TERAHERTZ DETECTORS BASED ON
SUPERCONDUCTING KINETIC INDUCTANCE

E.N. Grossman, D.G. McDonald, J.E. Sauvageau*

Abstract – The inductance of a superconducting stripline varies with the concentration of Cooper pairs in the superconductor. This inductance variation may be used as the basis for highly sensitive radiometers, bolometers, and heterodyne mixers. We describe recent progress on three kinetic inductance devices: a large-area, absolute radiometer intended for use in the NIST Low-Background Infrared Calibration Facility, and small, antenna-coupled devices used either in a bolometric mode as direct detectors, or in a non-equilibrium, “photoinductive” mode as heterodyne mixers. The photoinductive mixers are of particular interest because their frequency coverage starts at approximately the energy gap, $2\Delta$, and extends upward. The impedance matching concerns which make extension of SIS mixers to high frequencies so difficult in practice are greatly relaxed for photoinductors because they lack the large parasitic capacitances inherent in a junction-like geometry.

I. Introduction

Kinetic inductance is a property of superconducting thin films that has long been recognized as playing a major role in the propagation velocities and impedances of superconducting transmission lines. However, it was only fairly recently [1] that its potential usefulness in a host of other applications was recognized. As was emphasized in that original work, one of the main reasons kinetic inductance can serve as the basis for sensitive transducers of any kind is the fact that it is naturally compatible, both in terms of impedance matching and fabrication processes, with one of the most sensitive amplification devices known, namely the SQUID [2]. At NIST, we have lately begun a program to explore some of the applications of kinetic inductance devices in infrared and submillimeter wave detection. The goal of this presentation is to provide an overview of this work.

The (low-frequency) inductance per unit length of a superconducting stripline, when normalized to the ordinary, magnetic inductance, without field penetration, of a normal stripline with the

---

* The authors are with the Electromagnetic Technology Division of the National Institute of Standards and Technology, Boulder, CO 80303. This work is supported by the SDI Metrology Program, the Innovative Science and Technology Program of SDIO, and by NASA’s Office of Space Science and Applications. U. S. government publication, not subject to copyright.
same dimensions is given by [3]

\[
L \propto \left( 1 + \frac{\lambda_1}{d} \coth \frac{t_1}{\lambda_1} + \frac{\lambda_2}{d} \coth \frac{t_2}{\lambda_2} \right)
\] (1)

where \(t_1\) and \(t_2\) are the thicknesses of the two superconductors, \(d\) their separation, and \(\lambda_1\) and \(\lambda_2\) the two superconducting penetration depths. (The penetration depth is a characteristic length scale of a superconducting material, equal to the e-folding length over which the magnetic field decays due to Meissner effect at the surface of a bulk sample.) The reason far-i.r. radiation enters the picture is that the penetration depth is related to the density of superconducting pairs, \(N_s\),

\[
\lambda^2 = \frac{m}{\mu_0 e^2 N_s} \left( \frac{\xi_0}{l_e} \right)
\] (2)

and \(N_s\) can be altered by very small changes in temperature, or by direct application of pair-breaking radiation \((h\nu > 2\Delta(T))\), where \(\Delta\) is the temperature-dependent superconducting energy gap. In (2), \(m\) and \(e\) are the mass and charge of the electron, and \(\mu_0\) the permeability of free space. The term in parentheses is a correction to what would otherwise be just the conventional London penetration depth. It accounts for the effect of impure or otherwise "dirty" materials. \(\xi_0\) is the superconducting coherence length (roughly speaking, the "size" of a Cooper pair) and \(l_e\) the electronic mean free path. From (1) and (2) it is clear that, other things being equal, dirty materials, i.e. those with \(l_e\) small, will produce the greatest inductance changes for a given change in the pair density. As will be seen, this leads to our choice of metallic alloys such as NbTi and NbZr, in which \(l_e\) is comparable to the interatomic spacing, for the photoinductive elements in our devices.

