PHONON-MEDIATED PARTICLE DETECTION USING SUPERCONDUCTING TUNGSTEN TRANSITION-EDGE SENSORS

A DISSERTATION
SUBMITTED TO THE DEPARTMENT OF PHYSICS
AND THE COMMITTEE ON GRADUATE STUDIES
OF STANFORD UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Kent David Irwin
February 1995

Revised April 28, 1995

I certify that I have read this dissertation and that in my opinion it is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Blas Cabrera (Principal Advisor)

I certify that I have read this dissertation and that in my opinion it is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

William A. Little

I certify that I have read this dissertation and that in my opinion it is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Charles M. Marcus

Approved for the University Committee on Graduate Studies:

ABSTRACT

This thesis describes the development of several superconducting tungsten thin film based particle detector technologies. The initial motivation for this work was the construction of detectors sensitive to dark matter and neutrino scattering events. These technologies also show promise in other applications, including high resolution x-ray spectroscopy.

The detectors described here consist of a tungsten thin film deposited on a silicon substrate. When an incident particle scatters in the silicon crystal, it deposits energy in the form of phonons which propagate to the surface of the crystal where they are absorbed in the tungsten thin film. The superconducting film is biased at or near its transition temperature. Changes in the resistance of the film are measured.

The superconducting titanium transition-edge sensors previously developed by our group exhibit a threshold phonon energy density below which no signal is detectable. This threshold density poses severe restrictions on resolution, energy threshold, and absorber mass. In order to overcome these limitations, several new technologies were developed. In each case, a superconducting film with a sharp transition well below that of titanium (~ 380 mK) is necessary. To this end superconducting W films were developed with ~ 1 mK wide transitions at 70 mK. Before this work W thin films always exhibited transition temperatures > 600 mK.

The first technology described here consists of a W thin film patterned into a 1 μm wide line 1.6 m long in a meander pattern. The line is biased at a constant current, and is temperature biased near the middle of its superconducting transition. When an event deposits energy in the W film, the resulting voltage pulse is measured with a cryogenic FET. A quantum efficient sensor is also described in which the heat capacity of individual, thermally isolated film segments biased just below their transition have heat capacities small enough that individual phonons drive them normal.

The most promising technology discussed here is a novel sensor in which the temperature of the superconducting W film is held constant within its transition by an electrothermal feedback process. Energy deposited in the film by a particle interaction is removed by a reduction in the feedback Joule heating. This mode of operation leads to substantial improvements in resolution, linearity, dynamic range, and count rate. The sensor consists of a low impedance W film pad that is voltage biased. Particle interactions cause current pulses that are measured with a DC SQUID array. The fundamental limits on the energy resolution of this detector are analyzed, and found to be

below the rms thermodynamic energy fluctuations in the film, and better than any existing technology operating at the same temperature, count rate, and absorber heat capacity. The enhancement of this electrothermal feedback technology with quasiparticle trapping is also explored. In this approach, superconducting Al thin film pads are placed in electrical contact with W lines. When phonons enter the Al film, they create quasiparticles which diffuse into the W lines on times of ~ 100 ns. Once in the W films they are rapidly thermalized. This enhancement allows the instrumentation of large surface areas with smaller W heat capacity.

Using the quasiparticle trap enhanced electrothermal feedback technology, an energy resolution of < 400 eV FWHM is measured for 6 keV x-rays interacting on the backside of a 1mm thick silicon substrate. This sensitivity is sufficient for the construction of a dark matter detector, which will begin this year. Finally, the application of these technologies to other problems, including high resolution x-ray spectroscopy, infrared bolometry, and the resolution of individual low energy (~ 1 eV) photons is described.

ACKNOWLEDGMENTS

It has been a great pleasure and honor to learn from Blas Cabrera during my years at Stanford. His excitement and dedication to his research are contagious, and have greatly influenced me. Blas has tremendous scientific intuition, and a unique way of thinking about and exploring both theoretical and practical problems that always seems to lead to the correct solution. His breadth of knowledge is amazing. As an advisor, he has provided an excellent balance of guidance and freedom that combine to create very fertile ground for a young scientist.

Without the efforts of Sae Woo Nam in the last year of this work, the dramatic progress we have made during this time would have been impossible. He developed the SQUID instrumentation used with the low impedance tungsten detectors, took countless data sets, and has taken over running the experiment where I left off. Sae Woo has contributed endless energy and enthusiasm which have energized the whole group, and have helped to take the final phase of this research from concept to realization far faster than I would have thought possible.

Barron Chugg was an invaluable ally during the grueling, several year long battle with the demonic dilution refrigerator cryoleak, often physically restraining me from smashing an $80,000 instrument into tiny shards of frozen metal. He developed the technology for thick wafer processing necessary to create an actual dark matter detector. Barron did the mask layout for the final rounds of particle detectors, and, with Betty Young, fabricated all of the detectors on thick substrates. Finally, his sense of humor has always kept things sane.

I am extremely grateful to Professor Betty Young. While a graduate student in the Cabrera group, she developed the first titanium transition-edge sensors and taught me much of what I know about cryogenics and device fabrication. As if that weren't enough, her recent return to collaborate with the Cabrera group has proven invaluable. Working in the clean room with Barron, Betty has provided the experience and care necessary to fabricate these complicated detectors on thick substrates. The precision that she brings to any physics problem will set a standard for me throughout my career.

Adrian Lee helped keep my first several years in the group lighthearted. His further development of titanium transition edge sensors laid the groundwork for this thesis. Adrian designed and fabricated the cryogenic FETs used with the high impedance tungsten detectors. I am also grateful for many interesting discussions.

Michael Penn has developed the charge collection technology which, when combined with our tungsten sensors, will make a real dark matter detector possible. His taste in wine (Machiavelli '88) and music (Wagner) have infused my graduate experience with a leavening of culture and sophistication. Mike - thanks for being a fellow "dilettante."

Brian Dougherty has always been a voice calling our group back to its final purpose: a real dark matter experiment. His grasp of detector technology and fundamental physics have been extremely influential. The brainstorming sessions we have had were enjoyable and useful. I am also thankful for Brian's advice and encouragement about future career plans.

I am greatly indebted to Prof. Douglas Osheroff, who provided the dilution refrigerator for the first tungsten runs. His advice about cryogenic technology has provided guidance throughout this work. His students Ben Tigner and Dominic Salvino were good friends during my first year, and were very gracious in allowing me to be a stowaway in their refrigerator and in helping me with instrumentation.

Robin King was instrumental in the development of our tungsten films, and in the fabrication of all our detectors. My interactions with her in the last six years have always been delightful. Many other people at CIS, including especially Marnel King, Jim McVittie, and Pat Burke, have also played significant roles.

Countless parts used in the dilution refrigerator and instrumentation were machined with great precision in Wolfgang Jung's machine shop in Varian. Karlheinz Merkle and Dan Semides always crafted perfect parts with fast turnaround.

Mike Hennessy has brought great experience and ability to the construction of much of the support structure used in this work. No other member of this group has proven as willing as he has to sacrifice his life for the experiment. I'm glad it didn't go quite that far, Mike.

Prof. Barbara Neuhauser and her students have been invaluable collaborators. Her thin film laboratory at San Francisco State University deposited our gold films, did wire bonding on the not infrequent occasions that our bonder was out of order, and helped in many other ways.

I am grateful to John Martinis, Prof. Dan McCammon, and Prof. Bernard Sadoulet for many useful discussions about the noise calculations included here. John Martinis and Rick Welty, in conjunction with HYPRES, developed the SQUID arrays that are used in this work.

My wife, Joanne, has put up with far more than she should have during these last six months, which were our first six months. Thank you for the support and love, and for your perspective on what is truly important. I look forward to the next fifty years.

I am thankful to my mother and father, whose example I hope to emulate. They made it possible for me to be here, and encouraged me along the way. Thank you for always supporting me. Thank you especially for making it possible for me to know God, who, when all is said and done, deserves the final credit for anything I accomplish.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1: DARK MATTER AND NEUTRINO DETECTION

1.1. Dark Matter

This section presents a brief discussion of some of the evidence for dark matter, the prospectives for its detection, and some of the characteristics of the dark matter detector we would like to construct. For a more detailed discussion of dark matter and its detection, see, for example, the review by B. L. Dougherty [1], on which this author relies heavily.

The amount of matter in our universe may be expressed in the form of the dimensionless parameter Ω, which is the ratio of the density of matter to the critical density which would be sufficient for the universe to be just closed. Thus, if $\Omega < 1$, the universe will expand forever. If $\Omega > 1$, the universe will eventually collapse. If $\Omega = 1$, the universe is critically closed.

The matter that current instruments can detect, including stars and dust, is sufficient to set $\Omega > 0.01$. There is considerable evidence, however, that much of the matter in the universe is "dark matter," not visible to our instruments. A study of galactic rotation curves indicates the presence of sufficient gravitationally-interacting matter clustered in a roughly spherical halo around the galactic disk to set Ω as high as 0.1 [2]. As larger length scales are studied, evidence is found for ever larger amounts of dark matter. At the largest length scales a great deal of uncertainty exists in the measurements, but the results are at least consistent with $\Omega = 1$.

There is a strong theoretical prejudice for $\Omega = 1$. First, it is the simplest answer. Second, there is the "naturalness" argument. $\Omega = 1$ is an unstable equilibrium. If Ω deviates from unity even slightly, its deviation will rapidly increase. For the universe to deviate from $\Omega = 1$ now, and yet be greater than the 0.01 indicated by visible matter, there would have to be an extremely unlikely fine tuning near 1 in the early stages of the universe. Finally, the inflationary hypothesis would set $\Omega = 1$ as a direct result of inflationary mechanisms.

If $\Omega = 0.1$, it is possible that all of the dark matter is baryonic (consisting of black holes, brown dwarfs, etc.) If Ω is closer to unity, however, restrictions due to our understanding of Big Bang nucleosynthesis and the present abundance of light elements indicate that dark matter must be largely nonbaryonic.

Recently, microlensing experiments have found evidence for baryonic tenth solar mass Massive Compact Halo Objects (MACHOs) in the halo around the Milky Way.

The results are as yet preliminary, but the density of MACHOs does not seem to be large enough to provide the local $\Omega = 0.1$ densities we observe due to galactic rotation curves.

Nonbaryonic dark matter candidates are divided into two camps: cold and hot dark matter. Hot dark matter was relativistic when it fell out of equilibrium with the microwave background radiation. Cold dark matter was non-relativistic at the time of this decoupling. Computer models of galactic structure formation and COBE measurements of the cosmic microwave background anisotropy both suggest the presence of small-scale fluctuations due to the clumping of cold dark matter before decoupling. Hot (relativistic) dark matter would not clump on small enough length scales, and could not provide the necessary structure [3]. Note that there is some additional evidence for the need for an admixture of hot dark matter as well, which may be provided by neutrinos.

There is evidence that dark matter is weakly interacting. Dark matter is known to inhabit an approximately spherical halo around our galaxy. If this matter was strongly interacting it would have radiated and collapsed into the galactic disk.

Neutrinos seem to be ruled out as cold dark matter candidates by current restrictions on their mass [4]. Possible weakly interacting cold dark matter candidates include the lightest supersymmetric partner (LSP), cosmions, light Higgs, etc. Collectively, these candidates are referred to as Weakly Interacting Massive Particles (WIMPs). Non-WIMP cold dark matter candidates, which are beyond the scope of this discussion, include monopoles and axions.

The various different WIMP candidates under consideration have been proposed as possible extensions of the standard model. Their properties, including couplings, are sufficiently constrained by theory and accelerator experiments to make some statements about possible direct detection. Typical dark matter candidates might have mass of ~ 10 GeV/c^2 or greater, densities of ~ 0.3 GeV cm^{-3}, and velocities of $\sim 3 \times 10^5$ km sec^{-1}. Because of the weak interaction of these particles, scattering events would be quite rare. The detection of a sufficient number of scattering events in times of order 1 year requires the instrumentation of a fairly large absorber mass (~ 1 kg of Si or Ge). Because of the low energies of expected scattering events, energy thresholds ~ 1 keV FWHM are desirable.

Because of the large absorber mass, radioactive background event rates are quite high. It is imperative to both reduce the background rate and discriminate background events from candidate dark matter events. Dark matter detectors are constructed out of low radioactive background materials. Experiments are located below ground, where the only events of cosmic origin are expected to be some higher energy muons, neutrinos,

and perhaps WIMPs. Sites with low background radiation due to decay in the surrounding earth must be chosen. Passive shields will be used, including lead to stop photons and moderator to stop neutrons. Active detectors will be placed around the device, allowing for vetoing of simultaneous events in the device.

Certain detector characteristics would allow dramatic improvement in background vetoing capabilities. High device energy resolution would allow the removal of background peaks. Position resolution in the sensor would allow the vetoing of events that occur near the surface of the crystal (low energy background x-rays), and multiple scattering events from neutrons. Finally, the simultaneous measurement of charge and phonon signals allows the discrimination of electron and nuclear scattering events. (Electron scattering events show a proportionately larger ionization signal.) Since most background radiation will scatter off of the electrons, but dark matter (and neutrons) will often scatter off of nuclei, much of the background might be excluded by looking only at nuclear recoils.

These considerations allow the establishment of goals for the performance of our detector technology. First, event rate considerations require that a technology can be scaled up to a large (~ 1 kg) absorber mass. Event threshold and energy resolution should be < 1 keV. Event position resolution of ~ 1 mm would allow the vetoing of surface events and multiple scattering within the crystal. The technology should lend itself well to use in a hybrid phonon/ionization measurement configuration.

This thesis will focus on the development of a phonon-mediated detector technology with energy resolution < 1 keV, and yielding ~ 1 mm event position resolution. As will be described, since this technology is sensitive to athermal phonons, it will scale up to thicker absorbers without a large loss in the energy resolution.

In chapter 11, the adaptation of this technology for our first dark matter detector will be discussed. In this adaptation, an ionization collector will be instrumented on the back surface of the crystal, allowing for electron/nuclear recoil discrimination. M. J. Penn has demonstrated this adaptation with our previous titanium technology.

1.2. Neutrinos

The sensors developed in this thesis are also intended for use in a neutrino detector. Such a detector would, for the first time, be sensitive to coherent scattering of neutrinos off of nuclei. It also might set a better direct laboratory limit on the neutrino magnetic dipole moment, and could be used to search for neutrino oscillations [5].

To tree level (the first order Feynman diagrams), a neutrino with incident energy E will scatter off an electron with recoil energy T with a cross section [6]

$$\frac{d\sigma}{dT} = \frac{G_F^2 M}{2\pi}\left\{(C_V + C_A)^2 + (C_V - C_A)^2\left[1 - \frac{T}{E_\nu}\right]^2 - (C_V^2 - C_A^2)\frac{MT}{E_\nu^2}\right\}, \qquad (1.1)$$

for $0 \le T \le 2E_\nu^2/(2E_\nu + M)$, and where $C_V = 2\sin^2\theta_W + 1/2$ and $C_A = 1/2$ for electron neutrinos, $C_V = 2\sin^2\theta_W - 1/2$ and $C_A = -1/2$ for mu and tau neutrinos, and for anti-neutrinos (including those produced by a reactor) we substitute $-C_A$ for C_A. θ_W is the weak mixing angle.

Neutrino scattering off of a nucleus is of a similar form. Because the wavelength of the transfer momentum is much larger than the nucleus, the nucleus can be taken to interact as a point particle, and the scattering is coherent. Thus, the "weak charge" of the nucleus is summed before being squared, and the cross section is much larger than that of electron scattering. Unfortunately, this process is kinematically limited to below 10 keV, which makes its observation impossible with conventional semiconductor diodes, considering the small fraction of energy that is partitioned into electron-hole pairs. Since this process takes place through the neutral weak current, it is neutrino flavor blind, a fact which might prove useful in neutrino oscillation experiments, as a check for the total number of neutrinos independent of oscillation.

As the first step in these calculations, the recoil energy spectrum for $\bar{\nu}_e$ coherent nuclear scattering, $\bar{\nu}_e$ electron scattering, and $\bar{\nu}_\mu$ - electron scattering were computed. The reactor neutrino spectrum models were taken from the paper of Vogel and Engel [7]. The total neutrino flux was taken to be 6 x 10^{12} neutrinos cm^{-2} sec^{-1} (the flux expected at the San Onofre nuclear power plant in Southern California).

Calculations were performed for five reactor spectra: ^{235}U, ^{238}U, ^{239}Pu, ^{241}Pu, and ^{252}Cf. Fig. 1.1 shows that the recoil spectra are similar enough that we need only consider one model. Throughout the rest of the calculations, the ^{235}U model was used.

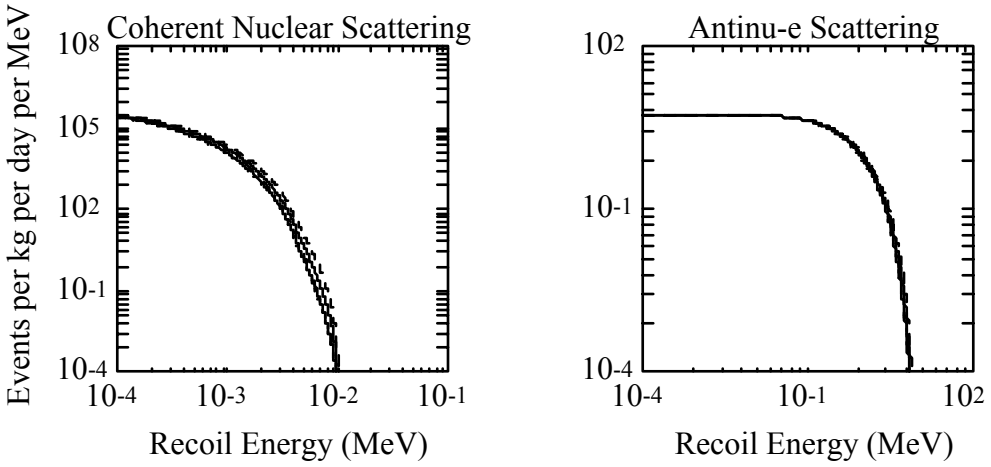


Fig 1.1 Neutrino-nuclear Recoil spectra for 5 nuclear reactor models. Each model is depicted by a different line. They are similar enough that they cannot be clearly resolved.

Fig. 1.2 shows the recoil spectra for the ^{235}U model. Coherent nuclear scattering and $\bar{\nu}_e$ - electron scattering are shown, as well as $\bar{\nu}_\mu$ - electron scattering for use in estimating the sensitivity to oscillation experiments.

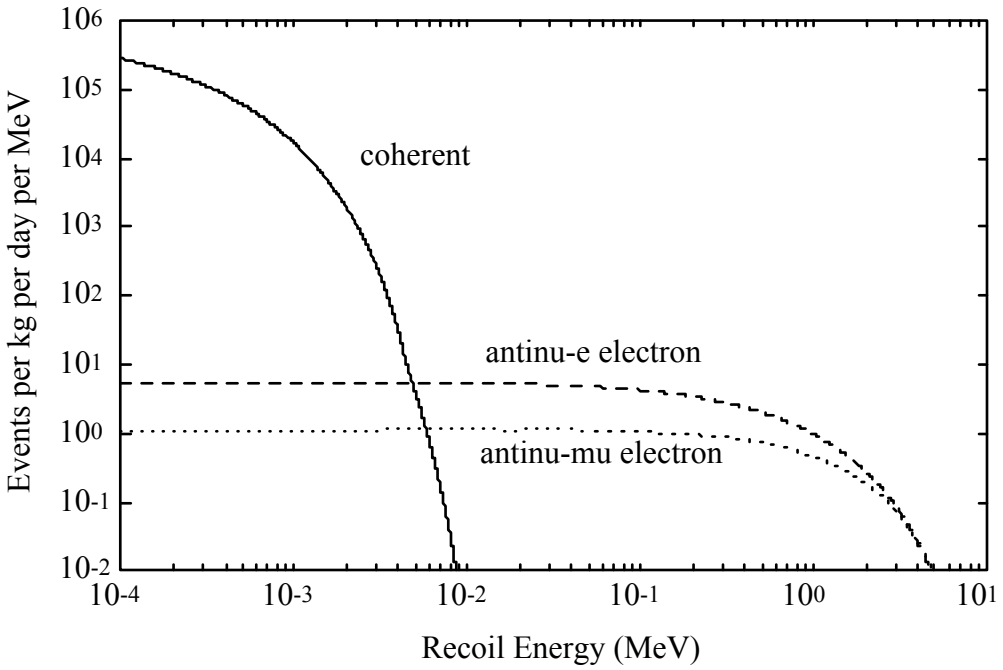


Fig 1.2 Event rate vs recoil energy for ^{235}U spectrum. The spectrum of coherent nuclear scattering and electron scattering are shown for reactor-produced anti-electron neutrinos. The spectrum that would be observed for electron scattering of anti-mu neutrinos is shown for estimating sensitivity to neutrino oscillation.

In order to look at the statistics a reactor experiment would likely see, we conducted Monte Carlo models of 100 day runs with 3 kg silicon crystals. Figure 3 shows the statistics for coherent nuclear scattering, with 0.1 keV bins. A .5 keV threshold would yield 100's of events per 0.1 keV bin at the lower energies. With our statistics, the recoil energy spectrum could be fit with some accuracy up to about 3 keV.

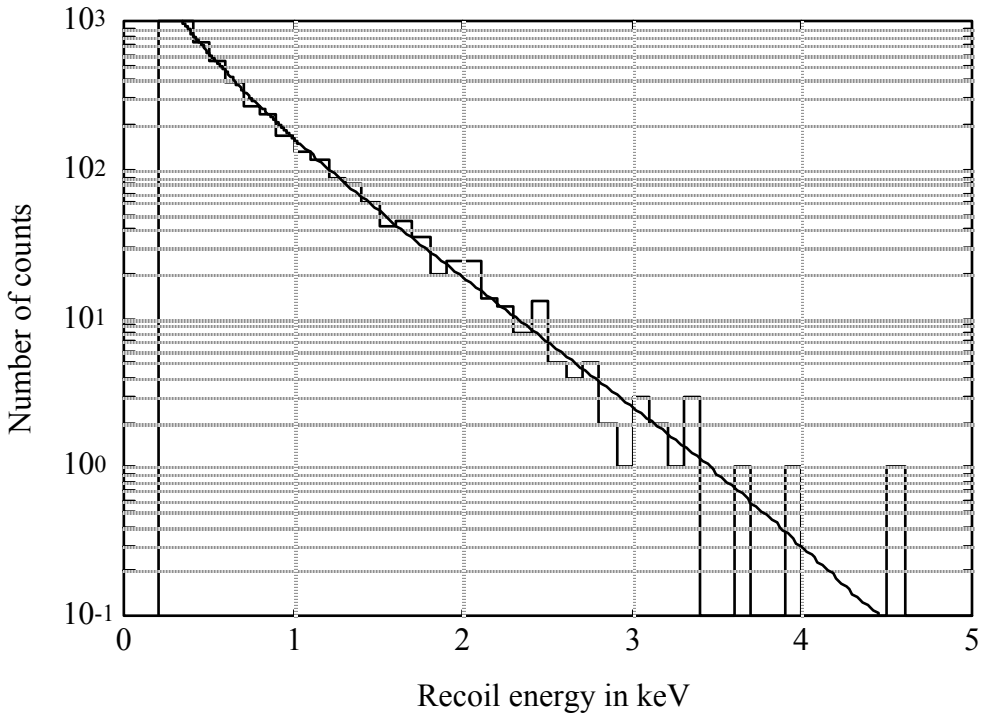


Fig 1.3 Coherent nuclear scattering statistics during a 100 day run with a 3 kg detector. The number of counts is shown on a log scale. The solid line shows theoretical spectrum. The bins show statistics that might be observed.

A reactor neutrino spectrum might also be used to measure the magnetic moment of the neutrino. A magnetic moment of 10^{-10} Bohr magnetons would explain the correlation between neutrinos and sunspot activity. Analyses of SN1987A limit the neutrino magnetic moment to about 10^{-12} Bohr magnetons, but the arguments are indirect, and based on various assumptions [7]. A reactor neutrino experiment would be a direct measurement.

To model the effect of a neutrino magnetic moment on a reactor experiment, the following cross-sectional terms were included [7].

For neutrino-electron scattering,

$$\frac{d\sigma_\mu}{dT} = \frac{\pi\alpha^2\mu_\nu^2}{m_e^2}\left(\frac{1 - T/E_\nu}{T}\right) \tag{1.2}$$

For coherent nuclear scattering on a spin-zero nucleus,

$$\frac{d\sigma_\mu}{dT} = \frac{\pi\alpha^2\mu_\nu^2}{m_e^2}\left(\frac{1 - T/E_\nu}{T} + \frac{T}{4E_\nu^2}\right)Z^2 \tag{1.3}$$

As can be seen in fig. 1.4, the magnetic moment effects on coherent nuclear scattering are too small to be measured with our detectors. But the effects on the electron scattering are substantial.

