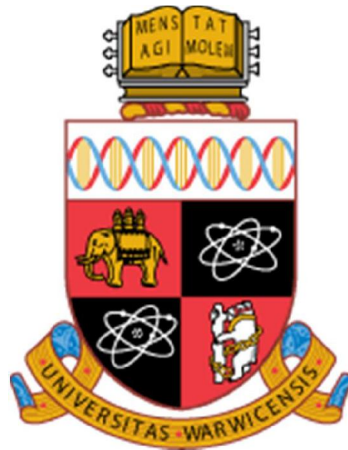
<u>USER'S DECLARATION</u>

1. I undertake not to quote or make use of any information from this thesis without making acknowledgement to the author.
2. I further undertake to allow no-one else to use this thesis while it is in my care.

DATE            SIGNATURE            ADDRESS

…………………………………………………………………………………………………

…………………………………………………………………………………………………

…………………………………………………………………………………………………

…………………………………………………………………………………………………

# Cooling and heat transport in low dimensional phonon systems, superconductors and silicon

*by*

*Juha Muhonen*

*Thesis*

Submitted to the University of Warwick

in partial fulfilment of the requirements

for admission to the degree of

Doctor of Philosophy

Department of Physics

THE UNIVERSITY OF

WARWICK

# Contents

# Acknowledgements

I started working in the PICO group of Low Temperature Laboratory (nowadays O.V. Lounasmaa Laboratory) at the Helsinki University of Technology (nowadays Aalto University) at the fall of 2007 as a Master's Thesis student and continued as a doctoral student from fall 2008 onwards. Although organisations' names have changed, and the people have changed, the PICO group has always stayed as a group full of enjoyable people who are also very hard working scientists.

I would like to thank the leader of the group and my instructor, Prof. Jukka Pekola, for letting me work in this great group and for instructing me throughout the years. I have been in the fortunate position where, in addition to my instructor, I have had the guidance of many senior and knowledgeable people, above all Dr. Matthias Meschke who has taught me most of what I know about low temperature experimental techniques. From our collaborators at VTT Technical research center of Finland, the guidance of first Dr. Antti Niskanen and in later years Dr. Mika Prunnila has been invaluable.

All the past and the present PICO group members Dr. Thomas Aref, Timothe Faivre, Anna Feshchenko, Simone Gasparinetti, Dr. Meri Helle, Dr. Tommy Holmqvist, Dr. Sergey Kafanov, Dr. Antti Kemppinen, Jonne Koski, Sarah MacLeod, Ville Maisi, Dr. Mikko Möttönen, Dr. Hung Nguyen, Dr. Joonas Peltonen, Olli-Pentti Saira, Dr. Andrey Timofeev and Dr. Youngsoo Yoon have also provided help and support in addition to the pleasant working environment. From Aalto and VTT I have also had the good fortune to collaborate with Prof. Sami Franssila, Dr. David Gunnarsson, Dr. Tero Heikkilä, Dr. Mika Sillanpää, Antti Peltonen and Dr. Lauri Sainiemi.

This work is very much an international collaboration. My research was partly funded by the Engineering and Physical Sciences Research

# List of Publications

This thesis consists of an overview and of the following publications which are referred to in the text by their Roman numerals.

**I** J. T. Muhonen, A. O. Niskanen, M. Meschke, Yu. A. Pashkin, J. S. Tsai, L. Sainiemi, S. Franssila, and J. P. Pekola. Electronic cooling of a submicron-sized metallic beam. *Applied Physics Letters* **94**, 073101, 2009.

**II** Mika A. Sillanpää, Jayanta Sarkar, Jaakko Sulkko, Juha Muhonen, and Pertti J. Hakonen. Accessing nanomechanical resonators via a fast microwave circuit. *Applied Physics Letters* **95**, 011909, 2009.

**III** J. T. Muhonen, M. J. Prest, M. Prunnila, D. Gunnarsson, V. A. Shah, A. Dobbie, M. Myronov, R. J. H. Morris, T. E. Whall, E. H. C. Parker, and D. R. Leadley. Strain dependence of electron-phonon energy loss rate in many-valley semiconductors. *Applied Physics Letters* **98**, 182103, 2011.

**IV** N. Vercruyssen, R. Barends, T. M. Klapwijk, J. T. Muhonen, M. Meschke, and J. P. Pekola. Substrate-dependent quasiparticle recombination time in superconducting resonators. *Applied Physics Letters* **99**, 062509, 2011.

**V** J. T. Peltonen, J. T. Muhonen, M. Meschke, N. B. Kopnin, and J. P. Pekola. Magnetic-field-induced stabilization of nonequilibrium superconductivity in a normal-metal/insulator/superconductor junction. *Physical Review B* **84**, 220502(R), 2011.

**VI** M. J. Prest, J. T. Muhonen, M. Prunnila, D. Gunnarsson, V. A. Shah, J. S. Richardson-Bullock, A. Dobbie, M. Myronov, R. J. H. Morris, T. E. Whall, E. H. C. Parker, and D. R. Leadley. Strain enhanced electron cooling in a degenerately doped semiconductor. *Applied Physics Letters* **99**, 251908, 2011.

**VII** Juha T Muhonen, Matthias Meschke and Jukka P Pekola. Micrometre-scale refrigerators. *Reports on Progress in Physics* **75**, 046501, 2012.

# Author's Contribution

**Publication I: "Electronic cooling of a submicron-sized metallic beam"**

The author had the main responsibility for sample fabrication, measurements, data analysis and writing of the manuscript.

**Publication II: "Accessing nanomechanical resonators via a fast microwave circuit"**

The author assisted in sample fabrication.

**Publication III: "Strain dependence of electron-phonon energy loss rate in many-valley semiconductors"**

The author had the main responsibility for measurements, data analysis and writing of the manuscript.

**Publication IV: "Substrate-dependent quasiparticle recombination time in superconducting resonators"**

The author assisted in sample fabrication, and took part in data analysis and writing of the manuscript.

**Publication V: "Magnetic-field-induced stabilization of nonequilibrium superconductivity in a normal-metal/insulator/superconductor junction"**

The author did part of the measurements and data analysis.

**Publication VI: "Strain enhanced electron cooling in a degenerately doped semiconductor"**

The author had the main responsibility of measurements and took part in designing the sample, data analysis and manuscript writing.

**Publication VII: "Micrometre-scale refrigerators"**

The author had the main responsibility of writing.

# Abstract

Temperatures below 0.1 kelvin can be nowadays routinely attained. The methods for achieving these temperatures rely on either mixing the rare and expensive isotope of helium with the more common isotope (dilution refrigerator) or on adiabatic demagnetisation of paramagnetic salt (ADR). Although both of these methods are mature, they still remain complicated enough to limit the usage only to specialized laboratories. The research done in this thesis revolves around a promising alternative to these techniques; using normal metal - insulator - superconductor (NIS) junctions.

One of the defining properties of a superconductor is a gap in its electronic density of states. This gap enables it to act as an energy filter for electrons. Because of this property, when a proper bias voltage is applied over a NIS junction the normal metal part will cool down as current passes the junction. The cooling properties of NIS junctions were demonstrated almost two decades ago with cooling powers of the order of one picowatt. At present cooling powers of few hundreds of picowatts have been achieved.

This thesis describes research on three areas related to NIS junctions. Firstly we use NIS junctions to cool low dimensional lattice systems, both 1D and 2D. The cooling of a 1D lattice (beam) is interesting for fundamental research. The 2D lattice cooling (membrane) is aimed at bringing NIS devices closer to more widespread use. An electronically cooled membrane would offer a platform on which applications, such as radiation detectors or superconducting electronics, could be integrated.

Secondly we focus on the limitations of NIS cooling. In all cooling, one of the main problems is the dissipation of the extracted heat. As the other side of the junction (normal metal) is cooled, the other side (superconductor) is heated with many times larger power. This heat can then weaken the superconducting properties and heat up the phonon system around the junction. These effects act to counter the cooling effect and have been one of the main obstacles in scaling up the cooling power of NIS devices. We study these effects both numerically and experimentally.

Thirdly, we study the cooling of silicon with superconducting tunnel junctions. In these superconductor - semiconductor structures the normal metal in a NIS structure is replaced with highly doped silicon. Specifically we study the effects of induced lattice strain to the electron-phonon coupling in silicon and hence to the cooling properties of these structures.

# 1. Introduction

The topic of this thesis lies at the crossroads of two broad research areas in condensed matter physics. The first one is low temperature physics studying matter at temperatures close to absolute zero. At those temperatures, the thermal noise always present at higher temperatures becomes negligible and the true ground states of systems can be studied. The other is physics at the mesoscopic level, popularly referred to as nanophysics. Mesoscopic is a concept used for the middleground between macroscopic (continuum) physics and microscopic (atomic) physics. It is loosely defined as the length scale where structures are large enough so that many concepts from continuum physics can be used to describe them but small enough so that new effects not present at bulk level appear. In many electric structures this happens when the dimensions of the structures start to be of the order of $\sim 100$ nm, this is what will be referred as the nanoscale.

Electric phenomena in nanoscale structures have been studied extensively in the last few decades (for evidence, one needs not to look further than to statistics of the production of sub 100 nm transistors or scientific publications mentioning nanoelectronics). The interest has been, of course, for great part driven by the fact that the ever continuing miniaturization of electronics and need for more computing power is reaching the point where, in even commercial devices, the designers have to take into account the peculiarities of the nanoscale emerging from quantum phenomena. Hence, today the research area of nanoelectronics is very mature and established.

The topic of this thesis, low temperature thermal phenomena at the nanoscale [1] have been much less studied. This is partly because much less commercial interest lies at very low temperature thermal phenomena but partly also because thermal measurements are not generally as

straightforward to perform as electric measurements. Whereas electronic signals are easy to amplify and measure with commercial devices, thermal signals always need a local thermometer, i.e., a device that transforms the thermal signal into something measurable, in practice usually an electric signal. For some rough comparison for the developments in the two fields, one could note that quantized electric conductance was demonstrated already at 1980s [2] whereas single quantum thermal conductance was demonstrated only in 2000 [3].

Even the definition of temperature at the nanoscale can spark debates. In classical thermodynamics temperature is defined at the thermodynamic limit where the volume of the system approaches infinity (keeping density of particles constant) [4]. Obviously this limit is very far away from the nanoscale. In this thesis, we take a very experimentalist view on temperature and always define it through the effective temperature entering the occupation factors (Fermi-Dirac or Bose-Einstein) of a system. It will become clear that this is an obvious choice for us as this is the quantity that we can measure.

This definition also means that even a simple piece of a metal such as copper does not have single temperature at low temperatures. The piece of copper is actually composed of different subsystems which each then have their own effective temperature. The subsystems considered in this thesis are the lattice (phonons) and the electrons. (In addition e.g. the nuclear system can have its own temperature but this is not considered here.) These two systems are connected through a very non-linear coupling, the electron-phonon coupling, which is an important topic in this work.

Another important concept for the thesis is superconductivity [5, 6]. Superconductivity is a phase of state that appears in certain materials below a material dependent transition temperature. Its hallmark is the absence of electrical resistivity, a fact that has sparked a lot of material research in hope of finding a material that would be superconducting at room temperature. In this thesis, we are not primarily interested in the electric resistivity but another feature appearing alongside the zero resistivity, namely the energy gap in the density of states. Specifically this thesis will deal with superconducting tunnel junctions and their use as coolers and thermometers. It is the energy gap that enables both usages.

The thesis is organised as follows. In the next subsections in this chapter we will give an introduction to the basic fabrication and measurement methods used in the thesis. In Chapter 2 we will discuss the basic theory

behind cooling with superconducting tunnel junctions. In Chapter 3 we discuss experiments and modelling on using these junctions to cool lattice systems where the dimensionality of phonons is effectively reduced. In Chapter 4 the properties of the superconductor important for the cooling applications (quasiparticle dynamics) are discussed in more detail and experiments and modelling related to this are presented. And in Chapter 5 we discuss how the electron-phonon coupling in silicon can be modified by inducing strain to the silicon layer.

## 1.1 Fabrication methods

The workhorse of nanofabrication is electron-beam-lithography (EBL). This lithography method is based on scanning a suitable resist with a focused electron beam. Typical spotsize of the beam is 1-2 nm. The resolution of EBL, however, is limited by the resist. The resist used throughout the work described in this thesis is called poly(methyl methacrylate) or PMMA. With dedicated EBL systems using very high voltage ($\sim 100$ kV), resolution of around 10 nm can be achieved with PMMA. For our 30 kV system the resolution limit is around 50 nm.

The workflow of a typical fabrication process is as follows: spinning the resist on a substrate, EBL, development in a suitable chemical, deposition of material through the resist mask and finally removal of resist and extra material on top of it in lift-off. This is depicted in Fig. 1.1. In this way, metallic structures with length scales below 100 nm can be routinely produced.

Making tunnel junctions complicates the process somewhat. The method used for the structures described in this thesis is the so-called shadow evaporation. This requires a two-layer resist where PMMA is on top and another resist with higher sensitivity lies below it. Most commonly used one is copolymer P(MMA-MAA). After exposure to electron beam, during development the underlayer is developed faster than the top layer, resulting in a cave like structure below the actual pattern in the top PMMA layer (known as undercut). Then depositing in different angles will produce copies of the mask at different locations (see Fig. 1.1). With proper mask design, this allows one to make structures where the two deposited metal layers intersect at specific locations. Making tunnel junctions is enabled by depositing some metal that can be oxidised first, letting some oxygen into the chamber and depositing the second metal in a different
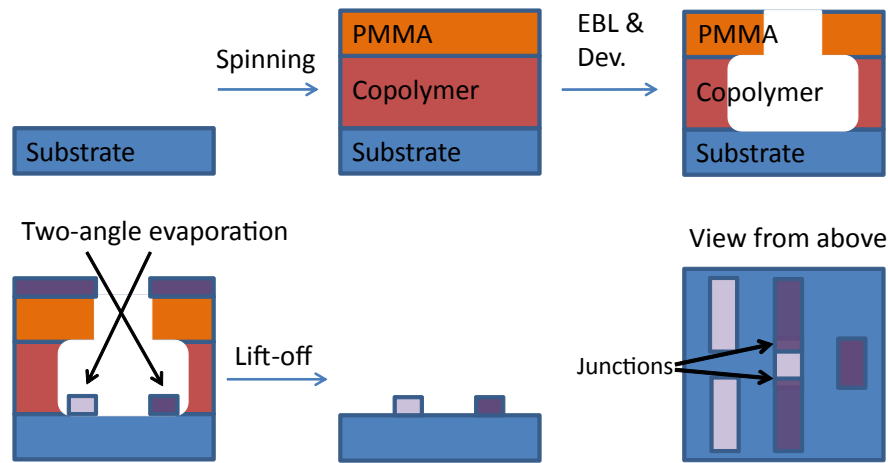
**Figure 1.1.** The fabrication process. See text for details.

angle. Hence, one can create junctions where the two metal layers are separated only by a thin oxide layer on a specified area.

The process described above obviously depends on a deposition done only on the line-of-sight. The method we used is electron beam evaporation which is a form of physical vapor deposition (PVD). In this method, an electron beam is used to heat up the source material to high enough temperatures so that it starts to evaporate. When this is done in high-vacuum, the evaporated atoms will fly in straight lines and condense into the cold surfaces that they land on. The substrate with the resist mask on top is placed on the line-of-sight of the heated source material on a tiltable plate, which enables the evaporation in different angles. A known amount of oxygen can also be introduced to the chamber to allow the oxidization.

The final step of the basic fabrication process is lift-off. At this stage the resist is dissolved and all the material deposited on top of the resist mask goes away. What is left is a substrate with the materials deposited through the mask on top to form the desired patterns. The lift-off of PMMA process is generally done in acetone heated slightly on a hot plate (to about 60 degrees Celsius). For more information about fabrication methods, see [7].

## 1.2 Experimental methods

This thesis discusses topics of low temperature physics. All the experiments are done at temperatures below 1 kelvin (-272.15 C), typically around 0.3 kelvin and below. To reach these temperatures specialized equipment is needed (one motivation for the topics of this thesis is to provide alternative methods for achieving low temperatures). The base plat-

form for low-temperature experiments is liquid helium, which provides a stable 4.2 K environment as the start point for the cooling. Liquid helium is commercially available. By pumping a helium bath, temperatures of around 1 K can be reached. To achieve lower temperatures one can use the rarer (and much more expensive) isotope of helium $^3$He (the more common variety is $^4$He). It has a boiling point of 3.2 K and by pumping it, temperatures around 0.3 K can be reached. For even lower temperatures, the chosen instrument in our laboratory is the dilution refrigerator. This kind of a refrigerator is based on the quantum properties of two different isotopes of helium mentioned above. Mixing the two isotopes is (in suitable conditions) an endothermic process and can be used to refrigerate objects. In a dilution refrigerator there is a continuous flow of $^3$He through a $^4$He phase resulting in continuous cooling power [8].

All the measurements discussed in this thesis are electronic. We mostly deal with low-frequencies and weak signals where minimising noise in the signal is of critical importance. This is taken into account at every step of the wiring and measurement. First, all the input signals going to the sample are fed through a large resistor, which will reduce the voltage noise from the voltage/current source. Secondly, all the wires to the sample are so-called twisted pairs (to reduce external noise from electromagnetic fields) and are thermalised carefully at several stages to reduce thermal noise. At the sample end of the wiring we have additionally very efficient low-pass filtering (below $\sim$ MHz) in the form of thermocoax cables. For more filtering there are surface mounted resistors on the sample stage, which form an RC filter with stray and wiring capacitance with cut-off frequency at around 10 kHz. For the amplification of output signals, specialised low-noise amplifiers powered by battery sources are used. At the backside of the amplifiers, opto-isolators are used to prevent any ground loops. A schematic of the wiring between the room temperature parts of the set-up and the sample at 50 mK is presented in Fig. 1.2.

To measurement electronics

BNC connectors

300 K (room temperature)

Twisted manganin
wire ~1 m

4 K ($^4$He bath)

1 K (Pot)

Thermocoax
~1.5 m

50 mK (Mixing Chamber)

Shielded
connector

Metallic sample
stage

Printed circuit
board

Surface mounted
resistors

Al bonding wire

Sample

Chip
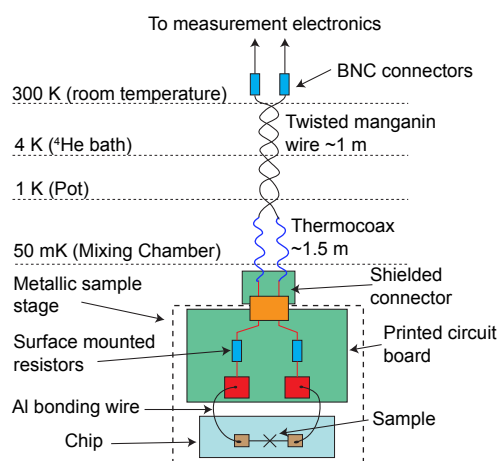
**Figure 1.2.** Schematic of the wiring in the cryostats.

# 2. Cooling with tunnel junctions

## 2.1 Basic principles

Electronic cooling with tunnel junctions is based on energy selective tunnelling. If one can filter electric current so that only low energy electrons are allowed to enter the structure and only high energy electrons are allowed to exit, this will lead to lowering of the average energy i.e., cooling. In order to realise this kind of "cooling current" an energy filter is required. In the low temperature experimental realisations of tunnel junction coolers this filter has been either the superconducting energy gap (first demonstration by [9] (NIS) and [10] (SINIS)) or a quantum dot [11]. In this thesis, we will focus on the coolers based on superconductors.

The appearance of the energy gap in a superconductor's density of electronic states is explained by the BCS theory of superconductivity [6]. The BCS theory states that when temperature becomes low enough, attractive interactions between the electrons due to their coupling to the lattice phonons can overcome their repulsive Coulomb interaction and lead to the formation of so-called Cooper pairs. These pairs are coherent quantum states of coupled electrons and can therefore flow without dissipation, i.e., with zero resistivity. Because a certain energy is needed to break these pairs, an energy gap (i.e. an energy region where no allowed states exist) appears into the density of electron (or more precisely quasiparticle) states of a metal when it enters the superconducting state at a critical temperature $T_c$. The width of this gap is defined by a material dependent constant called $\Delta$, which is the energy per electron required to break one Cooper pair. On the borders of this energy gap the density of states (DOS) diverges and far away from it, it reaches again the normal state value.

Quantitatively this BCS DOS is expressed as

$$N_S(E) = N_0 \frac{|E|}{\sqrt{E^2 - \Delta^2}} \Theta(|E| - |\Delta|), \qquad (2.1)$$

where $N_0$ is the DOS in the normal state, $\Delta$ is the energy gap of the superconductor and $\Theta(x)$ is the Heaviside step function. This is plotted in Fig. 2.1 **(a)**. The gap parameter $\Delta$ is temperature dependent, reaching its maximum value at zero temperature and vanishing at the critical temperature.

The basic principle of cooling with a NIS (normal metal - insulator - superconductor) tunnel junction can then be readily understood from an energy diagram presented in Fig. 2.1 **(b)**. In this so-called semiconductor picture, the occupation probability of electronic states is depicted as a function of energy. The normal metal and the superconductor are assumed to be separated from each other by a thin tunnel barrier and the distribution of quasiparticles on both sides follows the Fermi-Dirac distribution function

$$f(E, T) = [\exp(\frac{E - E_f}{k_B T}) + 1]^{-1}, \qquad (2.2)$$

where $E_f$ is the Fermi energy and $T$ the temperature. The DOS is assumed to be constant in the normal metal (we neglect the weak energy dependence because all relevant energies are close to the Fermi level) and follow BCS DOS in the superconductor. At low temperatures, the Fermi-Dirac function gets narrower and at temperatures $k_B T \ll \Delta$ in the superconductor all the states below the gap are filled and above it they are empty.

When only a small bias voltage is applied over the junction, no current can flow as the empty states above the gap are too high in energy for the normal metal electrons to reach. If, however, a sufficient voltage bias is applied over the junction so that the Fermi level of the normal metal is shifted upwards by energy $eV \approx \Delta$, current can flow. If this bias voltage is selected so that it is just below the gap, then electrons with high enough energy are preferentially removed from the normal metal part. This is the phenomenon of energy selective tunnelling, which enables cooling of the electron gas.