The microscopic mechanism by which pair-breaking radiation is converted into changes in \(N_s\) is indicated schematically in Fig. 1. Both direct pairbreaking and electromagnetic absorption by quasiparticles lead to an effective temperature \(T^*\) of the quasiparticle population which can, in general, be different from the bulk temperature of the lattice. Energy is exchanged between the quasiparticle population and the lattice through inelastic electron-phonon scattering, the same interaction as is responsible for the electron pairing of (conventional, low-\(T_c\)) superconductivity in the first place. The thermalization of the quasiparticle population through electron-phonon scattering takes place on a sub-nanosecond timescale; therefore, this "non-equilibrium" response in \(N_s\) has sufficient speed to be useful for a heterodyne mixer. The detailed dynamics of this response, and the way in which the quasiparticle and pairbreaking phonon populations are coupled, is a subject of some complexity which has been studied for many years [4]. In addition to this non-equilibrium response, there is also a slower, bolometric response due to the bulk heating of
the photoinductive element relative to the substrate. In this case, \( N_s \) is determined simply by equilibrium thermal generation of quasiparticles, with approximately the canonical \( 1 - (T/T_c)^4 \) dependence. Near \( T_c \) this leads to a very sensitive dependence of inductance on temperature. This is a conventional microbolometer response [5]. For the typical sizes of our devices, it can be expected to yield bandwidths of 100's of kHz, and phonon-limited noise-equivalent powers (NEP's) even at temperatures as high as 8 – 9 K.

The balance between the non-equilibrium and bolometric responses may be conveniently characterized by two thermal conductances in series, a microscopic \( G_m \) related to the electron-phonon scattering, and a macroscopic \( G_M \) determined by bulk and surface thermal conductivities. The details depend on the assumptions made for substrate material and superconducting film properties, but are typically as indicated by the dependence shown in Fig. 2, (which we have calculated using the formalism described in reference [4]). For a given detector geometry, the bolometric response dominates at temperatures near \( T_c \), and the non-equilibrium response dominates at temperatures well below \( T_c \). Early experiments on the response of superconducting films to fast laser pulses [6] showed this behavior very explicitly; near \( T_c \) the electrical pulses were heavily rounded due to the bulk heat capacity of the superconducting film, but as the temperature was lowered, the pulses grew progressively sharper until the rise times matched those of the laser pulses.

II. Radiometer

The first project undertaken at NIST which exploits kinetic inductance for infrared detection is an absolute radiometer intended for installation at the NIST Gaithersburg, Low-Background Infrared (LBIR) calibration facility [7]. The radiometer is intended for standards-grade intercomparisons of 300 K black-body emitters in the 10 – 150 THz frequency range, with spectral resolution \( \Delta \nu/\nu \approx 0.01 \). Its design is therefore optimized for high accuracy, i.e. large dynamic range, and not simply for high sensitivity. In other words, the important figure of merit for this application is not raw NEP, but rather the ratio of NEP to the largest signal power detectable without saturation, (which must be of the order of \( 10^{-7} \) W).

This leads to a somewhat different design strategy than that used in conventional i.r. bolometers. Since the NEP of a thermal detector may be expressed as the product of the bulk thermal conductance \( G \) [W/K] and a temperature noise spectral density \( S_T \) [K/Hz\(^{1/2}\)], lowering the thermal conductance of a conventional detector always improves the performance, even in the phonon-limited case, where \( S_T \propto G^{-1/2} \). Typically, bolometer designers do very extensive thermal engineering to obtain the minimum G possible at any given operating temperature. (Lower oper-
ating temperature can be thought of as a "brute force" method of attaining lower G.) However, saturation power also scales linearly with G, so in our case, in which high accuracy is the important figure of merit, lowering G does not improve performance at all. Rather, G is chosen so as to accommodate the largest signal powers employed, and all the effort is concentrated on lowering $S_T$, i.e. on building a sensitive thermometer.