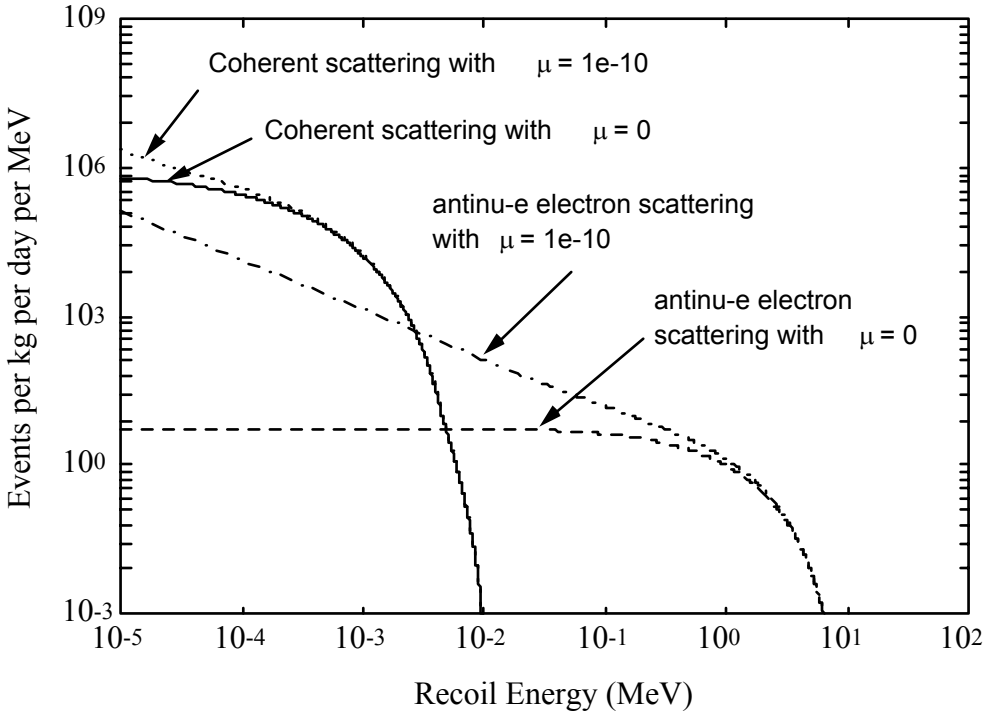


Fig 1.4 The effect of a neutrino magnetic moment on event rates. Rates are shown for 100 day run with a 3 kg. crystal with and without a neutrino magnetic moments.

The statistics of an experiment using nu-e electron scattering was modeled with a Monte Carlo calculation. Fig. 1.5 shows the statistics of a 100 day run with a 5 keV

threshold, just above the energies where coherent nuclear scattering begin to dominate. As will be discussed in later chapters, using simultaneous ionization and phonon measurements, electron and phonon recoils can be discriminated, so we could use portions of the electron recoil spectrum below this point. Even without this discrimination, the statistics are quite convincing below about 0.4 MeV. Note that the cross section contribution goes as the square of the magnetic moment, so contributions would be down by a factor of 100 for a moment of 10^{-11} Bohr magnetons. However, a direct upper limit on the value could be set by such an experiment that would exclude a neutrino magnetic moment as the explanation for the sunspot correlation.

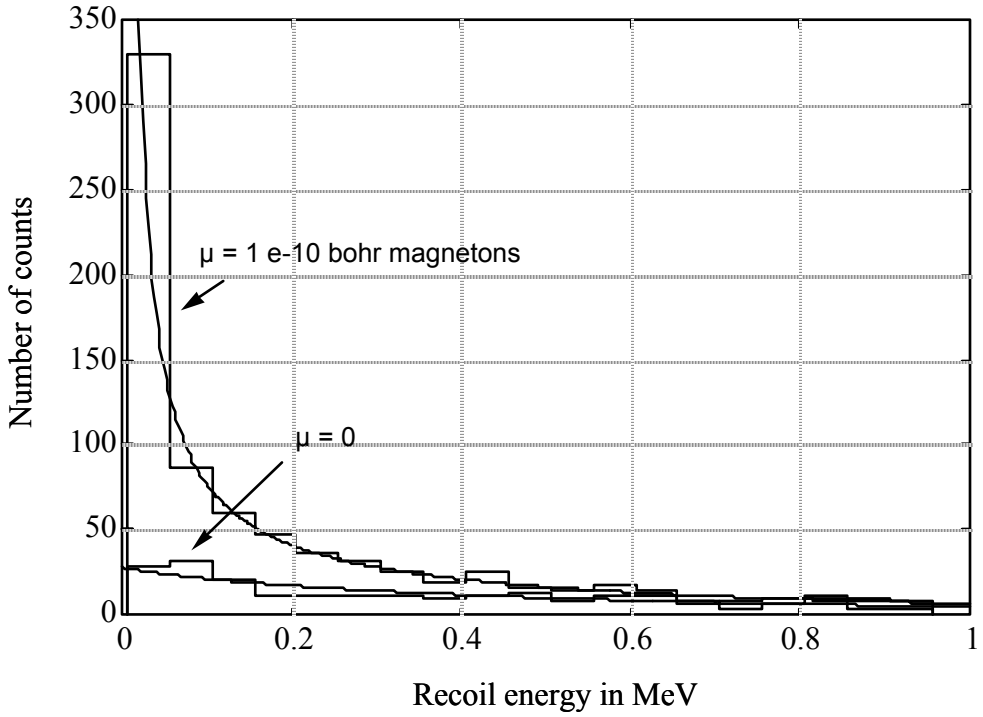


Fig 1.5 Neutrino magnetic moment statistics. A Monte Carlo is shown for a 100 day run showing the neutrino scattering with and without a neutrino magnetic moment.

In a neutrino experiment, the same background reduction considerations apply as in a dark matter experiment. These considerations are discussed in the previous section. It should be noted, however, that in the neutrino scattering experiment, a low energy threshold is more critical.

We are planning to use the San Onofre Nuclear Power Reactor in Southern California for an eventual neutrino experiment. It has a useful experimental site in the "tendon gallery" below the dome. This site is 25 m from the center of the core, and

would provide a flux of 6 x 10^{12} anti-neutrinos cm^{-2} sec^{-1}, as well as providing 20 meters of water equivalent overburden to reduce cosmic rays.

The detector goals described for a dark matter experiment apply equally to the case of the neutrino experiment. First, event rate considerations require that a technology can be scaled up to a large (~ 1 kg) absorber mass. Event threshold and energy resolution should be < 1 keV. Event position resolution of ~ 1 mm should allow the vetoing of surface events and multiple scattering within the crystal for background reduction. Finally, the technology should lend itself well to use in a hybrid phonon/ionization measurement configuration for discrimination of electron and nuclear scattering events.

CHAPTER 2: PHONON-MEDIATED PARTICLE DETECTION

When a low energy (~ 1 keV) particle interaction occurs in a silicon crystal, the energy is deposited in the form of high energy optical phonons, electron-hole pairs, and, for nuclear recoils, a small amount of lattice damage. The specific type and energy scale of the interaction determines the partioning of the energy.

For an electron recoil event, ~ 70% of the energy goes into phonons. For a 50 keV nuclear recoil, ~ 85% of the energy goes into phonons. For lower energy nuclear recoil events, less of the energy is partitioned into electron-hole pairs, so that at 10 keV ~ 90% of the energy goes into phonons. Since most of the energy goes into phonons, a phonon-mediated detector has the potential for better energy resolution than an ionization detector.

After several nanoseconds, these phonons relax to a roughly gaussian distribution of acoustic phonons with mean energy ~ 4 meV (1 THz). The relaxation rate scales as the fifth power of the frequency, so further relaxation proceeds very slowly. After a much longer time (typically tens of milliseconds, depending on the thermalizing properties of the crystal surface) the phonons thermalize, causing an elevation in the crystal lattice temperature.

Most phonon-mediated particle detectors are calorimeters, which operate by measuring the rise in the temperature of the absorber crystal. They typically consist of an absorber and a thermometer (usually a thermistor) which is thermally linked to the absorber. The detectors used by our group are sensitive instead to the athermal phonons. There are several advantages to this mode of operation.

First, since a calorimeter must wait for the phonon energy to thermalize in the absorber crystal, its speed is limited by the thermalization time in the absorber (typically ~ 10 ms). A sensor which detects athermal phonons is limited only by the time for the phonons to be collected in the sensor, and the time constants of the sensor itself. The detectors described in this work operate on timescales between 10 μs and 500 μs.

Secondly, since a calorimeter waits for the energy to thermalize in the absorber, the limiting heat capacity is that of both the thermometer and the absorber. An athermal phonon sensor absorbs the phonons before they thermalize in the absorber, so the limiting heat capacity of the detector is that of the thermometer alone, which can be a strong advantage for large absorber masses.

Finally, substantial improvements in background rejection and minimization of surface effects can be achieved by vetoing events that occur near the surface. A sensor

which detects athermal phonons can be made highly position sensitive. A calorimeter is not sensitive to the position information that is contained in the athermal phonons.

2.1. Superconducting transition-edge sensors

The detectors used by our group consist of a superconducting thin film deposited on a silicon substrate. The film is biased on or very near to its superconducting transition. When a particle interaction deposits energy in the silicon, athermal phonons from the interaction propagate to the surface of the crystal (fig 2.1). Phonons incident on a bare silicon surface reflect (mostly diffusely), while those incident on the superconducting film or a heat sink will either reflect or be absorbed. Any phonons that reflect will multiply scatter until absorbed in either the superconducting film or a heat sink. The higher energy phonons (> 1 THz) propagate diffusively, while some fraction have low enough energy to propagate ballistically.

The superconducting film is biased on or very near to its transition, so it contains many thermally excited quasiparticles. When an athermal phonon enters the film, it interacts strongly with the quasiparticles, and is rapidly thermalized, raising the temperature of the electrons in the superconductor. This temperature rise causes a resistive pulse.

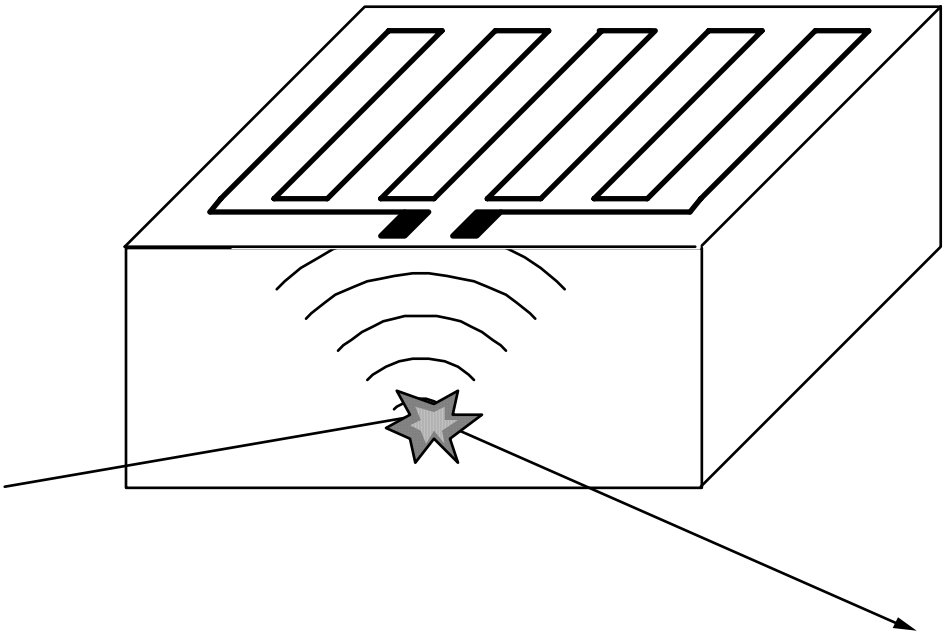


Fig 2.1 An athermal phonon-mediated transition-edge sensor. A particle scatters in a silicon absorber, creating athermal phonons which propagate to the surface and are absorbed in a superconducting thin film thermometer.

2.2. Limitations of aluminum and titanium transition-edge sensors

The first transition-edge sensors used by our group were fabricated with aluminum and titanium thin films. In these sensors, the film temperature is held below the superconducting transition temperature of the film (about 1.35 K for Al and 385 mK for Ti). The film is patterned into a meander line 2 μm wide, and several meters long, with a resistance of several MΩ in order to properly impedance match to our cryogenic FET preamplifiers. A constant bias current is applied to the film. When phonons from an interaction in the silicon are absorbed in the film, they are rapidly thermalized, and raise its local temperature. Wherever the temperature exceeds the transition temperature of the film, it is driven normal. A voltage pulse with a magnitude proportional to the length of line driven normal is measured with a FET.

Since the bias temperature of the film in Al and Ti sensors is below the transition temperature, only sections of the film that receive a critical density of incident phonons (about 400 per 2 μm x 2 μm square for Ti) are driven normal (fig 2.2). The detectors thus have an intrinsic threshold surface energy density below which they are insensitive.

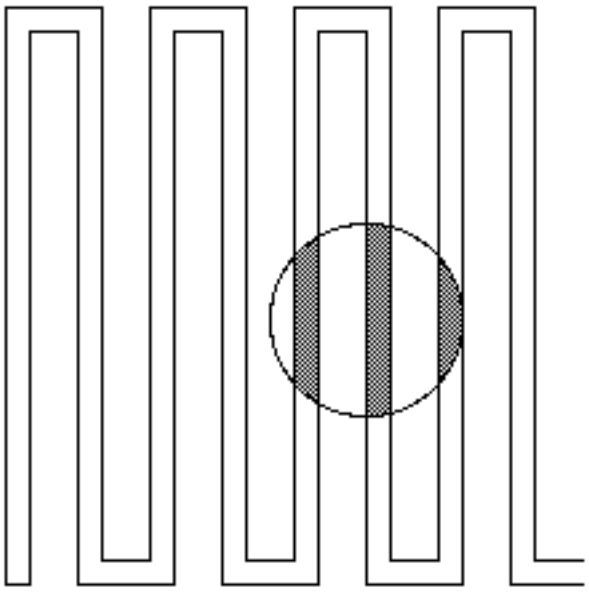


Fig 2.2 A TES with a threshold surface energy density. Near the interaction site, the phonon energy density exceeds the threshold density, and the film is driven normal. Elsewhere the film remains superconducting.

The short pulse duration in Al and Ti films (~ 40 ns for Al and ~ 3 μs for Ti) makes it impossible to bias in the middle of the transition, where there would be no intrinsic threshold energy density. When biased in the middle of the transition, our stray capacitance of ~ 15 pF leads to RC limited rise times of tens of microseconds, longer than the intrinsic pulse durations in our Al and Ti films.

There are both advantages and disadvantage to this mode of operation. The threshold energy density has the advantage of making transition nonlinearities and nonuniformities less important. The further away from T_c the substrate is biased, the higher the threshold energy density, and the less important are small variations and nonlinearities in transition properties.

One disadvantage of this mode of operation is that it is inherently nonlinear. The voltage pulse that occurs when a particle interacts with the silicon is proportional to the length of line that is driven normal, not to the total amount of energy deposited. This effect makes it important to understand the details of the phonon processes in the silicon, and to conduct extensive calibration.

A more fatal disadvantage of the threshold energy density for a large scale detector is the limit that it sets on incident particle energy threshold and absorber volume. The energy density scales linearly with the particle interaction energy. In an infinite crystal, the phonon energy density drops as the square of the distance from the interaction. Thus, the minimum energy detectable throughout the absorber volume drops as the square of the absorber thickness.

The first transition-edge sensors employed by our group were fabricated with Al thin films with T_c ~ 1.35 K. These thin films had low enough thresholds energy densities to resolve 5.5 MeV alpha particles through 300 μm of silicon [8]. Substantial reductions of threshold energy density were made by using Ti films with T_c ~385 mK. These films were sensitive enough to see 60 keV x-rays through 300 μm of silicon [9].

While these results are important, they are far from the threshold energy and active absorber volume we need to fabricate a larger scale dark matter search or neutrino detector. For these devices we would like to be able to see < 1 keV events through 1 cm of silicon.

2.3. Overcoming the limitations of titanium transition-edge sensors

In this work, we have explored three paths towards overcoming the limitations encountered with Al and Ti transition-edge sensors, in order to make a large scale detector feasible. One is to construct a detector slow enough that it can be biased in the middle of its transition. Such a device has no intrinsic threshold energy density, and is thus referred to here as a Zero Threshold Energy Density Transition Edge Sensor (ZTD-TES). The second is the construction of a "quantum efficient" sensor in which the phonon threshold exists, but is below the energy of a single athermal phonon. This device is referred to as a Quantum Efficient Transition Edge Sensor (QE-TES). Both of these approaches require the use of superconducting films with narrow transitions much colder than our titanium films. This requirement has led to the development of low T_c superconducting tungsten thin films, with transitions near 70 mK.

Finally, and most importantly, a new form of transition-edge sensor, the Electrothermal Feedback Transition-Edge Sensor (ETF-TES) has been developed with our W films. This device has no threshold energy density, has much better resolution than our previous sensors, and is much more linear.

The first two approaches, the ZTD-TES and QE-TES, were pursued in parallel during the first four years of this work. During this period, W films with appropriate superconducting transitions were developed, the Kelvinox-15 dilution refrigerator used in these studies was commissioned and instrumented, and high impedance tungsten transition-edge sensors were fabricated and tested. The ZTD-TES was successfully demonstrated. Major steps were taken towards the fabrication of a QE-TES.

The Electrothermal Feedback Transition-Edge Sensor was conceived, developed, and demonstrated during the final year. This approach was so promising that we pursue it to the exclusion of further development of the QE-TES. Our first large scale dark matter detector will be built with this technology.

CHAPTER 3: SUPERCONDUCTING TUNGSTEN FILMS

Tungsten thin films have been considered for use in cryogenic particle detectors for some time [6]. The reason for this interest is the ultra-low T_c of W in the bulk state. Pure, bulk samples of W go through a transition at ~ 15.5 mK, the lowest of any elemental superconductor.

This transition temperature is of interest to us for two reasons. First, it leads to a low heat capacity, and hence a low threshold energy density if the detector is biased below the transition. This condition is necessary for the creation of a "quantum efficient" device in which a single athermal phonon can drive a complete line segment normal. Secondly, operation at low temperatures leads to a very long intrinsic pulse duration. This is necessary to bias a high impedance device in the middle of the transition, where there is no intrinsic threshold. Finally, these films make the Electrothermal Feedback Transition-Edge Sensor (ETF-TES) possible.

3.1. The development of low T_c W films

Before this work, thin films of W were always found to have much higher transitions than bulk samples (between 0.6 and 4K). One reason for this increase in T_c is the crystal phase. Bulk W appears in the α phase (bcc). Thin films usually deposit in the β phase (an A15 structure) [10].

The crystal phase of our films was studied with x-ray diffraction. Auger analysis was used to measure film purity. To study the superconducting transition, we patterned the W films into a long meander line, and measured the line resistance with a four terminal measurement. These samples were cooled in a toploading Oxford 400 dilution refrigerator in a collaborative effort with Ben Tigner, Dominic Salvino, and Prof. Douglas Osheroff.

We have tried four different techniques to produce useful W films on silicon substrates. First, 40 nm films were deposited by sputtering with a Balzers DC Magnetron unit. When these films were cooled, a transition was seen at ~600 mK. Analysis of x-ray diffraction showed most of the characteristic peaks of the beta crystal phase (fig. 3.1a).
One large peak was missing, which may be explained by the randomization of one of the crystal axes between different granules.

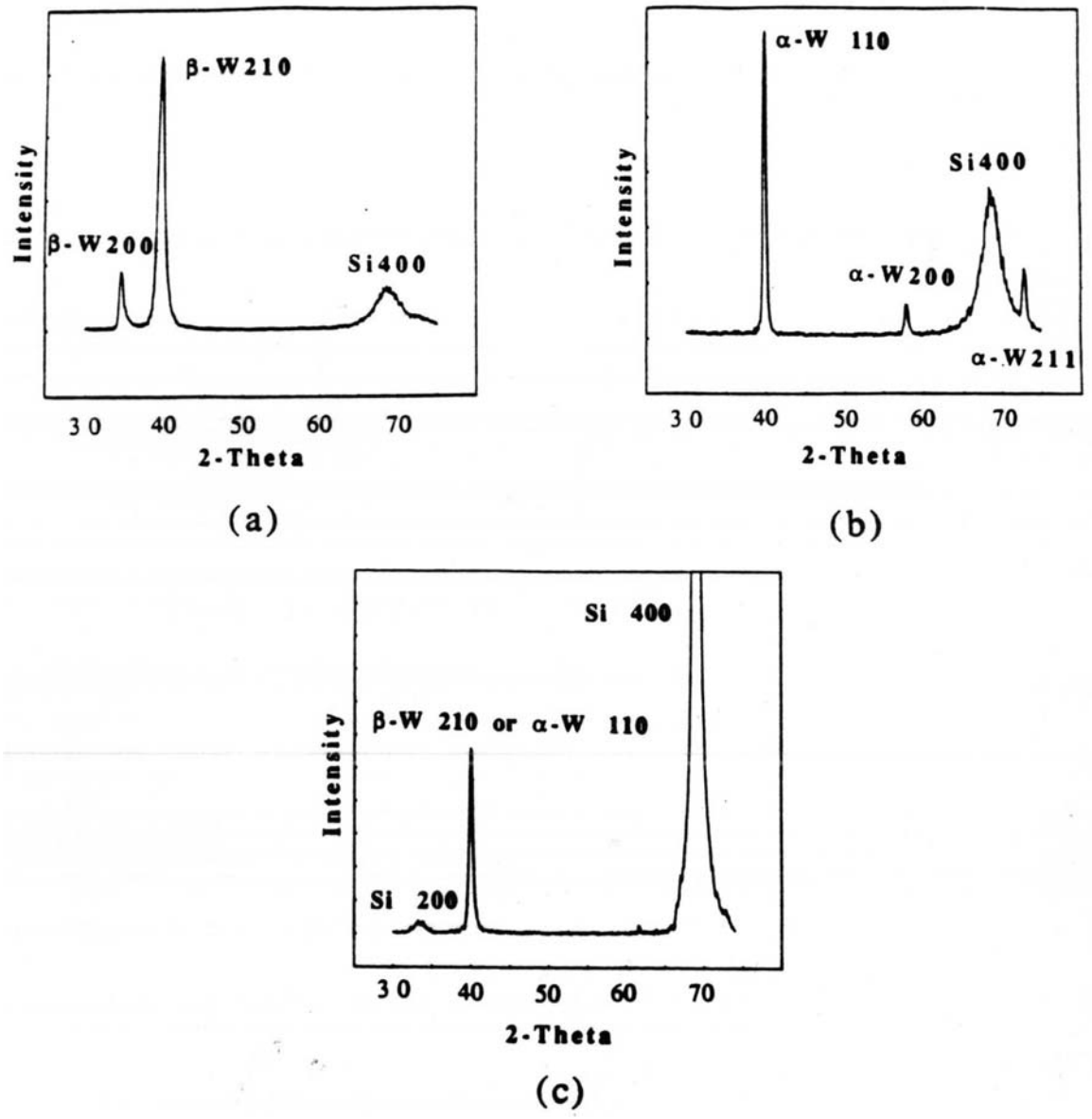


Fig 3.1 X-ray diffraction spectra of W thin films. (a) Room temperature sputtered film with no DC bias voltage. (b) LPCVD W film. (c) Heated substrate W film with only one indeterminate peak.

Analysis of x-ray diffraction has shown that some W films produced by low pressure chemical vapor deposition (LPCVD) deposit in the alpha crystal phase [11].

We examined a variety of LPCVD samples, and discovered that some did indeed have x-ray spectra characteristic of the alpha crystal phase (fig. 3.1b). Unfortunately, when most of these samples were cooled, no transition was observed down to ~ 5 mK. One film did, however, display a transition at 110 mK This film was deposited 60 nm thick by hydrogen reduction LPCVD using a hot wall reactor at 400 C. The transition width was 16 mK. This transition was disappointingly broad, as compared to the 4 mK widths we have seen with Ti films at 385 mK.

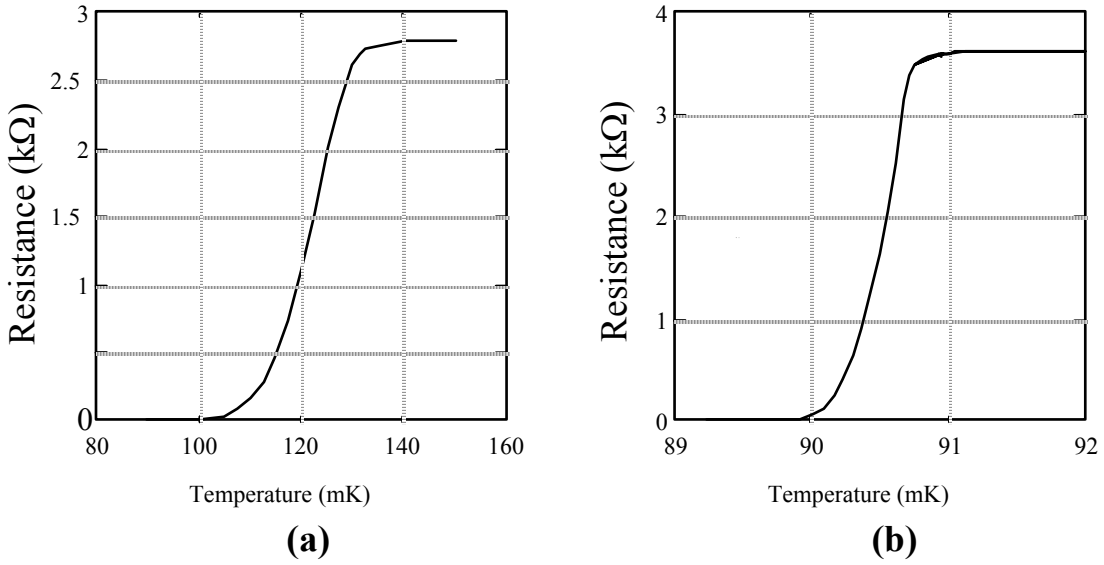


Fig 3.2 Superconducting transitions of W thin films. (a) 60 nm W, H_2 reduced LPCVD with hot wall reactor at 400 C. (b) 100 nm sputtered W with 350 C heated substrate.