## 2.2   Current and heat flow in a NIS junction

The current and heat flow through a NIS junction can be analyzed with a perturbation theory calculation for a low transparency junction. Gener-
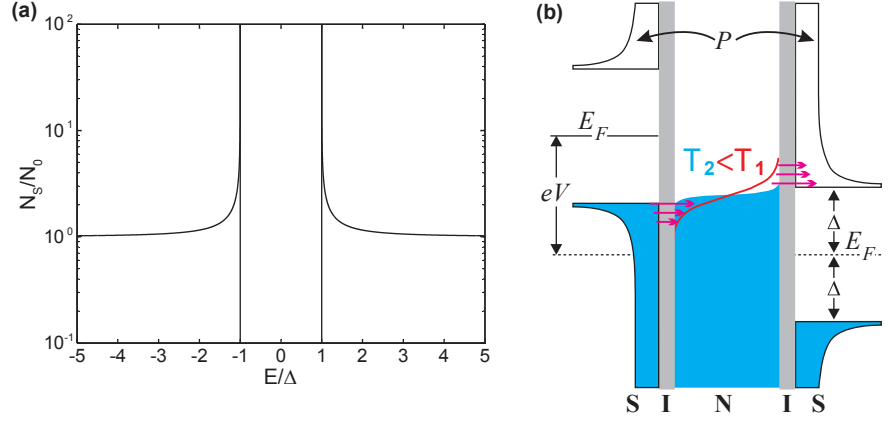
**Figure 2.1. (a)** The BCS density of states (Eq. (2.1)) normalised to the normal state density of states as a function of energy. The energy axis is normalised to the $\Delta$ parameter. **(b)** Energy diagram of a SINIS structure. The superconductors have a density of states as in (a) whereas in the normal metal it is assumed constant. Filled areas show the Fermi-Dirac distribution function (Eq. (2.2)). With proper applied bias voltage (as in figure) only electrons above (below) the Fermi energy can tunnel out from (into) the normal metal, leading to narrowing of the Fermi-Dirac distribution, i.e., cooling. The cooling power of this SINIS structure is double the cooling power of one NIS junction.

ally the first order calculation (Fermi's Golden Rule) gives results which are consistent with the experiments but when going to very high transparencies of the junction, the higher order processes can start to play a role. This issue is discussed in the next section. Calculating the exact matrix element for the tunnelling current would require detailed knowledge of the tunnel barrier but fortunately it can be related to an easily measurable quantity, the normal state resistance of the junction.

Let us assume we have a NIS junction with a bias voltage applied across the junction so that the Fermi levels of the quasiparticle systems on the two sides of the junction are shifted by $eV$. Assume also that the average transmission probability for an electron to tunnel through the tunnel barrier (i.e., the square of the tunnelling-matrix element) is $\mathcal{M}^2$. Fermi's golden rule tells us that the probability per unit time for tunnelling to happen at a certain energy is proportional to the densities of states on both sides and must obey the Pauli exclusion principle. In other words, the state from which the tunnelling happens must be occupied and the state to which the electron tunnels must be empty. The total current through the tunnel barrier (current from left to right minus that from right to left) from $N$ to $S$ can then be calculated by integrating over energy

$$
\begin{aligned}
I_{\text{NIS}} = {} & eA\mathcal{M}^2 \int_{-\infty}^{\infty} \mathrm{d}E\, N_0(E+eV)N_s(E) \\
& \times\, -\{f(E+eV,T_n)[1-f(E,T_s)] - f(E,T_s)[1-f(E+eV,T_n)]\}
\end{aligned}
$$
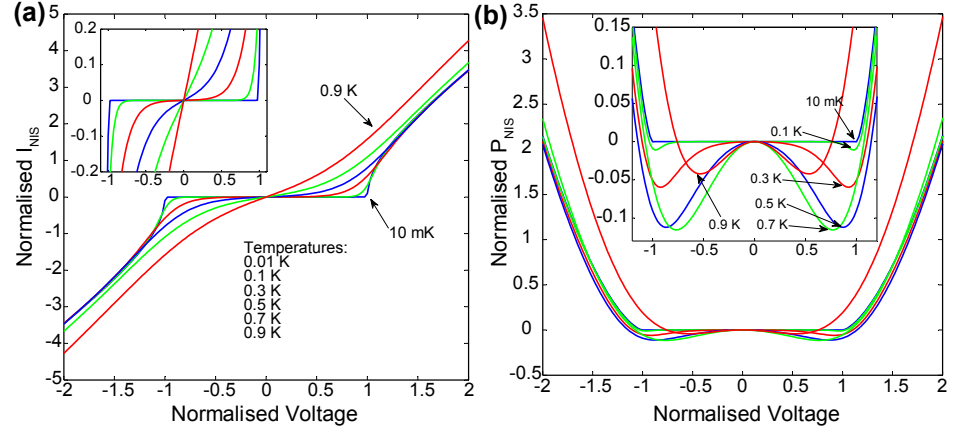
**Figure 2.2. (a)** Normalised current ($(eR_T/\Delta)I_{\text{NIS}}$) of a NIS junction, plotted as function of the normalised voltage ($eV/\Delta$) calculated at different bath temperatures. **(b)** Normalised cooling power ($(e^2R_T/\Delta^2)P_{\text{NIS}}$) of a NIS junction calculated at different bath temperatures. In both plots it is assumed that $T_n = T_s$ and that $\Delta(0) = 200\ \mu\text{eV}$ which corresponds to aluminium as a superconductor. Insets show a close-up to the small bias regime.

$$= \frac{1}{eR_T N_0}\int_{-\infty}^{\infty} \mathrm{d}E\, N_s(E)[f(E,T_s) - f(E+eV,T_n)], \qquad (2.3)$$

where $e$ is the electron charge, $A$ is the area of the junction, and $T_n$ and $T_s$ are the temperatures of the normal metal and the superconductor, respectively. $N_0$ and $N_s$ are the densities of states in the normal state (assumed constant) and in the superconductor Eq. (2.1), respectively. The last form comes from the fact that the normal state resistance ($R_T$) of the junction is $1/(e^2 A \mathcal{M}^2 N_0^2)$.

The heat flow (i.e., energy current) is calculated similarly, except now instead of the electron charge $e$ we have to weigh the integral with the energy the electron carries with it, $E + eV$,

$$P_{\text{NIS}} = \frac{1}{e^2 R_T N_0}\int_{-\infty}^{\infty} \mathrm{d}E\,(E+eV)N_s(E)[f(E,T_s) - f(E+eV,T_n)]. \quad (2.4)$$

This integral is negative near the gap (signifying heat flow out of the normal metal) and becomes positive with large bias voltages where the junction starts to behave as an ordinary resistor with the associated Joule heating. Note that the equation is symmetric between positive and negative voltages and hence putting two NIS junctions in series (i.e. making a SINIS structure) doubles the cooling power of the normal metal. The current and heat flow through a NIS junction calculated with Eqs. (2.3) and (2.4) are plotted in Fig. 2.2 **(a)** and **(b)**, respectively.

An important detail to note is that (although it is not obvious from the forms presented above) the current through the NIS junction depends only on the temperature of the normal metal (ignoring the temperature dependence of $\Delta$) whereas the heat flow is also affected by the tempera-
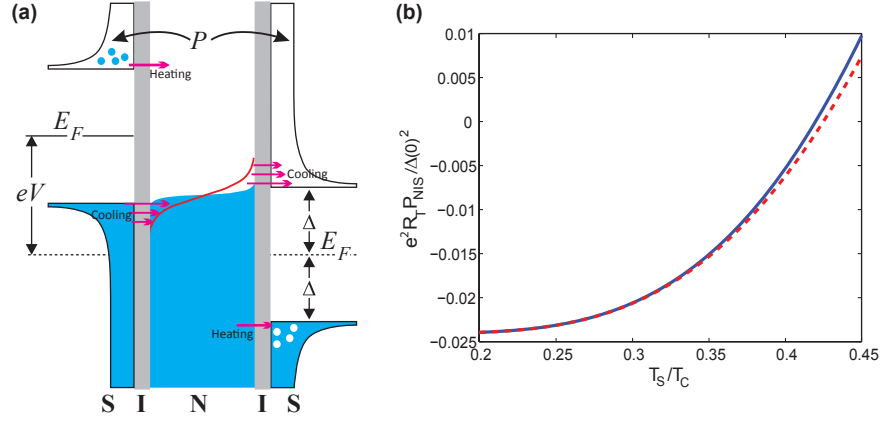
**Figure 2.3. (a)** A schematic diagram of how the heating of the superconductor (widening of $f_s$) affects the heat flow. Quasiparticles with energies above (below) the gap can now tunnel into (out from) the normal metal. **(b)** Calculated $P_{\mathrm{NIS}}$ at optimum bias voltage as a function of $T_s$. $T_n$ is assumed to be constant at 0.2 $T_c$. Dashed line calculated assuming constant $\Delta$, solid line with temperature dependent $\Delta$ from BCS theory.

ture of the superconductor. This is made more clear by writing Eqs. (2.3) and (2.4) in a symmetrised form

$$
\begin{aligned}
I_{\mathrm{NIS}} &= \frac{1}{2eR_T N_0} \int_{-\infty}^{\infty} \mathrm{d}E\, N_s(E)[f_n(E - eV) - f_n(E + eV)] && (2.5) \\
P_{\mathrm{NIS}} &= \frac{1}{2e^2 R_T N_0} \int_{-\infty}^{\infty} \mathrm{d}E\, N_s(E)\{eV[f_n(E - eV) - f_n(E + eV)] \\
&\quad + E[2f_s(E) - f_n(E + eV) - f_n(E - eV)]\}, && (2.6)
\end{aligned}
$$

where we have abbreviated $f_i(E) = f(E, T_i)$, $i = n, s$ for normal metal and superconductor, respectively. Comparing the two equations it is clear that in Eq. (2.6) the first term is just $I_{\mathrm{NIS}}V$, i.e., the Joule heating, and the second term (which depends also on $f_s$) is responsible for the cooling effect. Hence, although the electric current through the junction is unaffected by $f_s$, the cooling effect can be suppressed by the heating of the superconductor. This is because in addition to the filtered "cooling current" there will be also "heating current" by electron-like (hole-like) quasiparticles from the states that were empty (full) in the diagrammatic Fig. 2.1. A schematic of this and calculation of $P_{\mathrm{NIS}}$ as a function of $T_s$ are presented in Fig. 2.3.

From Eqs. (2.5) and (2.6) we can also see that

$$
\begin{aligned}
I_{\mathrm{NIS}}(V = 0) &= 0 && (2.7) \\
P_{\mathrm{NIS}}(V = 0) &= \frac{1}{e^2 R_T N_0} \int_{-\infty}^{\infty} \mathrm{d}E\, N_s(E) E[f_s(E) - f_n(E)]. && (2.8)
\end{aligned}
$$

This means that although net electronic current is always zero at zero voltage, there will be a net heat flow between $S$ and $N$ if there is a temperature difference between them even without any applied voltage. This

is quite natural as the two systems will try to equilibrate to the same temperature at a rate determined by the transparency of the barrier between them ($R_T^{-1}$). This will be important for the discussion of quasiparticle traps in Chapter 4.

It is also possible to obtain analytical estimates of the current and cooling power of a NIS junction for temperatures well below the critical temperature of the superconductor $k_B T \ll k_B T_C \simeq \Delta/1.76$ [12]. The cooling power maximizes at bias voltages $V = (\Delta - 0.66 k_B T)/e$ where it reaches

$$P_{\text{opt}} \simeq \frac{\Delta^2}{e^2 R_T}\Big[0.59\big(\frac{k_B T_N}{\Delta}\big)^{3/2} - \sqrt{\frac{2\pi k_B T_S}{\Delta}} e^{-\Delta/k_B T_S}\Big]. \tag{2.9}$$

At the optimal bias point, the current through the cooler junction is

$$I(V_{\text{opt}}) \simeq 0.48 \frac{\Delta}{e R_T}\sqrt{\frac{k_B T_N}{\Delta}}. \tag{2.10}$$

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

$$\eta = \frac{P_{\text{opt}}}{I(V_{\text{opt}})V} \simeq 0.7\frac{T}{T_C}, \tag{2.11}$$

where the last approximation applies again at $T \ll T_C$.

## 2.3  The sub-gap current

In practice, the behaviour of NIS junctions deviates somewhat from the ideal behaviour described above. Especially, at small bias voltages a leakage current exceeding that predicted by Eq. (2.3) is usually seen in experiments. This effect has been successfully modelled by modifying the ideal BCS DOS Eq. (2.1) by adding a so-called Dynes parameter $\gamma$ [13]

$$N_s(E) = N_0 \left| Re\big(\frac{E + i\gamma\Delta}{\sqrt{(E + i\gamma\Delta)^2 - \Delta^2}}\big)\right|. \tag{2.12}$$

In this form $\gamma$ is a phenomenological parameter describing finite amount of available states in the BCS gap. Experimentally $\gamma$ is the ratio between the resistivity of a NIS junction at $V = 0$ and the asymptotic resistance at high bias voltages. In effect the Dynes model amounts to viewing the subgap current of the NIS junction at very small voltages as if it were that of a fully normal (NIN) junction with tunnel resistance $R_T/\gamma$. Usually in good tunnel junctions gamma is around $10^{-4} - 10^{-5}$. From the outset the $\gamma$ parameter would seem to describe non-idealities of the superconducting

material which lead to a finite amount of states inside the BCS energy gap. This, however, might not be the case as other phenomena can create a sub-gap leakage current similar in form as one gets from Eq. (2.12) as has been highlighted recently [14].

Another cause for excess sub-gap current is the possibility of higher order processes, not taken into account when deriving Eq. (2.3). Although the influence of multi-electron-current compared to the first order current is usually small, a second order process called Andreev current [15] can dominate over the single-particle tunnellings at voltages $V \ll \Delta/e$. The Andreev current is essentially a process where a Cooper pair in the superconductor is transported into two quasiparticles in the normal metal or vice versa. This process is especially important in the cooling applications as the Andreev current, unlike the regular single-electron current, leads to dissipation in the N electrode at all bias voltages [16]. This can be understood from a simple energy diagram picture: a Cooper pair on the Fermi level in the superconductor creates two excitations in the normal conductor, whose average energy is $eV$, where $V$ is the bias voltage across the junction. Therefore all the power $P_{\mathrm{AR}}$ dissipates in N and it equals simply $P_{\mathrm{AR}} = I_{\mathrm{AR}}V$, where $I_{\mathrm{AR}}$ 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 $I_{\mathrm{AR}} = (16\mathcal{N}R_T^2)^{-1}R_K V$, where $\mathcal{N}$ is the number of conduction channels, and $R_K = h/e^2$ is the quantum resistance [17]. This ballistic description gives typically very small values for Andreev current with the transparencies common in NIS junctions. However, for larger diffusive junctions, typical for a NIS cooler, 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. Also the diffusive case can be analysed theoretically [18, 19, 20] and recent experimental results seem to agree with the predictions [21, 22].

According to [19] it can be written in a simple form at the limits where the junction dimensions are much bigger or much smaller than the coher-

ence length in the normal metal $L_E = \sqrt{\hbar D_n / \max(eV, k_B T)}$. At the small junction limit the Andreev current is

$$I_{\mathrm{AR}} = \frac{V}{\pi R_T^2} \left[ R_n \ln\left(\frac{\hbar}{eV\tau}\right) + R_s \ln\left(\frac{\hbar}{\Delta\tau}\right) \right]. \qquad (2.13)$$

Here $R_n$ ($R_s$) is the square resistance of the normal metal (superconductor) defined as $R_i = (e^2 N_i D_i d_i)^{-1}$, where $N_i$ is the density of states, $D_i$ is the diffusion constant and $d_i$ is the thickness of the metal film. The time constant $\tau$ is roughly the time the quasiparticle spends at the junction area $A/D$, where $A$ is the junction area. This formula was used in [21] to explain experimental current-voltage characteristics in the sub-gap regime. The other limit of a big junction, which is typically more relevant to the NIS cooler case, yields [19]

$$I_{\mathrm{AR}} = \frac{\hbar}{e^3 A R_T^2} \left[ \frac{1}{N_n d_n} \tanh\left(\frac{eV}{2k_B T}\right) + \frac{1}{N_s d_s} \frac{eV}{2\pi\Delta\sqrt{1 - eV/\Delta}} \right]. \qquad (2.14)$$

This expression is applicable only at voltages $V < \Delta/e$ as the latter term diverges when voltage equals the gap. This is the expression used in [22] to fit experimental data.

In our experiments, the hallmark of Andreev current is the appearance of a step at zero voltage in the I-V curve. This step can be understood from equation Eq. (2.14) where at zero temperature the normal metal part (first term) does not depend on the voltage and hence creates a jump from negative to positive Andreev current at zero voltage. This jump is then smoothened with temperature by the hyperbolic tangent term. In Fig. 2.4 we plot both the Andreev current and the sub-gap single-particle current (originating from $\gamma$). It can be seen that for a low leakage junction Andreev current dominates at low bias voltages but with voltages greater than $0.5\Delta/e$ its contribution becomes negligible.

## 2.4 The heat balance: diffusion and electron-phonon coupling in normal metals

A simple diagram of the relevant heat channels for NIS coolers is presented in Fig. 2.5. The external power $P_{\mathrm{NIS}}$ is exchanged with the electron system. Electron-electron ($\gamma_{e-e}$) interaction drives the electron subsystem towards an quasi-equilibrium (Fermi-Dirac) distribution, electron-phonon interaction ($\gamma_{e-p}$) couples it to the phonon bath at temperature $T_p$ and it is coupled via photons to the electromagnetic environment at $T_{env}$. The coupling to the electromagnetic environment might be responsible for the
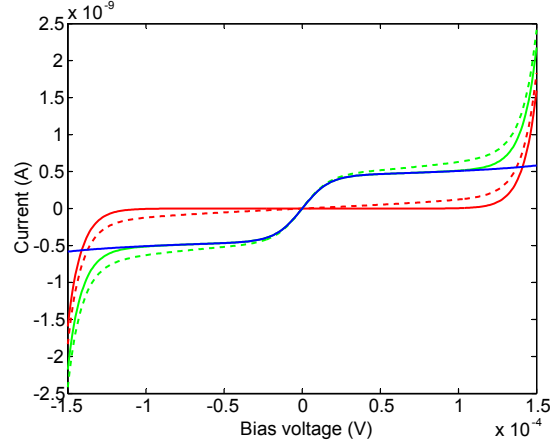
**Figure 2.4.** Andreev current (with step in the middle) and the single particle current ($I_{\mathrm{NIS}}$) with $\gamma = 0$ (solid line) and $\gamma = 10^{-4}$ (dashed line) as well as their sum at the low bias regime. Parameters $R = 100\ \Omega$, $\Delta = 200\ \mu$eV, $A = 1\ \mu$m$^2$, $d_n = 40$ nm and $d_s = 30$ nm.
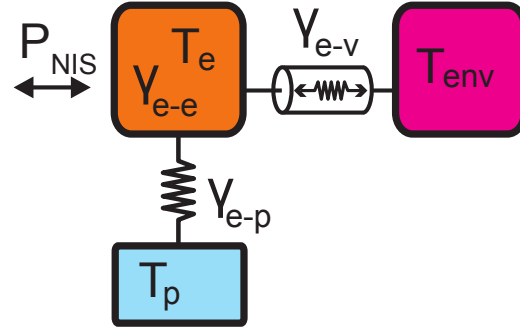


**Figure 2.5.** Simplified diagram of the relevant relaxation mechanisms.

excess sub-gap leakage current (see [14]) but otherwise its effects to cooling are small and we will neglect it from now on. Hence, the heat balance is determined by the equilibrium between $P_{\mathrm{NIS}}$ and electron-phonon coupling, meaning that temperature $T_n$ is found by solving

$$P_{\mathrm{NIS}}(T_n, T_s) = P_{e-p}(T_n, T_p), \tag{2.15}$$

self-consistently. Throughout we will assume $\gamma_{e-e}$ to be fast enough so that we can assume the electron distribution to follow the Fermi-Dirac distribution Eq. (2.2) with an effective temperature. This is generally a good assumption in systems that are only tunnel coupled to their environment such as SINIS structures [23, 24].

With NIS coolers it is often the case that the junction itself covers only a small fraction of the normal metal part to be cooled. This is especially true in the membrane cooling applications where large "cold fingers" are needed (see Sec. 3.2). In these cases the whole electron gas will not be at the same effective temperature but there will be a temperature profile inside the normal metal. Making the assumptions outlined above, so that

a position dependent temperature can be defined, then in normal metals, also at low temperatures, the thermal conductivity follows textbook models of the electron gas. The heat current density is related to the temperature gradient as $Q = -\kappa_n \nabla T$, where $\kappa_n$ is the thermal conductivity (in the normal state) and can be related to the electrical conductivity via Wiedemann-Franz law $\kappa_n = \mathcal{L}\sigma T$, where $\mathcal{L}$ is the Lorentz number and $\sigma$ the electrical conductivity. With these assumptions a steady-state diffusion equation can be written for a differential volume element

$$\nabla \cdot (-\kappa_n \nabla T_e) = \mathcal{P}_{e-p} + \mathcal{P}_{\text{ext}}, \tag{2.16}$$

where we have used $\mathcal{P}_{\text{ext}}$ as the power density from all possible external heating sources and $\mathcal{P}_{e-p} = P_{e-p}/\mathcal{V}$, where $\mathcal{V}$ is the volume. Solving this equation self-consistently and with boundary conditions will yield the temperature profile of the conductor.

### 2.4.1 Electron-phonon coupling

As electron-phonon coupling will be an important concept throughout this thesis, an overview of the related theory will be presented here. At low temperatures the two separate systems, electrons and phonons, are coupled through weak electron-phonon coupling. It is the main mechanism for heat flux to the electron system from the surrounding environment and hence an important conduction mechanism in electronic cooling applications. Because of the many-body nature of the electron wavefunctions in solid state systems, the coupling, although conceptually simple, is generally not possible to calculate directly and approximations are needed. The most used one, is the so-called deformation potential theorem [25, 26, 27]. Piezoelectric coupling plays no role in the materials discussed in the thesis and is neglected here.

Deformation potential coupling assumes first of all the Born - Oppenheimer approximation, i.e., that the electron system can respond instantaneously to changes in the ion configuration. The Hamiltonian of the solid can then be divided into three parts

$$\mathcal{H}(\mathbf{R}_i, \mathbf{r}_e) = \mathcal{H}_i(\mathbf{R}_i) + \mathcal{H}_e(\mathbf{R}_{i0}, \mathbf{r}_e) + \mathcal{H}_{e-i}(\delta\mathbf{R}_i, \mathbf{r}_e). \tag{2.17}$$

The position of ions $\mathbf{R}_i$ has been written as $\mathbf{R}_i = \mathbf{R}_{i0} + \delta\mathbf{R}_i$, where $\mathbf{R}_{i0}$ is the equilibrium position and $\delta\mathbf{R}_i$ the displacement. Here and in all that follows we use bold symbols to denote vector quantities, their absolute values will be denoted with the same symbol without boldfacing (i.e., $|\mathbf{q}| =$

$q$). The first part describes the lattice, the second part gives the electronic band structure (assumed to depend only on the equilibrium position of the ions) and the third part gives the coupling between the lattice and the electron system ($\mathbf{r}_e$ is the electron coordinate). The displacements are assumed to be small and we can linearise

$$\mathcal{H}_{e-i}(\delta\mathbf{R}_i, \mathbf{r}_e) \approx \left.\frac{\partial\mathcal{H}_e}{\partial\mathbf{R}_i}\right|_{\mathbf{R}_i=\mathbf{R}_{i0}} \cdot \delta\mathbf{R}_i = \left.\frac{\partial U(\mathbf{R}_i, \mathbf{r}_e)}{\partial\mathbf{R}_i}\right|_{\mathbf{R}_i=\mathbf{R}_{i0}} \cdot \delta\mathbf{R}_i, \qquad (2.18)$$

where in the last step we have assumed that the electronic Hamiltonian can be written $\mathcal{H}_e = -\frac{\hbar^2}{2m}\nabla^2 + U(\mathbf{R}_i, \mathbf{r}_e)$. The displacement of ions from their equilibrium positions causes shifts to the effective potential that the electrons feel. This shift includes both the changed position of the ions and the accompanying change of position of all the other electrons. When considering metals (large number of electrons) the latter effect is the dominant one, whereas in semiconductors and insulators the main effect comes from the shifted ion potential.