The method we employ is outlined schematically in Fig. 3 and has been described in further detail in Ref. 7. Four inductive superconducting striplines are arranged in a Wheatstone bridge configuration. The Nb striplines are approximately 3 μm wide and have a stretched-out length of 4 cm each, yielding inductances in the convenient range of 0.1 μH at our typical operating temperatures of 0.6 – 0.9 $T_c$. Two of the inductors are thermally isolated from the rest of the chip by etching a trench on the back side of the Si, surrounding a square island at the center of the chip, of 2 mm$^2$ area. The island is physically attached to the rest of the chip only by the 9 μm thick, boron-doped, etch stop layer. This etch stop layer, which forms a relatively flexible, thin membrane, determines the thermal conductance $G$ between the temperature-sensitive inductors and the substrate. A resistive heater lies on the island along with the thermally isolated inductors. A small inductance mismatch is built into the bridge to allow it to be balanced even with the maximum anticipated powers absorbed in the island. The bridge is driven with a large audio-frequency current, typically 5 mA at 3 kHz. The imbalance current is sensed by a commercial DC SQUID amplifier, located off chip and operated in flux-locked mode.

Thermal responsivity and noise measurements were performed on the bridge over an operating temperature range (of the chip, not the island) of 4.0 – 7.0 K. The thermal conductance was determined independently by using the resistive superconducting transition of the island inductors as a thermometer. Responsivity was measured from the change in in-phase and quadrature components of the demodulated SQUID output as small (10 nW) pulses of electrical power were applied to the island heater. Noise was determined from the RMS voltage noise level at the SQUID output measured directly on an audio-frequency spectrum analyzer. The responsivity measurements are shown in Fig. 4 as the filled circles. The curve is a theoretical model, with no free parameters except the geometric aspect ratio of the inductors (which is known from optical microscopy to \( \sim 10\% \)), and assuming the canonical $N_s \propto 1 - (T/T_c)^4$ dependence of the pair density. The measured NEP of the bridge is $\text{NEP} = 4.4 \times 10^{-11}$ W/Hz$^{1/2}$ at an operating temperature of 6.6 K. The measured thermal conductance is $G = 1.0 \times 10^{-4}$ W/K, which implies $S_T = 440$ nK/Hz$^{1/2}$. As discussed above, the high NEP, some 3 orders of magnitude higher (i.e. worse) than conventional liquid-He temperature bolometers optimized for sensitivity, is a direct result of the high value of $G$, which in
turn is the result of our design emphasis on absolute accuracy. In terms of temperature sensitivity, the result is in fact somewhat better than any other i.r. bolometer we know of (most of which are based on semiconductor thermistors). The most important conclusion from this experiment is perhaps the fact that no excess noise source is observable. That is, the measured voltage noise level agrees with that implied by the parameters of the SQUID to within the measurement accuracy.

III. Antenna-coupled Bolometers and Mixers

Given the impressive level of temperature sensitivity attainable with kinetic inductance elements coupled to SQUID amplifiers, it is natural to explore the possibilities of i.r. detectors when the relevant thermal conductances are made very small. The desirability of an entirely lithographic fabrication method and an interest in large arrays have led us to consider a strategy of obtaining low $G$ by antenna-coupling the incident radiation into a very small ($\ll \lambda$) inductor as opposed to using a large area, surface-absorbing detector similar to the radiometer. Although the short wavelength limit of lithographic antenna performance has not been well explored beyond 119 $\mu$m [8], there seems to be no fundamental obstacle besides lithographic resolution to extending high efficiency antennas to the thermal infrared, where many of our array applications exist. At 119 $\mu$m and longer, where our heterodyne mixer work is concentrated, on the other hand, high efficiencies have already been achieved [9].