The third technique tried was sputtering W films onto silicon with a heated substrate. This was done in a Balzers 450 DC Magnetron sputtering unit. W films on a Si substrate can react with the Si to form WSi_2, which is a normal metal. This is known to occur after deposition at ~ 600 C. During deposition, it has been seen to occur at temperatures as low as ~ 450 C.

Since oxygen is thought to stabilize the beta crystal phase, the wafer was etched before sputtering to remove all of the oxide. After the Balzers was pumped down, a short backsputter was done to remove any remaining oxide. The first film was deposited 100 nm thick at a substrate temperature of 350 C. For this film, we observed a transition at T_c = 90 mK (fig 3.2b) The transition was very narrow, about 0.4 mK. Indeterminate results were obtained from x-ray diffraction analysis, as the only W peak seen exists for both alpha and beta phases (fig 3.1c). This observation may by explained by a

randomization of a further crystal axis between granules, causing the disappearance of the other alpha peaks. It should be noted that a beta peak seen in our high T_c W films was also missing.

Deposition	Thickness	T_c	ΔT
Sputtered		600 mK	
LPCVD	60 nm	110 mK	16 mK
Sputter 350	100 nm	90.6 mK	0.4 mK
Sputter 350	40 nm	84 mK	3.5 mK
Sputter 450	40 nm	93 mK	2.2 mK

Table 3.1 Measured properties of superconducting W films

The success with this deposition technique led to the testing of further heated substrate films (Table 3.1). A 40 nm thick film was deposited at 350 C, and showed a transition at 84 mK, but with a width of 3.5 mK. Finally a 40 nm thick film was deposited at 450C. This film showed a transition at 93 mK, with a 2.2 mK width.

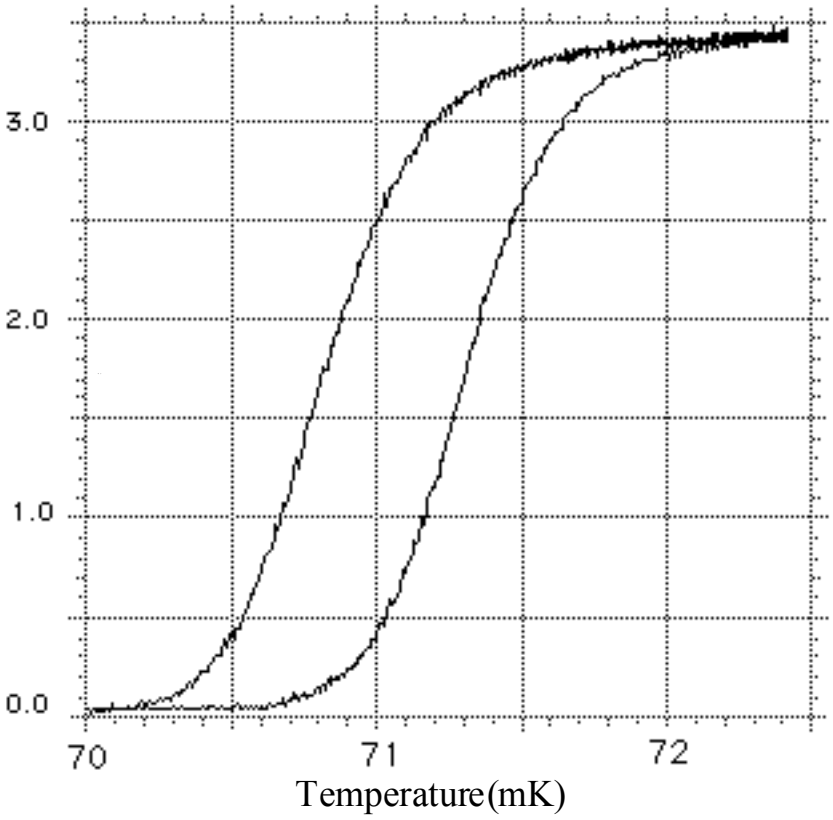


Fig 3.3 Superconducting transition of sputtered W film with -200V DC bias. The hysteresis seen is a function of the rate of temperature change.

One final technique was tested that led to our best results. We deposited a W film by DC sputtering with a substrate bias voltage of -200 V. As in the case of the heated substrate films, x-ray diffraction revealed only the one peak that could be either alpha or beta phase. When cooled, a transition was found at ~ 70 mK with a transition width of 0.7 mK (fig. 3.3). Films fabricated with this technique are used in all the results described in this thesis.

3.2. Superconducting W Film Deposition Details

The W films used in this work are deposited in a Balzers 450 DC magnetron sputtering system at CIS by Robin King. The wafers are pre-cleaned with a dirty sulfuric dip, followed by HF, and then a clean sulfuric dip, followed by HF to remove all remaining oxide. After the wafers are loaded in the Balzers, the system is baked out, and then an RF backsputter is done to remove all of the remaining oxide. The system is then pumped down again, typically to ~ 10^{-8} torr.

Deposition typically occurs at a power of 2.5 kW, and an argon pressure of 5 mbar. A DC bias voltage of -200 V is applied to the substrate during deposition. Typical drum rotation speed is 8 Hz. We find that the dc bias power slows down the deposition, such that deposition times are ~ 30% longer with the dc bias applied.

CHAPTER 4: EXPERIMENTAL APPARATUS

4.1. Kelvinox-15 Dilution Refrigerator

The previous detectors used in our group were cooled with a 1 K pumped ^{4}He cold plate probe (for the Al TES) or in a pump ^{3}He refrigerator (the Ti TES). The detectors used in this work, however, are fabricated with W films having superconducting transition temperatures near 70 mK. In order to attain these temperatures an Oxford Instruments Kelvinox-15 dilution refrigerator is used. This refrigerator has a base temperature of ~ 40 mK, and a cooling power of 15 μW at 100 mK. It is extremely small, and is lifted by hand to be placed in the helium dewar - a frightening task, considering the price tag. The dilution unit was specially fabricated with extra space between the 1K pot and the still, in order to fit the amplifier stage for our cryogenic FETs and SQUIDs. The IVC is attached via a tapered grease seal.

We struggled with a particularly difficult leak in the refrigerator for several years. This leak was only present at cryogenic temperatures, and then only intermittently. After two years, it finally leaked at room temperature, allowing us to track it down to a flaw in a solder joint in the still, hidden underneath a heater coil.

The sample stage and amplifier boards were manufactured at Stanford out of OFHC copper. The copper is gold plated with no nickel underplating, for fear of magnetic fields. The IVC is wrapped in several layers of Nb foil to reduce EMI. The dewar itself is placed in an annealed, degaussed μ-metal shield.

The low frequency electrical lines enter through two 24-pin Fischer connectors installed by Oxford Instruments. Access to the 24-pin Fischer connectors goes through two large filter boxes made at Stanford, and equipped with various low pass filters with a knee as low as 10 kHz in some cases. Eight high frequency lines go from the amplifier stage up to room temperature in stainless steel coaxes, ending in a home made brass plate with hermetically sealed SMA connectors.

The dewar sits on a cork board to reduce vibrations, and some attempt was made to isolate the pumping lines from pump vibration. Still, for our high impedance devices, microphonics was a limiting factor. Fortunately, in our low impedance technologies, microphonics is negligible.

4.2. Cryogenic MESFET preamplifiers

The high impedance devices used in this work are current biased, and use FETs to measure the voltage pulses resulting from particle interactions. It is important to keep the electronics-limited rise time of a detector small as compared to the event duration. When biased in the middle of the transition, our high impedance W sensors have resistances of order 1 MΩ. Electrical lines running all the way up to a room temperature amp tend to have capacitance of ~ 100 pF. This capacitance would lead to an RC time constant of ~ 100 μs, which is too long. It is thus important to have a cryogenic preamplifier stage on the 1K pot of the refrigerator, where the capacitance of a channel can be limited to ~ 15 pF, yielding RC ~ 15 μs, which is an acceptable rise time for our application.

Silicon JFETs have excellent noise performance, but must be heated to prevent carrier freeze out. Instead, two channels of Sony 3SK164 dual-gate GaAs MESFET preamplifiers were installed for use with our high impedance sensors, since they operate well at cryogenic temperatures. The amplifier boards were designed by A.T. Lee for use with our Ti TES [12].

4.3. DC SQUID arrays

The low impedance technologies described in this work are voltage biased, and use SQUIDs to measure the current pulses resulting from particle interactions. Because of the time constants of our pulses, and the fact that we are interested in pulse leading edge information, we need SQUIDs with bandwidths ~ 1 MHz. Conventional single SQUID readout systems are limited to bandwidths ~ 50 kHz by the electronics needed to impedance match to the extremely low impedance SQUIDs. Fortunately, DC SQUID arrays have been demonstrated that meet our bandwidth requirements [13-14]. Since they are made of several hundred SQUIDs in series, these arrays have much higher impedance, allowing simpler room temperature electronics. This matching leads to a high bandwidth (175 MHz at unity gain) and allows the use of less expensive electronics, which is important for a multi-channel detector.

We have entered into a collaboration with R. P. Welty and J. M. Martinis of NIST, and HYPRES, Inc. to develop the use of these SQUID arrays for our application. In our group, S. W. Nam has tested a range of SQUID arrays, and made suggestions to R. P. Welty about their improvements. He has installed two channels of these SQUID arrays onto the amplifier stage below the 1 K pot. The noise of the SQUID arrays used in this experiment is ~ 5 pA/√Hz, with a 0.3 μH input inductance.

In this work, these arrays were first run in an open-loop mode, with no feedback. This mode of operation limited the dynamic range of pulses, and introduced substantial nonlinearities. The last data sets with the W/Al ETF-TES were taken with feedback electronics, leading to substantial improvements in reproducibility.

CHAPTER 5: THE ZERO THRESHOLD ENERGY DENSITY TRANSITION EDGE SENSOR

The development of W thin films with transition temperatures near 70 mK has made it possible to create a high impedance transition edge sensor without an intrinsic threshold surface energy density. Operating at a lower film temperature leads to longer intrinsic pulse durations, allowing us to bias in the middle of the transition while keeping pulse durations longer than the RC limited electronics time constant.

5.1. Intrinsic pulse duration

The intrinsic pulse duration is defined as the pulse duration that will be observed in the limit of small joule heating. When phonon energy is absorbed in the film, it is rapidly thermalized, causing a rise in the temperature of the film. The rate at which heat is returned to the substrate is determined by the heat capacity of the film and the thermal impedance between the film and the substrate. The heat loss into the substrate is described by the differential equation

$$C\frac{dT}{dt} = -K(T^n - T_s^n) \tag{5.1}$$

where C is the heat capacity of the film, T is the temperature of the film, K is a material and geometry dependent parameter, and n is a number whose value depends on the dominant thermal impedance between the substrate and the electrons in the superconducting film. At higher temperatures and in thicker films, this impedance is set by the Kapitza boundary resistance between the film and the substrate, and $n = 4$. For thinner films, and at lower temperatures, the electron-phonon decoupling in the film dominates [15, 16]. In this case n is either 5 or 6, depending on the theory and the temperature range.

When the temperature of the silicon substrate is close to the temperature of the film, as it is in all of our high impedance detectors, this equation may be approximated as

$$C\frac{d\Delta T}{dt} = -g\Delta T \tag{5.1}$$

where $\Delta T = T - T_s$, and $g = dP/dT = nKT^{n-1}$ is the thermal conductance between the film and the substrate. The solution of this differential equation is a simple exponential, with a time constant of $\tau = C/g$, the intrinsic time constant for heat loss into the substrate. Since the film heat capacity scales approximately linearly with the transition temperature, and the thermal conductance scales as a higher power of the transition temperature, films with colder transitions will have longer pulse durations. We have observed $1/e$ pulse durations of less than a microsecond for our Al films, and ~ 3 μs for our Ti films. The pulse durations in the W thin films described in this work have been found to be hundreds of microseconds.

RC risetimes in the transition center for a high impedance meander are tens of microseconds, so only the W films can be biased in the center of the transition. When W TES sensors are biased in the middle of the transition, they have no intrinsic threshold energy density. This fact leads to improvement in signal to noise and resolution, and allows the instrumentation of larger active absorber volumes. This type of sensor is referred to here as the Zero Threshold Energy Density Transition Edge Sensor (ZTD-TES).

Note that the recent development of the low impedance SQUID readout technology described in chapter 7 has opened the possibility of the construction of a Ti TES without a threshold surface energy density. Due to their low impedance, the risetimes of these devices are limited to the L/R time constant (the ratio of the SQUID input coil inductance to the bias resistance), which can be < 1 μs. The ease of operation at the higher Ti transition temperatures (~ 385 mK) may make such devices worth investigating at a later time.

5.2. Noise limits of the ZTD-TES

The resolutions of high impedance Ti and Al transition edge sensors are limited by different factors than the W ZTD-TES.

Ti and Al films are biased below the transition. Unless a complete pixel is driven normal (a square of film with a length equal to the meander line width), no voltage pulse will be seen. If the film is biased far enough below the transition, then to first order the resistance of a pixel may be assumed to change discontinuously from zero to the normal resistance. These devices have small Johnson voltage noise contributions, as the resistance even at the peak is relatively small. Furthermore, they are biased at high enough current that the voltage signal associated with a single normal pixel may be resolved above the FET noise. Thus, the resolution is limited by statistical variations in

the length of line driven normal, not by the FET noise. In order to drive a section of film normal in a Ti TES, about four hundred 1 THz athermal phonons must be absorbed per pixel. The length of line driven normal is further complicated by the fact that heat flows some distance down the Ti line during the pulse duration.

In the ZTD-TES, there is no phonon threshold. Instead of changing discontinuously from the superconducting to the normal state, a section of film biased this way can remain on a linear portion of the transition during the event. In this case, the resistive signal size is locally proportional to the total energy absorbed. Unfortunately, nonuniformities in film T_c mean that the signal of the whole detector is dependent on the position of the phonon energy distribution.

The signal to noise of a resistive pulse in a ZTD-TES is fundamentally limited by the Johnson noise of the pattern and the phonon noise (thermodynamic energy fluctuations) associated with the thermal impedance between the film and the substrate. If the film is properly impedance matched to the FET (~ 1 MΩ), the Johnson noise will be about the same size as the FET noise ($V_n = \sqrt{4kTR} \approx 2\,nV/\sqrt{Hz}$ for 1 MΩ at 70 mK). The phonon noise causes temperature fluctuations in the W film, and hence resistance and voltage fluctuations. The maximum bias current in a high impedance W TES is small enough that the voltage fluctuations associated with the phonon noise are small as compared to the Johnson and FET noise. As such, the phonon noise may be neglected in a high impedance W TES. Thus, the spectral density of the noise observed during the pulse, if external noise sources are eliminated, will be a constant independent of the bias current. Maximizing the signal to noise ratio in this case is simply maximizing the signal size.

5.3. Optimizing the T_c of a ZTD-TES

The signal size of a voltage pulse in a W ZTD-TES is determined by the film sharpness, the heat capacity, the bias current, and the distribution of the incident energy. Films with lower transition temperatures tend to have lower heat capacities and sharper transitions, but also must be biased at smaller currents.

The heat capacity of a film biased on its transition is somewhere between the normal metal heat capacity, C_N, and the heat capacity of the superconductor just below the transition, C_{sc}. In the BCS theory near T_c, C_{sc} is about 2.4 times larger than C_N. Both C_N and C_{sc} are proportional to the transition temperature. Thus, if a TES is made with a film at a lower transition temperature, the same energy deposition will cause a larger temperature excursion, and larger resistance pulse sizes.

Films with lower transition temperatures also tend to be sharper, since ΔT_c tends to scale as T_c. This effect tends to cause larger resistive pulses for the same temperature excursion in films with lower T_c. Note, however, that at the very coldest transition temperatures, it becomes increasingly difficult to create films with proportionately sharper transitions. In the W films fabricated in Munich [17], the transition width is still greater than 4 mK, even though T_c is about 15 mK.

Finally, films with lower transition temperatures have smaller stable bias currents.

5.3.1. The maximum stable bias current

In order to optimize the T_c of a ZTD-TES, it is necessary to understand the maximum stable bias current. In steady state, the power loss to the substrate is equal to the Joule power:

$$I^2 R(T) = K\left(T^n - T_s^n\right) \approx g(T - T_s) \tag{5.2}$$

In order to increase the bias current while maintaining the film at a constant bias temperature, the substrate must be cooled. This process can be followed only so far before the steady state condition becomes an unstable equilibrium, and infinitesimal fluctuations in the film temperature cause the film to be driven either completely normal or completely superconducting. A small increase in the temperature will cause both the Joule heating and the heat flow to the substrate to increase. As long as the increase in the heat flow to the substrate is larger, the equilibrium is stable. When the increase in the Joule heating becomes larger, small fluctuations lead to a runaway condition. Thus, the condition for stability is that $\delta\left(I^2 R\right) < \delta\left(g\left(T - T_s\right)\right)$, or $I^2 (dR/dT)\delta T < g\delta T$. If we use the parameter $\alpha \equiv d\log R / d\log T$, a unitless measure of the sharpness of a superconducting transition, our condition becomes $I^2 R_0 \alpha / T_c < g$. This condition sets a limit on how large our bias current may be made before the film latches. The latching current is thus

$$I_{\max} = \sqrt{\frac{gT_c}{\alpha R_0}} \tag{5.3}$$

Note that the parameter α is inversely proportional to the film resistance, and is thus dependent on the bias temperature. αR_0, however, is constant throughout the linear portion of the superconducting transition. Since transition width tends to scale as T_c, the

achievable αR_0 tends to be independent of Tc. The thermal conductance g scales as T_c^{n-1}, where n = 4, 5, or 6, so $I_{\max}$ scales as $T_c^{n/2}$.

Note that if the film transition is not uniform, heat flowing down the W line will further affect the maximum stable bias current.

5.3.2. Signal size of a ZTD-TES

If a signal is assumed to cause a uniform temperature rise throughout a W ZTD-TES, the signal size due to a small energy input δE can be shown to be $\delta V = I(dR/dT)\delta T = I(dR/dT)(\delta E/C)$. Using $\alpha \equiv d\log R/d\log T$, we arrive at $\delta V = (I_{latch}\alpha R_0/T_c)(\delta E/C)$. It has been shown that the heat capacity scales as T_c, and that αR_0 tends to remain the same as T_c is varied. The latching current has been shown to scale as $T_c^{n/2}$. Thus, δV scales approximately as $T_c^{n/2-2}$. If n = 4, the signal size stays constant as Tc is reduced. If n = 5 or 6, the signal size is larger at higher Tc's.

In these devices the signal to noise ratio is maximized by maximizing the signal size. Thus, it is desirable to fabricate films with T_c's as high as possible, as long as the intrinsic pulse duration remains long compared to the RC time constant of the film biased in the middle of its transition. Our 70 mK W films have pulse durations of hundreds of μs, which is long enough to bias in the middle of the transition. It would thus be inadvisable to use films with lower Tc's for this application.

5.4. Fabrication of a W ZTD-TES

In each case, the device was fabricated on a 0.5 mm thick, 4" diameter Si wafer of [100] orientation. The wafers used in these experiments were off-the-shelf silicon from the stockroom of the Center for Integrated Systems at Stanford. The doping of the silicon varied, but standard wafers of higher resistivity were chosen (usually around 20 Ω cm). At this level of doping, isotopic scattering still dominates the phonon scattering, so paying more for higher purity silicon was unnecessary. When future experiments are done with charge collection, higher purity silicon will be important.

Our tungsten films are deposited as described in section 3.2. After deposition, the wafers are patterned photolithographically. Before coating, the wafers are baked in the singe furnace at 150 C for 30 minutes. The standard resist is applied in a Silicon Valley Group coater model #8626, which has its own singe and postbake.

The masks for the patterning were initially designed using VEM, and are now designed using Magic. The Dracula program is used to translate these files into OCT,

then CIF. Dracula then converts the file into the MEBES format necessary for the actual mask making. The MEBES files are transferred to magnetic tape to carry them to the mask making facilities.

The masks are made on the Center for Materials Research MEBES machine (Manufacturing Electron Beam Exposure System) by Paul Jerabek. The masks are made from a glass plate coated with chromium. The plates are covered with PMMA, a positive resist that is cross-linked with an electron beam. The MEBES machine projects an electron beam that is deflected by a magnetic field. It is capable of writing 0.25 μm features. After exposure in the MEBES, the PMMA is developed, and a chromium etch is used to create the final mask.

The aligner used is an Ultratech 1000 projection aligner in CIS. Unlike the Canon aligner that was used in previous work, the Ultratech projects a 1:1 pattern. The first several years, the aligner was run by Marnel King of CIS. Later, this author used the machine to pattern standard 500 μm thick 4" wafers. B. Chugg and B. Young have now modified this machine to allow the processing of thick Si and Ge wafers.

The resist is developed in a Silicon Valley Group developer model #8632, which does an automatic postbake.

Our W wet etch is done in the metal wet bench in CIS. A fresh solution of 30% hydrogen peroxide is always used. Etch times are typically 200 seconds for 30 nm thick films. Because surface residues vary, the activation time causes substantial variation in this number. Our results have always been excellent with 1 μm thick lines. The resist is stripped only with the clean PRS 1000 dip. It has been found that the EMT resist strip etches tungsten, so this strip is NOT used.

It should be noted that the dry etch that Adrian Lee developed for ultra-thin titanium lines [12] works well with W, but the success of our wet etch makes a dry etch unnecessary, and indeed inadvisable, as our bi-layer processes benefit from the smooth edges wet etching provides.

After processing, the wafers are diced into 1 cm x 1 cm dies using a Kulicke & Soffa wafersaw. Industry standard blue sticky tape (Semiconductor Equipment Corporation medium tack blue stickytape) is used to protect the wafer surfaces (both back and front) during the dicing. This tape leaves very little residue, although we have had some recent problems with a particular batch.

Our high impedance devices are patterned into a two channel detector. Each channel is 2 mm x 4 mm, and is composed of 400 lines on a 5 μm pitch. Each meander line is thus 1.6 m long. The channels are adjacent to each other in a 4 mm x 4 mm square. For this work, devices were patterned both as 1 μm wide and 2 μm wide lines

5.5. 6 keV x-ray Experiments

The ZTD-TES was first tested in an Oxford 400 toploading dilution refrigerator in collaboration with Prof. Doug Osheroff, Ben Tigner, and Dominic Salvino. After initial data was taken, devices were cooled in our Kelvinox 15 dilution refrigerator. As noted in Chapter 4, this refrigerator was wired with two channels of Sony 3SK164 FET preamps on the 1K pot.

An ^{55}Fe source was mounted on the backside of the crystal. It was collimated by a slit across the centers of both patterns. 5.89 keV and 6.49 keV x-rays were thus incident on the crystal from this source (fig 5.1).

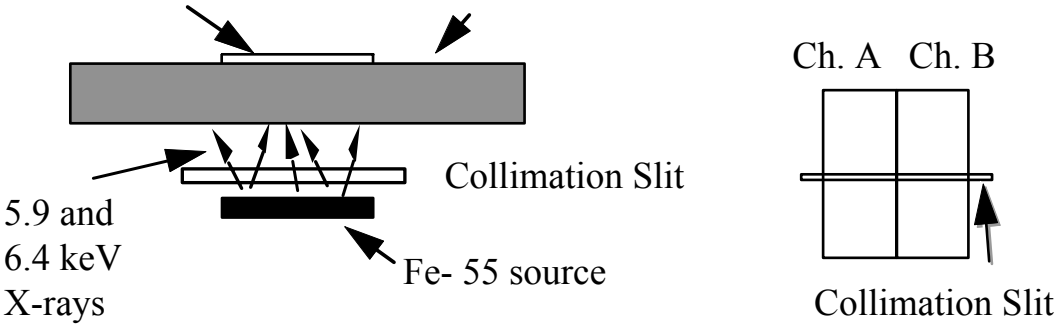


Fig 5.1 The source and detector geometry

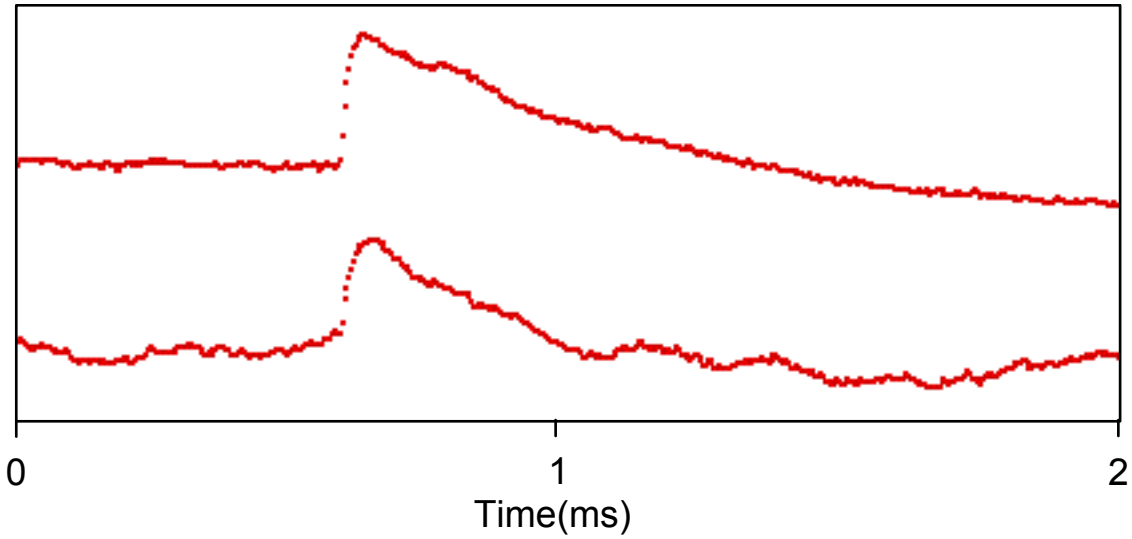


Fig 5.2 Coincident pulses in channel A and B.