The deformation potential theorem is then a tool to evaluate $\partial U/\partial\mathbf{R}_i$. In its most general form it is applicable to both metals and semiconductors although, as mentioned above, the microscopic mechanisms behind the shifts in the potential are different. In the deformation potential theorem the lattice vibrations are described as waves of elastic strain and the shift in the effective potential is written as

$$\left.\frac{\partial U(\mathbf{R}_i, \mathbf{r}_e)}{\partial\mathbf{R}_i}\right|_{\mathbf{R}_i=\mathbf{R}_{i0}} \cdot \delta\mathbf{R}_i \equiv \Xi_\mathbf{k} \cdot \epsilon, \qquad (2.19)$$

where $\epsilon$ is the strain tensor and $\Xi_\mathbf{k}$ is a matrix of deformation potential constants and have the dimension of energy. It is important to note that in general $\Xi_\mathbf{k}$ is a function of the electron wavevector $\mathbf{k}$, although we do not write this explicitly below. $\Xi$ can also depend on the band index.

The strain tensor can have a maximum of six independent components and is defined as

$$\epsilon = \begin{pmatrix} \frac{\partial\mathbf{u}_x}{\partial x} & \frac{1}{2}\left(\frac{\partial\mathbf{u}_x}{\partial y} + \frac{\partial\mathbf{u}_y}{\partial x}\right) & \frac{1}{2}\left(\frac{\partial\mathbf{u}_x}{\partial z} + \frac{\partial\mathbf{u}_z}{\partial x}\right) \\ \frac{1}{2}\left(\frac{\partial\mathbf{u}_x}{\partial y} + \frac{\partial\mathbf{u}_y}{\partial x}\right) & \frac{\partial\mathbf{u}_y}{\partial y} & \frac{1}{2}\left(\frac{\partial\mathbf{u}_y}{\partial z} + \frac{\partial\mathbf{u}_z}{\partial y}\right) \\ \frac{1}{2}\left(\frac{\partial\mathbf{u}_x}{\partial z} + \frac{\partial\mathbf{u}_z}{\partial x}\right) & \frac{1}{2}\left(\frac{\partial\mathbf{u}_y}{\partial z} + \frac{\partial\mathbf{u}_z}{\partial y}\right) & \frac{\partial\mathbf{u}_z}{\partial z} \end{pmatrix}, \qquad (2.20)$$

where $\mathbf{u}$ is the displacement and $x, y, z$ are the spatial coordinates. Here we assume that displacement can be written in the standard second quantised form for a 3D solid

$$\mathbf{u}(\mathbf{r}) = \sqrt{\frac{\hbar}{2\rho\mathcal{V}\omega_\mathbf{q}}}\hat{e}_\mathbf{q}(a_\mathbf{q}e^{i\mathbf{q}\mathbf{r}} + a_\mathbf{q}^\dagger e^{-i\mathbf{q}\mathbf{r}}), \qquad (2.21)$$

where $a_{\mathbf{q}}$ ($a_{\mathbf{q}}^\dagger$) is the annihilation (creation) operator for phonon with wavevector $\mathbf{q}$ and $\hat{e}$ is the polarisation vector.

In metals (especially at low temperatures considered here) it is common to consider only the coupling to the longitudinal modes because of momentum conservation considerations [28]. For longitudinal modes the off-diagonal elements of $\epsilon$ are zero. Assuming also isotropic material so that the diagonal elements in deformation potential tensor are all equal we are left with

$$\mathcal{H}_{e-i} = \Xi \cdot \epsilon = \Xi \nabla \cdot \mathbf{u} = i\Xi \sqrt{\frac{\hbar}{2\rho \mathcal{V} \omega_{\mathbf{q}}}} \mathbf{q}(a_{\mathbf{q}} e^{i\mathbf{qr}} - a_{\mathbf{q}}^\dagger e^{-i\mathbf{qr}}). \tag{2.22}$$

The scattering rates are then calculated from the standard first order perturbation theory (Fermi golden rule)

$$\Gamma_{e-p} = \frac{2\pi}{\hbar} |\langle \mathbf{k}', n_{\mathbf{q}}^f | \mathcal{H}_{e-i} | \mathbf{k}, n_{\mathbf{q}}^i \rangle|^2 \delta(E_{\mathbf{k}'} - E_{\mathbf{k}} - (n_{\mathbf{q}}^i - n_{\mathbf{q}}^f)\hbar\omega_{\mathbf{q}}). \tag{2.23}$$

As mentioned before, in general $\Xi$ is a function of $\mathbf{k}$ and hence in order to calculate the total scattering rate, we should know it at every point of the Fermi sphere. We have neglected the effects of screening so far. It can be shown that at the limit of good metal ,i.e. very efficient screening, ($q \ll \lambda$, where $\lambda$ is the screening vector, see for example [27]) $\Xi$ is simply a constant and has the value $\Xi = \frac{2}{3} E_F$. At his limit all the effects of screening can be included in $\Xi$, unlike in most calculations in Section 5.1. This limit is sometimes called the *scalar deformation potential* approximation.

We can now put Eq. (2.22) into Eq. (2.23) and use standard Bose relations for the phonon operators, which leaves us with two answers

$$\Gamma_{e-p}^+ = \frac{2\pi}{\hbar} |\mathcal{M}_q|^2 (n_q + 1) \left|\langle \mathbf{k}'|e^{-i\mathbf{qr}}|\mathbf{k}\rangle\right|^2 \delta(E_{\mathbf{k}'} - E_{\mathbf{k}} + \hbar\omega_q) \tag{2.24}$$

$$\Gamma_{e-p}^- = \frac{2\pi}{\hbar} |\mathcal{M}_q|^2 n_q \left|\langle \mathbf{k}'|e^{i\mathbf{qr}}|\mathbf{k}\rangle\right|^2 \delta(E_{\mathbf{k}'} - E_{\mathbf{k}} - \hbar\omega_q) \tag{2.25}$$

$$|\mathcal{M}_q| = \Xi \sqrt{\frac{\hbar}{2\rho \mathcal{V} \omega_q}} q = \frac{2}{3} E_F \sqrt{\frac{\hbar q}{2\rho \mathcal{V} c_l}}, \tag{2.26}$$

where we have also assumed linear dispersion relation $\omega_q = c_l q$ ($c_l$ is the longitudinal speed of sound). After inserting the second quantised wavefunctions for the electron states and integrating over $\mathbf{r}$

$$\Gamma_{e-p}^+ = \mathcal{M}_0^2 q (n_q + 1) f(E_{\mathbf{k}})(1 - f(E_{\mathbf{k}-\mathbf{q}}))\delta(E_{\mathbf{k}-\mathbf{q}} - E_{\mathbf{k}} + \hbar\omega_q) \tag{2.27}$$

$$\Gamma_{e-p}^- = \mathcal{M}_0^2 q\, n_q\, f(E_{\mathbf{k}})(1 - f(E_{\mathbf{k}+\mathbf{q}}))\delta(E_{\mathbf{k}+\mathbf{q}} - E_{\mathbf{k}} - \hbar\omega_q) \tag{2.28}$$

$$\mathcal{M}_0^2 = \left(\frac{2}{3} E_F\right)^2 \frac{\pi}{\rho \mathcal{V} c_l} \tag{2.29}$$

where $f(E)$ is the occupation probability of electron with energy $E$. The

total energy flow between phonon and electron systems can then be calculated with

$$P_{e-p} = \sum_{\mathbf{k},\mathbf{q}} \hbar\omega_q (\Gamma^+_{e-p} - \Gamma^-_{e-p}), \tag{2.30}$$

which can following [29] be divided into two parts

$$
\begin{aligned}
P_{e-p} &= P_0(T_e) - P_1(T_p) \tag{2.31}\\
P_0 &= \sum_{\mathbf{k},\mathbf{q}} \hbar\omega_q \mathcal{M}_0^2 q f(E_{\mathbf{k}})(1 - f(E_{\mathbf{k}-\mathbf{q}}))\,\delta(E_{\mathbf{k}-\mathbf{q}} - E_{\mathbf{k}} + \hbar\omega_q)\\
P_1 &= \sum_{\mathbf{k},\mathbf{q}} \hbar\omega_q \mathcal{M}_0^2 q f(E_{\mathbf{k}}) n_q [(1 - f(E_{\mathbf{k}-\mathbf{q}}))\,\delta(E_{\mathbf{k}-\mathbf{q}} - E_{\mathbf{k}} + \hbar\omega_q)\\
&\quad -(1 - f(E_{\mathbf{k}+\mathbf{q}}))\,\delta(E_{\mathbf{k}+\mathbf{q}} - E_{\mathbf{k}} - \hbar\omega_q)],
\end{aligned}
$$

where $P_0$ depends only on electron temperature and $P_1$ only on phonon temperature, although the latter is not obvious from the form.

In Eq. (2.31) the sums can then be replaced with integrals using the densities of states as was done in [29]. However, we will instead re-write it with the help of the electronic response function $\chi(\mathbf{q}, \omega)$ (also known as polarization function in some contexts). This will be convenient for the Section 5.1 where we want to introduce the effects of screening and disorder into the calculation as $\chi(\mathbf{q}, \omega)$. It was shown in [30] (and will be discussed in more detail in Sec. 5.1) that using $\chi$ the $e-p$ energy flow can be written as

$$P_{e-p} = \sum_{\mathbf{q}} \frac{\hbar\omega_q}{2\pi} - 2\mathcal{V}\mathrm{Im}\{\chi(\mathbf{q}, \omega)\}\mathcal{M}_0^2 q [n_q^{Te} - n_q^{Tp}], \tag{2.32}$$

where we have marked $n_q^T$ to represent Bose-Einstein distribution function with energy $\hbar c_l q$ and temperature $T$

$$n_q^T = [\exp(\frac{\hbar c_l q}{k_B T}) - 1]^{-1}. \tag{2.33}$$

Using the pure low temperature limit of $\chi(\mathbf{q}, \omega)$ (see Sec. 5.1) we will get

$$
\begin{aligned}
P_{e-p} &= \sum_{\mathbf{q}} \hbar\omega_q \mathcal{V} \frac{c_l n_F}{2v_F} \mathcal{M}_0^2 q [n_q^{Te} - n_q^{Tp}]\\
&= \sum_{\mathbf{q}} \mathcal{V} \frac{\pi}{2} \frac{n_F \hbar c_l}{\rho \mathcal{V} v_F} \left(\frac{2}{3} E_F\right)^2 q^2 [n_q^{Te} - n_q^{Tp}]. \tag{2.34}
\end{aligned}
$$

Finally converting the sum to a spherical integration with density of states $(2\pi)^{-3}$ we will arrive at the result

$$
\begin{aligned}
P_{e-p} &= \Sigma\mathcal{V}(T_e^5 - T_p^5) \tag{2.35}\\
\Sigma &= \frac{1}{4\pi} \frac{N_F \hbar c_l}{\rho \mathcal{V} v_F} \left(\frac{2}{3} E_F\right)^2 \frac{\Gamma(5)\zeta(5)k_B^5}{(\hbar c_l)^5}, \tag{2.36}
\end{aligned}
$$

where $v_F, N_F$ are the velocity and density of electron states at the Fermi energy, respectively, and $\Gamma$ is the gamma function and $\zeta$ is the Riemann-Zeta function $\Gamma(5)\zeta(5) \approx 24.89$. This result is identical to the one in [29]. In practice, $\Sigma$ is a material dependent constant determined experimentally, for values in different materials see for example [1].

# 3. Cooling and heat transport in reduced dimensions

NIS junctions cool directly the electron system and can take advantage of the fact that at low temperatures the coupling between the electron system and the environment gets increasingly small allowing significant temperature reductions even with modest cooling powers. However, when considering applications in the fields of radiation detectors and/or superconducting electronics, this can also be a disadvantage. In many applications it would be desirable that the cooler would not be electronically connected to the element to be cooled as this can produce noise and and it restricts the design of the experimental circuitry.

In order to achieve electronic isolation one has to somehow thermally couple the element to be cooled to the normal metal side of the NIS junction without electrically coupling it. In practice this means coupling through a phonon system. However, in order to cool a phonon system with an electronic cooler one has to make the coupling from the phonon system to the environment smaller than the electron-phonon coupling. As mentioned before, $e - p$ coupling gets very weak at low temperatures and hence this is a very challenging condition. In practice, fulfilling this condition calls for reducing the effective dimensions of the phonon system, i.e., making it in a form of a beam or a membrane. In this Chapter we will discuss cooling these low dimensional objects with NIS coolers. Experiments as well as modelling with finite-element-method are discussed. This Chapter is related to Publications I, II and IV.

The length scale on which the effective dimensions of some systems are reduced is related to the dominant wavelength of the system. This dominant wavelength in turn is related to the quantum statistical nature of the system and is very different for bosons and fermions. For phonons (bosons) it is roughly the wavelength of the dominant thermal mode $\lambda_T = \hbar c/(k_B T)$ where $c$ is the speed of sound. For electrons (fermions) it is the

Fermi wavelength which for a free electron gas is $\lambda_F = 1/k_F = (\mathcal{V}/(3\pi^2\mathcal{N}))^{1/3}$ where $\mathcal{N}$ is the number of electrons and $\mathcal{V}$ is the volume. This quantity is independent of temperature. The order of magnitude for speed of sound for materials dealt with in this thesis is $\sim 5000$ m/s. This gives a $\lambda_T$ of $50/T$ nm whereas the Fermi wavelength in a typical metal (copper) is of the order of 0.1 nm. For the structures in this Chapter with dimensions around 100 nm, the electron system is always 3D but the phonon system dimensionality gets reduced at the lowest temperatures.

## 3.1   Cooling nanomechanical modes

One system where the coupling between the (local) phonon system and the environment can be made extremely small is nanomechanical oscillator. The $Q$-values, describing the dissipation of the local mechanical vibrations to the environment, can be as high as hundreds of thousands in single crystal nanoscale beam oscillators [31]. Although integrating any other cooled element into a nanosized beam is generally not very convenient, cooling down the beam itself can be of fundamental interest.

For some time it has been appreciated that cooling down the mechanical modes of high frequency mechanical resonators down to their quantum ground state should be possible [32, 33, 34], and this objective has been now recently achieved [35, 36, 37]. The quantum ground state of a mechanical vibration means that the vibration will then be quantized (i.e., consist of only one phonon mode). Such a system can then provide a test bed for fundamental experiments on the limits of quantum mechanics as well as enable high-precision sensors. 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 difficult as there is inevitably some dissipated power generated by the measurement of the vibrations. This difficulty can be overcome to some extent by just using very high frequency oscillators where the temperature needed to reach the ground state is high (as in [35]) or by directly cooling the mechanical modes with e.g. radiation pressure (as in [37]). Nevertheless, making the beam out of normal metal connected to NIS junctions would also circumvent this problem as the local modes would then be directly cooled through the (local) electron system.

In [38] it was suggested that NIS coolers could be used for cooling down metallic nanosized beams. In addition to the connection of studies of the quantum ground state of the mechanical oscillator, this kind of a system should exhibit a different powerlaw for the electron-phonon coupling if the transversal dimensions are made much smaller than $\lambda_T$, as then the dimensionality of the phonon system would be reduced. It is straightforward to calculate the heat-flow from 3D electron system to a 1D phonon system based on the equations presented in Sec. 2.4.1, if we assume the continuum limit where the length of the beam $L \gg \lambda_T$. As the electron system is still assumed three dimensional, we only need to replace $\mathbf{r}$ with a scalar $x$ in $\mathbf{u}$ and divergence with 1D partial derivation. The form will stay the same (aside of normalisation factor $\mathcal{V}$ replaced with $L$) and we then need to replace the spherical integration with a 1D integration in Eq. (2.34). This will lead to

$$
\begin{aligned}
P_{e-p}^{\mathrm{1D}} &= \Sigma_{1D} L(T_e^3 - T_p^3) \\
\Sigma_{1D} &= \frac{1}{4} \frac{N_F \hbar c_l}{\rho \mathcal{V} v_F} \left(\frac{2}{3} E_F\right)^2 \frac{\Gamma(3)\zeta(3)k_B^3}{(\hbar c_l)^3}.
\end{aligned}
\tag{3.1}
$$

As said in Sec. 2.4.1, $\Sigma$ is often considered an experimental parameter. Hence, it is important to note the $\Sigma_{1D}$ can be related to the bulk $\Sigma$ with

$$
\Sigma_{1D} = \pi \left(\frac{\hbar c_l}{k_B}\right)^2 \frac{\Gamma(3)\zeta(3)}{\Gamma(5)\zeta(5)} \Sigma = \frac{\pi}{12} \left(\frac{\hbar c_l}{k_B}\right)^2 \frac{\zeta(3)}{\zeta(5)} \Sigma.
\tag{3.2}
$$

In Publication I a beam cooler was demonstrated (see also [39, 40]). The fabrication process is the same as described in Section 1.1 except for an additional etching step as the last processing phase. In our work, isotropic reactive ion etching was used to etch the substrate underneath the structure. The etch was timed so that only the narrowest structure, i.e., the beam will be released. However, difficulties were encountered with failure of tunnel junctions during the etching step. In most tested structures the tunnel junctions were shorted. These failures were eventually traced back to charge collection by the metallic structures from the ion etching. As the area of the pads is much larger than the area of the beam, a significant voltage can develop over the oxide layers, assuming that each area collects the same amount of excess charge. A simple calculation shows that this voltage can be several volts. In order to prevent this voltage build up, fabrication was moved to slightly doped substrates. In this way when etching at room temperature the substrate provides a short over the tunnel junction. The doping was chosen light enough so that at the measurement temperatures below 1 K, the conductivity of the substrate
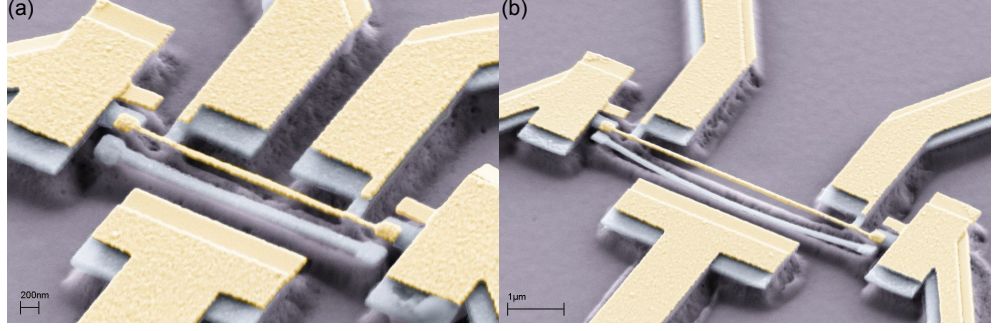
**Figure 3.1.** Suspended metallic beams connected to NIS junctions. Light material AuPd (normal metal) and darker material Al (superconductor). The beam lengths are 2 $\mu$m (**a**) and 4 $\mu$m (**b**). Width is roughly 100 nm. The bigger junctions at the ends are designed for cooling and the smaller junctions in the middle for thermometry. The gate is not used in the experiments.

was negligible (which was confirmed by the low sub-gap leakages of the junctions). This approach lead to considerable improvement of yield in fabrication. Representative beam cooler samples are presented in Fig. 3.1.

The heating and cooling properties of these structures were studied in order to see the effects of phonon dimensionality to the electron-phonon coupling. The measurement consists of heating/cooling the beam with the tunnel junctions at the end and monitoring the temperature of the beam. In effect we are then measuring the coupling from the electron system of the beam to the environment (meaning bulk phonons). There are two thermal bottlenecks on this heat conduction path: the $e - p$ coupling and the coupling between the local phonon modes and the bulk phonon modes. The heat flux from the local modes to the bulk phonons can be calculated assuming the bulk-beam interface to be a tunnel barrier with transmissivity $\omega/Q$ [38] (for more detailed account about phonon transmission through interfaces, see [41])

$$
\begin{aligned}
P_{\text{1D}-\text{3D}} &= \sum_q \hbar\omega_q \frac{\omega_q}{Q_q}(n_q^{T_q} - n_q^{T_{\text{bath}}}) \\
&= \int_{-\infty}^{\infty} \mathrm{d}q \frac{L}{2\pi} \frac{\hbar c_l^2}{Q_q} q^2 (n_q^{T_q} - n_q^{T_{\text{bath}}}) \\
&= \frac{\zeta(3)k_B^3}{Q\pi\hbar^2 c_l} L(T_p^3 - T_{\text{bath}}^3) \equiv KL(T_p^3 - T_{\text{bath}}^3),
\end{aligned}
\tag{3.3}
$$

where in the last forms we have again assumed the continuum limit. In the integration we additionally assume that all the modes are at same temperature $T_q = T_p$ and have the same $Q$-value. At steady state the energy flows in and out of a system are equal, i.e., $P_{\text{1D}-\text{3D}} = P_{e-p}^{\text{1D}}$ and we can eliminate $T_p$ (which is not probed in the experiment) in order to solve

the heatflow from the electron system to the bath

$$P^{1D}_{e-\text{bath}} = \frac{1}{1 + \frac{\Sigma_{1D}}{K}}\Sigma_{1D}L(T_e^3 - T_{\text{bath}}^3). \qquad (3.4)$$

The limits of this are quite obvious, if $\Sigma_{1D} \gg K$ then $K$ becomes the bottleneck and $P^{1D}_{e-\text{bath}} = KL(T_e^3 - T_{\text{bath}}^3)$ whereas at the other limit $\Sigma_{1D} \ll K$ we recover Eq. (3.1).

Exprimental data is shown in Fig. 3.2. Both a bulk sample (without the etching step) and a suspended sample were measured. In both cases the normal island was made out of Au:Pd (3:1 by mass) alloy. This material handled the etching step more reliably than copper. The results were not in complete agreement with the theory. From Fig. 3.2 it can be seen that in both samples the I-V curves can be fitted relatively well producing almost same parameters, except for a slight difference in the junction resistance. However, the cooling data of the bulk sample follows the calculated curve only at the low bias regime whereas the cooling data from the suspended sample seems to fit relatively well the calculated curve for bulk phonons.

These results would seem to indicate that the $e-p$ coupling was surprisingly not modified in the suspended beam. In fact, the suspended sample follows the bulk prediction more precisely than the bulk sample. Because of this, one would be tempted to speculate that in the bulk sample some local phonon heating effect (which will be a recurring theme in the dissertation) would be playing a role. This kind of a substrate related effect would then intuitively play more minor role in the suspended case. However, as the bulk data indicates that the normal metal heats up slower than expected in the high bias regime, this is very hard to reconcile with the data presented later in the thesis on other systems.