The basic physical structures of our microbolometers and our heterodyne mixers are identical. This structure is illustrated in Fig. 5. The thin film photoinductive element lies at the feed of a lithographic antenna. The size of the photoinductive element is taken to be $1 \mu$m $\times$ $1 \mu$m $\times$ 10 nm in our calculations. In general, it should be as small as possible, but submicron lithography is not required in order to achieve very interesting levels of sensitivity. The photoinductive element forms one section of the loop of a DC SQUID. The SQUID loop, as shown in the lower, cross-sectional view of the feed region, consists of the lower photoinductive element, which contacts the two antenna feed points, the two Josephson junctions, and an upper photoinductive film, identical to the lower one, which closes the loop. I.r. radiation concentrated by the antenna is dissipated in the photoinductive elements, where, either through non-equilibrium effects or through bolometric heating, the pair population is reduced, increasing the inductance of the two photoinductive films. The DC SQUID is operated in a somewhat unusual way, as a 3-terminal inductance transducer rather than as a magnetic flux transducer. As shown in the inset of Fig. 6, a large DC bias or “control” current, $I_c$, analogous to the magnetic bias of a conventional SQUID magnetometer, is passed through the photoinductive film. This is in addition to the conventional bias current,
$I_b$, passed through the two Josephson junctions in parallel. The control current serves to set the operating point on the steep section of a high-order lobe in the SQUID’s voltage-flux (V - $\Phi$) pattern (Fig. 6). As the loop inductance varies with incident i.r. radiation, the V - $\Phi$ pattern accordions in and out, modulating the SQUID output voltage. This turns out to be an extremely sensitive way to read out very small changes in inductance.

The fundamental reason this device is expected to be sensitive can be summarized in one word: impedance-matching. At frequencies above the energy gap, $\nu > 2\Delta(T)/\hbar$, which is the range for which the detector is intended, the thin film superconductor has the surface impedance of a normal metal. Thus, for $\sim 10$ nm thicknesses, sheet resistances will be $10 - 40$ $\Omega$/square for Nb or Nb alloys. This is ideal for matching to lithographic antennas. Note that, unlike SIS mixers, Schottky diodes, or other devices with a junction-like geometry, there will be no large parasitic capacitance associated with the photoinductive element. On the other hand, at i.f. frequencies and below, the thin film is a superconducting (i.e. inductive) short. This is ideal for matching to SQUID amplifiers. Thus, without any special matching structures or tuning elements, the device is naturally impedance-matched to the rest of the circuit at both the RF and the i.f. frequencies.

Given that we have not discussed anything regarding non-linearities in the I-V curve, there might be some who would ask how such a device can operate as a sensitive heterodyne mixer. This is a parochial viewpoint, however, arising out of the accidental fact that Schottky diodes and SIS mixers happen to dominate the mixer world at millimeter frequencies, and both of those devices mix through the non-linearity in their I-V curves. In fact, any sensitive direct detector can also act as a sensitive heterodyne mixer. This occurs because the electromagnetic field amplitudes (signal and local oscillator (LO)) add, while the detector senses total power, which is proportional to the square of the total field amplitude. In other words, $E = E_{LO} \cos \omega_{LO} t + E_s \cos \omega_s t$, (where $E_s \ll E_{LO}$), which implies that $P = P_{LO} + 2\sqrt{P_{LO}P_s} \cos \omega_{IF} t + P_s + P_{2\omega} + ...$. Heterodyne mixers of this type, which includes the InSb hot-electron mixer and the 2D electron gas mixer [10] are often called “total power” or “envelope” mixers. Although it is true that any direct detector will also function as a heterodyne mixer, the question of whether it can serve as useful heterodyne mixer is more complicated. In particular, the device’s usefulness strongly depends on its speed. Unlike diode-like mixers, the available i.f. bandwidth of total power mixers is limited by the basic physics of the detection process. In the case of the photoinductive mixer, the bandwidth is limited by the microscopic processes of quasiparticle recombination, quasiparticle scattering with pairbreaking phonons, and pairbreaking phonon escape. The size of our SQUID is small enough that self-resonance effects should be a less restrictive limit than the quasiparticle
recombination.