The film was biased with a 1 nA current. When the film was biased in temperature within its superconducting transition, voltage pulses were observed with typical heights of 20μV, and durations on the order of 500μs (fig 5.2).

All pulses were seen to be coincident in both channel A and B. In fig 5.3, the pulse height in channel A is plotted against the pulse height in channel B for a set of several thousand coincident pulses. Triggering is done on either pulse A or pulse B. As can be seen in the figure, the data is clearly pulled away from the noise events near the origin. This implies that events throughout the pulse volume were observed with a certain minimum pulse height. This result is important for instrumenting large mass scale detectors. Note that in our Ti detector, only events directly underneath a channel yielded a signal in that channel. Only a small fraction of the pulses were coincident (those just on the boundary between channel A and channel B.)

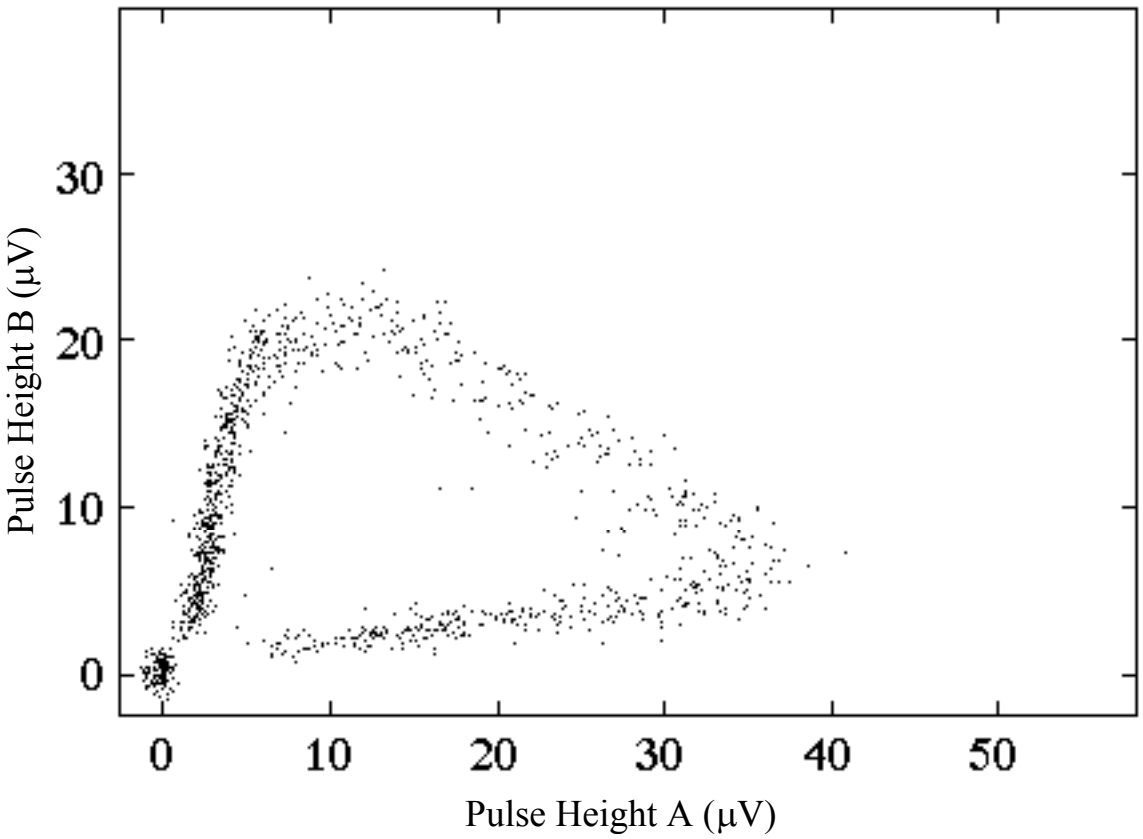


Fig 5.3 A plot of the pulse height in channel A verses channel B. All events are coincident.

5.6. A Monte Carlo model

To explain this data we have created a simple model of phonon propagation throughout the crystal (see Appendix A). Phonons are assumed to propagate ballistically through the crystal. At surfaces, the direction of phonons is randomized (fig 5.4). The geometry of the detector is modeled (fig 5.5a). When a phonon impinges on the W film or a heat sink, it is absorbed with a probability of 0.3. The phonon continues to scatter until it is absorbed either in a heat sink or a TES. The signal in each channel is assumed to be proportional to the total number of phonons it absorbs.

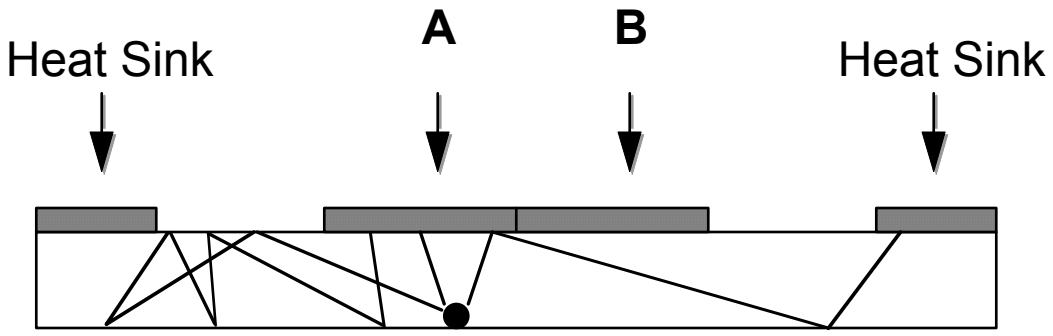


Fig 5.4 The Monte Carlo model of phonon propagation. A side view of the sensor is shown. The interaction occurs near the backside. Phonons scatter until absorbed in a sensor or heat sink.

In fig 5.5b, a Monte Carlo model for an x-ray interaction in the center of the chip is shown. 10,000 phonons are modeled, and a dot is placed at the site of the absorption of each phonon. The full trajectory of several phonons is traced for illustration.

To model our pulse height A vs. B results, 1,500 x-ray events were modeled. The x-ray events were evenly distributed throughout the collimation slit. The y-position of the event was taken to be zero, and the x-position was a random number across the 10 mm wide dimension of the crystal. For each x-ray event, the trajectories of 1,000 phonons were followed. The pulse height in each channel was assumed to be proportional to the number of phonons absorbed. The results of this Monte Carlo model are compared to the data in fig 5.6. The model succeeds in reproducing the qualitative features of the data. It should be noted that the spread of the data in the Monte Carlo is statistical in origin, depending on the number of phonons modeled. The spread in the real data is due primarily to amplifier noise and microphonics.

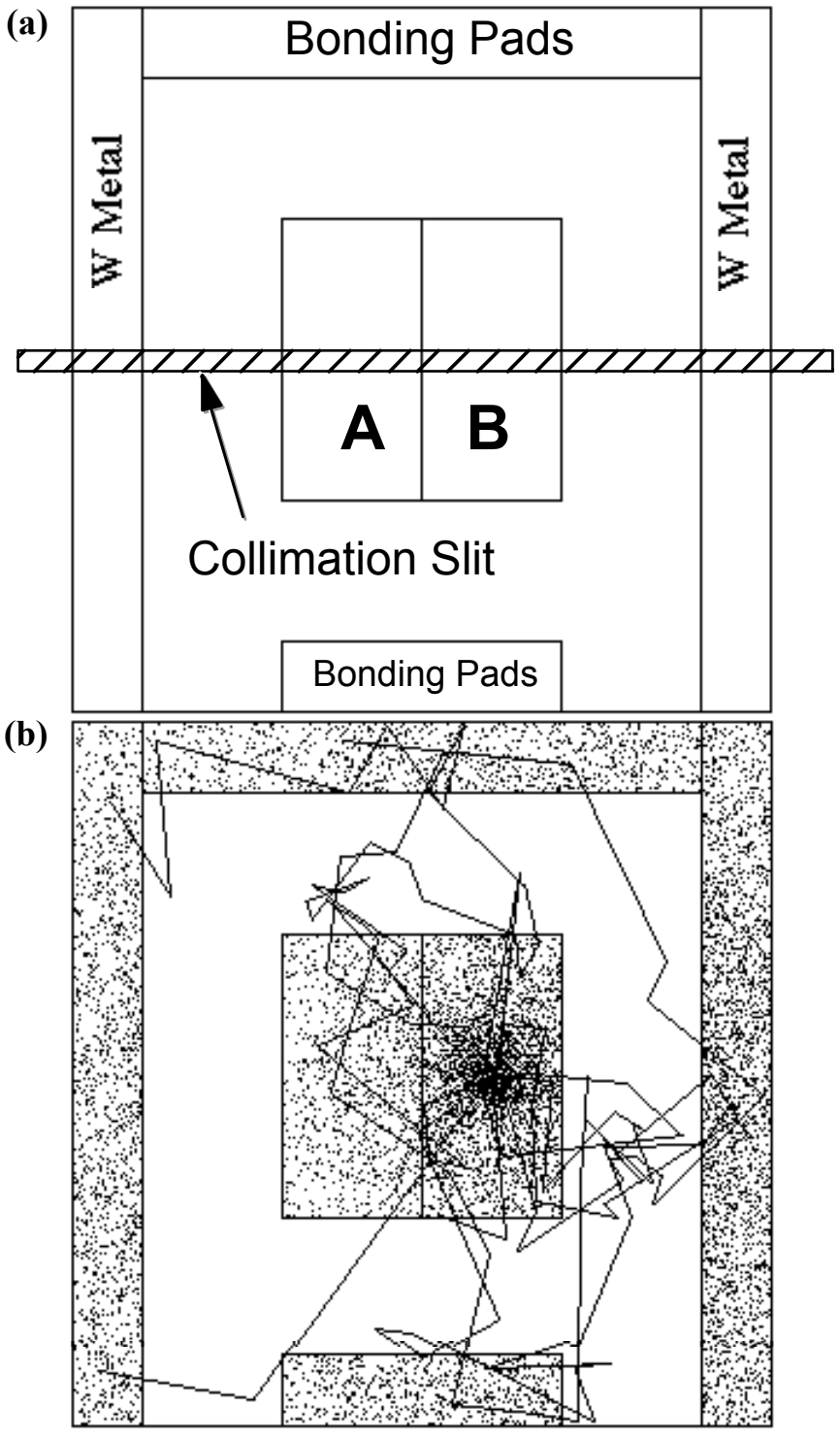


Fig 5.5 A top view of (a) the detector geometry and (b) the Monte Carlo. Every point represents the absorption of a phonon in a sensor or heat sink. Several complete phonon trajectories are shown.

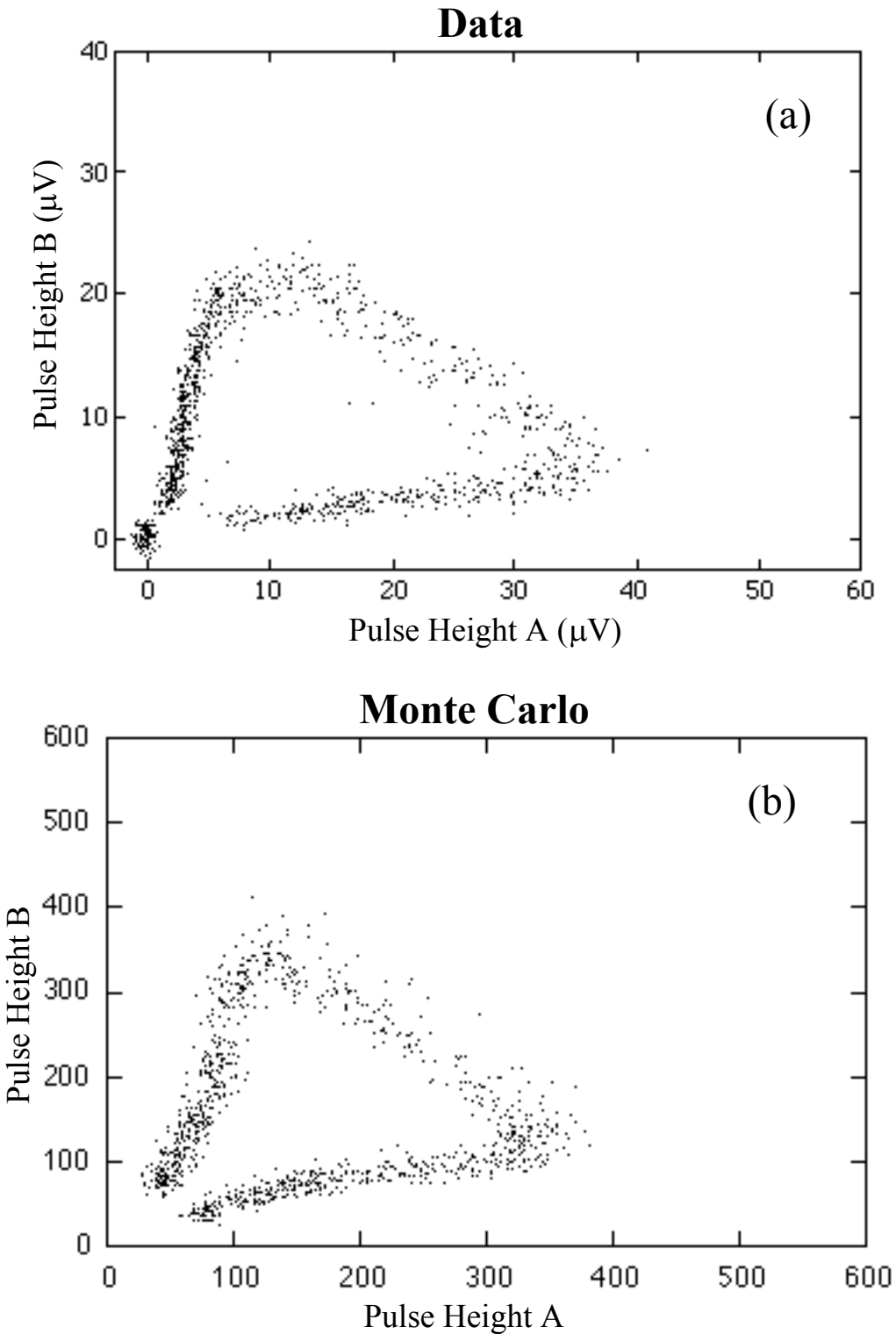


Fig 5.6 Pulse height A vs. B plots. (a) Experimental data. (b) A Monte Carlo phonon scattering simulation incorporating the detector geometry, and both k_α and k_β x-rays.

5.7. Deconvolution of the event position and energy

The pulse height A vs. B plot contains both event position and energy information. In order to calculate the event energy it must first be deconvoluted from the event position.

Events occurring near the center of the crystal cause large pulse heights in both channels. If an event occurs closer to the center of one channel, the pulse height in that channel becomes larger, and the pulse height in the other channel becomes proportionately smaller. Events occurring between the centers of the two patterns lie on the linear portion of the data with large pulse height in both channel A and B. If events occur closer to the edge of the crystal, more phonons are absorbed in the heat sinks instead of the sensors, and the pulse height in both channels becomes smaller. These are the branches of the data that curl back towards the origin. By plotting pulse height A vs B, we have event position resolution in one dimension to about a millimeter in the region close to the sensors.

It is reasonably uncomplicated to deconvolute the position from the energy for events in the 2 mm wide region between the centers of the two patterns. In this region, the data falls approximately along a line, so the event energy is just the sum of the pulse height in channel A and B, once gain differences in the two channels are adjusted. An energy histogram for events between the centers of the two patterns can thus be made by cutting the data that lies outside of this region, adjusting for gain differences between the two channels, and adding the pulse heights.

In fig 5.7a, a data set is shown with two lines separating the regions of data that will be cut, and a 45˚ line that lies along the linear portion of the data once the gains are properly adjusted. Fig 5.7b shows a histogram of the gain-adjusted sum of the two pulse heights, without data cuts. Fig 5.7c shows the data that remains after the cuts are done, and fig 5.7d shows the final histogram of the gain-adjust pulse height sum after the two cuts. In the histograms the peak is normalized to a 6 keV event energy.

Once the position is deconvoluted from the energy, this ZTD-TES sensor thus measures the energy of the event with a FWHM of 1 keV, and measures the event position between the centers of the two sensors to within ~ 1 mm. There is some sign of a high energy shoulder which may be associated with a k_β x-ray peak.

Application of our Monte Carlo model to different detector geometries shows that careful design of our sensor geometry and heat sinks will make it possible to deconvolute position and energy over the entire crystal, and reduce the signal that is lost to heat sinks.

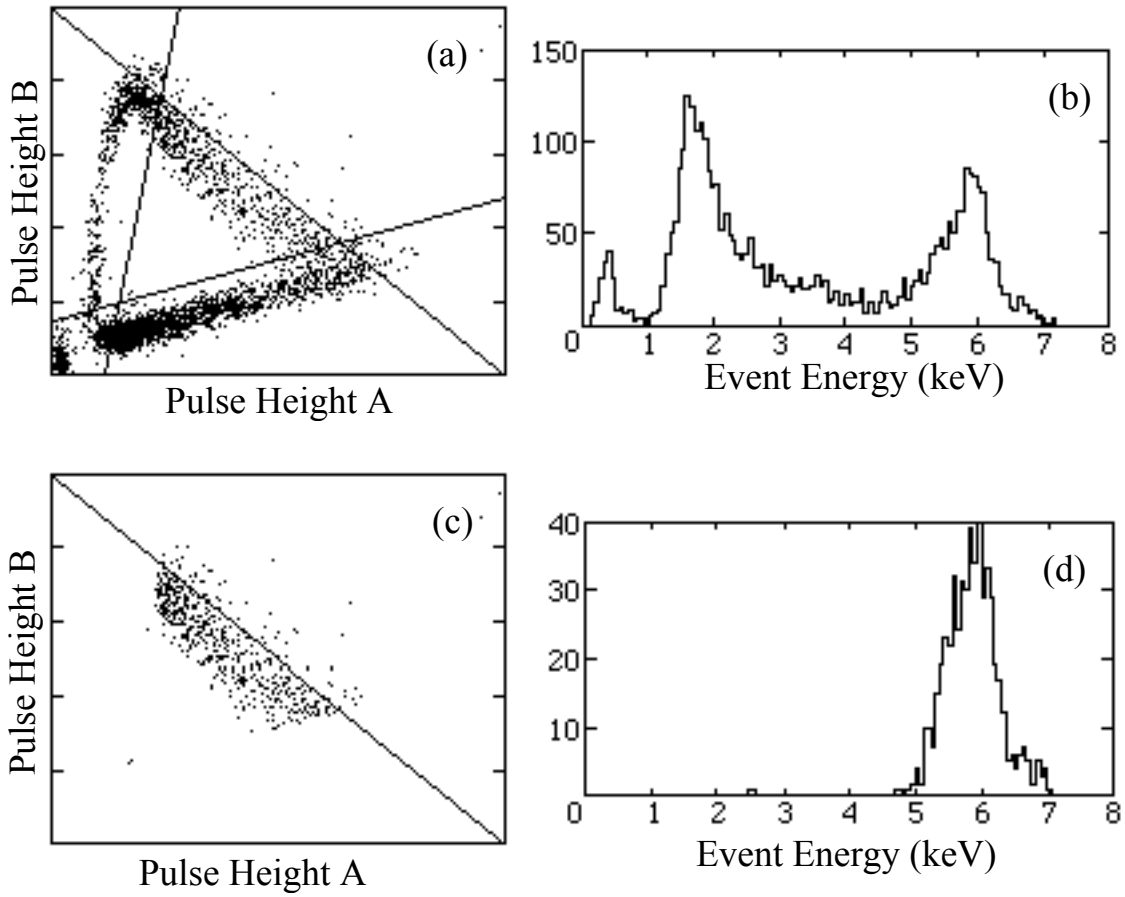


Fig 5.7 The energy histogram for a ZTD-TES data run. (a) A data set with lines showing the cuts that will be made, and the direction of the histogram projection. (b) The histogram of pulse height A + B without cuts normalized to 6 keV at the peak. (c) The data set after cuts. (c) The pulse height A+B histogram with data cuts, normalized to 6 keV at the peak.

5.8. Success of the W ZTD-TES

We have successfully created a W ZTD-TES, in order to instrument larger absorber volumes. This detector enables us to observe 6 keV x-rays through a 500 μm thick silicon crystal with a 1 keV FWHM resolution - a major improvement over our Ti TES.

A figure of merit for a large absorber volume phonon sensor is the square of the length scale of the absorber divided by the minimum detectable energy. For Ti TES, 60 keV x-rays could just be resolved through 300 μm of silicon, for a figure of merit of 1500 mm^2/keV. The W ZTD-TES can resolve 6 keV x-rays through 500 μm of silicon,

implying a figure of merit of greater than 41,700. Thus, the W ZTD-TES can be said to be at least 28 times more sensitive than our previous Ti TES.

Unfortunately, the signal to noise ratio of these devices, and hence resolution, is not as good as we would like. Limitations on the latching current force us to bias these detectors at ~ 1 nA, as compared to hundreds of nA for our Ti sensors. This effect leads to a very small signal size. In addition, the ZTD-TES proved to be much noisier than the Ti TES. Since the ZTD-TES is biased in the middle of its superconducting transition, its impedance during an event is ~ 1 MΩ. The Ti TES has an maximum impedance of ~ 10 kΩ during an event. This causes the ZTD-TES to couple much more strongly to microphonics and other external sources of noise.

Two other approaches to overcoming the limitations associated with Ti transition-edge sensors have been explored in this work. The first was the construction of a "quantum efficient" sensor, the QE-TES. In this device, the W film is biased below its superconducting transition temperature. As such, its operating impedance should be low enough to be much less sensitive to microphonics and other sources of noise. This approach will be described next.

CHAPTER 6: THE QUANTUM EFFICIENT TRANSITION EDGE SENSOR

As has been discussed, since the bias temperature of the film in Ti sensors is below the transition temperature, only sections of the film that receive a critical density of incident phonons (about 400 per 2 μm x 2 μm square for Ti) are driven normal. Ti TES detectors thus have an intrinsic threshold energy density below which they are insensitive. This threshold places an intrinsic limit on the detectable incident particle interaction energy, and the absorber volume. The phonon density scales linearly with the particle interaction energy. In the approximation of an infinite crystal, the phonon energy density drops as the square of the distance from the interaction. Thus, the minimum energy detectable throughout the absorber volume drops as the square of the absorber thickness.

This intrinsic limit can be avoided by biasing in the middle of the transition, as in the ZTD-TES. Alternatively, this limit can be made unimportant by reducing the heat capacity of a film to the point where a single athermal phonon drives a complete line segment normal. When this performance is reached, the fact that phonon energy is quantized allows much thicker absorbers to be instrumented, limited only by phonon down-conversion and the detector area that must be covered. Such a device is called a Quantum Efficient Transition-Edge Sensor (QE-TES).

When a film is biased sufficiently far below its superconducting transition temperature, to first order the resistance of a pixel may be assumed to change discontinuously from zero to the normal resistance. The threshold surface energy density (E_σ) is the energy required to drive a unit area of film normal under these conditions. Near T_c, according to BCS theory,

$$E_\sigma \approx 5N(0)\Delta_0^2 d\,\Delta T/T_c \tag{6.1}$$

where $N(0)$ is the density of states, $\Delta_0 \approx 1.76kT_c$ is the gap, d is the film thickness, and $\Delta T = T_c - T_{bias}$ is the temperature excursion necessary to drive the film normal (typically at least the transition width). Assuming that the film is biased one transition width below T_c, and that the density of states is the same as in bulk Ti, our best Ti film to date has a calculated threshold energy density $E_\sigma \approx 130$ meV/μm^2 for a 40 nm thick film [12]. Assuming the density of states in our W films to be the same as in bulk W, our best W film to date has a calculated $E_\sigma \approx 2$ meV/μm^2. The mean phonon energy in our detectors is about 4 meV (1 THz). Thus, to create a quantum efficient detector with a W

film, the W segment size would have to be reduced to dimensions of less than 1 μm x 2 μm. To create a quantum efficient detector with Ti, the size of a segment would have to be reduced to dimensions of less than 0.17 μm x 0.17 μm.