A probable explanation is uncertainty in the applied voltage (x-axis of the figure). In these measurements (unlike later ones in this thesis) the voltage was not directly measured but only deduced from the applied voltage and known voltage division. However, the input resistance of the current amplifier used can be of the same order of magnitude as the dynamical resistance of samples at the gap edge and this creates some uncertainty to the calculated applied voltage around the gap edge. In effect this would mean that the voltage over the sample was less than what was deduced and hence push the bulk data to the correct direction. Suspended sample data would also be pushed more towards the $T^3$ curves. Nevertheless, the amount of cooling achieved is not sensitive to this experimental flaw and that clearly remains incompatible with the reduced $e-p$ coupling
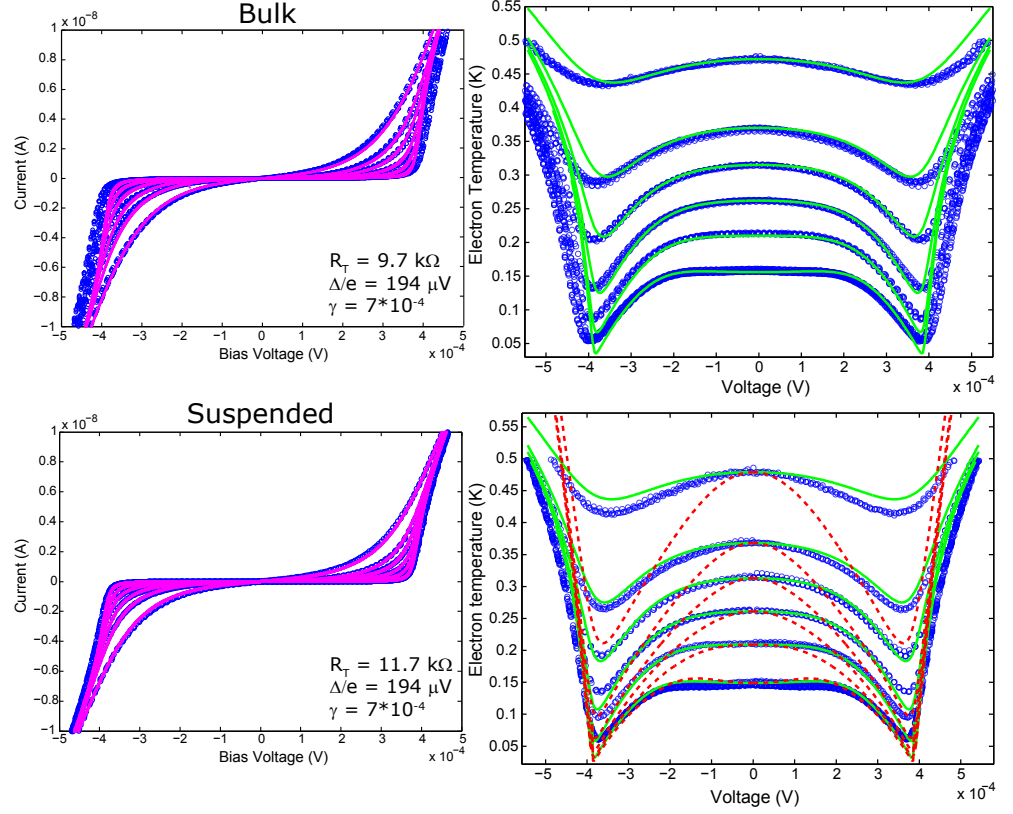
**Figure 3.2.** Data from a bulk control sample (top) and the sample shown in Fig. 3.1 (a) (bottom). Left panels show I-V curves and fits giving parameters mentioned in the figures. Right panels show cooling data and calculated curves with parameters from the I-V fit and $\Sigma = 1.7 \times 10^9 \ \mathrm{WK^{-5}m^{-3}}$. Dashed line in lower panel is calculated with Eqs. (3.1) and (3.2), i.e., the highest coupling limit at the 1D theory. Speed of sound is assumed to be 5000 m/s. The volumes of the two samples are different as the bulk control sample is a 4 $\mu$m long wire. This is taken into account in the calculations. The cooling curves have been corrected for asymmetry in the thermometer voltage measurement, coming from the resistance of the wires. The resistances were 350 $\Omega$ and 197 $\Omega$ for the bulk and suspended sample, respectively. Joule heating from these resistances is taken into account when calculating cooling curves.

hypothesis.

## 3.2 Membrane cooling

The most straightforward way to achieve the needed isolation from the environment and to enable integration of NIS coolers with applications 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 islands extending from a NIS junction to the membrane. The junctions itself need to be located on the bulk, in order to dissipate the excess heat on the "backside" of the cooler.

This kind of a 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 aluminium as a superconductor are facing a technological challenge in providing below 0.1 K temperatures 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 $^3$He refrigerator, or even better, a pumped $^4$He bath, combined with a NIS cooler. The first applications to benefit would be the ones where the fabrication onto a membrane is straightforward. This group includes especially radiation detectors, which are by default often fabricated on a membrane.

The membrane cooling was first demonstrated in [42] and significantly improved in [43] with a small membrane volume coupled to the bath through four few hundred micrometer long and $\sim$5 $\mu$m wide bridges. A considerable temperature decrease was achieved (from 200 mK to 100 mK), although the actual cooling power was modest ($\sim$ pW). However, actual application demonstrations have been done recently by the group of Ullom at NIST. They demonstrated first the cooling of a macroscopic size Ge cube [44] and then an aluminium transition-edge detector, designed for X-ray sensing [45]. In the latter experiment, an effective temperature reduction from 300 mK to 190 mK was achieved in the noise properties of the detector, presenting a significant technological advance. 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 few hundreds of picowatts.

### 3.2.1 Modelling and experiments on thermal conductivity

In order to test how much effective cooling power is needed for the platform, a finite-element-method (FEM) model of the temperature profile of the platform was devised [46]. In the model we assume that the edges of the membrane stay always at bath temperature, and that the at one end of the cold finger heat is removed with some fixed power (simulating a NIS cooler biased at the optimum cooling point). The diffusion of heat in the membrane and in the cold finger are modelled with three temperatures: $T_m$ for the membrane, $T_p$ for the phonons in the cold fingers and $T_e$ for the electrons in the cold finger. These are coupled through electron-phonon coupling and a thermal boundary resistance $\mathcal{K}$. Hence, in a 2D model we have three coupled diffusion equations for the temperatures

$$\nabla \cdot (-d_m \kappa_m T_m^n) \nabla T_m = \mathcal{K}(T_m^4 - T_p^4) \tag{3.5}$$

$$\nabla \cdot (-d_{cf} \kappa_m T_p^n) \nabla T_p = \mathcal{K}(T_p^4 - T_m^4) + d_{cf} \Sigma(T_p^5 - T_e^5) \tag{3.6}$$

$$\nabla \cdot (-d_{cf} \kappa_0 T_e) \nabla T_e = d_{cf} \Sigma(T_e^5 - T_p^5), \tag{3.7}$$

where $d_{cf}$ and $d_m$ are the thickness of the cold finger and the membrane, respectively, and $\kappa_m$ and $\kappa_0$ are the thermal conductance prefactors of the membrane and the metal. All the temperatures depend on the two spatial variables but other parameters are constant. The thermal conductance of the membrane is assumed to follow a powerlaw with exponent $n$ whereas the thermal conductance of the metal is from Wiedemann-Franz law. Here we assume the standard 3D $e-p$ coupling, which would seem logical based on the results in the previous section. In any case, although some difference between the 2D and 3D case have been seen [47], the difference is not huge and has only minor effect for this particular model. The thermal boundary resistance $\mathcal{K}$ is assumed large enough not to play any role. The phonons in the metal layer are assumed to behave similarly to the phonons in the membrane. In Fig. 3.3, an example of the heat profiles from FEM simulation is shown and in Fig. 3.4 the calculated temperature of the membrane as a function of the resistance of the cooling arrays is presented.

In order to make the model more precise, the exact thermal conductivities of the copper and the membrane should be known. The thermal conductivity of copper is related to its electrical conductivity through
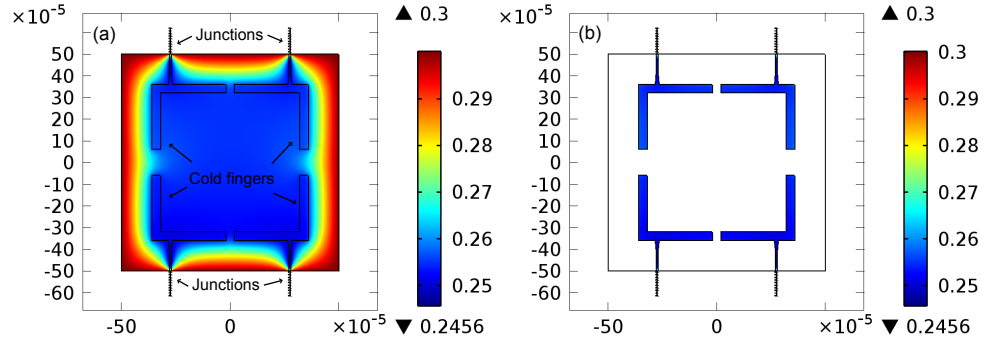
**Figure 3.3.** Finite-element-modelling of the membrane platform with commercial software COMSOL. **(a)** The membrane temperature profile. **(b)** The electronic temperature of the normal metal. Here we assume a bath temperature of 0.3 K and that each junction array shown in the edges have total cooling power of a SINIS structure with asymptotic resistance of 20 $\Omega$ and is biased at the optimum cooling point. Other parameters: $d_m = 100$ nm, $d_{cf} = 40$ nm, $\kappa_m T^n = 8T^{2.1}$ mWm$^{-1}$K$^{-1}$, $\kappa_0 = 1.03$ mWm$^{-1}$K$^{-2}$ and $\Sigma = 2$ GWm$^{-3}$K$^{-5}$.
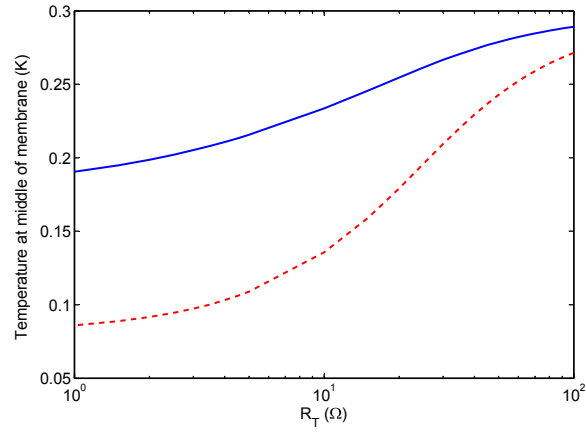


**Figure 3.4.** Calculated temperature at the middle of the membrane as a function of the resistance of one junction array. For solid line parameters and geometry are the same as in Fig. 3.3. Dashed line calculated with otherwise same parameters except $\kappa_m$ is divided by a factor of 20. This gives a rough idea of the performance if the membrane would be perforated around the edges. Already with quite high $R_T$:s a significant temperature reduction can be achieved but reaching very low temperatures requires highly transparent junctions and reduction of the thermal heatload with, e.g. membrane perforations.

Wiedemann-Franz law and so electrical measurements were carried out at 4 K for copper wires deposited with the same evaporator, resulting to measured square resistance of 1.8 Ω for 20 nm thick wires and 0.59 Ω for 40 nm thick wires at 4.2 K and a residual resistivity ratio (RRR) of around 2.

For the thermal conductivity of the membrane, various measurements exist [48, 49, 50, 51]. Common for all these measurements is that the heat conduction powerlaw has been consistently measured to be around 2 (i.e., $\kappa_m \propto T^2$) which might indicate a 2 dimensional phonon system. However, as silicon nitride is an amorphous material and the same powerlaw has been measured also for bulk amorphous materials at low temperatures [52], this is not totally clear. In any case, variations in material quality and stresses on the membrane might have big influence on the heat conductance. Hence, we have measured the thermal conductance from several silicon nitride membranes bought from the commercial supplier [53], which were used in the membrane cooler samples.

The thermal conductivity measurement is done as a heating measurement. Multiple SINIS structures are fabricated on the membrane (fabrication details are discussed in Sec. 3.2.3), to use one as a heater and other ones as thermometers in order to extract the temperature profile. For the full membranes, some approximations are usually done in order to extract an analytical result. In practice, we assume a circular boundary condition instead of a square. Assuming that at a distance of $R$ from the heater the temperature is at bath temperature and knowing the temperature at distance $r$ as well as heating power $P$, we can extract the thermal conductivity $\kappa_m$ as [49]

$$\kappa_m = \frac{\ln(R/r)}{2\pi d} \left(\frac{\Delta T}{\Delta P}\right)^{-1},$$
(3.8)

where $d$ is the thickness of the membrane. In practice we heat up the membrane from a known bath temperature and extract the linear response part of $T$ vs. $P$ curve to deduce $(\Delta T/\Delta P)$. The circular approximation was also checked against the results from the FEM simulations with the full geometry and (assuming point heater and point thermometer) works almost exactly as long as the thermometer is not too close to the edge of the membrane.
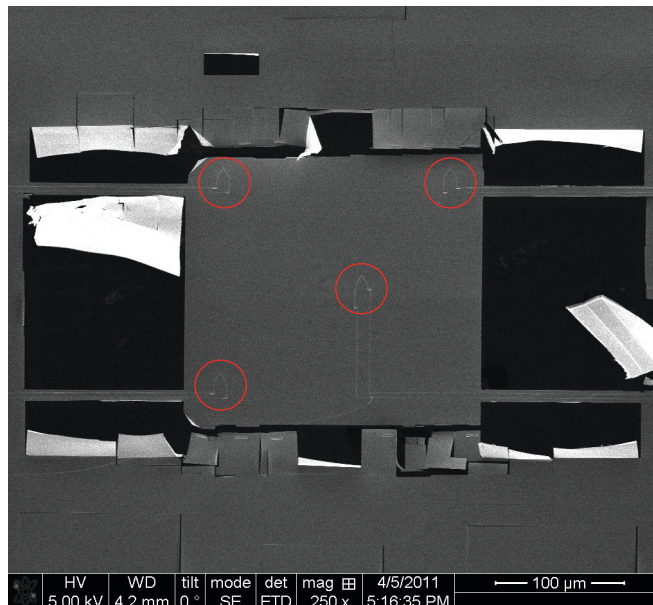
The results from the measurements performed on full membranes are shown in Fig. 3.5. We have measured $\kappa_m$ on three different membranes of 2 different sizes and from two distances in one membrane. There is some
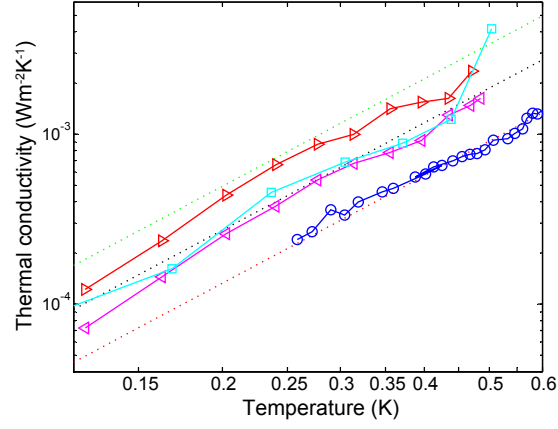
**Figure 3.5.** Measured thermal conductivity from silicon nitride membranes provided by commercial supplier [53]. The lowest data (circles) is measured from a $0.5 \times 0.5$ mm$^2$ membrane, other measurements are from $1 \times 1$ mm$^2$ membranes. Triangles pointing to different direction are measured from the same membrane with two different thermometers at different distances from the heater. Dotted lines are from top to bottom: 14.5 $T^{2.1}$ (value measured in [49]), 8 $T^{2.1}$ and 3.9 $T^{2.1}$ mWm$^{-1}$K$^{-1}$. The errorbars are about the size of the symbols, see text for details.

variation in the results, which is much larger than the expected uncertainty in the measurements. The 95 % confidence bounds of the linear fits to the heating data are about the size of the symbols in Fig. 3.5. These confidence bounds represent the uncertainty coming from the noise in the voltage measurement of the SINIS thermometers, which is the dominant error source here. The error coming from voltage to temperature conversion is small at this temperature range and possible error in the absolute temperature scale (estimated at around 1 %) plays no role in this measurement. The reasons for the variation between measurements are not known. One possibility is that during the fabrication process some impurities are introduced to the membrane surface and this then affects the thermal properties. Nevertheless, although there is some variation between measurements, they all are consistently showing lower thermal conductivity than the membranes employed in the past experiments.

It has been also reported before [54] that thermal conductivity in bridges with width of the order of the thermal wavelength of the phonons would be even lower than in membranes. Intuitively this should be the case as then the phonon system might be effectively one-dimensional instead of 2 dimensional as in the membrane. Related effect was also reported recently in crystalline silicon, where the thermal conductivity of a silicon beam was seen to depend on its shape on the scale of the thermal wavelength of the phonons [55, 56]. We tested this phenomenon by measuring the thermal conductivity in perforated geometries, where part of the membrane (con-

| HV | WD | tilt | mode | det | mag ⊞ | 4/5/2011 | ⸻ 100 µm ⸻ |
|---|---|---|---|---|---|---|---|
| 5.00 kV | 4.2 mm | 0 ° | SE | ETD | 250 x | 5:16:35 PM | |

**Figure 3.6.** A SEM image of one of the samples for the thermal conductivity measurement in restricted geometries. In this sample the bridges are 10 $\mu$m wide and 100 $\mu$m long. The four faint SINIS structures on the middle part are marked with circles (only two are used in the measurement). The triangle shaped part is the normal metal island.

taining heater and thermometer) was connected to the bulk of the membrane only through four 100 $\mu$m long bridges (see Fig. 3.6). Bridges of width $\sim 60\mu$m, $25\mu$m and $10\mu$m were measured. The results are shown in Fig. 3.7. The measurement reproducibility is much better than for the full membranes. Because of fabrication issues a 20 nm thick layer of Al was deposited on the membranes before perforation (see Sec. 3.2.3) and this shows up as an increase of the thermal conductivity at temperature above $\sim 400$ mK. At the lower temperatures the electronic contribution from Al is frozen out. The thermal conductivities measured, however, follow almost exactly the conductivities measured for the full membrane. The reason for this discrepancy to the previous measurements is not known but one possible explanation is lattice thermal conductivity of the Al layer. As the speed of sound of evaporated Al is probably lower than of that of the SiN membrane, this might move the transition from 2D to 1D to smaller dimensions. A sample with 5 $\mu$m bridges and no Al on the backside was also measured but because of new fabrication process needed here the thermometer leads had an unwanted copper shadow that dominated the thermal conductivity (see Sec 3.2.3).
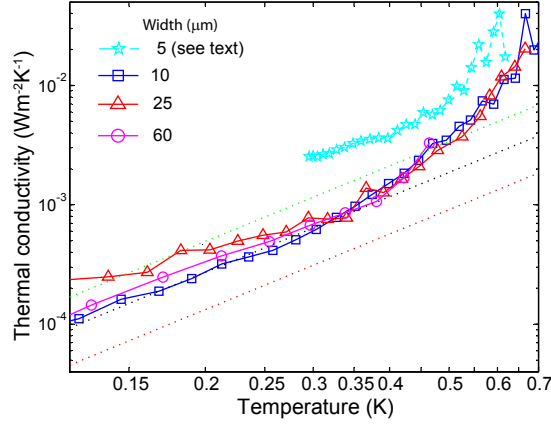
**Figure 3.7.** Measured thermal conductivity from perforated silicon nitride membranes. Different data sets are measured from samples with different width of bridges (shown in legend) connecting heated part to the bulk of the membrane. Dotted lines are the same as in Fig. 3.5. All samples except the 5 $\mu$m sample have 20 nm thick Al layer on the backside of the membrane, dominating the thermal conductivity above $\sim 400$ mK. In the 5 $\mu$m sample, the thermal conductivity is probably dominated by unintentional copper shadows of the Al leads of the thermometers. The errorbars are about the size of the symbols, see above for details.

### 3.2.2 Cooler experiments

Based on the FEM modelling presented before, a NIS cooler with moderately low cooling powers should be enough to demonstrate significant cooling of the membrane. This was tested experimentally with various membrane coolers in the course of the work described in this thesis, but invariably the results have been much worse than what the initial modelling suggested. The reasons are suspected to lie in local phonon heating induced by the heat load dissipated on the superconducting side of the NIS cooler. This same effect will be described in Section 4.3. In essence, the local phonon heating couples to the cold finger through electron-phonon coupling and creates parasitic heating power that compensates the cooling power of the junction.

Images of the first prototype are shown in Fig. 3.8. The four junction arrays were biased pairwise, each junction array had an asymptotic resistance of 100 $\Omega$, corresponding to junction resistivity of 400 $\Omega\mu$m$^2$. According to the FEM simulations these kind of coolers (even allowing for higher thermal conductivity of the membrane) should be able to refrigerate the membrane by 20 mK starting from 0.25 K (or 15 mK from 0.3 K). In the experiment, cooling of only 2-3 millikelvins was seen on temperature range 0.2 K - 0.3 K.

To diagnose the reasons for this discrepancy, cooler samples with extra
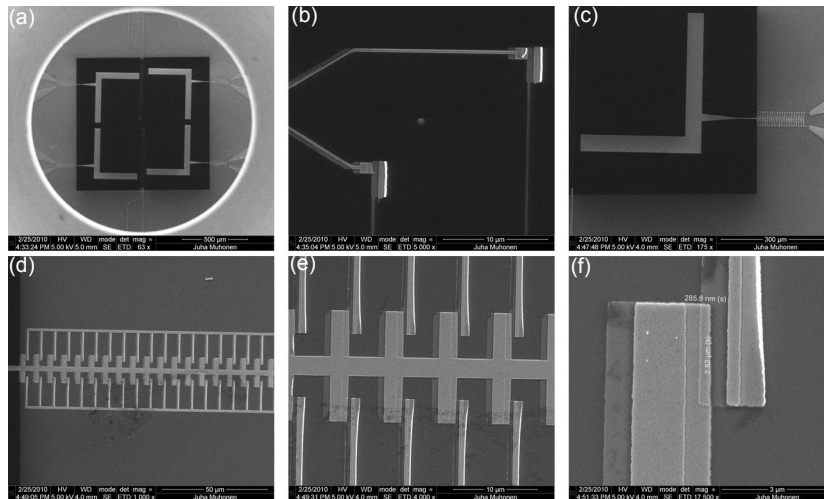
**Figure 3.8.** SEM images of the membrane cooler. **(a)** Overview of the sample. Junction arrays are located on the sides of the membrane and L-shaped cold fingers extend from them to the membrane. Two thermometers are located in the center of the membrane. **(b)** Close up on the membrane thermometer. **(c-f)** The cold finger and the junction array in progressively more detail. With this sample a temperature drop of around 2-3 mK starting from 0.2 - 0.3 K was seen with the thermometers in the middle, when all junction arrays were in use. Each junction array had asymptotic resistance of 100 $\Omega$.

thermometers attached to the cold fingers were fabricated. Measuring the temperature of the cold finger close to the cooler array and at the far end showed that the problem was two-fold: we both had lower cooling than expected close to the junction array and we had an unexpected thermal gradient over the cold finger so that at the membrane almost no cooling was seen (see Fig. 3.9). Experiments after this were focused on improving the thermal conductivity of the cold fingers. Thicker and wider copper fingers were tried but no real improvement on the membrane cooling was achieved although some effects were seen, see Fig. 3.10.