We have done a thorough design analysis of the photoinductive mixer which has been described in more detail elsewhere [11]. The heart of the analysis is calculation of the device's voltage responsivity. This has three main parts: calculation of the SQUID's sensitivity as an inductance transducer (in V/H), calculation of the photoinductive film's inductance sensitivity, $\delta L/\delta N_s$, (which follows from Eqns. 1 and 2), and calculation of the quasiparticle population's response to incident i.r. photons, which includes all the complicated dynamics of quasiparticle and phonon recombination and scattering. These dynamical processes are still the subject of active research; nonetheless, a basic theoretical framework does exist and has been, to some extent, confirmed by experiment. The Rothwarf-Taylor equations are a pair of coupled rate equations for the quasiparticle and pairbreaking phonon populations, and they represent the simplest level of approximation for treating these dynamical processes. By performing a small-signal stability analysis of these rate equations, we have made predictions (shown in Fig. 7) of the response of the quasiparticle density perturbation to a harmonically varying (at frequency $\omega_{RF}$) quasiparticle injection rate. This is proportional to the device's responsivity as a function of i.f. frequency. The response is nearly Lorentzian, and the 3 dB bandwidths are strongly dependent on temperature, varying from 200 MHz to 1 GHz as $T/T_c$ varies from 0.5 to 0.9. The absolute levels of the response shown in Fig. 7 assume the LO power level is held fixed, which is not realistic. In fact, at the higher temperatures, the optimum LO power also rises rapidly, so that there is a tradeoff between LO power requirement and available i.f. bandwidth. However, the levels of LO power required, even at the higher temperature, are still very low, only 10's of nW, many orders of magnitude less than Schottky diodes, and comparable to the levels required for mm-wave SIS mixers. In short, the bandwidths are large enough, 200 - 1000 MHz, and the required LO powers low enough, 10's of nW, to make the device extremely interesting as a practical THz mixer.

It is unknown whether the complicated scattering and recombination processes of the quasiparticle/phonon system will introduce excess noise into the device's response. So long as this recombination noise is no more than a factor of a few greater than the shot noise limit, however, the total noise should be dominated by the noise of the SQUID amplifier. The noise of well designed SQUIDs is tantamount to just the Johnson noise of the two junction shunt resistors in parallel. For our design, this amounts to a Johnson noise of roughly 1 $\Omega$ at $T = 4 - 8$ K, appearing at the SQUID output. Using this, and the estimate of voltage responsivity described above, the heterodyne noise performance of the device may be estimated, as shown in Fig. 8. Well above the gap frequency, noise temperatures should be approximately $10h\nu/k$, comparable to the best SIS
receiver results at frequencies well below the energy gap. As shown however, this performance is only achieved by using metallic alloy superconductors such as NbZr and NbTi, which have both high critical current densities (allowing large control currents to be used in the SQUID transducer), and large penetration depths, which yield larger inductance sensitivities (Eqns. 1 and 2). Pure niobium photoinductors are expected to yield performance comparable to current Schottky diodes. In both cases, however, the SQUID junctions can and will be made of conventional Nb/Al-AlOₓ/Nb trilayers.

IV. Recent Experimental Results

These theoretical estimates are sufficiently promising to have encouraged us to begin an experimental program to investigate antenna-coupled kinetic inductance bolometers and mixers. Initially, we are confining ourselves to pure Nb devices for ease of fabrication. Fig. 9 shows an early device, fabricated as indicated by the schematic in Fig. 5, but which lacks the SQUID junctions. The rectangular areas on each side of the antenna feed are windows in the SiO insulation layer where the SQUID junctions would lie. In this device, the sloped edges of the 100 nm thick antenna layer, needed for obtaining contact with the thin photoinductor, were fabricated by reactive-ion etching at high O₂ concentration. In later devices, we have switched to a different geometry, in which the photoinductive layer is deposited first, and a Au antenna deposited on top of it. A typical low temperature I-V curve of such a device is shown in Fig. 10. The measurement is performed in vacuum, and the I-V curve is therefore highly hysteretic due to self-heating. This fact can be used to estimate, in a very unambiguous way, the bulk thermal conductance from the device to the substrate. First the T_c of the film is measured by heating the entire substrate, and then, at any given operating temperature T < T_c, the I-V curve is measured. The power dissipated at the point just before the device switches back into the zero-voltage state yields G. The basic parameters of the device as derived from these measurements are listed in table I.