6.1. Attempts to construct a titanium QE-TES

To create a Ti segment that is quantum efficient for a typical athermal phonon, the size of a "segment" would have to be reduced to dimensions of less than 0.17 μm x 0.17 μm. If the heat were to escape into the substrate on time scales short compared to the time for heat to flow 0.17 μm down the Ti line, then each 0.17 μm x 0.17 μm square "pixel" of line would be thermally isolated on time scales of the pulse duration. Under this assumption, a Ti line of width < 0.17 μm would be expected to be borderline quantum efficient.

Our group has fabricated a Ti TES detector with lines ~ 0.15 μm wide as a first step towards such a detector. Quantum efficiency was not observed[12]. Even when the film was biased in the middle of its transition, where some portion of the line would be expected to be quantum efficient, no sign of approaching quantum efficiency was evident. This observation was understood with the realization that adjacent 0.25 μm x 0.25 μm pixels are not, in fact, thermally isolated on timescales of the pulse duration. The exponential time constant for heat flow between adjacent pixels is, to first order, the same as that for a thermal mass connected to a heat bath through a thermal impedance, $\tau = C/g$. For two adjacent Ti pixels, this time constant is ~ 10 ns, much shorter than our 3 μs pulse durations, and shorter than our electronics limited risetime.

Thus, in order to achieve quantum efficiency, superconducting film segments must be separated by thermally insulating segments to prevent the energy of an incident phonon from escaping down the line before the signal can be measured. When a superconductor is cooled to well below its transition temperature, effectively all of the electrons near the Fermi surface are bound in Cooper pairs. Since Cooper pairs do not scatter, they do not conduct heat. Thus, a superconductor with a Tc much higher than the bias temperature can be used to provide a thermally insulating, yet electrically conducting line segment.

Unfortunately, it is prohibitively difficult to create a quantum efficient detector with Ti using alternating segments of different material. The Ti segments would have to have areas smaller than 0.17 μm x 0.17 μm, and have good electrical connection to the high Tc segments, a difficult feat for a two layer process. This size segment would also lead to small Ti area coverage.

Fortunately, as shown above, a segment of W film with an area of 2 μm^2 would be borderline quantum efficient. The construction of such a two layer device is feasible with technology available to us at Stanford.

6.2. The QE-TES and the proximity effect

One possible problem with the QE-TES is the proximity effect. In the QE-TES, the thermally insulating, higher T_c superconducting segments are in electrical contact with W film segments. Thus, there will be a leakage of pairs from the higher T_c material into the W segment, and a leakage of quasiparticles from the W segment into the higher T_c material.

The Fermi velocity of electrons in W is $v_f = 1.35 \times 10^8 cm/s$ [18]. The density of states in W is $N(0) = 0.16 \times 10^{23} eV^{-1} cm^{-3}$ [18]. The effective density of electrons in W is thus $n = N(0) m_e v_f^2 / 3 = 5.49 \times 10^{22} cm^{-3}$, where m_e is the mass of the electron. The mean free path of an electrons gas is

$$l = \frac{m_e v_f}{n e^2 \rho} \tag{6.2}$$

where ρ = 6 μΩcm is the measured resistivity of our W films, and e is the charge of the electron. Eqn 6.2 yields a mean free path in our W films of ~ 15 nm, which is half of the thickness of our W films.

The effective coherence length in a material sets the scale of the proximity effect. The coherence length for a normal metal in the dirty limit is [19]

$$\xi_n = \sqrt{\frac{\hbar v_F l}{6 \pi k_B T}} \tag{6.3}$$

Thus, at 70 mK, the coherence length in normal W metal is calculated to be $\xi_n \approx 0.34 \mu m$. This parameter sets the length scale of the proximity effect.

The W film segments must be kept longer than the length scale of the proximity effect, or the W film transitions will be elevated, and possibly broadened. In order to fabricate films with length scales long compared to the proximity effect, and yet keep the film segment heat capacity small, it may be necessary to fabricate devices with 10 μm long, with lines ~ 0.1 μm wide. Our group has fabricated 0.15 μm wide Ti lines [12] using a technique that could be extended up to 0.1 μm wide lines.

Note also that, since the W film is actually a superconductor slightly above its transition (where the electron-electron interaction potential is non zero) the scale of the proximity effect may be extended. This point will be discussed further in chapter 10.

6.3. Fabrication of a candidate QE-TES

The simplest metal to use as a thermal insulator would be aluminum. Aluminum films superconduct at ~ 1K, and at 70 mK are far enough below their transition to be thermally insulating. Aluminum is a metal commonly in use in the electrical engineering industry, and so is readily processed in the Center for Integrated Systems at Stanford, where our devices are made.

We fabricated a candidate QE-TES at CIS that consisted of a series array of 3 μm wide W segments that were connected by 3 μm x 3 μm Al pads. The techniques used in this fabrication are described fully in section 5.4 and 10.6. Because of a slight misalignment, this sensor was not continuous, and could not be tested. The alignment problem was solved, but another candidate QE-TES was never developed due to our more promising development of a competing technology, the Electrothermal Feedback Transition-Edge Sensor, which will be discussed in the following chapters.

6.4. Avoiding the precision alignment step with the proximity effect

We have examined four different techniques to create a quantum efficient detector without the necessity of a high precision alignment step. Eliminating the precision alignment step (substituting a step where precision is less important) might allow the low T_c W segments to be much shorter, making submicron processing unnecessary, and improving their area coverage by reducing the size of the high T_c segments.

6.4.1. Using the proximity effect

The simplest technique would be to make use of the proximity effect. In this technique, the W film is patterned into a single W line, 1 μm wide, and about 2 m long. On top of this W line, an aluminum film is deposited, and patterned into bars of Al 1 μm wide, 3 μm long, and spaced apart 2 μm center to center (fig 6.1).

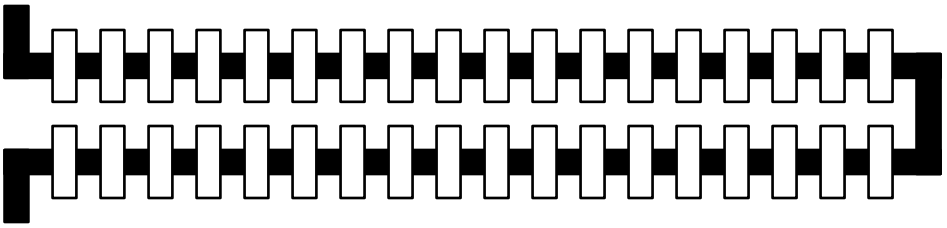


Fig 6.1 A proximity effect based QE-TES. A continuous W line is covered by short Al segments. The proximity effect raises the transition temperature of the W film segments they cover.

As shown previously, a simple calculation indicates that the appropriate coherence length for our W films is about 340 nm. Since our films are only 30 nm thick, the T_c of the W film under the aluminum film should be raised to well above 70 mK by the proximity effect. If this effect is pronounced enough, it might cause the W films under the Al to be thermally insulating. Thus the 1 μm long W segments might be thermally isolated, even though the W line remained physically and electrically continuous.

Our first round of candidate QE-TES detectors included several patterns to test this technique. Time constraints prevented these detectors from being tested, as we turned our attention to the more promising direction of the ETF-TES detectors.

6.4.2. Using selective CVD deposition of the high T_c segment

The second way to avoid the alignment step would be to use the high-T_c, or β crystal phase of W as the higher temperature superconductor [20]. In this process, α phase (low-T_c) films are covered with an oxide layer, then patterned. β phase (high-T_c) films are then selectively deposited between the low-T_c segments by low pressure chemical vapor deposition (LPCVD). Since the process is selective, the beta W deposits only on the bare silicon and the exposed alpha tungsten, not on the portions of alpha tungsten covered by the oxide layer. Films with transitions as high as 2 K can be deposited selectively [21].

A possible problem with this technique is that an oxide layer may form between the alpha and beta tungsten segments. This technique has not been tested, because of time constraints.

6.4.3. Using electron beam annealing

It has been observed that analysis of W film crystal phase by electron beam diffraction can actually cause the beta phase of W to convert to the alpha phase. If, as expected, the phase as measured with electron diffraction turns out to correspond to the high-T_c and low-T_c phases of W, this technique might be used to create a W QE-TES.

In this approach, a W film would be deposited in the high-T_c, beta crystal phase. Such a film can be created by sputtering with the Balzers DC Magnetron with a room temperature substrate and no DC bias voltage. This film would then be patterned into a long meander line 1 μm wide and about 2 m long. An electron beam would then be used to transform alternating 1 μm segments of W film into the low-T_c alpha phase by annealing. A 1 μm wide beam would be moved in stripes perpendicular to the direction of the W line, on a spacing of 2 μm. Any alignment to the W lines could be quite rough. The alternating T_c segments would automatically be in intimate electrical contact, with no overlap.

The MEBES machine which is used to make masks for CIS might be used for this purpose. Since this is the machine used to make masks, the electron beam exposure pattern can be specified in the same way as it is for masks. A set of MEBES files were created to test this possibility. They included a large square to test if the MEBES would indeed convert the high T_c phase to the low T_c phase, and a set of alternating stripes to fabricate a candidate QE-TES detector.

This possibility was never tested, again due to time constraints, and the pursuit of the ETF-TES. A number of questions about it remain. First, does the conversion of crystal phase as seen by electron diffraction correspond to a conversion of W film T_c? Would the MEBES machine be powerful enough to anneal the beta phase into the alpha phase? Would the effect be localized, or would the heat from the electron beam anneal adjacent segments as well? Would the heat from the annealing convert the tungsten film to tungsten silicide?

An effort is currently under way to examine some of these questions at thin film lab at San Francisco State University [22]. W. Owens and B. Neuhauser are converting an electron microscope to attempt to use it to anneal W film. These questions will hopefully be answered in the future.

6.5. QE-TES as a phonon energy spectrometer

The minimum phonon energy needed to drive a line segment normal is controllable. A quantum efficient device might therefore be used as a phonon energy spectrometer. The threshold energy density of the film, near T_c, is proportional to $T_c - T_{bias}$. Thus the phonon energy spectrum of a series of pulses might be measured by varying the bias temperature. The energy needed to drive a line segment normal is proportional to the area of the segment. A single detector could have line segments of different areas, allowing the measurement of the phonon energy spectrum of a single particle interaction.

If electron and nucleus scattering events do indeed lead to different phonon energy spectra, this effect might be used as a way to discriminate between electron and nucleus scattering events, without the necessity of charge collection.

CHAPTER 7: THE ELECTROTHERMAL FEEDBACK TRANSITION EDGE SENSOR

Some progress has been made towards the fabrication of a dark matter or neutrino detector based on either the ZTD-TES or the QE-TES. Unfortunately, there are several limitations that would make it difficult to scale these technologies up for a practical large absorber volume detector.

The first limitation is a disappointing signal to noise ratio. Limitations on the latching current forced us to bias these detectors at ~ 1 nA, as compared to hundreds of nA for our Ti sensors. This effect leads to very small signal size. At the same time, the high bias impedance of the ZTD-TES W sensors strongly couple to microphonics and other external sources of noise. While we expect the noise to be less in a QE-TES, we do not expect the signal size to be greatly increased.

A second, more critical limitation of these sensors is the gradient in superconducting transition temperature that they exhibit. The silicon crystal can only be biased at one temperature. Unfortunately, our W films have a significant variation in T_c across the wafer. In fig 7.1, the transition of two side by side W meanders is plotted. In this case, each meander forms a 2 mm x 4 mm rectangle. The two channels are side by side so as to form a 4 mm x 4 mm square.

As can be seen, the transitions in this case are about 1.0 mK wide. More significantly, due to a gradient in the Tc, they are offset by 0.6 mK. In order to run this detector in a coincident mode, the silicon must be biased in temperature in the top third of the transition of one channel, and in the bottom third of the transition of the second channel. One channel is thus much noisier than the other, as it has a higher impedance and couples better to microphonics and other external sources of noise.

If this difference in T_c between the two channels is caused by a gradient in T_c, there is a very large variation in sensitivity at different points in the W film. Sections further below the transition take more energy to be driven normal. Even worse, sections biased above the transition are normal, and completely insensitive to phonon signals. They act merely as phonon absorbers, without returning any useful information. At these bias conditions, these completely insensitive sections make up a full half of the W coverage.

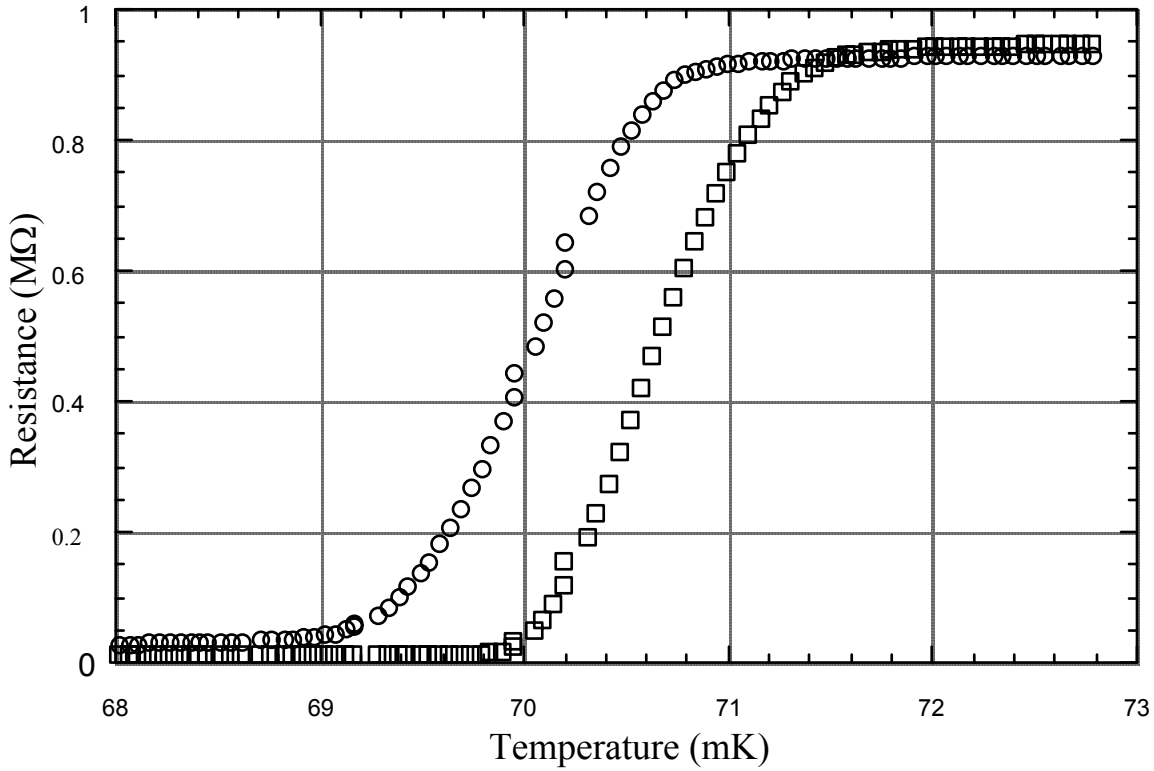


Fig 7.1 Superconducting transitions of two tungsten pads on the same chip. The transition width is narrow (~1.0mK), but the transitions are offset by ~ 0.6 mK.

The variation in T_c between these two channels is bad enough, but if an entire 3" wafer were instrumented, the gradient in T_c might be much worse than that across 4 mm. If the difference in T_c in different channels across the detector is larger than the transition width, it will be completely impossible to bias them all in the transition, with no phonon threshold, ruling out a W ZTD-TES with widely spaced channels. It will likewise be impossible to bias below the transition, but close enough to be quantum efficient, ruling out a W QE-TES with widely spaced channels.

Concerns about signal to noise ratio and possible area coverage led to a search for a detector technology without these limitations. This search led to the Electrothermal Feedback Transition Edge Sensor (ETF-TES).

7.1. A self-biasing technology

The ETF-TES uses electrothermal feedback to hold the temperature of a superconducting film constant within its transition, while the substrate is cooled to well

below the transition temperature of the film. In addition to many other improvements, this technique allows different channels across a wafer to be biased near the center of each of their transitions, even though the transition temperature varies between channels.

In the ETF-TES, a low impedance superconducting film is voltage biased, and the substrate is cooled to well below the transition temperature. As the film cools, its resistance drops, and the Joule heating increases. When the Joule heating matches the heat loss into the substrate, a stable equilibrium is established. The bias resistance can be chosen by varying the bias voltage. The detector self-biases on its transition.

A change in temperature causes a change in Joule power dissipation, which in turn affects the temperature. This kind of process is referred to as electrothermal feedback. Since temperature excursions are reduced in the voltage biased case, this effect is a negative feedback process. In the current biased case, Joule power increases deviations from equilibrium, and the electrothermal feedback is positive.

Clearly, this technique makes differences in T_c between different channels less important, as each channel is biased at a different temperature. It should also be noted that this self-biasing behavior can be used to make a gradient in transition within a channel less important as well. While the ZTD-TES and QE-TES are inherently series devices, the ETF-TES works in a parallel configuration. If the W film is separated into many short segments that are electrically connected in parallel, but thermally isolated, each will independently bias in the middle of its transition. Since each of these segments will cover a very small area, any gradient in T_c will be unimportant. This technique is used in our W/Al hybrid quasiparticle trap assisted devices, which will be discussed later.

The self-biasing nature of this detector technology makes scaling up to a larger detector much more straightforward. There are, however, other important advantages to this technology. Since the substrate temperature is allowed to cool to the base temperature of the refrigerator, temperature regulation is unnecessary, eliminating the EMI noise that is coupled into the probe by the temperature controller. Instabilities in the temperature are also less important, as long as $T_{sub}^n << T_c^n$. This technology also has better performance in resolution, linearity, dynamic range, and count rate.

7.2. Advantages of running with negative electrothermal feedback

Superconducting transition edge sensors suffer from limitations due to film nonuniformity, transition nonlinearity, and limited dynamic range. In electrical circuits, analogous limitations are often avoided by using feedback techniques. The Electrothermal Feedback Transition Edge Sensor implements this scheme by using

feedback to hold the temperature of the film constant. The feedback signal is the required heat input.

Joule heating is used to provide this negative feedback. As described above, when a film is voltage biased, and the substrate is cooled to well below the transition temperature, the film can be made to self-regulate in temperature within its transition. The feedback signal is the change in Joule power in the film, the product of the bias voltage and the current, which is measured with a SQUID (Superconducting Quantum Interference Device) amplifier.

7.3. Pulse shortening

It has been noted that electrothermal feedback can effect the pulse duration [23,24,17]. Under voltage bias, pulses are shortened. If film transitions are sharp and the substrate is cooled to well below the transition of the film, this effect is quite large.

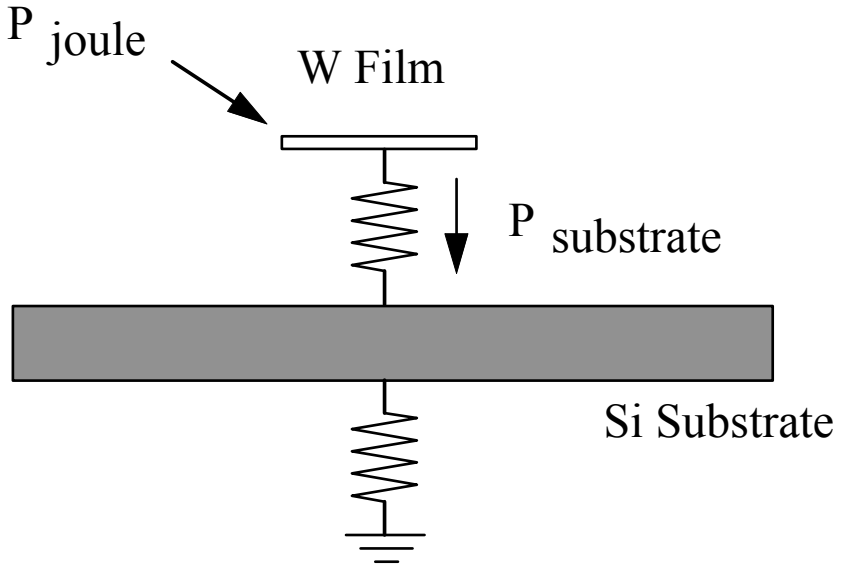


Fig 7.2 The power flow model for the ETF-TES. Joule power flows into the film, through a thermal impedance into the substrate, and finally into the heat sink.

In order to study the effect of feedback on pulse duration, we model the detector as a film of heat capacity C connected to a substrate at a constant temperature T_s. The rate of change of film temperature, T, is determined by the film heat capacity and the net heat flow.

$$C\frac{dT}{dt} = P_{in} - P_{out} \tag{7.1}$$

The film is biased at a constant voltage, V. The power input is the Joule power, V^2/R. The heat loss to the substrate goes as $K(T^n - T_s^n)$, where K is a material and geometry dependent parameter and n is a number whose value depends on the dominant thermal impedance between the substrate and the electrons in the superconducting film. This thermal impedance is set at higher temperatures and in thicker films by the Kapitza boundary resistance between the film and the substrate. For thinner films, and at lower temperatures, the electron-phonon decoupling in the film dominates [15,16]. If the Kapitza resistance dominates, n is 4. If electron-phonon decoupling in the film dominates, n is either 5 or 6, depending on the theory and the temperature range.

Thus, the power flow equation is

$$C\frac{dT}{dt} = \frac{V^2}{R(T)} - K\left(T^n - T_s^n\right) \tag{7.2}$$

For events heating the film a small amount, ΔT, above the equilibrium temperature, the power flow equation may be expanded to first order. In this expansion, the zeroth order (equilibrium) terms cancel. We are left with the first order terms

$$C\frac{d\Delta T}{dt} = -\frac{V^2}{R_0^2}\frac{dR}{dT}\Delta T - g\Delta T \tag{7.3}$$

where the thermal conductance $g = dP/dT = nKT^{n-1}$. Now we substitute $\alpha = (T/R)(dR/dT)$, a unitless measure of the sharpness of the superconducting transition, and P_0, the equilibrium Joule power.

$$C\frac{d\Delta T}{dt} = -\frac{P_0\alpha}{T_0}\Delta T - g\Delta T \tag{7.4}$$

The solution of this differential equation is a simple exponential, with a time constant τ_{eff}. Substituting $P_0 = K\left(T_0^n - T_s^n\right)$ and $g = nKT^{n-1}$,

$$\text{for negative feedback:} \quad \tau_{eff} = \frac{\tau_0}{1 + \frac{\alpha}{n}\left(1 - \frac{T_s^n}{T_0^n}\right)} \tag{7.5}$$

where $\tau_0 = C/g$ is the intrinsic time constant of the film (the pulse recovery time constant in the absence of Joule heating).

Note that in the case of positive electrothermal feedback (current bias) the time constant is lengthened instead of shortened:

$$\text{for positive feedback:} \quad \tau_{eff} = \frac{\tau_0}{1 - \frac{\alpha}{n}\left(1 - \frac{T_s^n}{T_0^n}\right)} \tag{7.6}$$

When the substrate temperature is much colder than the film, $T_s^n << T_0^n$, and we enter the extreme negative feedback regime. The exponential pulse recovery time constant for the voltage biased case is seen to be

$$\tau_{eff} = \frac{\tau_0}{1 + \alpha/n} \approx \frac{\tau_0 n}{\alpha} \tag{7.7}$$

in the extreme negative feedback regime

We have fabricated W films with α as high as 1000 [20]. Thus, electrothermal feedback can shorten pulse duration by two orders of magnitude. Under these conditions, major simplifications are possible in interpreting detector pulses.

7.4. Self calibration

When the pulse duration is much shorter than the intrinsic time constant, the detector is operating in a mode where the pulse energy is removed by a reduction in Joule heating, instead of an increased heat flow to the substrate. Thus, the energy deposited in the film is simply the integral of the change in the Joule power, or the voltage bias times the integral of the change in SQUID current:

$$E = \int \Delta P_{Joule} dt = V_0 \int \Delta I dt \tag{7.8}$$

This observation leads to a simplification in pulse analysis. Nonuniformities in the film properties and nonlinearities in the resistive transition will effect the pulse shape, but not the pulse integral. In addition, the dynamic range is improved. If an energetic event drives the film completely normal, the total energy is still measured. The saturation causes the pulses to lengthen, but energy loss is small as long as the pulse duration is still much smaller than τ_0.

This technique allows a direct measurement of energy collection efficiency. The energy is simply the voltage bias multiplied by the integral of the current change, as measured by the SQUID. Since there is no proportionality constant, this measurement has no free parameters, and can be compared directly to the incident energy.

7.5. Instrumentation of the ETF-TES

In its simplest form, the Electrothermal Feedback Transition Edge Sensor (ETF-TES) is a superconducting film on a silicon substrate. The film is voltage biased, and the current is measured with a SQUID (fig 7.3).