One possibility considered was that we had a parasitic heat conduction channel through the Al layer located underneath the copper layer in the cold finger. This extra layer is unavoidable with the shadow evaporation process if only one lithography step is used. Hence, we moved to using two-step lithography where the Al layer could be avoided. In two-step lithography, additional problems arise not only from the obvious need for alignment between the two lithography steps but also from having a contact resistance between the copper deposited in the first step to form the NIS coolers and the cold finger deposited on the second step. To avoid this, cleaning *in situ* is needed before deposition of the cold finger. A sample was fabricated and measured where a cold finger made of Au was deposited in the second lithography step. The temperatures in the cold
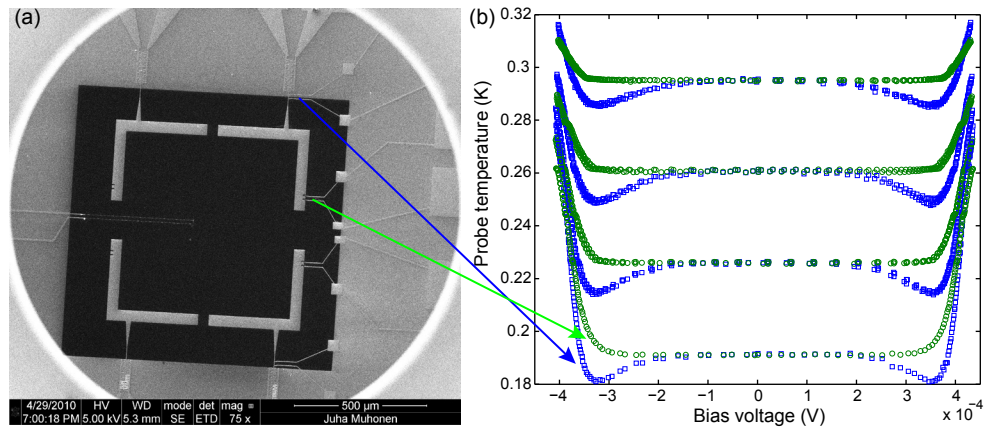
**Figure 3.9.** **(a)** SEM image of the sample. Two extra thermometer junction pairs are integrated to the cold finger, one close to the cooling junctions, the other one at the far end. **(b)** Data from sample presented in (a) at four different bath temperatures. Squares show the measured temperature from the thermometer junctions close to the cooling array and circles at the other end. Asymptotic resistance of each cooling array was 65 $\Omega$. From FEM simulations the closer thermometer would be expected to reach 0.21 K and the farther thermometer 0.22 K starting from 0.25 K. The possible errors in the absolute temperature scale are of the order of 1 %, coming from the uncertainty in the calibration of the cryostat thermometer. The biggest source of error in the amount of cooling (relative temperature drop), is noise in the voltage measurement seen as point-to-point scatter in the plots.
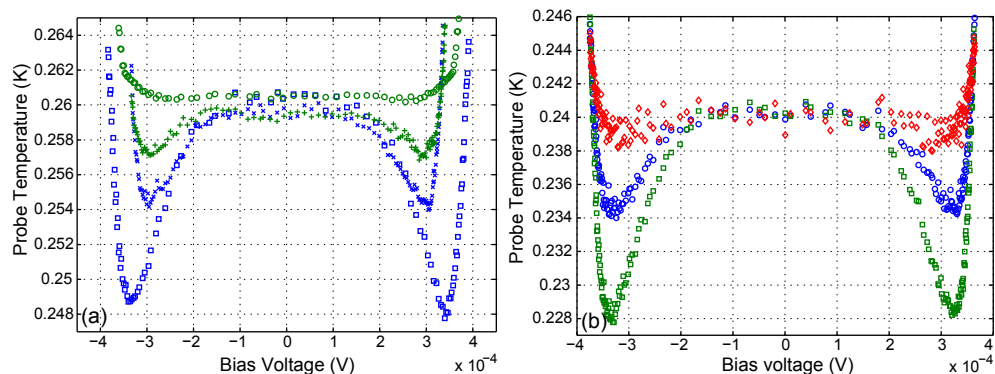


**Figure 3.10.** Effects of thicker and wider Cu layer. **(a)** Squares and circles show part of the same data which is presented in Fig. 3.9 (b), crosses and plusses show data of a similar sample with a 60 nm thick (instead of 30 nm) copper layer in the cold fingers. In the sample with thicker copper layer, the thermal gradient over the cold finger is reduced, but at the same time the overall cooling effect decreases due to the extra heatload. The cooling effect is again much smaller than expected. **(b)** Data from a sample with 50 nm thick cold finger with a geometry where the narrowest part of the cold finger was widened. The two lowest curves correspond to similarly placed thermometers as in Fig. 3.9 and diamonds correspond to a membrane thermometer (faintly visible in the centre of Fig. 3.9 (a)). This cooler had lower resistance ($R_T = 40$ $\Omega$ for the junction array) so that the cooling at the nearest thermometer is roughly the same as in Fig. 3.9 (b). The thermal gradient over the cold finger is reduced but membrane cooling is not enhanced. Expected minimum temperature from the simulations would be 200 mK, 198 mK and 190 mK for the three thermometers. Assuming only fifth of the expected cooling power would produce the measured cooling for the thermometer nearest to the cooling junctions, but the membrane would still be expected to cool to 0.23 K. Error assessment same as in Fig. 3.9.

44
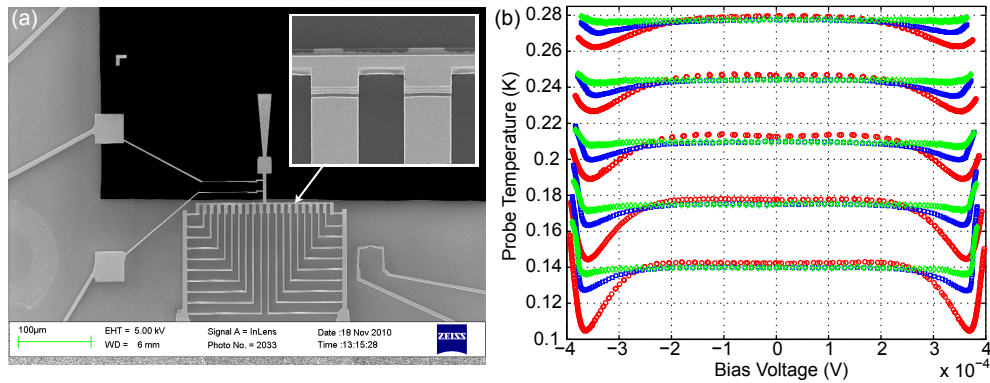
**Figure 3.11. (a)** SEM image of the other cooler array design, inset shows a close-up to the junction area. The superconducting leads are now designed so that they are fully covered with the normal metal. This sample was measured with and without the cold finger, this image was taken before the cold finger is in place. **(b)** Cooling performance of the sample in (a). The circles (lowest) points show the performance before the integration of the cold finger. The squares and diamonds show the cooling performance after the fabrication of the cold finger at two thermometers, the one near the junctions (same as circles) and the one at the end of the cold finger, respectively. The $R_T$ of the junction array was 70 $\Omega$. Error assessment same as in Fig. 3.9.

finger could not be probed here as tunnel junction fabrication with gold was not successful. The membrane temperature was probed and cooling effect was very similar to the previous measurements, i.e., of the order of few millikelvins.

All the data collected from the measurements seem to indicate that the real problem did not lie in the properties of the cold finger. Rather it seemed probable that we had an unexpected heatload coupling to the cold finger and this combined with the fact that our coolers were performing below expectations caused the effective cooling power to be reduced dramatically. The most likely cause for this extra heatload is thought to be local phonon heating created by heatload to the superconducting side of the NIS cooler. Before any real modelling, we tried to minimize the effects of this heatload by maximising the cross-sectional area of the superconducting lead close to the junction and optimising the contact area between the superconducting lead and the normal metal layer on top of it (acting as a quasiparticle trap, see Chapter 4). These coolers were fabricated with two-step lithography and measured both without the cold finger part and with it. However, the cooling performance was not considerably improved, see Fig. 3.11.

The discrepancy between FEM simulations and the experimental results highlighted the need for more accurate modelling and experiments on the heat dissipation on the superconducting side, this is the topic of the

next Chapter. The effects of local phonon heating should be minimised in the final design by separating the heated phonon system from the cold finger more effectively. In practice this will mean perforating the membrane around the edges.

### 3.2.3 Fabrication issues

There are many aspects on the fabrication process on the 100 nm thick silicon nitride process that need not to be generally considered when fabricating on a bulk substrate (the basics of which are described in Section 1.1). The first issue is the amount of undercut achieved in the basic copolymer-PMMA process. When fabricating on a bulk substrate, the largest part of the dose for the underlayer (copolymer) comes from backscattering of the electrons from the substrate. This backscattering is not very localized and hence produces a dose pattern that extends on all sides of the primary beam. Combined with a proper developer, this is the phenomenon that enables the undercut in copolymer and is essential for the shadow evaporation process. When fabricating on a thin membrane, most of this backscattering is lost as the majority of electrons can pass through the thin membrane without any scattering. This means that not only making EBL on a membrane requires doses many times higher than on a bulk substrate but in addition achieving a nice undercut profile becomes exceedingly difficult with just MIBK development. Moving to an extra developer that develops the copolymer much faster than the PMMA enables making large undercuts without the extra dose from backscattered electrons. The developer for this purpose has been discussed before in another PhD dissertation [57], and is called methylglycol. An example of the effect of this developer to tunnel junction fabrication on a membrane is presented in Fig. 3.12.

For the thermal conductivity measurements, additional requirements for the undercut arise. To measure the small thermal conductivity of the thin dielectric membrane precisely, one has to be careful not to introduce any parallel conduction channels. A long copper wire (the shadow of the superconducting Al lead of the thermometer) can conduct heat better than the perforated membrane even if made relatively thin. Hence, getting rid off this shadow was very important for those measurements. This can be done if the undercut is small enough so that evaporating in a large angle will produce copper only on the junction area (with higher dose) but not on the leads.
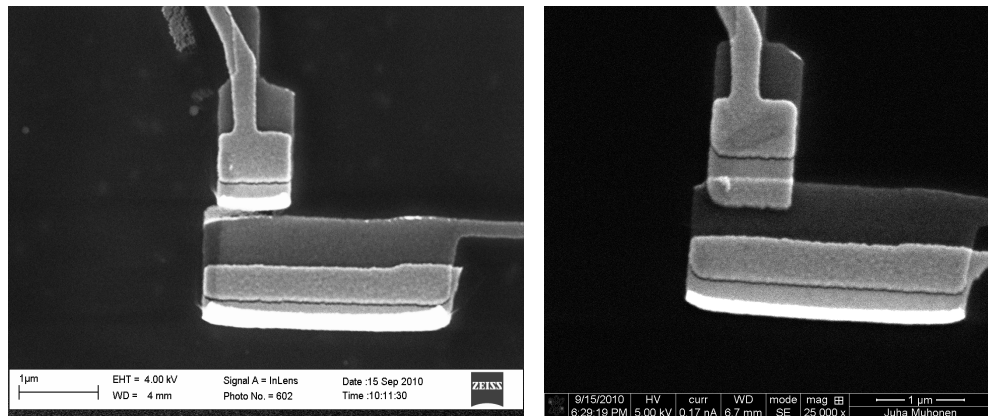
**Figure 3.12.** Effects of methylglycol to sample fabrication on the membrane. The sample on the left is done with just MIBK development, not enough undercut is achieved to produce tunnel junctions. On the example on right, methylglycol is used to extend the undercut.

For the perforation of the membranes in the thermal conductivity measurements with the bridge geometries, a dual-beam system from FEI (Helios 2000) was used [58]. Dual-beam refers to a machine having both electron beam (for imaging and possibly lithography) and an ion-beam (for milling, induced deposition and lithography). Perforating the membrane with the focused ion-beam (FIB) offers the maximum flexibility for the process as the cutting profile can be adjusted "on-the-fly" and imaging can be done simultaneously with milling. The downside is that the process is relatively slow for large areas and at the moment cannot be automated easily.

For our silicon nitride membranes the yield of the perforations was in the beginning of the process eminently low. The membranes tended to crack from the edges of the perforations and sometimes just flew away completely. The origin of these problems was ultimately traced to charging of the membrane. Our perforation process was to mill small lines into the membrane in the form of a square or circle so that the middle of the pattern will then "fall" away. However, when milling this kind of small lines, both sides of the line tend to gather some amount of the $Ga^+$ ions used in the milling and hence charge up. These sides then start to repel one another and once the line is deep enough this force can be strong enough to cause the pattern to tear off violently causing damage to the rest of the membrane.

Charging of the substrate is known to be a problem in FIB processing and the conventional method is to scan the working area with the electron beam sequentially with the milling in order to neutralise the ion charges with the electron charges. This approach was tested without con-
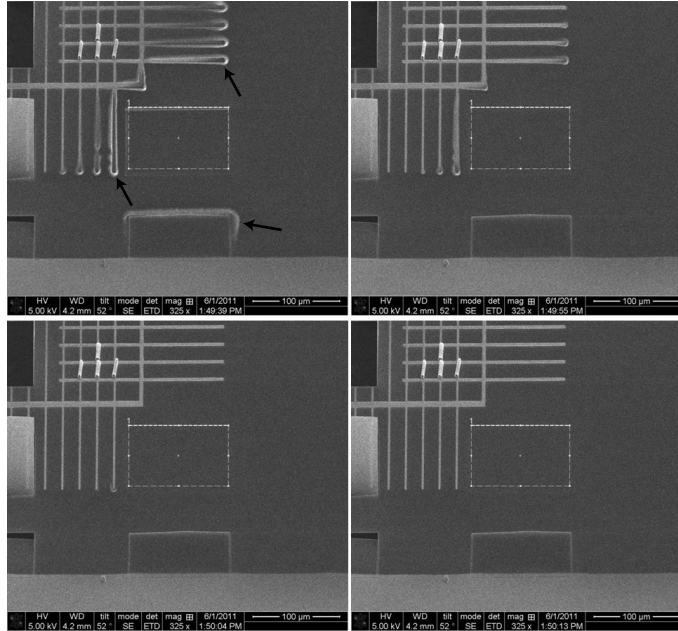
**Figure 3.13.** SEM image slide show taken after ion milling. The time order of the figures is from left to right and from top to bottom. The charge polarisation is seen as "flames" around the edges of the metallic structures after milling and it then fades away with time. The arrows point to locations where this can be most clearly seen.

siderable improvement. A working solution (visible in the thermal conductivity measurements above) was to cover one side of the membrane with aluminium layer before the milling. The aluminium layer is a good conductor and prevents charge build-up in the vicinity of the milled area.

As a superconductor, Al has a very small effect to the thermal conductivity at the lowest temperatures (at least electronically), enabling the thermal conductivity measurements to be made. Nevertheless, for achieving the smallest possible thermal conductivity in the narrow bridge geometries, one needs to abandon the Al layer. For this purpose a metallic pattern on top of the membrane but away from the "active" area was tested. This kind of "lightning rod" can be introduced to the areas where the milling is made as long as care is taken that no metal is left on the bridges. In Fig. 3.13 some SEM images of this kind of a structure during milling is shown. The charging can be seen as flame-like pattern around the metal and the milled shape. The charge is seen escaping via the metallic wires. A narrow bridge structure without the Al layer was realized using these lightning rods. Unfortunately, in this structure the copper shadows of the thermometer leads were not completely eliminated and they dominated the thermal conductivity.

# 4. Quasiparticle dynamics in a superconductor

For the practical applications of NIS cooling, understanding the thermal dynamics of the quasiparticle excitations in the superconductor is of crucial importance in order to know how much power can be dissipated into the superconducting lead without critically increasing the quasiparticle population (see Chap. 2). In this Chapter we will go through the relevant theory, modelling and experiments, relating to the behaviour of quasiparticles in the superconducting side of a NIS cooler. This Chapter is related to Publications V and VII.

## 4.1 Thermal properties of a BCS superconductor

In a BCS superconductor, the thermal properties of the electron system are modelled with quasiparticle excitations. The Cooper pair condensate itself carries no entropy and has no explicit role in the thermal properties. The dominant relaxation mechanisms are analogous to the normal metal case: there are the quasiparticle heat conductivity along the superconductor, and quasiparticle-phonon relaxation, which in a superconductor 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 $\Delta$. Combined, these effects lead to exponentially suppressed heat conductivity and coupling to the phonon system at low temperatures $[\kappa, \Sigma \propto \exp(-\Delta/(k_B T))]$.

The BCS-theory predicts the quasiparticle density $n_{\mathrm{qp}}$ in thermal equilibrium to be

$$
n_{\mathrm{qp}} \;=\; \int_{-\infty}^{\infty} \mathrm{d}E \, N_S(E) f(E) = 2N_0 \int_{\Delta}^{\infty} \mathrm{d}E \, \frac{E}{\sqrt{E^2 - \Delta^2}} f(E)
$$

$$\approx \quad N_0\sqrt{2\pi k_B T \Delta}\, e^{-\Delta/k_B T}. \tag{4.1}$$

The last step is an approximation that applies when $\Delta \gg k_B T$. For illustration, one can apply the parameters of aluminium in Eq. (4.1). Assuming $\Delta/k_B = 2.4$ K and a temperature of $T = 100$ mK the quasiparticle density is very low $n_{\rm qp} \sim 10^{-5}\ (\mu{\rm m})^{-3}$. However, invariably experiments have shown quasiparticle densities above what is predicted by Eq. (4.1) at the lowest bath temperatures. The lowest ever observed density is about $3 \cdot 10^{-2}\ \mu{\rm m}^{-3}$ and the source of excess quasiparticles has been demonstrated to be noise radiated from the environment [59].

The reduction in thermal conductivity at the superconducting state has been calculated theoretically soon after the BCS theory appeared [60]. 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

$$\gamma(T) = \frac{3}{2\pi^2} \int_{\Delta/k_B T}^{\infty} \frac{x^2}{\cosh^2(x/2)} dx \simeq \frac{6}{\pi^2}\left(\frac{\Delta}{k_B T}\right)^2 e^{-\Delta/k_B T}, \tag{4.2}$$

where the approximation shown on the right applies again for $k_B T \ll \Delta$. The suppression factor $\gamma(T)$ is plotted in Fig. 4.1 (a). Note that we assume everywhere that the superconducting gap $\Delta$ has the temperature dependence given by BCS theory. Experimental evidence for Eq. (4.2) was added recently in [61].

The coupling between the quasiparticle system and the phonon system is mainly mediated by the quasiparticle recombination, which is a process where two quasiparticles of opposite momenta (k and $-$k where $\hbar v_f k \approx \Delta$) recombine to form a Cooper pair and emit a phonon with energy equal to $2\Delta$. The recombination rate was studied several decades ago [62] but the associated heat flux from quasiparticles to the phonon system has been experimentally determined only very recently [63]. From quasiclassical theory the heatflux reads

$$
\begin{aligned}
P_{qp-p}^s \;=\; & -\frac{\Sigma \mathcal{V}}{2\Gamma(5)\zeta(5)k_B^5} \int_{-\infty}^{\infty} {\rm d}E\, E \int_{-\infty}^{\infty} {\rm d}\epsilon\, \epsilon^2 {\rm sqn}(\epsilon) N_S(E) N_S(E+\epsilon) \left(1 - \frac{\Delta^2}{E(E+\epsilon)}\right) \\
& \times \; \left\{ \coth\left(\frac{\epsilon}{k_B T_p}\right) [f(E+\epsilon) - f(E)] + f(E) + f(E+\epsilon) - 2f(E)f(E+\epsilon) \right\}. \tag{4.3}
\end{aligned}
$$

Which at the limit $T_p = 0$ gives

$$
\begin{aligned}
P_{qp-p}^s \;=\; & -\frac{\Sigma \mathcal{V}}{\Gamma(5)\zeta(5)k_B^5} \int_{-\infty}^{\infty} {\rm d}E\, E \int_{-\infty}^{\infty} {\rm d}\epsilon\, \epsilon^2 N_S(E) N_S(E+\epsilon) \left(1 - \frac{\Delta^2}{E(E+\epsilon)}\right) \\
& \times \; \begin{cases} f(E+\epsilon)(1-f(E)) & \text{if } \epsilon > 0 \\ f(E)(f(E+\epsilon)-1) & \text{if } \epsilon < 0 \end{cases}. \tag{4.4}
\end{aligned}
$$
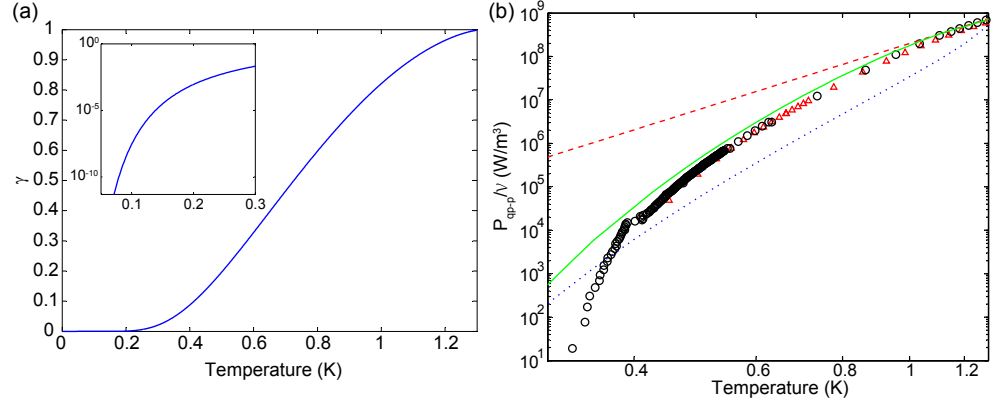
**Figure 4.1.** **(a)** The reduction factor of thermal conductivity Eq. (4.2) plotted as a function of temperature. Inset shows a close-up to the low temperature regime on logarithmic scale. **(b)** Quasiparticle-phonon coupling. Symbols showing the experimental data from [63], dashed line the normal state $e-p$ coupling (with $\Sigma = 2\times10^8$ WK$^{-5}$m$^{-3}$), solid line is from Eq. (4.4) and dotted line is analytical approximation valid at low temperatures $P_{qp-p}^s = 0.98e^{-\Delta/k_B T}P_{e-p}$. In both (a) and (b), $T_c$ of 1.3 K is assumed for the calculations.

The experimental data from [63] is shown in Fig. 4.1 (b) along with a calculated curve from Eq. (4.4), demonstrating orders of magnitude weaker coupling in the superconducting state as compared to the normal state already at temperatures of the order $0.3T_c$ (0.4 K).

A diffusion relation for the temperature profile in the superconductor can be constructed by inserting Eq. (4.4) and Eq. (4.2) to a conventional diffusion equation

$$\nabla \cdot [-\kappa_s(\mathbf{x}, T_{qp})\nabla T_{qp}(\mathbf{x})] = \mathcal{P}_{qp-p}^s(\mathbf{x}, T_{qp}, T_p) + \mathcal{P}_{\text{ext}}(\mathbf{x}), \qquad (4.5)$$

where $\mathcal{P}_{\text{ext}}$ is the power density from possible external sources, $\mathcal{P}_{qp-p}^s = P_{qp-p}^s/\mathcal{V}$ is the power density and x describes the spatial coordinate.