<table>
<thead>
<tr>
<th>Table I. Cryogenic Properties</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thickness</td>
</tr>
<tr>
<td>Length × Width</td>
</tr>
<tr>
<td>T_c</td>
</tr>
<tr>
<td>Z_{normal}(DC)</td>
</tr>
<tr>
<td>G_M (6 K)</td>
</tr>
<tr>
<td>NEP_{phonon} = (4kT^2G)^{1/2}</td>
</tr>
</tbody>
</table>

Finally, we have measured the optical responsivity of the large-area, radiometer chip using a
1.3 μm wavelength laser diode coupled into the cryostat with an optical fiber. The silicon chip is transparent at 1.3 μm and the fiber illuminates the chip from the back side. The purpose of this experiment was to try to find a definite non-equilibrium response of the photoinductive stripline when it is illuminated with pairbreaking radiation. (1.3 μm is about 300 times the gap energy.) Unfortunately, we have not yet been able to set up the appropriate driver electronics to allow us to modulate the laser diode. A component of the response that remained essentially flat at frequencies well above the thermal time constant of the island (which corresponds to about 600 Hz) would be an unambiguous sign of non-equilibrium response. The optical responsivities we have so far measured at d.c. are shown as a function of temperature in Fig. 4 alongside the thermal responsivity data. Over the temperature range covered, they track the thermal responsivities to within the experimental error. This is expected on the basis of the very large volume of the photoinductor (the microscopic G_m scales with detector volume) and the fact that the fiber illuminated the island approximately uniformly. We hope to separate out the thermal and non-equilibrium responses in future experiments of this type on the radiometer chip.

V. References


[8] The only quantitative work on antennas at 10 μm, A. Sanchez, C. F. Davis, K. C. Liu, and A.


Figure 1  Thermal circuit showing how the far IR power flows through the system. The device is a conventional bolometer if $G_M \ll G_m$.

Figure 2. Distinction between bolometric and hot-electron regimes, in terms of microscopic and macroscopic thermal conductances.
Figure 3A: Thermometer electrical circuit schematic.

Figure 3B: Typical radiometer thermometer chip fabricated and tested. The central region in the figure is the thermally isolated island. This region is surrounded by the boron-doped membrane and contains the two kinetic inductors $L_1$ (left) and $L_3$ (right). The chip is 1 cm on a side.

Voltage Responsivity

Figure 4. Measured thermal and optical responsivities of the radiometer. The theoretical model assumes a conventional $N_s \propto 1 - (T/T_c)^4$ dependence.
Figure 5  Top: Example antenna with detector.
Bottom: Schematic drawing of the detector, the kinetic inductance element, between the terminals of the antenna. The SQUID loop is shown connecting the Josephson junctions (JJ). The side view is a magnified cross sectional view at the center of the antenna. The length of the antenna load is about 1 µm. Other features are not to the same scale.
Figure 6  SQUID voltage response as a function of control current $I_c$. The inset shows the SQUID schematic, with Josephson junctions denoted by $\times$. 
Figure 7. Frequency response of hot-electron photointuctive mixer.

Figure 8. Comparison of predicted photointuctive mixer performance with current SIS and Schottky mixers.
Figure 9. SEM photograph of an early device fabricated without SQUID junctions.

Figure 10. Hysteretic current-voltage curve of a photoinductive element at the feed of a lithographic antenna, at 5 K.