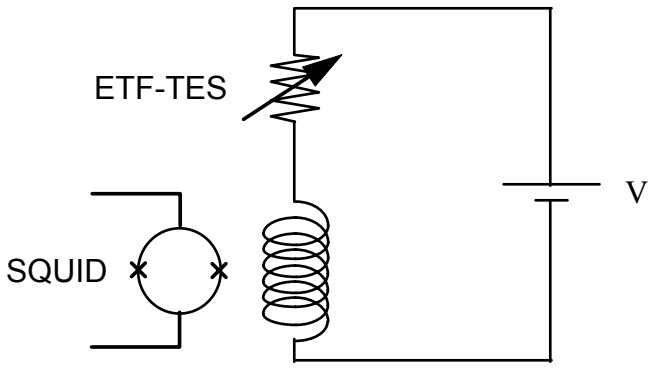


Fig 7.3 Instrumentation of the ETF-TES

The inductance of the SQUID sets a practical lower limit on the pulse duration. The L/R time constant (from the SQUID inductance and the resistance of the film) must be kept short compared to the effective pulse duration. If this is not done, resonant oscillations can occur which involve the electrothermal feedback and the SQUID inductance, as will be discussed in more detail later. This effect limits the minimum pulse duration to ~ 1 μs. Thus, to operate deep in the feedback regime, intrinsic time constants must be at least 100 μs.

In order to achieve intrinsic time constants on the 100 μs time scale, the film transition must be below about 100 mK. Our W films have transitions at about 70 mK. Intrinsic time constants we have measured for these films were experimentally found to be ~ 500 μs.

Conventional SQUID systems do not have large enough bandwidth to resolve pulses with effective time constants in the μs range. Fortunately, SQUID arrays have been developed that have large enough bandwidths for this purpose. These arrays have been installed on the 1K pot of our dilution refrigerator (see section 4.3).

CHAPTER 8: FUNDAMENTAL LIMITS ON THE ETF-TES ENERGY RESOLUTION

The energy resolution of an ETF-TES phonon-mediated particle detector will likely be limited by the phonon physics and energy collection statistics, rather than Johnson noise and phonon noise in the transition-edge sensor. Nevertheless, it is important to understand fundamental limitations on the transition-edge resolution itself. Also, in further applications of this technology where energy collection and thermalization is improved (for example, in high resolution x-ray detection) the actual resolution may approach the limit set by phonon and Johnson noise.

8.1. RMS Thermodynamic Energy Fluctuations

The rms thermodynamic energy fluctuations in the film may be calculated by considering it to be a system with a weak connection to a large heat reservoir (the substrate). The heat capacity of the system is $C = \partial\overline{E}/\partial T$. Now the mean energy $\overline{E}$ can be found from the first derivative of the partition function $\overline{E} = -\partial \ln Z/\partial\beta$, so

$$C = -\frac{\partial^2 \ln Z}{\partial\beta^2}\frac{\partial\beta}{\partial T} = \frac{\partial^2 \ln Z}{\partial\beta^2}\frac{1}{kT^2} \tag{8.1}$$

where we have used $\beta = 1/kT$. The dispersion of the energy can be found from the second derivative of the partition function $\langle \Delta E^2 \rangle = \partial^2 \ln Z/\partial\beta^2$. From Eqn. 8.1 and the expression for the dispersion we have the expression for the RMS thermodynamic energy fluctuations $\langle \Delta E^2 \rangle = kT^2 C$.

It is widely, and erroneously, believed that these thermodynamic fluctuations set a hard "thermodynamic limit", $\Delta E_{rms} = \sqrt{kT^2 C}$, as the ultimate possible energy resolution of a bolometric detector. It should be noted, however, that these fluctuations occur on the timescale of the thermal time constant of the system, $\tau = C/g$. If energy is deposited in the film on timescales fast as compared to the thermal time constant, it can be measured with an accuracy below the rms energy fluctuations in the system, as long as information contained in frequencies faster than the thermal time constant is taken into account [24]. For most technologies, however, the Johnson noise or other noise sources drown out information in frequencies much faster than the thermal time constant. It has been noted that a superconducting transition-edge sensor with a high enough α should theoretically

have information in these high frequencies that is above the Johnson noise [25]. This goal has not been achieved, however, because of difficulties in impedance matching a suitable amplifier to this type of thermometer, and, more fundamentally, because such a detector would have to be made extremely slow [25]. The ETF-TES is well impedance matched to a SQUID, and by its use of electrothermal feedback pulse shortening avoids this rate limit.

To understand these considerations, we will first analyze the case of a transition-edge sensor running with negligible feedback, and then extend the calculation to the extreme negative feedback regime. This procedure will allow us to independently discuss the advantages of a superconducting transition edge sensor, and the further advantages of the ETF-TES.

8.2. A transition edge sensor with negligible feedback

Even with perfect shielding and noiseless electronics, the energy resolution of a transition edge sensor is ultimately limited by thermodynamic energy fluctuations in the detector (phonon noise), and the Johnson noise of the film.

First it is instructive to consider the resolution limitations imposed by the phonon noise in the absence of Johnson noise. A simple order of magnitude estimate is useful here. The phonon noise has power spectral density $S = 4kT^2g$. For this estimate we are concerned with times of order Δt, and bandwidth of order $\Delta B = 1/\Delta t$. The power noise folded into this measurement is thus $P_n \sim \sqrt{4kT^2g\Delta B} \sim \sqrt{4kT^2g/\Delta t}$. The rms energy fluctuations associated with this power noise are $\Delta E_{rms} \sim P_n \Delta t \sim \sqrt{4kT^2g\Delta t}$. These fluctuations will set the final limit on energy resolution. Thus, until either the thermalization time of the detector is encountered, or Johnson and other sources of noise become comparable to the phonon noise, it pays to look at higher frequencies and shorter times. While using higher frequencies folds in more rms power fluctuations, it folds in less rms energy fluctuations, which set the fundamental limit on energy resolution. It is also worth noting that if the timescale considered is the intrinsic time constant of the detector, $\Delta t = \tau = C/g$, the resolution limit becomes

$$\Delta E_{rms} \sim \sqrt{4kT^2C} \tag{8.2}$$

which is close to the rms energy fluctuations calculated above.

Consider a transition edge sensor operating with the substrate temperature very close to the bias temperature of the film. The Joule heating is assumed to be small enough to be neglected.

The power flow equation for this sensor is

$$C\frac{dT}{dt} = -g(T - T_0) + P(t) \tag{8.3}$$

where $P(t)$ is the power flow from the substrate due to phonon noise. Note that the expression for the heat flow to the substrate, $-g(T - T_s)$, is only valid for $T - T_s << T$, which holds in the negligible feedback regime. If we consider small signals and look at one Fourier component, we find that

$$i\omega C T_\omega = -gT_\omega + P_\omega \tag{8.4}$$

Thus, the temperature fluctuations due to the phonon noise are

$$T_\omega = \frac{P_\omega}{g + i\omega C} \tag{8.5}$$

Now small changes in temperature cause small changes in current

$$\Delta I = \Delta\left(\frac{V}{R}\right) = -\frac{V}{R_0^2}\frac{dR}{dT}\Delta T = -\frac{I_0\alpha}{T_0}\Delta T \tag{8.6}$$

We now assume that the detector is perfectly shielded, so that there are no external sources of noise. We also assume that the detector has been impedance matched so that the amplifier noise is small as compared to the Johnson and phonon noise. Then total current fluctuations due to phonon noise and Johnson noise are

$$I_\omega = \frac{V_\omega}{R_0} - \frac{I_0\alpha}{T_0}\frac{P_\omega}{g + i\omega C} \tag{8.7}$$

The phonon noise and Johnson noise are uncorrelated, and hence will add in quadrature. The power spectral density of the current noise is

$$I_n^2 = \frac{V_n^2}{R_0^2} + \frac{I_0^2 \alpha^2}{T_0^2} \frac{P_n^2}{g^2 + \omega^2 C^2} \tag{8.8}$$

The spectral density of the Johnson noise is $V_n^2 = 4kT_0R_0$. The spectral density of the phonon noise is $P_n^2 = 4kT_0^2 g$. Thus,

$$I_n^2 = \frac{4kT_0}{R_0} + \frac{4kI_0^2\alpha^2}{g} \frac{1}{1 + \omega^2 \tau^2} \tag{8.9}$$

Where $\tau = C/g$, the intrinsic time constant of the detector. Now we use the fact that, at equilibrium, the Joule power is equal to the power flowing to the substrate, or $I_0^2 R = g(T - T_s)$. Then,

$$I_n^2 = \frac{4kT_0}{R} \left[1 + \alpha^2 \frac{(1 - T_s/T_0)}{1 + \omega^2 \tau^2} \right] \tag{8.10}$$

The first term here is the Johnson current noise, and the second is the phonon noise. As can be seen, the Johnson noise is white. The phonon noise is flat at low frequencies, but rolls off at higher frequencies with a knee at $\omega = 1/\tau$.

When a particle interaction causes heat to be dumped into the film, the resulting current pulse has the form

$$I = -\frac{I_0 \alpha}{T_0} \frac{E}{C} e^{-t/\tau} \tag{8.11}$$

with power spectral density

$$I_\omega^2 = \frac{\alpha^2 I_0^2 E^2}{T_0^2 C^2} \frac{\tau^2}{1 + \omega^2 \tau^2} \tag{8.12}$$

The signal has a form similar to that of the phonon noise, rather than the Johnson noise. It is flat at low frequencies, but rolls off at higher frequencies with a knee at $\omega = 1/\tau$. Both the signal and phonon noise are proportional to α. (See Fig. 8.1)

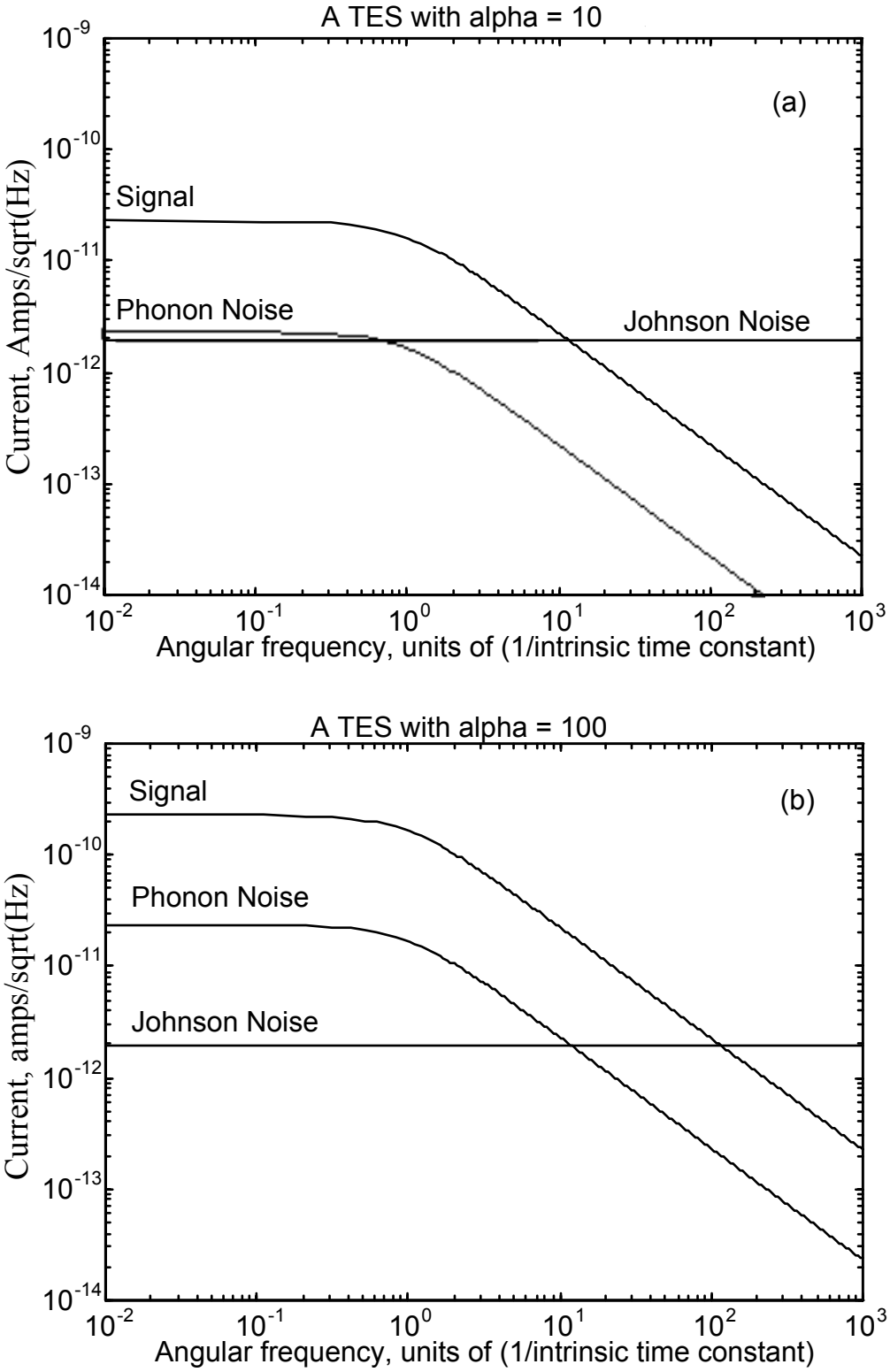


Fig 8.1 The signal and noise in the negligible feedback regime. The signal current, phonon noise current, and Johnson noise current are plotted as a function of frequency for an example TES with (a) $\alpha = 10$. (b) $\alpha = 100$.

For superconductors with a large α, the phonon noise dominates the Johnson noise at low frequencies. In this case, the signal to noise ratio at low frequencies is independent of α. If only frequencies below the knee were taken into account, the resolution of this detector would indeed be worse than the "thermodynamic limit" $\Delta E = \sqrt{kT^2C}$. At higher frequencies, however, the signal to noise ratio still contains information, until the signal drops below the white Johnson noise. The frequency at which this occurs increases linearly with α. Indeed, if infinite bandwidth is used, as α diverges the true thermodynamic limit on the energy resolution approaches zero [25]. Since the information that allows a detector to operate below the rms energy fluctuations in the sensor is in the higher frequencies, it may be said to be contained in the leading edge of the pulse.

As observed by McCammon [25], these bandwidth requirements pose a problem for a transition edge sensor intended to be used below $\Delta E = \sqrt{kT^2C}$. In order to take advantage of the information provided by a large α, the bandwidth used must increase linearly with α. But the thermalization time is a fundamental limit on how large a bandwidth can be used. Thus, in order to take advantages of frequencies much faster than the knee, the knee itself must be made lower. The intrinsic time constant of the detector, $\tau = C/g$, must be made extremely slow. This consideration prevents practical use of the simple transition edge sensor below $\Delta E = \sqrt{kT^2C}$.

To complete the analysis of the noise limits we now introduce optimal filtering theory.

8.3. The optimal filter

As has been shown, when an infinitesimal amount of heat is dumped into a superconducting film, an exponential current pulse occurs, with a characteristic time constant $\tau = C/g$. The amplitude of this pulse is proportional to the energy deposited. If the amount of energy deposited causes a temperature rise that is a significant fraction of a transition width, however, the pulse is no longer exponential, and we may no longer neglect the higher order terms in our expansions. Then the calculation of fundamental limits becomes much more complicated. For the purposes of this discussion, we will assume a small pulse.

Many different filters may be used in order to estimate the amount of energy that a pulse represents. In a case where all pulses may be assumed to have the same shape, however, there exists one optimal filter which will minimize the error in the energy

estimate. A superconducting transition-edge sensor in the small pulse limit is such a case.

The optimal filter is derived by minimizing the error in the energy estimate in a least squares sense. We do this calculation in the frequency domain so that non-white noise terms may be used. When a pulse occurs, $D(f) = AM(f) + N(f)$ where the data, $D(f)$, is made up of the normalized model pulse $M(f)$, multiplied by the amplitude A, plus the noise $N(f)$. The chi-square to minimize the noise is [26]

$$\chi^2 = \int_{-\infty}^{\infty} \frac{[D(f) - AM(f)]^2}{N^2(f)} df \tag{8.13}$$

Setting $d\chi^2/dA = 0$ yields an estimation of the amplitude

$$A = \frac{\int_{-\infty}^{\infty} \frac{DM^* + MD^*}{2|N|^2} df}{\int_{-\infty}^{\infty} \left|\frac{M}{N}\right|^2 df} \tag{8.14}$$

This equation can be simplified further by exploiting symmetry properties. Since $D(f)$ and $M(f)$ are Fourier transforms of real quantities, $D(-f) = D(f)^*$ and $M(-f) = M(f)^*$,

$$\int_{-\infty}^{\infty} \frac{D(f)M(f)^*}{2|N|^2} df = -\int_{\infty}^{-\infty} \frac{D(-f)M(-f)^*}{2|N|^2} df = \int_{-\infty}^{\infty} \frac{M(f)D(f)^*}{2|N|^2} df \tag{8.15}$$

Thus, the first term in Eqn. 8.14 is equal to the second term. Since the first and second terms are complex conjugates, they must be real. We may thus write the optimal estimate of the pulse amplitude as

$$A = \frac{\int_{-\infty}^{\infty} \frac{DM^*}{|N|^2} df}{\int_{-\infty}^{\infty} \left|\frac{M}{N}\right|^2 df} \tag{8.16}$$

This result is the equation for the optimal filter. Conceptually, this equation says that for each frequency, an estimate is made of the amplitude $A = D(f)/M(f)$. The estimates at each frequency are averaged with a weight of the square of the signal to noise ratio at that frequency $\left[M(f)/N(f)\right]^2$.

In order to calculate the fundamental resolution limit attainable by the optimal filter, one calculates the rms amplitude that would be measured if no pulse were present, and $D(f) = N(f)$. When the data is in units of measured power, and the model pulse is correctly normalized, the amplitude is the energy deposited. Then it can be shown [24] that

$$\Delta E_{rms} = \left(\int_0^\infty \frac{4\,df}{NEP^2(f)} \right)^{1/2} \tag{8.17}$$

Where $NEP(f)$ is the noise equivalent power, the equivalent power fluctuations that would cause the current noise calculated in eqn. 8.10, including that which is actually due to the voltage noise.

8.4. The Noise Equivalent Power

From eqn 8.9 we have the power spectral density of the current noise

$$I_n^2 = \frac{4kT_0}{R_0} + \frac{4kI_0^2\alpha^2}{g}\frac{1}{1+\omega^2\tau^2} \tag{8.18}$$

But

$$\Delta I = \Delta\left(\frac{V}{R}\right) = -\frac{V}{R_0^2}\frac{dR}{dT}\Delta T = -\frac{I_0\alpha}{T_0}\Delta T \tag{8.19}$$

and from eqn 8.5

$$NEP^2 = T_n^2 g^2 (1+\omega^2\tau^2) \tag{8.20}$$

So

$$NEP^2(f) = \frac{T_0^2 g^2 \left(1+\omega^2\tau^2\right)}{I_0^2 \alpha^2} I_n^2 = 4kT_0^2 g\left(1+\frac{gT_0\left(1+\omega^2\tau^2\right)}{\alpha^2 P_0}\right) \quad (8.21)$$

And, finally from eqn 8.17,

$$\Delta E_{rms} = \sqrt{\frac{4kT_0^2 C/\alpha}{\sqrt{1-T_s/T_0}}} \quad (8.22)$$

It has been shown that for the general bolometer, the fundamental resolution limit is $\Delta E_{rms} = \xi\sqrt{kT^2C}$ [24]. Eqn 8.22 sets the value $\xi = \left(\alpha\sqrt{1-T_s/T_0}/4\right)^{-1/2}$ for the transition-edge sensor in the non feedback limit. As can be seen, as α becomes larger, the fundamental resolution limit becomes smaller, and drops below the rms energy fluctuation in the sensor, $\Delta E_{rms} = \sqrt{kT^2C}$, for $\alpha > 4/\sqrt{1-T_s/T_0}$.

This result highlights the advantage of using superconducting transition-edge sensors. The TES provides an α that is extremely high, allowing an energy resolution below the thermodynamic limit.

It should be noted that this resolution limit is independent of whether the film is current biased, and the voltage measured with a FET (the conventional conditions), or voltage biased, and the current measure with a SQUID (as is the case in the ETF-TES.)

8.5. A transition edge sensor in the extreme negative feedback regime

We have calculated the fundamental resolution limits on a superconducting transition-edge sensor in the non-feedback limit. This result is valid whether the film is current or voltage biased.

We now move on to consider the Electrothermal Feedback Transition-Edge sensor. This sensor operates in an extreme negative feedback regime. The results we discuss here will only apply to our voltage biased, negative feedback case: conventional, current biased sensors with positive electrothermal feedback are unstable in the extreme feedback regime.

Here we will demonstrate two key advantages of the ETF-TES over a conventional TES. The first is an effect first pointed out by Mather [23] for bolometers with high α. For low frequencies, electrothermal feedback suppresses the Johnson

noise. For our extreme feedback case, the Johnson noise is actually made negligible for frequencies below the knee.

The second effect is that the electrothermal feedback makes pulses much shorter than the intrinsic time constant. As explained previously, a TES can only have resolutions below the rms thermodynamic energy fluctuation if frequencies much faster than the intrinsic time constant are used. Since no information exists faster than the thermalization time, for a conventional TES, this means that the device must be made extremely slow. In the ETF-TES, however, the intrinsic time constant can be made extremely slow, while the effective time constant is close to the thermalization limit. Then, even though only frequencies below the knee are taken into account, these frequencies are much faster than the "intrinsic" time constant. Resolutions well below the rms energy fluctuations may be attained without increasing the bandwidth far past the knee.

8.6. A simple estimate of the ETF-TES fundamental energy resolution

Once again, a simple, order of magnitude estimate of detector resolution is useful before a detailed calculation. In this estimate, we will only take into account frequencies up to the knee. As will be discussed, the Johnson noise of the film in an ETF-TES is highly suppressed by the negative electrothermal feedback in frequencies below the knee. Thus the primary fundamental source of noise in this device is the phonon noise. The power spectral density of the phonon noise is $S = 4kT^2 g$. In this estimate we will only make use of frequencies up to the knee. Thus, the time constant we are interested in is the effective time constant, τ_{eff}. The bandwidth we are concerned with for this order of magnitude calculation is $\Delta B = 1/\tau_{eff}$. The effective time constant of this device is

$$\tau_{eff} = \frac{\tau_0}{1 + \alpha/n} \approx \frac{Cn}{g\alpha} \tag{8.23}$$

Thus, the rms energy fluctuations in the ETF-TES in the bandwidth of the signal are $\Delta E_{rms} \sim P_n \sqrt{\Delta B}\, \tau_{eff} = \sqrt{4kT^2 g}\sqrt{Cn/(g\alpha)}$, or

$$\Delta E_{rms} \sim \sqrt{4kT^2 C n/\alpha} \tag{8.24}$$

This result is below the rms energy fluctuations in the system by a factor of $\sim 2\sqrt{n/\alpha}$, even though only frequencies below the knee are used. It is useful to compare

this result to the order of magnitude calculation of the resolution limit in the conventional TES for frequencies below the knee. From Eqn. 8.2 we have $\Delta E_{rms} \sim \sqrt{4kT^2C}$ for a conventional TES, taking into account only frequencies below the knee. Thus, for a detector with speed comparable to the thermalization time, where frequencies above the knee cannot be used, the ETF-TES has fundamental limits about an order of magnitude better than a conventional TES.

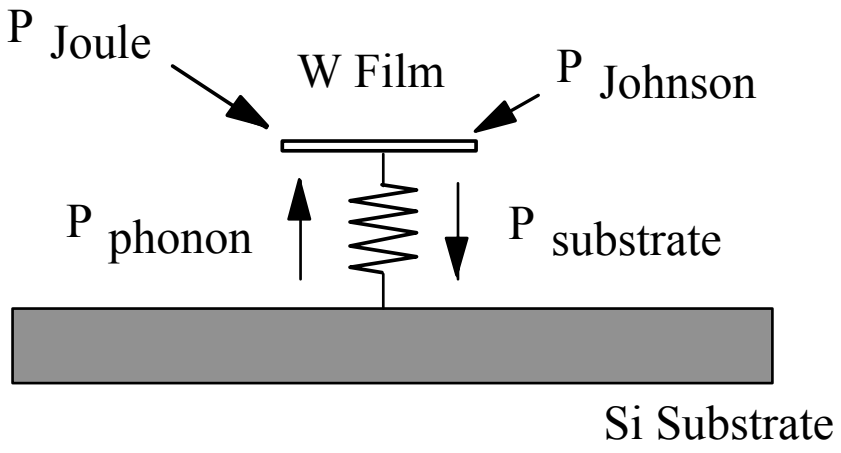


Fig 8.2 A power flow diagram for the ETF-TES. Sources of power into the film are the Joule power dissipation, and $P_{Johnson}$, the work that the bias voltage does on the Johnson noise current. This power escapes into the substrate through a thermal impedance. Finally, there is a power fluctuation, P_{phonon}, across the thermal impedance due to the phonon noise.