## 4.2 Thermal model of the quasiparticles for NIS cooling applications

As mentioned in Chapter 2, the efficiency of a NIS cooler (the ratio of the cooling power over the input power) is roughly $0.7\ T/T_C$. At 0.3 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 is 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, as mentioned above, the dissipated power can create a high density of non-equilibrium quasiparticles on the super-

conducting side of the cooler, i.e. heat it up. This tends to suppress the cooling power of a NIS junction.

In many instances [64, 65],III the degradation of the cooling power of a NIS junction due to the overheating of the superconductor has been successfully modelled 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 $P_{bf} = \beta(I_{\mathrm{NIS}}V - P_{\mathrm{NIS}})$. Typically $\beta$ found experimentally lies between 1 and 10 %. Although this kind of a 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 to model this effect more precisely, based on diffusion equations [66, 67, 68, 69]. Assuming a diffusion of quasiparticles away from the junction area as well as their relaxation, one can self-consistently calculate the quasiparticle distribution of the superconductor and hence the cooling power of the junction.

Since the 70s [70, 71], there has been a discussion if the non-thermal quasiparticles can be most accurately described with an effective temperature ($T^*$) or with an effective chemical potential ($\mu^*$) and this issue has also been revisited in recent paper [72]. In all models discussed in this dissertation the quasiparticle distribution is described by a Fermi function with a position dependent effective temperature. This is a convenient choice as it allows a straightforward calculation of the tunnelling rates based on the extracted effective temperature.

On the other hand, an argument can be made that the quasiparticle distribution will not follow a Fermi distribution at all and then the correct way forward is to model the diffusion of the quasiparticle number $n$ directly by using an equation [66, 67]

$$D_s \frac{\partial^2 n(x)}{\partial x^2} = \Gamma_{qp-p} + \Gamma_{\mathrm{ext}}, \tag{4.6}$$

where $\Gamma_{qp-p} + \Gamma_{\mathrm{ext}}$ are now the relaxation (scattering) rates to phonons and external environment, respectively, and can be converted to power by multiplying with the energy exchanged in each scattering event. $D_s$ can be related to the normal state diffusion constant $D_n$ by $D_s = \sqrt{1 - (\Delta/E)^2}D_n$, where $E$ is the energy of the quasiparticle and $D_n$ is related to normal state heat conductivity through $\kappa_n = \mathcal{L}D_n N_F e^2 T$ ($\mathcal{L}$ is the Lorentz number). In this way, one does not need to make the assumption of a thermal distribution but one now needs to consider explicitly the energies of the quasiparticles. In practice, one of two assumptions is then needed. One can either assume that the quasiparticles follow a thermal distribution

and, in fact, it is straightforward to show that with this assumption the left-hand-side (LHS) of Eq. (4.6) will be equivalent to LHS of Eq. (4.5) (the connection between $n$ and $T_s$ is from Eq. (4.1)). The other option (adapted in [66] and [67]) is to replace the $E$ in Eq. (4.6) with an approximate average energy of quasiparticles in the sample $\langle E \rangle$. In the previous examples it was assumed that $\langle E \rangle = \Delta$. However, in those articles, in order to calculate the effects on the NIS junction, an effective temperature $T_s$ is still assumed at the junction area and calculated from the solved $n$. Because of this, the difference between the two approaches is small.

To make some simple estimates with the effective temperature 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 (we neglect the $P_{\text{ext}}$ term and the prefactor of the order of unity in Eq. (4.4))

$$\frac{\partial}{\partial x}\left[\frac{6}{\pi^2}\left(\frac{\Delta}{k_B T}\right)^2 e^{-\Delta/(k_B T)}\kappa_n \frac{\partial T}{\partial x}\right] = e^{-\Delta/(k_B T)}\Sigma(T^5 - T_p^5), \qquad (4.7)$$

where we have written $T = T_{qp}$ for clarity. Linearising Eq. (4.7) for small temperature differences $\delta T(x) = T(x) - T_p$, one obtains a simple expression for the temperature profile in a uniform one-dimensional wire. For a wire extending to positive $x$, we have then $\delta T(x) = \delta T(0)e^{-x/\ell_0}$, where $\delta T(0)$ is determined by the heat input at the end of the wire and the relaxation length is given by

$$\ell_0 = \frac{\Delta}{\pi k_B}\sqrt{\frac{6\mathcal{L}\sigma}{5\Sigma}}T_p^{-5/2}. \qquad (4.8)$$

Putting the parameters of aluminium in Eq. (4.8), we find that $\ell_0 T^{5/2} \simeq 50$ $\mu$m·K$^{5/2}$. 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 linearised approximation by employing the boundary condition $P = -\kappa_s A \frac{d\delta T(x)}{dx}|_{x=0}$, where $A$ is the cross-sectional area of the wire. Inverting this for the temperature rise for a given heat input, we find

$$\delta T(0) = \frac{\ell_0}{\kappa_s}\frac{P}{A} = \frac{\pi k_B}{\Delta\sqrt{30\mathcal{L}\sigma\Sigma}}e^{\Delta/(k_B T)}T^{-3/2}\frac{P}{A}. \qquad (4.9)$$

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

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

Fortunately, the situation is usually not as bad as this result would seem to suggest, as the effects of so-called quasiparticle traps [73, 74, 75, 76] have been neglected so far. 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 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. A safe way of introducing moderately efficient traps 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. As explained in Sec. 1.1 this fabrication procedure produces first the superconducting (e.g. aluminium) lead, which is subsequently oxidized, and thereafter a metal layer (e.g. copper) is deposited at another angle, forming the NIS junction. In the process, this kind of overlap structure is also created outside the junctions and with proper mask design can be used to cover the superconducting leads. The mechanism of the quasiparticle thermalization in this structure is via tunnelling of hot quasiparticles into the normal metal, i.e., a NIS junction with zero voltage (see Eq. (2.8)). A metal layer in direct contact with the superconductor would be more efficient trap but 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, optimising the distance of these kinds of direct traps needs care.

By similar arguments used in obtaining Eq. (4.8), we can obtain a thermal relaxation length in a structure with a trap separated by an insulating layer,

$$\ell_1 = \left( \frac{2\sqrt{2}d\rho_T\sigma}{\sqrt{\pi}} \right)^{1/2} \left( \frac{k_BT}{\Delta} \right)^{1/4} \tag{4.10}$$

and for the temperature rise for a given heat input

$$\delta T(0) = \frac{\ell_1}{\kappa_s}\frac{P}{A} = \sqrt{\frac{\sqrt{2}d\rho_T\pi^4}{18\sqrt{\pi}\mathcal{L}^2\sigma}} \left( \frac{k_B}{\Delta} \right)^{9/4} e^{\Delta/(k_BT)}T^{5/4}\frac{P}{A}. \tag{4.11}$$

Here $d$ is the thickness of the superconducting lead and $\rho_T$ is the specific resistivity of the trap barrier. For a relatively resistive barrier $\rho_T = 1$ k$\Omega\mu$m$^2$ with $d = 30$ nm and $T = 200$ mK, we obtain $\ell_1 \approx 20$ $\mu$m, which

is about two orders of magnitude shorter than in a bare superconducting wire. Both this value and the relaxation length for bare Al wire are in accordance with the values obtained in [77]. Similarly the heating power required for certain temperature rise is two orders of magnitude larger than for a bare wire $\delta T(0) \simeq 0.15P/A \simeq (1.5 \cdot 10^{13} \text{ K} \cdot \text{W}^{-1})P$.

Some numerical results in 1D with the full equation are shown in Fig. 4.2. They have been calculated solving Eq. (4.5) using Matlab PDE solver *pdepe*. A fixed temperature boundary condition is assumed at one end of the wire ($x = 0$) and at the other end ($x = 0.5$ mm) a fixed power input (a-c) or a NIS junction (d) is applied. $\mathcal{P}_{\text{ext}}$ is now the heat flow to the trap. The numerical results show that effective (transparent) quasiparticle traps are needed for a NIS cooler to work at all. Also notable is that the effectiveness of the trap is not very dependent on the distance of the trap from the junction, as long as the separation is of the order of few micrometers. The fact that the transparent trap case seems to be more efficient at high powers than a normal state wire is partly a by-product of our assumption of an idealised trap fixed at temperature 0.2 K. At high enough transparencies the trap will also heat-up as the $e - p$ coupling in the trap will become the thermal bottleneck. However, it should be noted that as Cu has an order of magnitude higher $\Sigma$, a transparent trap might be even more efficient in heat removal than normal state Al. This would be emphasized if the trap would be thicker than the Al wire.

## 4.3 Experimental results and comparison to 2D simulations

In order to experimentally test the effects of excess quasiparticles to NIS cooling, a set of samples with differing geometries (different cross-sectional area $A$) and trap transparencies (parameter $\rho$) was fabricated [78]. The properties of these are summarised in Table 4.1. The sample geometry and their cooling characteristics are shown in Fig. 4.3. Because of the differing junction resistances $R_T$, this "raw data" of the normal metal temperature cannot, however, be compared directly. In order to extract the effects of different dimensions and $\rho$ we have to normalise the data. The normalisation is achieved by comparing the measured temperature reduction of the normal metal $\delta T$ to the optimal temperature reduction $\delta T_{\text{opt}}$ calculated by solving the heat balance equation Eq. (2.15) with the parameters $(R_T, \Delta, \gamma, \mathcal{V})$ of the corresponding sample. Note that the effects of sub-gap leakage are included in $\delta T_{\text{opt}}$ as a finite $\gamma$ and hence these
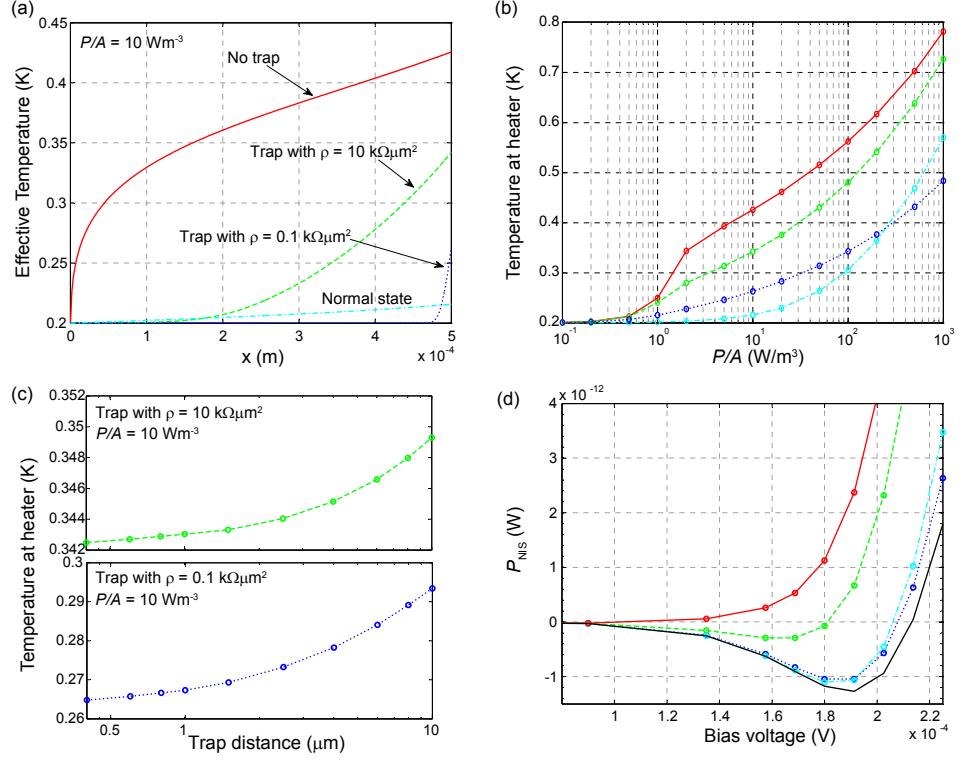
**Figure 4.2.** Results from the 1D effective temperature diffusion model.

**(a)** Temperature profile of a superconducting wire, with a power input of $P/A = 10$ W/m$^2$ at the $x = 0.5$ mm and a fixed temperature of 0.2 K at $x = 0$. $A$ is the cross-sectional area of the wire, thickness is 35 nm (required for the calculations with trap). Phonons and trap temperature are assumed to be fixed at 0.2 K. Solid curve is a pure superconducting wire, dashed and dotted curve with a quasiparticle trap assuming oxide resistivity of 10 and 0.1 k$\Omega\mu$m$^2$, respectively. Dash-dotted curve assumes a normal Al wire.

**(b)** Temperature of the wire at $x = 0.5$ mm as a function of the power input with same assumptions as in (a). Curve legends same as in (a).

**(c)** Temperature of the wire at $x = 0.5$ mm as a function of trap distance, with power input $P/A = 10$ W/m$^2$. Trap distance of 10 $\mu$m means that the trap is assumed to be located at the interval $0 \leq x \leq 0.49$ mm. Upper and lower curve with oxide resistivity of 10 and 0.1 k$\Omega\mu$m$^2$, respectively. Note, that in other plots the trap distance is assumed to be zero.

**(d)** Illustration of the effects of superconductor heating to NIS cooling. Assumptions and curves same as in (a) but now the power input is from a NIS junction with $R_T$ of 500 $\Omega$ and $\Delta = 200$ $\mu$eV. The plot represents the power to the normal metal part, calculated self-consistently. In this case we also have to fix the width of the wire, assumed to be 2 $\mu$m. $T_n$ is fixed at 0.2 K. Lowest solid curve is with $T_s$ also fixed at 0.2 K.

In all plots we assume that $\kappa_n = 0.9$ WK$^{-1}$m$^{-2}$ and $\Sigma = 2 \times 10^8$ WK$^{-5}$m$^{-3}$ (in normal state).
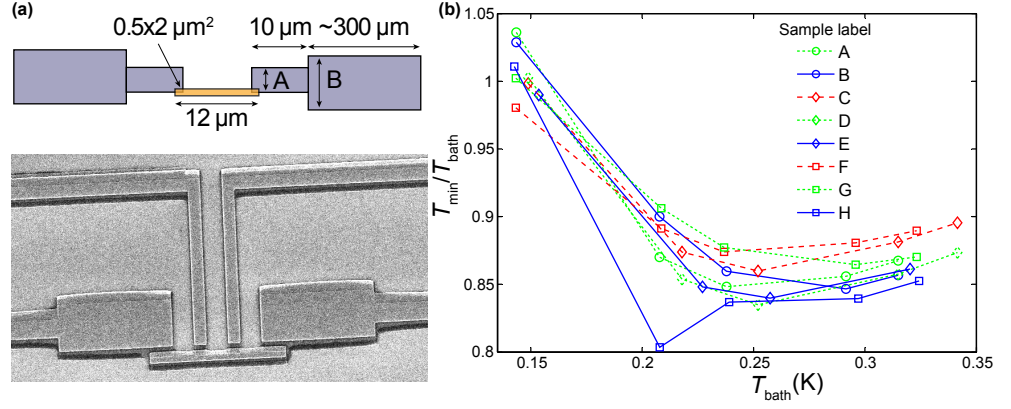
**Figure 4.3. (a)** A schematic of the geometry of samples and a scanning electron micrograph of one representative sample. The thermometer junctions in the middle of the island are not shown in the schematic. Lighter material is Cu and darker Al. **(b)** The ratio of achieved minimum temperature of normal metal $T_{\min}$ and bath temperature $T_{\mathrm{bath}}$ as a function of $T_{\mathrm{bath}}$ in samples A-H. The same labelling will be used throughout so that the symbol will refer to dimension A (circle = 2.5, diamond = 5 and square = 10 $\mu$m) and line to the dimension B (dashed red = 2.5, dotted green = 5 and solid blue = 10 $\mu$m).

do not contribute to the difference of the calculated and measured case. The influence of Andreev current is not included in the calculation but as shown in Sec. 2.3, its contribution is expected to be negligible at the optimum cooling voltage.

In all the samples the variations in $\Delta$ and $\gamma$ are very small. This can be seen in Fig. 4.4 where the current-voltage (I-V) curves measured at $\sim 50$ mK of samples A-H are plotted at low bias voltages with I-V curves calculated with the extreme parameters. The variations in $\Delta$ are in between 194-200 $\mu$eV and $\gamma$ values are in the range 2-3 $\times 10^{-4}$. These changes have very little effect to the calculated $\delta T_{\mathrm{opt}}$ and hence we will assume that all the samples have $\Delta = 200$ $\mu$eV and $\gamma = 2 \times 10^{-4}$. The $\delta T / \delta T_{\mathrm{opt}}$ ratio is presented in Fig. 4.5 for samples A-H as a function of the bath temperature.

The data presented in Fig. 4.5 demonstrates that the results are in qualitative agreement with the results from 1D diffusion model. If we reject the samples where A>B, the coolers get more efficient as the lead is made wider although the differences between different samples are small. To make more accurate comparisons between the effective temperature diffusion model of quasiparticles and our experiments, we have done 2D finite-element-method (FEM) simulations of the temperature profiles of our structures. The meshing and solving of the differential equations were done with the commercial program COMSOL (numerical values for the NIS heatflow integrals and the thermal conductivity of the superconductor were evaluated in Matlab). In the 2D model, we had three effective,
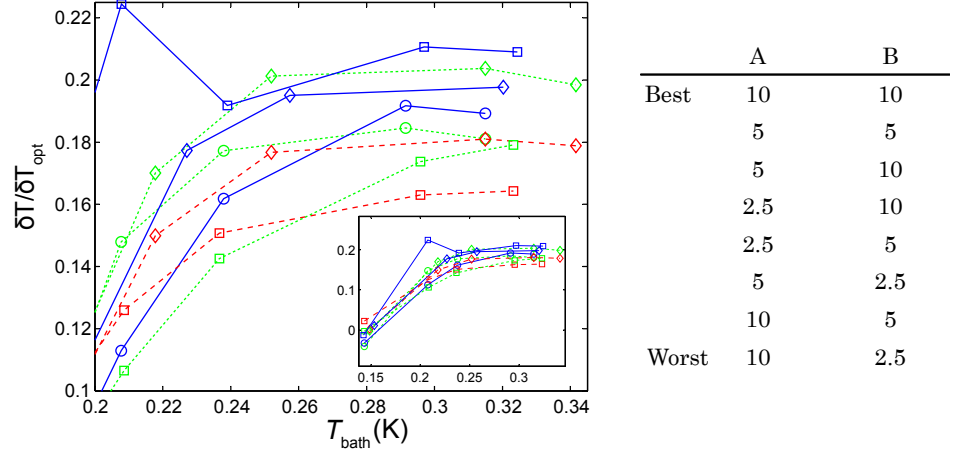
**Table 4.1.** Sample parameters. A and B refer to the dimensions in Fig. 4.3 (a), $d_{Al}$ is the thickness of the Al layer, $d_{\mathrm{trap}}$ is the thickness of the Cu layer on top of Al, $\mathcal{V}$ is the volume of the normal metal island, $R_T$ is the total resistance of the sample (i.e., 2 times the $R_T$ of a single junction), and $\rho$ is the resistivity of the oxide between the Al layer and the trap (calculated from the measured $R_T$ and junction area). OGP refers to "on the ground plane", NGP to "not on the ground plane" and DT to direct trap, see text.

| Sample | A ($\mu$m) | B ($\mu$m) | $d_{Al}$ (nm) | $d_{\mathrm{trap}}$ (nm) | $\mathcal{V}$ ($10^{-19}$m$^3$) | $R_T$ ($\Omega$) | $\rho$ ($\Omega\mu$m$^2$) |
|---|---|---|---|---|---|---|---|
| A | 2.5 | 5 | 18 | 27 | 3.08 | 277 | 139 |
| B | 2.5 | 10 | 18 | 27 | 3.08 | 232 | 116 |
| C | 5 | 2.5 | 18 | 27 | 3.08 | 452 | 226 |
| D | 5 | 5 | 18 | 27 | 3.08 | 350 | 175 |
| E | 5 | 10 | 18 | 27 | 3.08 | 320 | 160 |
| F | 10 | 2.5 | 18 | 27 | 3.08 | 370 | 185 |
| G | 10 | 5 | 18 | 27 | 3.08 | 260 | 130 |
| H | 10 | 10 | 18 | 27 | 3.08 | 294 | 147 |
| J (OGP) | 5 | 2.5 | 18 | 27 | 3.08 | 230 | 115 |
| K (NGP) | 5 | 2.5 | 18 | 27 | 3.08 | 230 | 115 |
| L (OGP) | 10 | 10 | 30 | 30 | 1.96 | 650 | 285 |
| M (NGP) | 10 | 10 | 30 | 30 | 1.96 | 600 | 255 |
| O (DT) | 10 | 10 | 25 | 20 | 3.44 | 165 | 0 |
| P (DT) | 10 | 10 | 25 | 20 | 3.44 | 186 | 0 |
| Q (DT) | 10 | 10 | 50 | 50 | 3.44 | 182 | 0 |
| R (DT) | 10 | 10 | 50 | 50 | 3.44 | 203 | 0 |
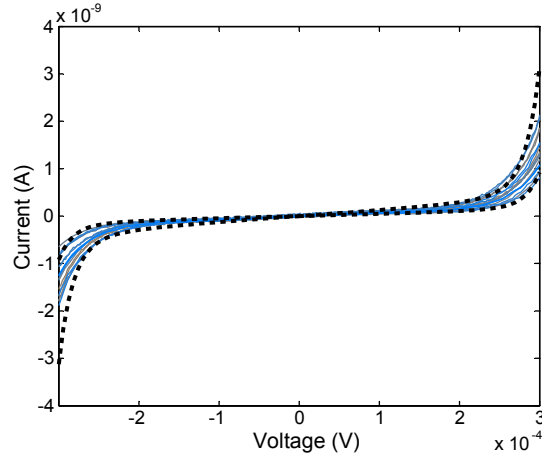


**Figure 4.4.** Measured current-voltate characteristics of samples A-H (thin lines) and two calculated I-V curves (solid lines) with parameters $R_T = 250\ \Omega$, $\gamma = 3 \times 10^{-4}$, $\Delta = 194\ \mu$eV for one and $R_T = 450\ \Omega$, $\gamma = 2 \times 10^{-4}$, $\Delta = 200\ \mu$eV for the other. At this low bias regime, we can neglect the thermal effects, as well as series resistance of our experimental wiring in this 2-point measurement. All the data falls in between the two calculated curves demonstrating that the variation in $\Delta$ and $\gamma$ is small between samples.

| | A | B |
|---|---|---|
| Best | 10 | 10 |
| | 5 | 5 |
| | 5 | 10 |
| | 2.5 | 10 |
| | 2.5 | 5 |
| | 5 | 2.5 |
| | 10 | 5 |
| Worst | 10 | 2.5 |

**Figure 4.5.** The $\delta T/\delta T_{\mathrm{opt}}$ ratio for samples A-H. Inset shows the data on larger range. The table shows the dimensions of samples from best to worst according to (a) at $T_{\mathrm{bath}} = 0.3$ K. The samples where dimension A was larger than dimension B perform worst. Otherwise the trend follows first dimension A and then dimension B. The differences are however small.

position dependent temperatures: the temperature of the cooled normal metal volume $T_n$, the effective temperature of the quasiparticle distribution in superconductor $T_s$ and the temperature of the normal metal acting as a quasiparticle trap $T_{\mathrm{trap}}$. In the normal metals relaxation was assumed to be from the electron system to the phonon system (see Eq. (2.36)) and in the superconductor the quasiparticle relaxation to the trap. Calculations were done to compare the relaxation to the trap to the quasiparticle-phonon coupling (recombination), showing the latter to be negligible with any tunnelling barrier resistivity and/or thickness of the superconducting film relevant to our experimental conditions.

In Fig. 4.6 (a) the resulting cooling curves from the 2D FEM simulations are shown. On the first sight there seems to be a large discrepancy between the simulations and experiments, the experiment being considerably worse than simulations would predict. We speculate this discrepancy to be due to local phonon heating. In the simulations it is assumed that the phonons stay all the time at the same temperature as the cryostat. As our insulating substrate is not a good heat conductor, this assumptions might not be a good one, especially with the relatively transparent junctions used in the experiment. We account this in our model phenomenologically by letting the phonon temperature $T_p$ at the normal metals (cooled volume and quasiparticle trap) to be a free (fitting) parameter. In Fig. 4.6 (b) the phonon temperatures required to reproduce the experiment are shown as a function of bath temperature. The required phonon temperatures are remarkably similar for all samples.
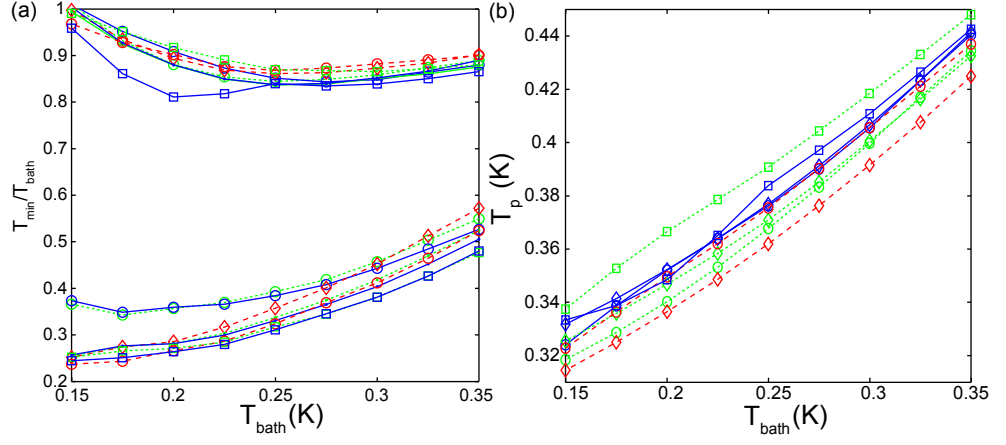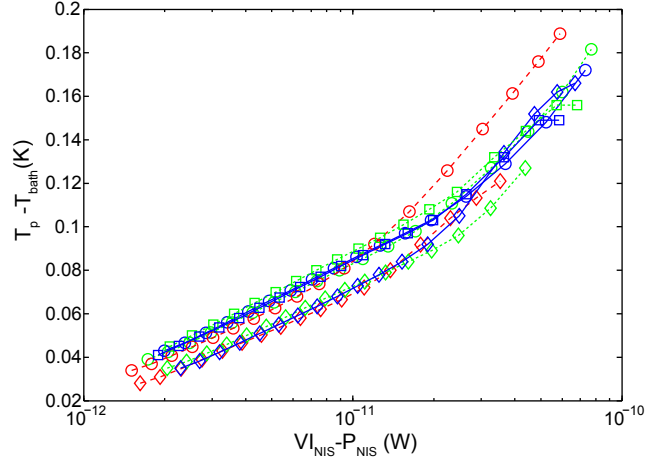
**Figure 4.6.** Experiment compared with 2D FEM simulations. **(a)** The lower lines show the result of simulations assuming $T_p = T_{\text{bath}}$, upper lines show the fit to the experimental results with varying $T_p$. **(b)** Phonon temperatures $T_p$ required to reproduce experimental results with the simulations.

In Fig. 4.6 the achieved cooling points are fitted without considering whether the optimal cooling voltages match in the simulations and the experiments. In most cases they do not. Another, maybe more appropriate way is to match the whole $V - T_e$ curve by letting the $T_p$ to be a voltage dependent fitting parameter. This was also done for the same dataset at one bath temperature around 0.3 K (the exact temperature varied a bit between measurements). In Fig. 4.7 we show the resulting $T_p$ from this kind of fitting procedure as a function of the input power to the superconductor. It is notable that they all seem to follow quite the same curve as would be expected on a similar substrate.

The results shown in Fig. 4.7 suggest also a physical explanation for the phenomenological $\beta$ parameter described in Sec. 4.2. Our results indicate that the $\beta$ generally observed (1-10 %) cannot be explained with just the excess quasiparticle population in the superconductor. This overheating of the superconductor would lead to only very small corrections to the cooling power with our relatively transparent ($\rho \approx 100\ \Omega\mu\text{m}^2$) quasiparticle traps. Local phonon heating resulting from quasiparticle energy dissipation in the trap, however, provides a straightforward explanation for the large $\beta$.

We tested the local phonon heating hypothesis experimentally by varying the heat conductivity of our substrate. If the assumption is correct, then changing the heat conductivity of the substrate should have an effect to the performance of the coolers. To vary the heat conductivity of the substrate we introduce a so-called "ground plane" (GP). By this we mean a copper layer deposited first onto the substrate and then covered

**Figure 4.7.** Experiment compared with 2D FEM simulations. The cooling curve on whole voltage bias range is fitted to experiment by allowing $T_p$ to vary with voltage. The plot shows the $T_p - T_{\mathrm{bath}}$ that reproduced the experimental curves as a function of the power into the superconducting lead. The two bunches are a result of slightly differing bath temperatures (295 mK vs. 315 mK) between the experiments. Otherwise the data seem to be following one curve, as would be expected on a similar substrate.

with relatively thin insulating layer (few tens of nanometers) before the depositions of the actual cooler sample. The copper layer is a good heat conductor and hence increases the heat conductivity of the substrate. In Fig. 4.8, we present the same figure as in Fig. 4.5 from samples C,H and now including also samples J,K from Table 4.1 (which is the same geometry as sample C but on a ground plane chip) and L,M (which is the same geometry as sample H but on a ground plane chip). The ground plane was not made continuous but copper was deposited as stripes. The samples J,L lie directly on top of the ground plane copper layer but samples K,M are in between the stripes few hundred micrometers off the copper layer.

The improvement in samples made on the ground plane chip is readily observed. It is notable that the improvement is considerably larger than the differences between the samples A-H, showing that properties of the substrate are indeed playing a role in NIS cooling. As it was recently reported [79] the substrate properties do not affect the $e - p$ coupling in metals. The remaining explanation is then heat conductivity through the substrate. Explanation for the fact that the samples that do not lie directly on the ground plane would seem to perform better is not totally clear. Our FEM modelling suggests there is an optimum thermal conductivity (compromise between removing excess heat efficiently and coupling it too strongly to the island) and this might play a role here.

Although introducing the ground plane improved the performance, the $\delta T/\delta T_{\mathrm{opt}}$ ratio is still nevertheless considerably below 1, even below 0.5.
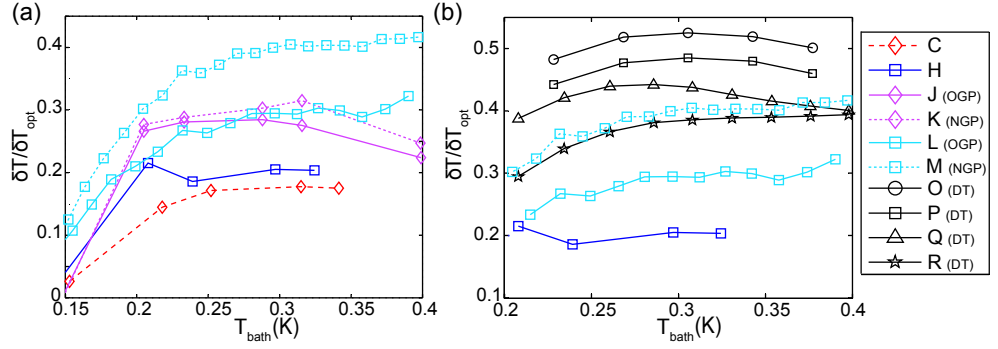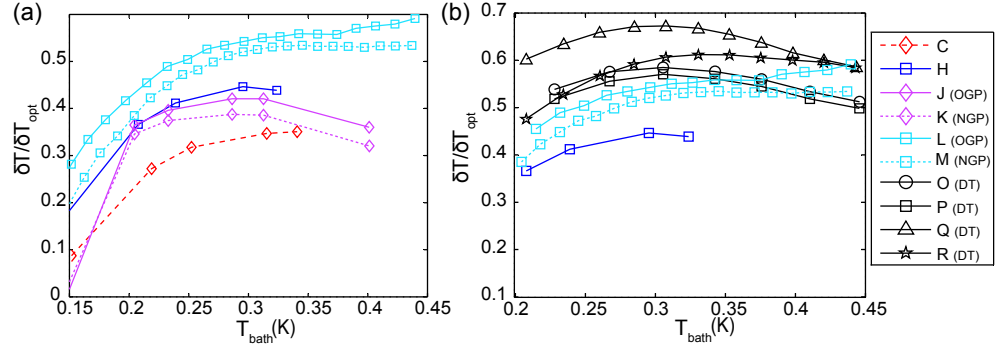
**Figure 4.8. (a)** The $\delta T/\delta T_{\rm opt}$ ratio of samples C,J,K (5-2.5 geometry on different substrates) and H,L,M (10-10 geometry on different substrates). The samples on ground plane chip perform significantly better than the similar samples without ground plane demonstrating that substrate properties play a crucial role in the efficiency of NIS coolers. **(b)** The $\delta T/\delta T_{\rm opt}$ ratio of samples H,L,M (same as in (a)) and samples O-R (with direct trap). In the direct trap samples performance is enhanced, although there is some variance in results.

To get closer to unity we have to still improve the thermalisation. One option for this is to use an additional normal metal layer, directly in contact with the Al layer to act as direct quasiparticle trap and heat conductor. We will call this kind of structure a "direct trap" (DT) structure. As now there is no tunnel barrier but only a highly transparent superconductor-normal metal (SN) interface, the thermalisation of quasiparticles to the normal metal temperature should be orders of magnitude faster. In Fig. 4.8 also the $\delta T/\delta T_{\rm opt}$ ratio of the direct trap samples (O-R) is presented. Here the geometry is the same as in H and the other curves from H geometry samples are also presented for comparison. The improvement is again very clear.

### 4.3.1 Enhancement of quasiparticle relaxation in small magnetic fields

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 or film 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

**Figure 4.9. (a)** The $\delta T/\delta T_{\mathrm{opt}}$ ratio of samples C,J,K (5-2.5 geometry on different substrates) and H,L,M (10-10 geometry on different substrates) at the optimum magnetic field. **(b)** The $\delta T/\delta T_{\mathrm{opt}}$ ratio of samples H,L,M (same as in (a)) and samples O-R (with direct trap) at the optimum magnetic field.

not 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) [80].

It was reported in Publication V that small magnetic fields can enhance performance of NIS junction coolers. This was attributed to vortices entering the superconducting lead acting as quasiparticle traps [81]. The vortices create areas where the superconducting gap is locally suppressed. Because of this suppression, the coupling between the quasiparticle system and the phonon system is enhanced. This leads to enhancement of the cooling effect as long as the vortices are not created in close proximity to the junction. When, at large enough magnetic field, vortices start to be created also in the junction area the cooler performance degrades.

This effect was also tested in our batch of samples. Improvement in cooling was invariably seen in small enough magnetic fields but the optimum field depended on the lead geometry as would be expected from the theory presented in V. Figure 4.9 presents the same plots as in Fig. 4.8 but now measured in the optimal magnetic field for each sample (this field is different for different samples but is always less than 5 G). The performance is enhanced in all samples but the amount of enhancement varies between samples. It is notable from Fig. 4.9 (b) that at the optimum field all the direct trap and ground plane samples seem to be bunching together. This might indicate the extreme limit that can be reached just by improving the quasiparticle thermalisation in our samples.

In summary, the experiments performed verify that the dissipation of the extracted heat is the main cause for sub-optimal performance of high power NIS coolers. This is not, however, only because increasing quasi-

particle population in the superconductor, as this effect can be largely alleviated by quasiparticle traps. Another crucial factor is heating of the phonon system which can then couple back to the cooled electron system. This gives a natural explanation to the often used backflow parameter $\beta$. We verified that by varying the heat conductivity of the substrate we could improve the cooling performance, showing that the substrate phonons are playing a role. The performance of the coolers can be enhanced further by introducing quasiparticle traps in a transparent contact with the superconducting layer, and with small magnetic fields.

# 5. Electron-phonon coupling and tunnel junction cooling in silicon

The basic principles of NIS cooling (presented in Chap. 2) apply also if the normal metal island is replaced by a heavily doped semiconductor. The superconductor - semiconductor (S-Sm) cooler presents some benefits compared to NIS cooler: (i) The electron-phonon coupling strength is generally weaker in semiconductors than in 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 semiconductor. 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 subgap currents are typically observed in S-Sm junctions leading to non-ideal cooler performance. Also typical junction resistivities are high leading to modest cooling powers.

The cooling effect in S-Sm structures was first presented in [82] and extended in [83, 84]. In [82], a cooling power of roughly 0.5 pW was achieved with two 5x18 $\mu$m$^2$ junctions having $R_T$ of 800 $\Omega$ (total for the two junctions in series). This lead to 30 % drop in temperature from 175 mK because of the small $e - p$ coupling. The doping level of the n$^+$ silicon was $4 \times 10^{19}$ cm$^{-3}$. In [83, 84], 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 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. How-

ever, 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 mK) and increasing doping to above $1 \times 10^{20}$ 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 Eq. (2.12). 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 this Chapter we discuss the electron-phonon coupling in silicon and how the coupling can be modified by inducing strain to the silicon film. This can then be used to enhance the cooling effect in S-Sm coolers. This Chapter is related to publications III and VI.

## 5.1  Electron-phonon coupling in Si and effects of strain

The correct form of electron-phonon coupling in silicon has been a topical question recently [85, 86, 30]. In highly doped silicon the effects of disorder and screening are important and the many-valley character of the electron system can play a crucial role. However, for clarity we will discuss first the pure, weak screening limit as in this limit in direct gap semiconductors the form of electron-phonon coupling is the same as for metals (Eq. (2.22)) with a different deformation potential constant $\Xi$. In semiconductors $\Xi$ is a separately measurable quantity as it describes simply how much the bottom of the conduction band and the top of the valence band will move in response to applied stress. This can be measured, for example, with photon absorption/emission experiments while applying mechanical stress. Generally in semiconductors only a small part of the k-space is occupied and it is a good approximation to assume that all important electronic states are very close to the minimum of the conduction band. Then the k dependency of $\Xi$ can be neglected from the outset and as extensive approximations as in the metal case are not needed.

In silicon, however, the situation is more complicated as it is a indirect gap semiconductor and the bottom of the conduction band does not lie at the center of the Brilloun zone (k = 0). Using the notations from crystallography, the bottom is instead located at a point (0.83 times the distance to the zone edge) at the [1 0 0] axis or equivalently on [0 1 0], [0 0 1] and the corresponding negative axes, making it six-fold degenerate. For describing the coupling between stress and the movement of

66

these band edges two coupling constants are needed: one describing the dilatational (uniform) stress (commonly notated $\Xi_d$) and other describing uniaxial stress that will affect two of the band edges (commonly referred as valleys) differently than the other two ($\Xi_u$). The deformation potential Hamiltonian will the be written

$$\mathcal{H}_{e-i} = \Xi_d(\epsilon_{xx} + \epsilon_{yy} + \epsilon_{zz}) + \Xi_u\epsilon_{xx} = \Xi_d(\nabla \cdot \mathbf{u}) + \Xi_u\frac{\partial \mathbf{u}_x}{\partial x}, \qquad (5.1)$$

for the valley at [1 0 0] and similarly for other valleys. We have marked $\epsilon_{ii}$ to describe the diagonal terms in the stress tensor Eq. (2.20). To calculate the total effective coupling constant one has to take the average of this over the solid angle (the analytical results were first calculated in [26]).

Everything discussed so far would have effect only in the $\Xi$ entering the $e - p$ calculation as compared to the calculation done for metals. Differences in the temperature dependence however arise when the effects of disorder, screening and the valley degeneracy are included. In order to include these to the $e - p$ calculation we introduce the electronic response function $\chi(\mathbf{k}, \omega)$, also known as polarization function. We use the version derived with random phase approximation (RPA) (for details see for example [87]). Within RPA a perturbing potential $U(\mathbf{q})$ will induce a change in the electron density $\varrho$

$$\delta\varrho(\mathbf{q}, \omega) \quad = \quad \sum_p c_{p+q}^\dagger c_p = U(q)\chi(\mathbf{q}, \omega) \qquad (5.2)$$

$$\chi(\mathbf{q}, \omega) \quad = \quad \frac{1}{\mathcal{V}}\lim_{\alpha \to 0}\sum_p \frac{f(E_p) - f(E_{p+q})}{E_p - E_{p+q} - \hbar\omega + i\hbar\alpha}, \qquad (5.3)$$

where $f$ is the fermi-dirac distribution function and $E_p$ the energy of an electron with wavevector $p$. At zero temperature in 3 dimensions $\chi$ can be evaluated analytically (for pure systems) and gives [87]

$$\chi_0(\mathbf{q}, \omega) \quad = \quad \frac{-N_F k_F}{4q}\left[H(\frac{\omega}{qv_F} + \frac{q}{2k_F}) - H(\frac{\omega}{qv_F} - \frac{q}{2k_F})\right] \qquad (5.4)$$

$$H(x) \quad = \quad 2x + (x^2 - 1)\ln\left(\frac{x - 1}{x + 1}\right), \qquad (5.5)$$

which at the limit $q \ll k_F$ becomes

$$\chi_0(\mathbf{q}, \omega) \approx -N_F\left[1 + i\frac{\omega}{qv_F}\right]. \qquad (5.6)$$

This is the form used in Sec. 2.4.1.

With the help of the response function, the effects of screening can also be straightforwardly included. The effective dielectric constant is

$$\epsilon(\mathbf{q}, \omega) = \epsilon_0\left(1 + \frac{e^2}{\epsilon_0 q^2}\chi(\mathbf{q}, \omega)\right). \qquad (5.7)$$

At the $q \ll k_F$, zero temperature and zero frequency $\omega = 0$ limit, this reduces to the familiar Thomas-Fermi screening. When screening is important (see below) this has to be included into the calculation by replacing $\chi$ with $\chi/\epsilon$.

In highly doped semiconductors also the effects of disorder must be taken into account as the electrons mean free path is typically very short. This dirty or diffusive limit (as opposed to the pure limit discussed above), is defined by the product of the (thermal) phonon wavelength $q$ and the mean free path of electrons in the material $l$. Material is said to be dirty when $ql \ll 1$. In order to include the effects of disorder, a phenomenological relaxation time $\tau$ has to be introduced and with this the response function will be modified to [88]

$$\chi_\tau(\mathbf{q}, \omega) = \frac{\chi_0(\mathbf{q}, \omega + i\tau^{-1})}{1 - \frac{i\tau^{-1}}{\omega + i\tau^{-1}}\left[1 - \chi_0(\mathbf{q}, \omega + i\tau^{-1})/\chi_0(\mathbf{q}, 0)\right]}, \qquad (5.8)$$

which at the diffusive limit ($\omega\tau$, $qv_F\tau \ll 1$) will fortunately reduce to a simple form

$$\chi_\tau(\mathbf{q}, \omega) = -N_F \frac{iD_0 q^2}{\omega + iD_0 q^2}, \qquad (5.9)$$

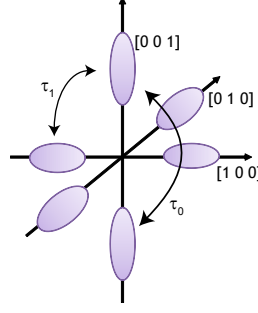where $D_0 = v_F^2 \tau/3$ is the diffusion constant.

As shown in [30], using then the form

$$P_{e-p} = \sum_{\mathbf{q}} \frac{\hbar\omega_q}{2\pi} - 2\mathcal{V}\text{Im}\{\chi(\mathbf{q}, \omega)\}\mathcal{M}_0^2 q[n_q^{Te} - n_q^{Tp}], \qquad (5.10)$$

with the pure or diffusive response function and including or excluding the dielectric function we can derive the $e-p$ coupling at the pure/diffusive strong/weak screening limit, respectively. The results have been presented in [30] and are exactly the same as the ones calculated with Green's function methods in [85]. As mentioned before, at the pure weak screening limit both also reproduce the results from [29]. The resulting powerlaws without screening are $T^5$ and $T^4$ at the pure and diffusive limit, respectively, and $T^9$ and $T^8$ when including screening.

The role of valley degeneracy in $e - p$ coupling has been also studied recently [86, 30]. It was shown that as the (uniaxial) phonons can lift the degeneracy of the valleys one should not use one but two different relaxation times in the calculation at the diffusive limit, one for scattering between the valleys on different axes ($\tau_1$) and other inside one valley (or between the valleys on the same axis) ($\tau_0$), i.e., $\tau^{-1} = \tau_0^{-1} + (L - 1)\tau_1^{-1}$, where $L$ is the number of valleys (see Fig. 5.1). This then leads to a response function of the form [30]

$$\text{Im}(\chi_A) = -N_F \frac{(\Gamma_{iv} + D_0 q^2)\omega}{(\Gamma_{iv} + D_0 q^2)^2 + \omega^2}, \qquad (5.11)$$

**Figure 5.1.** Schematic of the energy valleys in silicon in momentum space. Also shown are the two different relaxation times for $e - e$ scattering assumed in Eq. (5.11)

.

where $D_0$ is the diffusion constant and $\Gamma_{iv}$ is the total intervalley scattering rate $((L/2)\tau_1^{-1})$. Note that cases where an electron would be directly transferred by a phonon to a different valley are not considered here as this would require a large momentum phonon. Rather the effects of phonon lifting the degeneracy of valleys (to a singlet and doublet) are considered by assuming different $e - e$ scattering rates between the valleys that are not degenerate. The single-valley diffusive response function can be restored by simply setting $\Gamma_{iv} = 0$ above. The limit of interest here is $\omega \ll D_0 q^2 \ll \Gamma_{iv}$. The first limit is quite general [85] and the second should apply if the intervalley scattering rate is comparable to intravalley scattering [86]. Then

$$\text{Im}(\chi_A) = -N_F \frac{\omega}{\Gamma_{iv}}. \tag{5.12}$$

Putting this to Eq. (5.10)

$$P_{e-p} \;\; = \;\; \sum_{\mathbf{q}} \mathcal{V} \frac{N_F \hbar c_l}{\rho \mathcal{V} \Gamma_{iv}} \Xi^2 q^3 [n_q^{Te} - n_q^{Tp}], \tag{5.13}$$

which after integration results in

$$P_{e-p} = \frac{4\pi^4}{63} \frac{N_F k_B^6}{\rho \mathcal{V} \Gamma_{iv} \hbar^5 c_l^5} \Xi^2 \mathcal{V} (T_e^6 - T_p^6). \tag{5.14}$$

To get the total heat flow one still has to sum this over the phonon modes (longitudinal and two transversal) using correct speed of sound and the deformation potential constants. This is the same result presented in [86], where it was calculated using Boltzmann transport equations and was shown to fit experimental data very well. It is worth noting that none of the limits without the intervalley scattering produce the correct $T^6$ powerlaw at 3D. In 2D the diffusive strongly screened limit produces $T^6$ powerlaw and this was argued in [85] to be the correct limit of experiments performed before [89]. However, the prefactor of that term is orders of magnitude smaller than the one found in experiments.