We now proceed to a more precise calculation of the fundamental resolution limits of an ETF-TES. Consider the power flows in the ETF-TES (Fig. 8.2) The power flow equation for the ETF-TES is

$$C\frac{dT}{dt} = \frac{V^2}{R(T)} - K\left(T^n - T_s^n\right) + P_J(t) + P_{phon}(t) \qquad (8.25)$$

Here $P_J(t) = I_J(t)V_0$ is the work that the bias voltage does on the Johnson noise current, and $P_{phon}(t)$ is the power flow across the thermal impedance due to phonon noise. If we consider small signals and look at one Fourier component, we find that

$$i\omega C T_\omega = -\frac{V^2}{R_0^2}\frac{dR}{dT}T_\omega - gT_\omega + I_\omega V_0 + P_\omega \tag{8.26}$$

Thus, the temperature fluctuations due to the phonon noise are

$$T_\omega = \frac{I_0 V_n + P_n}{P_0\alpha/T_0 + g + i\omega C} \tag{8.27}$$

Here we have used $I_n V_0 = V_n V_0/R = I_0 V_n$. Now small changes in temperature cause small changes in current

$$\Delta I = \Delta\left(\frac{V}{R}\right) = -\frac{V}{R_0^2}\frac{dR}{dT}\Delta T = -\frac{I_0\alpha}{T_0}\Delta T \tag{8.28}$$

We again assume that the detector is perfectly shielded, so that there are no external sources of noise, and that the detector has been impedance matched so that the amplifier noise is small as compared to the Johnson and phonon noise. The total current fluctuations due to phonon noise and Johnson noise are

$$I_\omega = \frac{V_n}{R_0} - I_0\alpha\frac{I_0 V_n + P_n}{P_0\alpha + gT_0 + i\omega C T_0} \tag{8.29}$$

The phonon noise and Johnson noise are uncorrelated, and hence will add in quadrature.

$$I_\omega^2 = V_n^2\left|\frac{1}{R_0} - \frac{\alpha I_0^2}{P_0\alpha + gT_0 + i\omega C T_0}\right|^2 + P_n^2\left|\frac{\alpha I_0}{P_0\alpha + gT_0 + i\omega C T_0}\right|^2 \tag{8.30}$$

We consider first the Johnson noise term. We take the absolute value squared of this term by multiplying by its complex conjugate. Using the fact that in the extreme feedback regime, $P_0\alpha \approx gT_0\alpha/n >> gT_0$, we may neglect certain terms. Note that using this approximation to neglect the gT_0 term *before* the absolute value is taken leads to the loss of the dominate term at zero frequency, so this approximation cannot be used until after the absolute value is taken. The noise current due to the Johnson noise is

$$I_\omega^2 = V_n^2\left|\frac{1}{R_0} - \frac{\alpha P_0/R_0}{P_0\alpha + gT_0 + i\omega CT_0}\right|^2 = \frac{V_n^2\left(g^2+\omega^2C^2\right)T_0^2}{\left(P_0^2 + 2P_0\alpha gT_0 + g^2T_0^2 + \omega^2C^2T_0^2\right)R_0^2} \quad (8.31)$$

$$I_\omega^2 \approx \frac{V_n^2\left(g^2+\omega^2C^2\right)T_0^2}{\left(P_0^2 + \omega^2C^2T_0^2\right)R_0^2} \quad (8.32)$$

The spectral density of the Johnson noise is $V_n^2 = 4kT_0R_0$. Also, in the extreme feedback limit, $P_0 = gT_0/n$ and $\tau_{eff} = (C/g)(n/\alpha)$. The Johnson noise current is thus

$$I_J^2 = \frac{4kT_0}{R_0}\frac{\left(n^2/\alpha^2 + \omega^2\tau_{eff}^2\right)}{1+\omega^2\tau_{eff}^2} \quad (8.33)$$

This result is significant. It was first pointed out by Mather that, in a bolometer with a high α, the electrothermal feedback suppresses the Johnson noise [23], as is seen in Eqn. 8.33. For frequencies well below the knee, $\omega^2\tau_{eff}^2 << 1$, and the Johnson noise is suppressed by the factor n^2/α^2. For our sharpest films, this factor is ~ 10,000. Since the electrothermal feedback cannot respond to frequencies much faster than the effective time constant, at high frequencies the Johnson noise goes to its white noise value, $I_J = \sqrt{4k\,T_0/R_0}$.

We now return to the phonon noise, which is evidently dominant for low frequencies. Once again, using the fact that in the extreme feedback regime $P_0\alpha \approx gT_0\,\alpha/n >> gT_0$, the current noise due to the phonon noise is seen to be

$$I_\omega^2 = P_n^2\left|\frac{\alpha I_0}{\alpha P_0 + gT_0 + i\omega CT_0}\right|^2 = \frac{P_n^2\alpha^2I_0^2}{\alpha^2P_0^2 + 2P_0\alpha gT_0 + g^2T_0^2 + \omega^2C^2T_0^2} \quad (8.34)$$

$$I_\omega^2 \approx \frac{P_n^2\alpha^2I_0^2}{\alpha^2P_0^2 + \omega^2C^2T_0^2} \quad (8.35)$$

In the extreme feedback limit, $P_0 = gT_0/n$ and $\tau_{eff} = (C/g)(n/\alpha)$. Thus, the phonon noise current is

$$I_\omega^2 = \frac{P_n^2I_0^2}{P_0^2\left(1+\omega^2\tau_{eff}^2\right)} = \frac{P_n^2}{V_0^2\left(1+\omega^2\tau_{eff}^2\right)} \quad (8.36)$$

When the substrate temperature is close to the temperature of the film, the spectral density of the phonon noise is $P_n^2 = 4kT_0^2 g$. In the extreme feedback limit, however, $T_s^n << T_0^n$, so that $P_n^2 = 2k\left(T_0^2 g + T_s^2 g_s\right) \approx 2kT_0^2 g$. This implies a phonon noise current of

$$I_\omega^2 = \frac{2kT_0^2 g}{V_0^2\left(1+\omega^2\tau_{eff}^2\right)} = \frac{2kT_0 nP_0}{V_0^2\left(1+\omega^2\tau_{eff}^2\right)} = \frac{4kT_0}{R_0}\frac{n/2}{1+\omega^2\tau_{eff}^2} \tag{8.37}$$

Leading to a total current noise of

$$I_J^2 = \frac{4kT_0}{R_0}\frac{\left(n^2/\alpha^2 + \omega^2\tau_{eff}^2\right)}{1+\omega^2\tau_{eff}^2} + \frac{4kT_0}{R_0}\frac{n/2}{1+\omega^2\tau_{eff}^2} \tag{8.38}$$

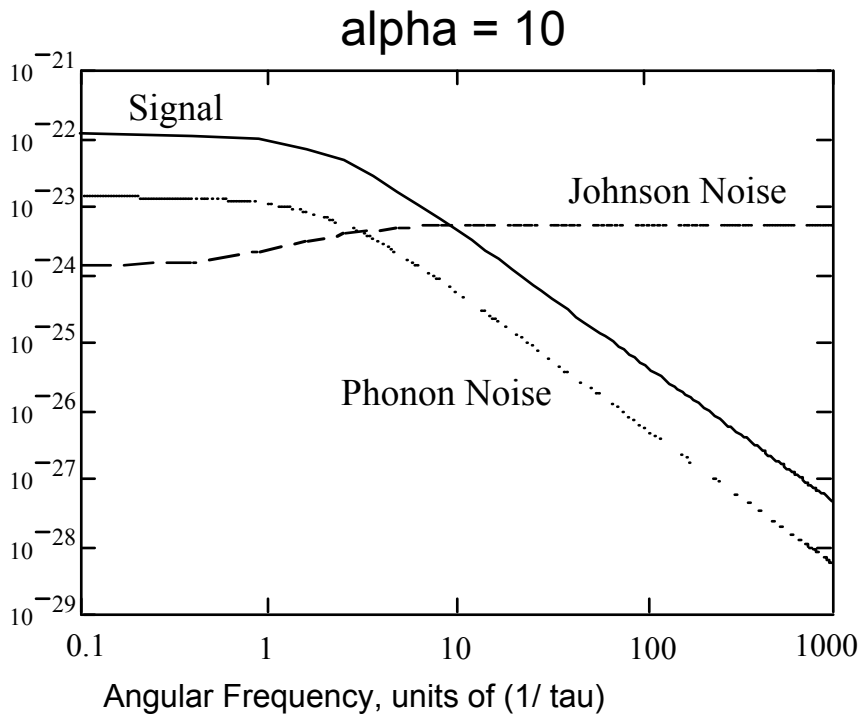
Where the first term is the Johnson noise, and the second term is the phonon noise. Now when an energy E is dumped into the film, the resulting current pulse of the ETF-TES has the form

$$I = -\frac{I_0\alpha}{T_0}\frac{E}{C}e^{-t/\tau_{eff}} \tag{8.39}$$

with power spectral density

$$I_\omega^2 = \frac{\alpha^2 I_0^2 E^2}{T_0^2 C^2}\frac{{\tau_{eff}}^2}{1+\omega^2{\tau_{eff}}^2} \approx \frac{I_0^2 E^2 g^2}{T_0^2 n^2}\frac{1}{1+\omega^2{\tau_{eff}}^2} \tag{8.40}$$

Where we have used $\tau_{eff} \approx Cg/n\alpha$. Once again, the signal is flat at low frequencies, and rolls off in the same manner as the phonon noise. For any given frequency below the knee of the ETF-TES, the phonon noise and Johnson noise are seen to have no dependence on α (Fig 8.3). The knee is, however pushed up to higher frequencies by electrothermal feedback. In contrast, in the conventional TES, both the signal and phonon noise power spectral densities were seen to be proportional to α^2 below the knee (Fig 8.1), while the frequency of the knee remains the same. This result is important - it means that, in the ETF-TES, the fundamental resolution limits may be improved by increasing α, even if only frequencies to the knee are used.



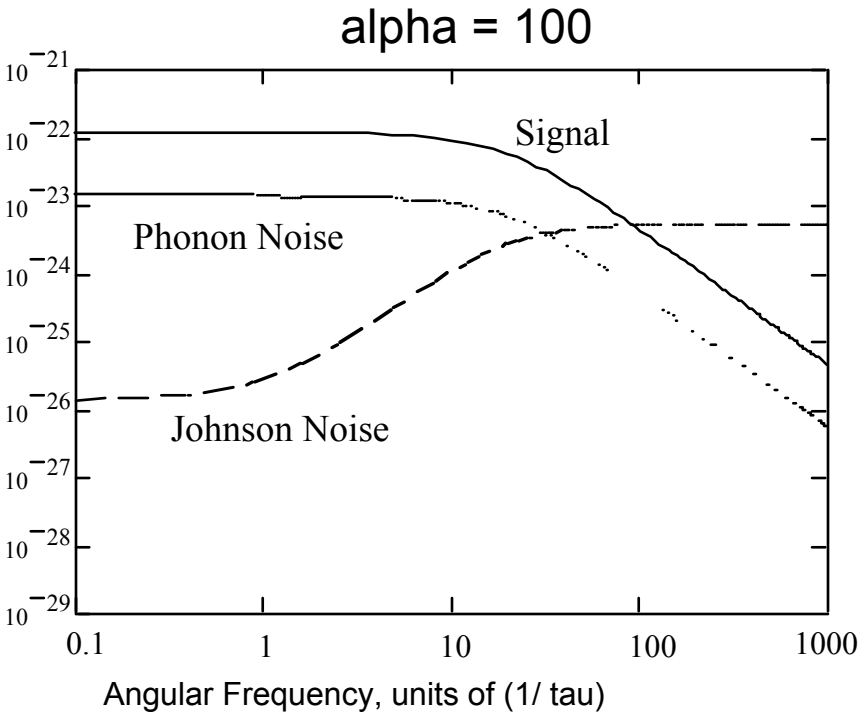


Fig 8.3 The signal and noise in the extreme feedback regime. The signal current, phonon noise current, and Johnson noise current are plotted as a function of frequency for an example TES with (a) $\alpha = 10$ (b) $\alpha = 100$.

8.7. The ETF-TES Noise Equivalent Power

For the ETF-TES,

$$\Delta I = \Delta\left(\frac{V}{R}\right) = -\frac{V}{R_0^2}\frac{dR}{dT}\Delta T = -\frac{I_0\alpha}{T_0}\Delta T \tag{8.41}$$

and

$$T_\omega = \frac{P_\omega}{g\,\alpha/n + i\omega C} \tag{8.42}$$

So

$$NEP^2(f) = \frac{T_0^2 g^2\left(\alpha^2/n^2 + \omega^2\tau_0^2\right)}{I_0^2\alpha^2} I_n^2 = \frac{T_0^2 g^2\left(1+\omega^2\tau_{eff}^2\right)}{I_0^2 n^2} I_n^2 \tag{8.43}$$

$$NEP^2(f) = \left(1+\omega^2\tau_{eff}^2\right) V_0^2 I_n^2 \tag{8.44}$$

Substituting the noise current from Eqn 8.38, we arrive at the noise equivalent power for the ETF-TES

$$NEP^2(f) = 4kT_0P_0\left(\alpha^2/n^2 + \omega^2\tau_{eff}^2 + n/2\right) \approx 4kT_0^2 g\left(\left(2\pi f\right)^2\tau_{eff}^2/n + 1/2\right) \tag{8.45}$$

Substituting the noise equivalent power into eqn. 8.17, we have

$$\Delta E_{rms} = \frac{1}{\sqrt{\int_0^\infty \frac{4df}{4kT_0^2\, g/n\left(\left(2\pi f\right)^2\tau_{eff}^2 + n/2\right)}}} = \sqrt{4kT_0P_0\,\tau_{eff}\sqrt{n/2}} \tag{8.46}$$

Finally, using $P_0 = gT_0/n$ and $\tau_{eff} = Cn/g\alpha$, and multiplying by 2.36 to get the FWHM resolution instead of the rms resolution, we arrive at the fundamental thermodynamic noise limitation of the small signal ETF-TES

$$\Delta E_{FWHM} = 2.36\Delta E_{rms} = 2.36\sqrt{4kT_0^2 C(1/\alpha)\sqrt{n/2}} \tag{8.47}$$

As has been mentioned, the fundamental resolution limit for the general bolometer is $\Delta E_{rms} = \xi\sqrt{kT_0^2 C}$ [24]. Thus, for an ETF-TES detector in the extreme feedback regime, $\xi = 2\sqrt{\sqrt{n/2}/\alpha}$, which is below the rms thermodynamic energy fluctuations in the sensor ($\Delta E = \sqrt{kT^2 C}$) by a factor ~ $2.5/\sqrt{\alpha}$. For α=1000, this factor can be < 0.1.

Note that if film nonuniformities, transition nonlinearities, or saturation of the film cause significant differences in the pulse shape, it is difficult to apply the optimal filter. The resolution calculated above is not strictly applicable in these cases. Fortunately, unlike conventional transition edge sensors, most of the information in an ETF-TES is contained in frequencies below $\omega = 1/\tau_{eff}$, so good results should be obtainable with non-optimal filtering techniques.

In summary, *the ETF-TES has fundamental resolution limitations lower than any existing technology operating at the same temperature, count rate, and absorber heat capacity.*

CHAPTER 9: ETF-TES EXPERIMENTAL RESULTS

The ETF-TES has been intended since its inception to be used in conjunction with quasiparticle trapping from aluminum films. Without Al films it would be difficult to scale up to a practical dark matter or neutrino detector. Nevertheless, in order to explore the operation and stability of electrothermal feedback in isolation from the quasiparticle trapping mechanism, our first experiments involved two simple W thin film pads.

The sensors used here consist of a 60 nm thick W film deposited on a [100] silicon substrate with dimensions 0.5 mm x 10 mm x 10 mm. The film was photolithographically patterned into two 1.8 mm x 1.8 mm squares spaced 5 mm apart, center to center.

The clean room fabrication of our two pad W devices was identical to the fabrication of our high impedance W devices (see section 5.4).

9.1. The gold heat sink

In order to operate in the extreme electrothermal feedback regime, the silicon substrate must be cooled to well below the transition temperature of the film. This cooling requires much better heat sinking between the silicon and the cold stage than we have previously used.

A simple consideration of the heat flow shows that, in order to cool the silicon to well below the film temperature, the thermal impedance between the silicon lattice and the cold stage must be smaller than the thermal impedance between the tungsten electrons and the silicon lattice. Unfortunately, as the substrate is cooled to well below the film transition, the thermal impedance between the silicon and the cold stage increases sharply. In order to provide better heat sinking, a gold film was evaporated onto the back of the substrate. The Au film was thicker than the W film, and was deposited over a larger area than the W film covered. Thus, both the Kapitza boundary resistance and the electron-phonon decoupling were kept small.

There is a drawback to using a heat sink with a large area coverage. The ETF-TES technology is intended to absorb both prompt and multiply scattered athermal phonons. As the area coverage of the heat sink is increased, more of the multiply scattered phonons are lost in the heat sink, reducing the signal size. As will be seen later, the heat sink can also make it more difficult to deconvolute position and energy in an event. One reason for choosing a gold film for the heat sink is that gold has a very poor

acoustic impedance match with silicon. Thus, athermal phonons incident on the Au film are more likely to reflect, and be absorbed later in the ETF-TES.

As will be seen in the following chapter, the problem of losing signal to the heat sink can be largely avoided when aluminum quasiparticle traps are used to absorb the phonons. The heat sink must have larger area coverage than the W, but may still have coverage small compared to the Al since the Al films do not dissipate heat. As long as the heat sink coverage is kept small as compared to the Al, most of the athermal phonons will be absorbed in the quasiparticle traps and measured. This advantage will be discussed more in the next chapter.

For the simple W pad detectors described here, the gold film was evaporated onto the backside of the silicon substrate. The evaporation was done through a mask with two slits 2.5 mm wide, 8.9 mm long, and spaced 6.35 mm apart, center to center. A strip of silicon directly underneath the W squares was thus left bare in an effort to increase the fraction of athermal phonons absorbed in the W films instead of the gold heat sink.

After gold plating, the chip was placed into a G10 chip holder much like those used with our high impedance devices. Gold plated beryllium copper springs provided the thermal connection between the gold film and the cold plate.

A constant bias current was applied to a parallel arrangement of the W film and a 20 mΩ shunt resistor (Fig 9.1). As long as the W bias resistance is kept large as compared to the shunt resistor, this circuit provides a constant bias voltage.

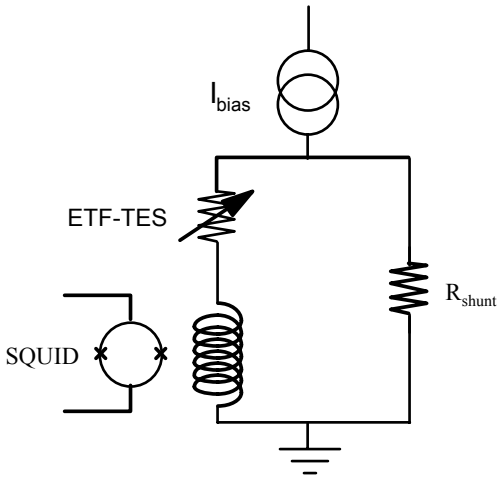


Fig 9.1 ETF-TES bias circuit

We used an array of DC SQUIDs to measure the current flowing through the W films (see section 4.3).

9.2. Self-biasing of the W film

Before the ETF-TES technology was tested, our computer models indicated that the devices would stably self-bias as long as the L/R time constant was short as compared to τ_{eff}. We were concerned, however, that the ETF-TES would contain some instability that had somehow escaped our modeling efforts.

Our first devices did, indeed, exhibit substantial instability. In these devices, electrical connections were made to the W film through only one wire bond on each side. As a result, the current density was extremely nonuniform. We later bonded to each device with many (~20) wire bonds on each side in order to improve the current uniformity. Devices bonded this manner were found, even in the extreme feedback regime, to self-bias quite stably. All results in this section were generated by devices bonded with many wire bonds.

A better way to make the current density uniform would be to connect an Al film along the entire side of the pad, and bond to the Al. Since Al is well below its superconducting transition at operational temperatures, it would create a very uniform current density through the W film. We have fabricated sensors with Al bars, but have not tested them. Note that in the quasiparticle-trap-assisted sensors that will be described in Chapter 10, the W film is patterned into long W lines connected in parallel with Al bars, so the current density is quite uniform.

The W films were biased at constant voltage, and the silicon was allowed to cool to well below the transition of the film. The W films were observed to stably self-bias on their superconducting transitions. As the voltage was reduced, the film resistance was reduced accordingly. Finally, when the film resistance was of the same order as the shunt resistor, the film was no longer voltage biased, and it went completely superconducting, or "unlatched." Before operation could continue, the film had to be driven normal again by heating the substrate up through the transition, or by putting a current pulse through that exceeded the critical current of the film.

Recall that the heat flow from the tungsten film to the silicon substrate scales as $T^n - T_s^n$. If $T_s^n << T^n$, and the film transition is narrow, the heat flow to the substrate becomes a constant, P_0. At equilibrium, this heat flow is matched by the Joule heating, V^2/R, and

$$V^2 = RP_0 \tag{9.1}$$

Thus, the resistive bias point can be chosen by changing the voltage bias. If the square of the voltage bias is plotted against the resistance, the result should be a line, with a slope equal to the equilibrium power.

In Fig. 9.2, the film resistance is plotted against the voltage bias squared. The resistance was determined by taking the ratio of the voltage bias to the current through the film (measured by the SQUID array). At the lower bias voltages, the data is approximately linear, with a slope of ~ 33 pW. This linearity indicates that the detector is on a sharp part of its transition, and equilibrium heat loss to the substrate is constant. At higher bias voltages, the data is no longer linear, and the heat loss to the substrate is rising, indicating that the film is off of the sharp part of its transition, and larger temperature excursions and power flows are occurring.

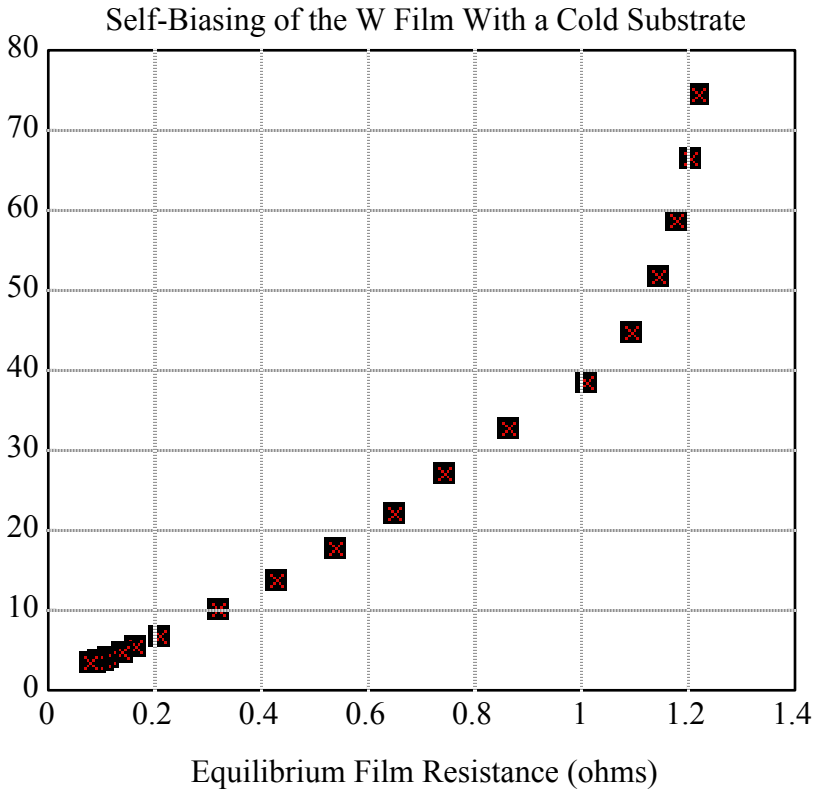


Fig 9.2 The bias resistance of the W film, as a function of V^2

9.3. Pulses

As shown in Fig. 9.3, an ^{55}Fe source was placed behind the silicon substrate. A copper collimator was used, with a thin slit oriented so that 6 keV x-rays were incident on the backside of the crystal, on a line passing through the centers of both W films. The x-rays were incident on a portion of the silicon with no gold heat sinks.

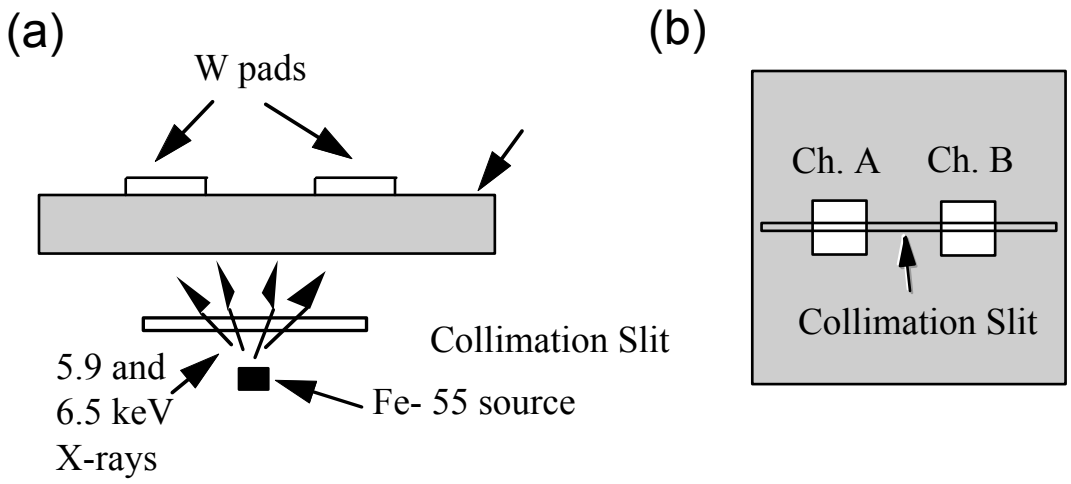


Fig. 9.3 Detector Geometry. (a) Cross sectional view of the detector with a collimated Fe-55 source incident on the back side. (b) Top view of the detector. The line of collimation going through the centers of both channels.