The above discussed intervalley scattering should exist in silicon alongside the intravalley component. However, in [30] it was argued that the intravalley scattering is at highly doped low temperature limit strongly screened and this is the reason that the intervalley case dominates in our experiments. (If both would be unscreened the intravalley scattering would be stronger with powerlaw $T^4$.) The screening is included into the calculations with the dielectric function $\epsilon(\mathbf{q}, \omega)$ as described above. We replace $\mathrm{Im}\{\chi(\mathbf{q}, \omega)\}$ with $\mathrm{Im}\{\chi(\mathbf{q}, \omega)/\epsilon(\mathbf{q}, \omega)\}$ in Eq. (5.10) and set $\Gamma_{iv} = 0$. This leads to

$$P_{e-p} = \sum_{\mathbf{q}} \mathcal{V} \frac{N_F \hbar c_l}{\rho \mathcal{V} D_0 q^2} \frac{1}{(1 + \kappa^2/q^2)^2} \Xi^2 q^3 [n_q^{Te} - n_q^{Tp}], \qquad (5.15)$$

where $\kappa^2 = e^2 N_F/\epsilon_0$ is the screening wavevector. At the strong screening limit $\kappa \gg q_T$ and we can keep only the highest power of the screening term

$$
\begin{aligned}
P_{e-p} &= \sum_{\mathbf{q}} \mathcal{V} \frac{N_F \hbar c_l}{\rho \mathcal{V} D_0 \kappa^4} \Xi^2 q^5 [n_q^{Te} - n_q^{Tp}] \\
&= \frac{4\pi^6}{5} \frac{N_F k_B^8}{\rho \mathcal{V} v_f l \kappa^4 \hbar^7 c_l^7} \Xi^2 \mathcal{V}(T_e^8 - T_p^8), \qquad (5.16)
\end{aligned}
$$

where we have used $D_0 = v_F l/3$. This is by a factor of $\frac{T^2}{l\kappa^4} \frac{\Gamma_{iv}}{\tau^{-1}}$ smaller than Eq. (5.14). Assuming the intervalley scattering rate is the same order of magnitude with the intravalley rate, this component is negligible for the total $e - p$ coupling.

In Publication III the effects of lattice mismatch induced strain to the $e - p$ coupling were experimentally studied. As was argued already above, strain is known to lift the degeneracy of the valleys along the axis the strain is applied. When a thin layer of Si is grown on a lattice mismatched substrate there will be biaxial permanent stress applied on the layer, see Fig. 5.2 (a). If this stress is strong enough the valleys on two of the axes become depopulated (Fig. 5.2 (b)). This will then eliminate the unscreened heat conduction channel (between valleys whose degeneracy the phonon could lift) and the single-valley strongly screened result should be recovered Eq. (5.16). In the experiment, it was found that the $e - p$ coupling was indeed reduced in the strained sample but not as much as would be predicted according to Eqs. (5.14) and (5.16). The reasons for this discrepancy are unclear. Nevertheless, the reduced $e - p$ coupling in strained Si was used to demonstrate enhanced cooling effect in Si in Publication VI.
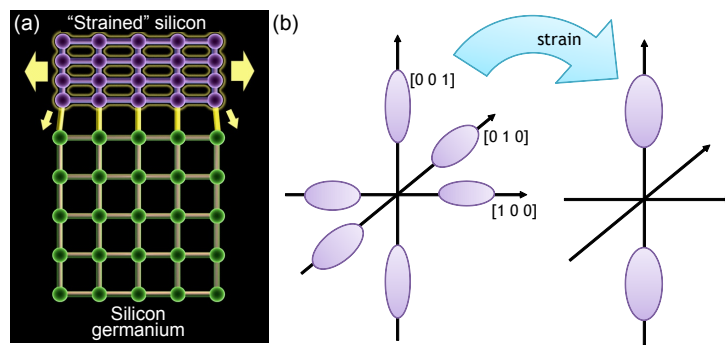
**Figure 5.2. (a)** Diagram showing how the lattice mismatch between the underlying substrate (silicon germanium) and silicon produces stress to the silicon layer. **(b)** Schematic of how the stress should affect the population of the energy valleys. The valleys on the strain axes become depopulated as they are lifted higher in energy.

# Bibliography

[1] Giazotto, F., Heikkilä, T. T., Luukanen, A., Savin, A. M. & Pekola, J. P. Opportunities for mesoscopics in thermometry and refrigeration: Physics and applications. *Rev. Mod. Phys.* **78**, 217–274 (2006).

[2] van Wees, B. J., van Houten, H., Beenakker, C. W. J., Williamson, J. G., Kouwenhoven, L. P., van der Marel, D. & Foxon, C. T. Quantized conductance of point contacts in a two-dimensional electron gas. *Phys. Rev. Lett.* **60**, 848–850 (1988).

[3] Schwab, K., Henriksen, E. A., Worlock, J. M. & Roukes, M. L. Measurement of the quantum of thermal conductance. *Nature* **404**, 974–977 (2000).

[4] Reichl, L. E. *A modern course in Statistical Phys* (John Wiley & Sons, 1998).

[5] Onnes, H. K. The resistance of pure mercury at helium temperatures. *Comm. Phys. Lab. Univ. Leiden.* **120** (1911).

[6] Bardeen, J., Cooper, L. N. & Schrieffer, J. R. Theory of superconductivity. *Phys. Rev.* **108**, 1175– (1957).

[7] Franssila, S. *Introduction to microfabrication* (John Wiley & Sons, 2010), 2nd edn.

[8] Pobell, F. *Matter and Methods at Low Temperatures* (Springer-Verlag Berlin Heidelberg, 2007), 3rd edn.

[9] Nahum, M., Eiles, T. M. & Martinis, J. M. Electronic microrefrigerator based on a normal-insulator-superconductor tunnel junction. *Applied Physics Letters* **65**, 3123–3125 (1994).

[10] Leivo, M. M., Pekola, J. P. & Averin, D. V. Efficient peltier refrigeration by a pair of normal metal/insulator/superconductor junctions. *Applied Physics Letters* **68**, 1996–1998 (1996).

[11] Prance, J. R., Smith, C. G., Griffiths, J. P., Chorley, S. J., Anderson, D., Jones, G. A. C., Farrer, I. & Ritchie, D. A. Electronic refrigeration of a two-dimensional electron gas. *Phys. Rev. Lett.* **102**, 146602 (2009).

[12] Anghel, D. V. & Pekola, J. P. Noise in refrigerating tunnel junctions and in microbolometers. *Journal of Low Temperature Physics* **123**, 197–218 (2001). 10.1023/A:1017589828739.

[13] Dynes, R. C., Narayanamurti, V. & Garno, J. P. Direct measurement of quasiparticle-lifetime broadening in a strong-coupled superconductor. *Phys. Rev. Lett.* **41**, 1509–1512 (1978).

[14] Pekola, J. P., Maisi, V. F., Kafanov, S., Chekurov, N., Kemppinen, A., Pashkin, Y. A., Saira, O.-P., Möttönen, M. & Tsai, J. S. Environment-assisted tunneling as an origin of the dynes density of states. *Phys. Rev. Lett.* **105**, 026803 (2010).

[15] Andreev, A. F. The thermal conductivity of the intermediate state in superconductors. *SOVIET PHYSICS JETP-USSR* **19**, 1228 (1964).

[16] Bardas, A. & Averin, D. Peltier effect in normal-metal - superconductor microcontacts. *Phys. Rev. B* **52**, 12873– (1995).

[17] Averin, D. V. & Pekola, J. P. Nonadiabatic charge pumping in a hybrid single-electron transistor. *Phys. Rev. Lett.* **101**, 066801 (2008).

[18] Hekking, F. W. J. & Nazarov, Y. V. Interference of two electrons entering a superconductor. *Phys. Rev. Lett.* **71**, 1625–1628 (1993).

[19] Hekking, F. W. J. & Nazarov, Y. V. Subgap conductivity of a superconductor–normal-metal tunnel interface. *Phys. Rev. B* **49**, 6847–6852 (1994).

[20] Vasenko, A. S., Bezuglyi, E. V., Courtois, H. & Hekking, F. W. J. Electron cooling by diffusive normal metal - superconductor tunnel junctions. *Phys. Rev. B* **81**, 094513 (2010).

[21] Rajauria, S., Gandit, P., Fournier, T., Hekking, F. W. J., Pannetier, B. & Courtois, H. Andreev current-induced dissipation in a hybrid superconducting tunnel junction. *Phys. Rev. Lett.* **100**, 207002 (2008).

[22] Lowell, P., O'Neil, G., Underwood, J. & Ullom, J. Andreev reflections in micrometer-scale normal metal-insulator-superconductor tunnel junctions. *Journal of Low Temperature Physics* 1–6 (2011). 10.1007/s10909-011-0425-2.

[23] Pothier, H., Guéron, S., Birge, N. O., Esteve, D. & Devoret, M. H. Energy distribution function of quasiparticles in mesoscopic wires. *Phys. Rev. Lett.* **79**, 3490–3493 (1997).

[24] Pekola, J. P., Heikkilä, T. T., Savin, A. M., Flyktman, J. T., Giazotto, F. & Hekking, F. W. J. Limitations in cooling electrons using normal-metal-superconductor tunnel junctions. *Phys. Rev. Lett.* **92**, 056804 (2004).

[25] Bardeen, J. & Shockley, W. Deformation potentials and mobilities in non-polar crystals. *Phys. Rev.* **80**, 72–80 (1950).

[26] Herring, C. & Vogt, E. Transport and deformation-potential theory for many-valley semiconductors with anisotropic scattering. *Phys. Rev.* **101**, 944–961 (1956).

[27] Ziman, J. M. *Principles of the theory of solids* (Cambridge University Press, 1964).

[28] Gantmakher, V. & Levinson, Y. *Carrier scattering in metals and semiconductors* (Elsevier Science Publishers, 1987).

[29] Wellstood, F. C., Urbina, C. & Clarke, J. Hot-electron effects in metals. *Phys. Rev. B* **49**, 5942–5955 (1994).

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

[31] Ekinci, K. L. & Roukes, M. L. Nanoelectromechanical systems. *Review of Scientific Instruments* **76**, 061101 (2005).

[32] Braginsky, V. B. & Khalili, F. Y. *Quantum Measurement* (Cambridge University Press, 1992).

[33] LaHaye, M. D., Buu, O., Camarota, B. & Schwab, K. C. Approaching the quantum limit of a nanomechanical resonator. *Science* **304**, 74–77 (2004).

[34] Schwab, K. C. & Roukes, M. L. Putting mechanics into quantum mechanics. *Physics Today* **58**, 36–42 (2005).

[35] O'Connell, A. D. *et al.* Quantum ground state and single-phonon control of a mechanical resonator. *Nature* **464**, 697–703 (2010).

[36] Teufel, J. D., Donner, T., Li, D., Harlow, J. W., Allman, M. S., Cicak, K., Sirois, A. J., Whittaker, J. D., Lehnert, K. W. & Simmonds, R. W. Sideband cooling of micromechanical motion to the quantum ground state. *Nature* **475**, 359–363 (2011).

[37] Chan, J., Alegre, T. P. M., Safavi-Naeini, A. H., Hill, J. T., Krause, A., Groblacher, S., Aspelmeyer, M. & Painter, O. Laser cooling of a nanomechanical oscillator into its quantum ground state. *Nature* **478**, 89–92 (2011).

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

[39] Paraoanu, G. S. & Halvari, A. M. Suspended single-electron transistors: Fabrication and measurement. *Appl. Phys. Lett.* **86**, 093101–3 (2005).

[40] Koppinen, P. J. & Maasilta, I. J. Phonon cooling of nanomechanical beams with tunnel junctions. *Phys. Rev. Lett.* **102**, 165502 (2009).

[41] Cross, M. C. & Lifshitz, R. Elastic wave transmission at an abrupt junction in a thin plate with application to heat transport and vibrations in mesoscopic systems. *Phys. Rev. B* **64**, 085324– (2001).

[42] Manninen, A. J., Leivo, M. M. & Pekola, J. P. Refrigeration of a dielectric membrane by superconductor/insulator/normal-metal/insulator/superconductor tunneling. *Applied Physics Letters* **70**, 1885–1887 (1997).

[43] Luukanen, A., Leivo, M. M., Suoknuuti, J. K., Manninen, A. J. & Pekola, J. P. On-chip refrigeration by evaporation of hot electrons at sub-kelvin temperatures. *Journal of Low Temperature Physics* **120**, 281–290 (2000). 10.1023/A:1004693929689.

[44] Clark, A. M., Miller, N. A., Williams, A., Ruggiero, S. T., Hilton, G. C., Vale, L. R., Beall, J. A., Irwin, K. D. & Ullom, J. N. Cooling of bulk material by electron-tunneling refrigerators. *Applied Physics Letters* **86**, 173508 (2005).

[45] Miller, N. A., O'Neil, G. C., Beall, J. A., Hilton, G. C., Irwin, K. D., Schmidt, D. R., Vale, L. R. & Ullom, J. N. High resolution x-ray transition-edge sensor cooled by tunnel junction refrigerators. *Applied Physics Letters* **92**, 163501 (2008).

[46] Acknowledgements to Tero Heikkilä for the first model.

[47] Karvonen, J. T. & Maasilta, I. J. Influence of phonon dimensionality on electron energy relaxation. *Phys. Rev. Lett.* **99**, 145503 (2007).

[48] Holmes, W., Gildemeister, J. M., Richards, P. L. & Kotsubo, V. Measurements of thermal transport in low stress silicon nitride films. *Applied Physics Letters* **72**, 2250–2252 (1998).

[49] Leivo, M. M. & Pekola, J. P. Thermal characteristics of silicon nitride membranes at sub-kelvin temperatures. *Applied Physics Letters* **72**, 1305–1307 (1998).

[50] Woodcraft, A. L., Sudiwala, R. V., Wakui, E., Bhatia, R. S., Bock, J. J. & Turner, A. D. Thermal conductance measurements of a silicon nitride membrane at low temperatures. *Physica B: Condensed Matter* **284-288, Part 2**, 1968 – 1969 (2000).

[51] Yefremenko, V. *et al.* Low temperature thermal transport in partially perforated silicon nitride membranes. *Applied Physics Letters* **94**, 183504 (2009).

[52] Pohl, R. O., Liu, X. & Thompson, E. Low-temperature thermal conductivity and acoustic attenuation in amorphous solids. *Rev. Mod. Phys.* **74**, 991–1013 (2002).

[53] Norcada inc. 4465 - 99 Street , Edmonton, AB T6E 5B6, Canada.

[54] Pekola, J. P., Manninen, A. J., Leivo, M. M., Arutyunov, K., Suoknuuti, J. K., Suppula, T. I. & Collaudin, B. Microrefrigeration by quasiparticle tunnelling in nis and sis junctions. *Physica B: Condensed Matter* **280**, 485 – 490 (2000).

[55] Heron, J. S., Fournier, T., Mingo, N. & Bourgeois, O. Mesoscopic size effects on the thermal conductance of silicon nanowire. *Nano Letters* **9**, 1861–1865 (2009).

[56] Heron, J.-S., Bera, C., Fournier, T., Mingo, N. & Bourgeois, O. Blocking phonons via nanoscale geometrical design. *Phys. Rev. B* **82**, 155458 (2010).

[57] Karvonen, J. *Thermal properties in low dimensional structures below 1K*. Ph.D. thesis, University of Jyväskylä, (2009).

[58] Most perforations presented in this thesis were done by Antti Peltonen.

[59] Saira, O.-P., Kemppinen, A., Maisi, V. F. & Pekola, J. P. Vanishing quasiparticle density in a hybrid al/cu/al single-electron transistor. *Phys. Rev. B* **85**, 012504 (2012).

[60] Bardeen, J., Rickayzen, G. & Tewordt, L. Theory of the thermal conductivity of superconductors. *Phys. Rev.* **113**, 982–994 (1959).

[61] Timofeev, A. V., Helle, M., Meschke, M., Möttönen, M. & Pekola, J. P. Electronic refrigeration at the quantum limit. *Phys. Rev. Lett.* **102**, 200801 (2009).

[62] Kaplan, S. B., Chi, C. C., Langenberg, D. N., Chang, J. J., Jafarey, S. & Scalapino, D. J. Quasiparticle and phonon lifetimes in superconductors. *Phys. Rev. B* **14**, 4854–4873 (1976).

[63] Timofeev, A. V., García, C. P., Kopnin, N. B., Savin, A. M., Meschke, M., Giazotto, F. & Pekola, J. P. Recombination-limited energy relaxation in a bardeen-cooper-schrieffer superconductor. *Phys. Rev. Lett.* **102**, 017003 (2009).

[64] Fisher, P. A., Ullom, J. N. & Nahum, M. High-power on-chip microrefrigerator based on a normal- metal/insulator/superconductor tunnel junction. *Applied Physics Letters* **74**, 2705–2707 (1999).

[65] Clark, A. M., Williams, A., Ruggiero, S. T., van den Berg, M. L. & Ullom, J. N. Practical electron-tunneling refrigerator. *Applied Physics Letters* **84**, 625–627 (2004).

[66] Ullom, J. N., Fisher, P. A. & Nahum, M. Energy-dependent quasiparticle group velocity in a superconductor. *Phys. Rev. B* **58**, 8225– (1998).

[67] Rajauria, S., Courtois, H. & Pannetier, B. Quasiparticle-diffusion-based heating in superconductor tunneling microcoolers. *Phys. Rev. B* **80**, 214521– (2009).

[68] Rajauria, S., Pascal, L. M. A., Gandit, P., Hekking, F. W. J., Pannetier, B. & Courtois, H. Efficiency of quasiparticle evacuation in superconducting devices. *Phys. Rev. B* **85**, 020505 (2012).

[69] O'Neil, G. C., Lowell, P. J., Underwood, J. M. & Ullom, J. N. Observations and modeling of large area normal-metal/insulator/superconductor refrigerator cooling from 300 mk to below 100 mk. *ArXiv e-prints* (2011). 1109.1273.

[70] Owen, C. S. & Scalapino, D. J. Superconducting state under the influence of external dynamic pair breaking. *Phys. Rev. Lett.* **28**, 1559–1561 (1972).

[71] Parker, W. H. Modified heating theory of nonequilibrium superconductors. *Phys. Rev. B* **12**, 3667–3672 (1975).

[72] Palmer, B. S., Sanchez, C. A., Naik, A., Manheimer, M. A., Schneiderman, J. F., Echternach, P. M. & Wellstood, F. C. Steady-state thermodynamics of nonequilibrium quasiparticles in a cooper-pair box. *Phys. Rev. B* **76**, 054501 (2007).

[73] Joyez, P., Lafarge, P., Filipe, A., Esteve, D. & Devoret, M. H. Observation of parity-induced suppression of josephson tunneling in the superconducting single electron transistor. *Phys. Rev. Lett.* **72**, 2458–2461 (1994).

[74] Ullom, J. N., Fisher, P. A. & Nahum, M. Measurements of quasiparticle thermalization in a normal metal. *Phys. Rev. B* **61**, 14839–14843 (2000).

[75] Pekola, J. P., Anghel, D. V., Suppula, T. I., Suoknuuti, J. K., Manninen, A. J. & Manninen, M. Trapping of quasiparticles of a nonequilibrium superconductor. *Applied Physics Letters* **76**, 2782–2784 (2000). Journal article.

[76] Court, N. A., Ferguson, A. J., Lutchyn, R. & Clark, R. G. Quantitative study of quasiparticle traps using the single-cooper-pair transistor. *Phys. Rev. B* **77**, 100501 (2008).

[77] Martinis, J. M., Ansmann, M. & Aumentado, J. Energy decay in superconducting josephson-junction qubits from nonequilibrium quasiparticle excitations. *Phys. Rev. Lett.* **103**, 097002 (2009).

[78] The sample fabrication and measurements presented in this Section were done in collaboration with Dr. Hung Nguyen and Dr. Thomas Aref.

[79] Underwood, J. M., Lowell, P. J., O'Neil, G. C. & Ullom, J. N. Insensitivity of sub-kelvin electron-phonon coupling to substrate properties. *Phys. Rev. Lett.* **107**, 255504 (2011).

[80] Stan, G., Field, S. B. & Martinis, J. M. Critical field for complete vortex expulsion from narrow superconducting strips. *Phys. Rev. Lett.* **92**, 097003 (2004).

[81] Ullom, J. N., Fisher, P. A. & Nahum, M. Magnetic field dependence of quasiparticle losses in a superconductor. *Applied Physics Letters* **73**, 2494–2496 (1998). Journal article.

[82] Savin, A. M., Prunnila, M., Kivinen, P. P., Pekola, J. P., Ahopelto, J. & Manninen, A. J. Efficient electronic cooling in heavily doped silicon by quasiparticle tunneling. *Applied Physics Letters* **79**, 1471–1473 (2001).

[83] Savin, A., Prunnila, M., Ahopelto, J., Kivinen, P., Törmä, P. & Pekola, J. Application of superconductor-semiconductor schottky barrier for electron cooling. *Physica B: Condensed Matter* **329-333**, 1481 – 1484 (2003). Proceedings of the 23rd International Conference on Low Temperature Physics.

[84] Savin, A., Pekola, J., Prunnila, M., Ahopelto, J. & Kivinen, P. Electronic cooling and hot electron effects in heavily doped silicononinsulator film. *Physica Scripta* **2004**, 57 (2004).

[85] Sergeev, A., Reizer, M. Y. & Mitin, V. Deformation electron-phonon coupling in disordered semiconductors and nanostructures. *Phys. Rev. Lett.* **94**, 136602 (2005).

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

[87] Mahan, G. D. *Condensed Matter in a Nutshell* (Princeton University Press, 2011).

[88] Mermin, N. D. Lindhard dielectric function in the relaxation-time approximation. *Phys. Rev. B* **1**, 2362–2363 (1970).

[89] Kivinen, P., Savin, A., Zgirski, M., Törmä, P., Pekola, J., Prunnila, M. & Ahopelto, J. Electron phonon heat transport and electronic thermal conductivity in heavily doped silicon-on-insulator film. *Journal of Applied Physics* **94**, 3201–3205 (2003).