As expected, the pulse durations were found to be dependent on the biasing conditions. When the substrate temperature was biased in the superconducting transition, and very low bias voltages were used, the current pulse was a simple exponential, with ~1 ms duration (Fig 9.4a). This result implies that the intrinsic time constant τ_0 is of order 1 ms.

When the substrate temperature was cooled to well below the superconducting transition of the film, pulses were observed to shorten, and have two time constants. The fast time constant got faster as the substrate cooled, and was measured to be as short as 10 μs. The longer time constant was ~300 μs (Fig 9.4b), and did not scale with the bias conditions.

Recall that when the substrate temperature is much colder than the film, the pulse recovery time constant due to electrothermal feedback, τ_{eff}, is calculated to be $\tau_0 n/\alpha$

(eqn. 7.7) where τ_0 is the intrinsic time constant of the film (the pulse recovery time constant in the absence of Joule heating), $\alpha = (T/R)(dR/dT)$, and n is either 4, 5, or 6 depending on the dominant thermal impedance between the electrons in the tungsten and the substrate.

The fast, 10 μs time constant can be explained as the prompt arrival of energy from the particle interaction. If $\alpha = 500$, the electrothermal feedback response to prompt energy would have a time constant of about 10 μs. Simple models of electrothermal feedback do not, however, predict a longer, second time constant response to prompt energy.

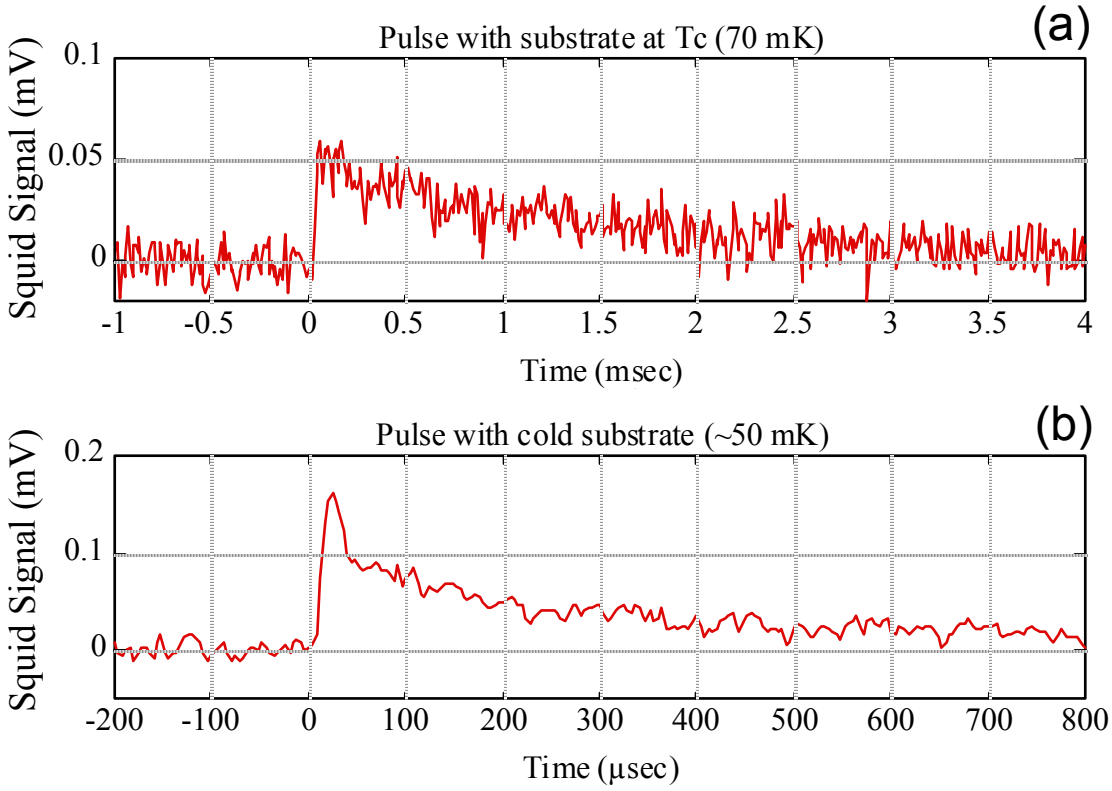


Fig 9.4 Observed ETF-TES pulses. (a) The film is biased in the transition, with little electrothermal feedback. (b) The film is biased in the extreme feedback regime (~50 mK). A 350 kHz low pass filter was used in both cases.

To investigate the origins of the second time constant, a small heat pulse was injected into the W by putting a fast (< 1 μs) voltage pulse across the film. At the end of the pulse, the voltage was returned to the bias voltage. On longer time scales, pulses of

reduced current in the SQUID were observed as electrothermal feedback returned the film temperature to the bias temperature. These pulses were simple exponentials with time constants equal to the fast time constant in x-ray induced pulses. The longer time constant was not observed on these events (Fig 9.5). This result is consistent with the identification of the 300 μs time constant in x-ray pulses as energy arriving in the film over a longer time scale.

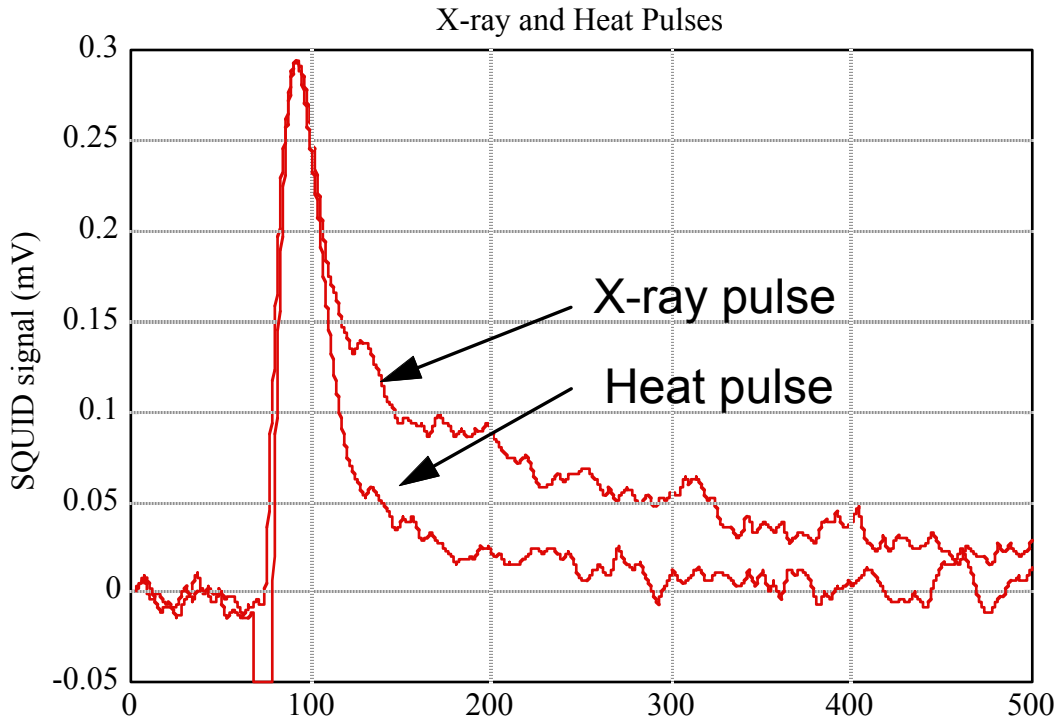


Fig 9.5 X-ray and injected heat pulses. The upper pulse is the response to a 6 keV x-ray in the substrate. The lower pulse is a heat pulse injected in the film by Joule heating. The negative going portion before the heat pulse is the voltage spike used to inject the heat. The pulse itself is the electrothermal feedback response to this heat.

9.4. Electrothermal Oscillations

When an event heats up the W film, electrothermal feedback causes the bias current to drop, decreasing the Joule power. Once the current drops sufficiently for the Joule power to decrease below the bias power, the temperature of the film will begin to return to its equilibrium value. If the inductance is large, however, the bias current will remain depressed for a time after the equilibrium temperature is reached, causing the

temperature to drop below its equilibrium value. In order to compensate, the current will have to rise above its equilibrium value, and oscillations will occur.

Before the construction of the first ETF-TES, our computer models (see Appendix B) predicted that these devices would exhibit electrothermal oscillations when the effective time constant became of the same order as the electronics time constant, $\tau_{el} = L/R$, which is determined by the inductance of the SQUID input loop and the bias resistance of the ETF-TES. The results of such a calculation with our device parameters is shown in fig. 9.6a.

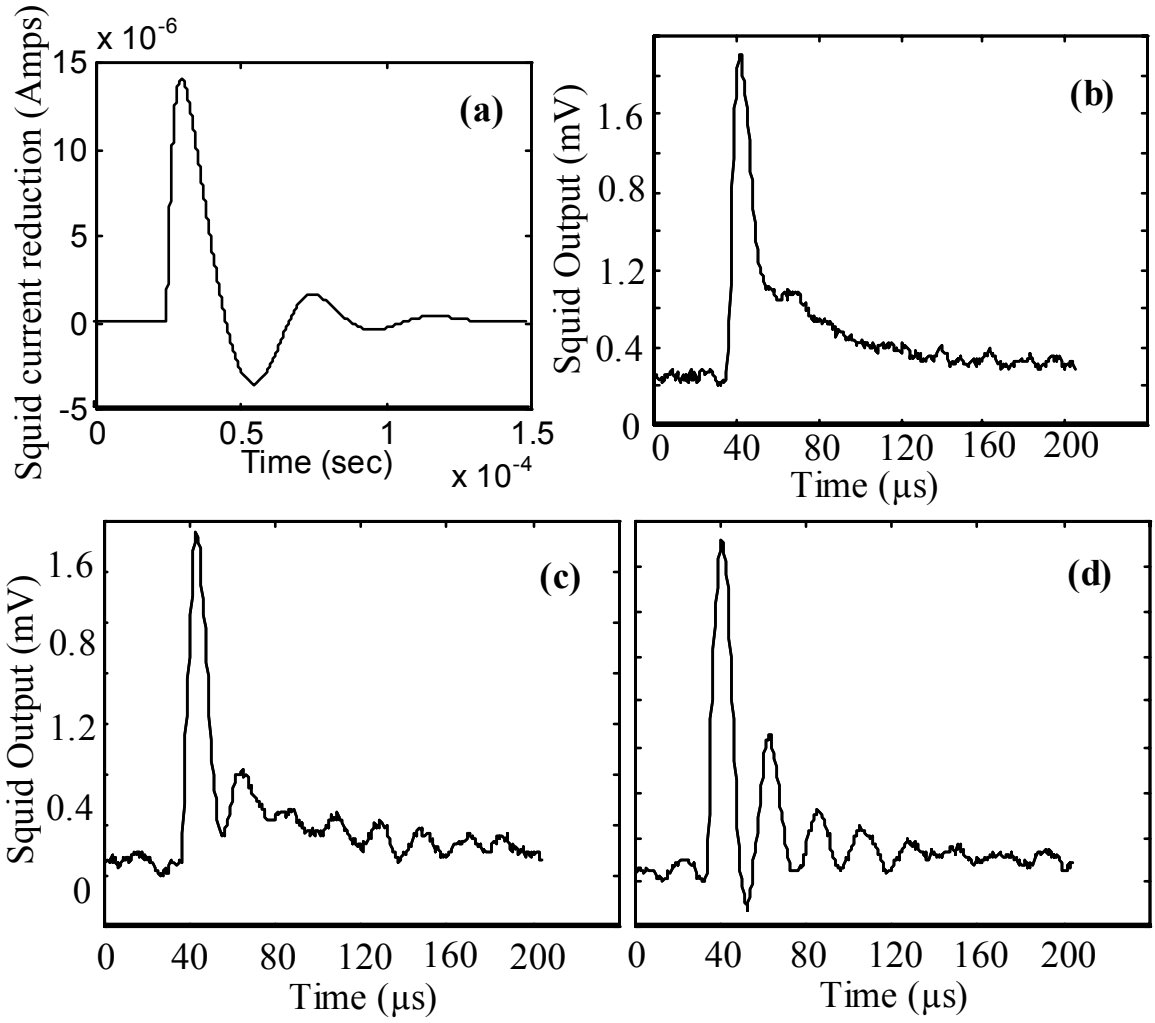


Fig 9.6 The onset of electrothermal oscillations. (a) A computer simulation of the ETF-TES performance, predicting electrothermal oscillations. Observed oscillations: (b) At higher bias resistance, electrothermal oscillations are damped. (c) As the bias resistance is

lowered, oscillation sets in. (d) As the resistance is lowered further, oscillations become large.

The effective time constant, $\tau_{eff} \approx \tau_0 n/\alpha$, can be shortened by biasing at low bias resistance, where the value of α is largest. The electronic time constant, $\tau_{el} = L/R$, can be increased by biasing at low bias resistance. When our W ETF-TES was biased at very low bias resistance, electrothermal oscillations were observed (Fig 9.6b - 9.6d).

Electrothermal oscillations will decrease the resolution of an ETF-TES. It is important, therefore, to bias the sensor at a resistance where the electronic time constant is fast as compared to the effective time constant, and electrothermal oscillations are small.

9.5. Optimal Voltage and Temperature Bias Conditions

The voltage and temperature must be adjusted to an optimal point for detector operation. Calculating this optimal point is a complex process, which is different for different conditions of detector operation. In each case both fundamental considerations (the fundamental limits on the energy resolution) and practical conditions (SQUID noise and other external noise sources) must be taken into account.

9.5.1. Optimal Bias Conditions: Delta Function

First we consider the case where the energy is collected in the film, and thermalized, on time scales short compared to any achievable effective time constant. The best resolution limit will occur in the extreme feedback regime, where nearly all of the energy is being removed (and hence measured) by electrothermal feedback. If time constants are not in the extreme regime, some of the heat will escape back into the substrate and not be measured.

The fundamental resolution limit for small pulses in the extreme feedback regime has been shown in eqn. 8.47 to be $\Delta E_{rms} = \sqrt{4kT_0^2 C(1/\alpha)\sqrt{n/2}}$. Thus, the fundamental limit will be best for a bias condition that maximizes α. Recall that $\alpha = (T/R)(dR/dT)$. To a very rough approximation, dR/dT is constant in the 10-90% region of the transitions of our film. Thus, α may be considered approximately inversely proportional to R in this region, and τ_{eff} is roughly proportional to R. Although α may be somewhat increased by biasing below the 10-90% range, such bias conditions are inadvisable due to the large temperature excursions that will occur. Thus, the fundamental limitations will

be best when the film resistance is at about 10% of the normal resistance (the low end of the approximately linear portion of the transition). As long as the substrate temperature is cold enough that the effective time constant is short compared to the intrinsic time constants, adjusting the temperature bias further will not have a major effect on the fundamental resolution limit.

The fundamental resolution limit takes into account the Johnson noise (which is suppressed) and phonon noise. Other noise sources, including SQUID noise, microphonics, and EMI pickup from the environment, pose further practical limitations. In the phonon noise case, both the noise and the signal scale in the same way with bias conditions. This is not necessarily true for external noise sources. Thus, for external noise sources, we must worry about the absolute magnitude of the current pulse.

Recall that the magnitude of a particle-induced current pulse is

$$\Delta I = \frac{-I_0 \alpha}{T_0} \frac{\Delta E}{C} \tag{9.2}$$

Reducing the temperature of the substrate leads to a shorter effective time constant. From eqn 7.5, we see that this time constant changes as the substrate temperature is adjusted

$$\tau_{eff} = \frac{\tau_0}{1 + \frac{\alpha}{n}\left(1 - \frac{T_s^n}{T_0^n}\right)} \approx \frac{\tau_0 n P_x}{\alpha P} \tag{9.3}$$

where P is the bias power and P_x is the bias power in the extreme feedback regime, as the substrate temperature goes to zero. It should also be noted that adjusting the bias resistance of the film changes the effective time constant by changing the value of α.

Combining eqns. 9.2 and 9.3, and using $P = I_0^2 R$, we find that the pulse magnitude scales as $\Delta E \sqrt{\alpha} / \sqrt{R \tau_{eff}}$.

Some of these components of external noise are characteristically white (such as SQUID noise), and some (such as microphonics) are characteristically pink ($1/f$). Consider first a white external noise source whose current equivalent noise (the current noise in the ETF-TES that would yield the same noise spectrum at the readout) has power spectral density I_n^2. The characteristic bandwidth appropriate for the effective time constant is $1/\tau_{eff}$, so the rms current fluctuations in our pulse bandwidth will be

$I_n\sqrt{B} = I_n/\sqrt{\tau_{eff}}$. The rms uncertainty in our peak height measurement should thus scale as

$$\Delta I_{rms} \sim \frac{\Delta E\sqrt{\alpha}/\sqrt{R\tau_{eff}}}{I_n/\sqrt{\tau_{eff}}} = \frac{\Delta E\alpha}{I_n\sqrt{R}} \tag{9.4}$$

So we find, once again, that the best signal to noise ratio occurs when α is maximized. Also, once again, the substrate temperature does not affect the resolution limit as long as we are in the extreme feedback regime. Thus, reducing the film resistance without changing α (by fabricating devices with lower normal impedance) will improve the resolution limit from constant white external noise sources. Unfortunately, reducing the normal impedance also has the effect of increasing pickup from the environment in the detector itself. Optimizing this parameter must be done by trying devices of different impedances.

Finally, consider a pink ($1/f$) external noise source (such as microphonics). In the white case, the resolution was independent of substrate temperature. Thus, in the $1/f$ case, where the noise rises at lower frequencies, the resolution will improve as the substrate temperature gets colder, the time constant gets shorter, and less low frequency noise is folded into the measurement.

Thus, in the case where energy arrives and is thermalized on time constants shorter than any achievable effective time constant, it is advisable to run about 10% of the way up the transition (to maximize α), and with the substrate as close as possible to base temperature, to minimize the contribution of $1/f$ noise components. The optimization of the normal impedance must be done by trial and error.

9.5.2. Optimal Bias Conditions: Energy Arriving Slowly

If the energy arrives (or is thermalized) on timescales long compared to the effective time constant, the optimal biasing conditions will change. In this case, our calculation of fundamental resolution limits (eqn. 8.47) is no longer valid. Instead, we use the fact that the dominant fundamental noise contribution is the phonon noise, which is $P_n\sqrt{\Delta B} = \sqrt{2kT_0^2 g\Delta B}$ for the ETF-TES. Since the electrothermal feedback cannot reduce the pulse duration past the energy arrival time, the appropriate bandwidth for this measurement is $1/\tau_{arr}$, the inverse of the energy arrival time. The fundamental limit on the rms energy resolution is $\Delta E_{rms} \sim P_n\sqrt{\Delta B}\,\tau_{arr} = \sqrt{2kT_0^2 g\tau_{arr}}$. Note that this equation is

independent of the biasing conditions, indicating that the bias conditions do not effect the fundamental limits on the resolution in this case.

External sources of noise set more stringent limitations. When energy arrives more slowly than the effective time constant, the signal size may be found by assuming that the device remains in quasi-equilibrium: the power removed by electrothermal feedback is at all times equal to the arriving power (assuming that we are in the feedback regime). The power removed by electrothermal feedback is the bias voltage multiplied by the change in current. Thus, the magnitude of the current signal is $\Delta I = P_{arr} / V$. Consider a white external noise source whose current equivalent noise has power spectral density I_n^2. The characteristic bandwidth appropriate for the problem is $1/\tau_{arr}$, so the rms current fluctuations in our pulse bandwidth will be $I_n\sqrt{B} = I_n/\sqrt{\tau_{arr}}$. The rms uncertainty in our peak height measurement should thus scale as

$$\Delta I_{rms} \sim \frac{P_{arr}\sqrt{\tau_{arr}}}{I_n V} \tag{9.4}$$

The voltage bias is reduced as the substrate temperature is increased. The effective time constant is increased as the substrate temperature is increased. Both of these changes improve the resolution. Thus, as long as the device remains in the feedback regime, the signal to noise is improved by increasing the effective time constant (by raising the substrate temperature.)

9.5.3. The General Case

Thus, if the energy arrives (or is thermalized) on timescales short as compared to the effective time constant, it pays to make the device faster, by lowering the substrate temperature. If the energy arrives on timescales long as compared to the effective time constant, it pays to make the device slower, as long as it is kept in the feedback regime.

In conclusion, the optimal bias condition will occur when α is maximized, and the energy arrival time is about the same as the effective time constant. α is maximized by biasing at about 10% up the transition, and the effective time constant is adjusted by changing the substrate temperature. The substrate temperature can be changed by either adjusting the cold stage temperature, or by adjusting the thermal conductance between the substrate and the cold stage.

9.6. ETF-TES Energy Resolution for Heat Pulses

In order to place a limit on the resolution obtainable by this ETF-TES if all the energy was completely thermalized in the W films, a set of 1 keV heat pulses were injected directly into the W. The energy in the heat pulse was self-calibrated by measuring the integral of the SQUID signal for an averaged pulse (fig 9.7). The FWHM resolution of the resulting spectrum was found to be ~20 eV. This resolution was limited by the digitization noise of our 8 bit oscilloscope. Future experiments will set more stringent limits.

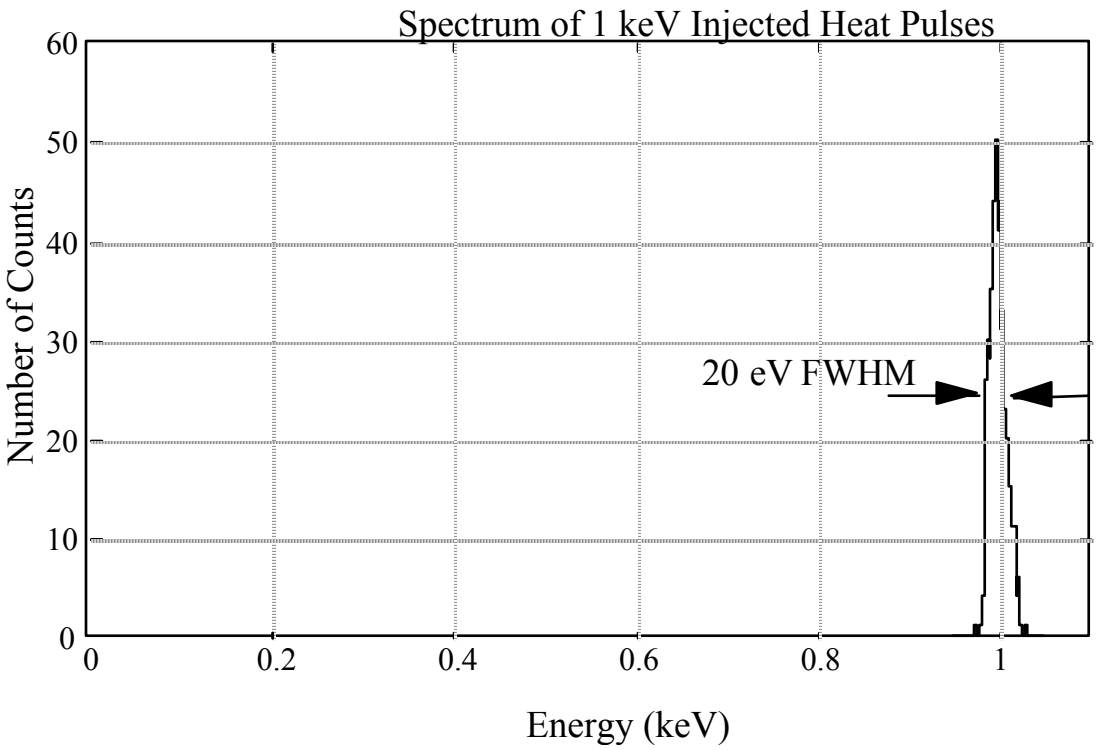


Fig 9.7 Resolution of the ETF-TES for 1 keV injected heat pulses.

9.7. Measurement of coincident events

To investigate the collection of energy from the silicon, we instrumented both W pads on the substrate of our device, and searched for coincident events. When our first W ETF-TES was run at the base temperature of the refrigerator, only a small percentage of the events were coincident. This observation made interpretation of the longer, 300 μs time constant in the ETF-TES difficult. If this time constant was a thermal signal associated with a temperature rise in the substrate, it would be detected in both channels.

If higher energy ballistic phonons multiply scattered within the crystal to be absorbed over these longer time constants, the event would also be coincident. Both the 300 μs tail and the lack of coincidences suggested that the phonon energy in the substrate was propagating very diffusively, and being absorbed elsewhere than the two channels.

Since our energy arrival time was apparently longer than τ_{eff}, we biased the temperature of the substrate up closer to the transition temperature of the W, closer to the optimal bias conditions discussed in section 9.5. Signal to noise improved and small coincident signals were observed above the noise (fig. 9.8).

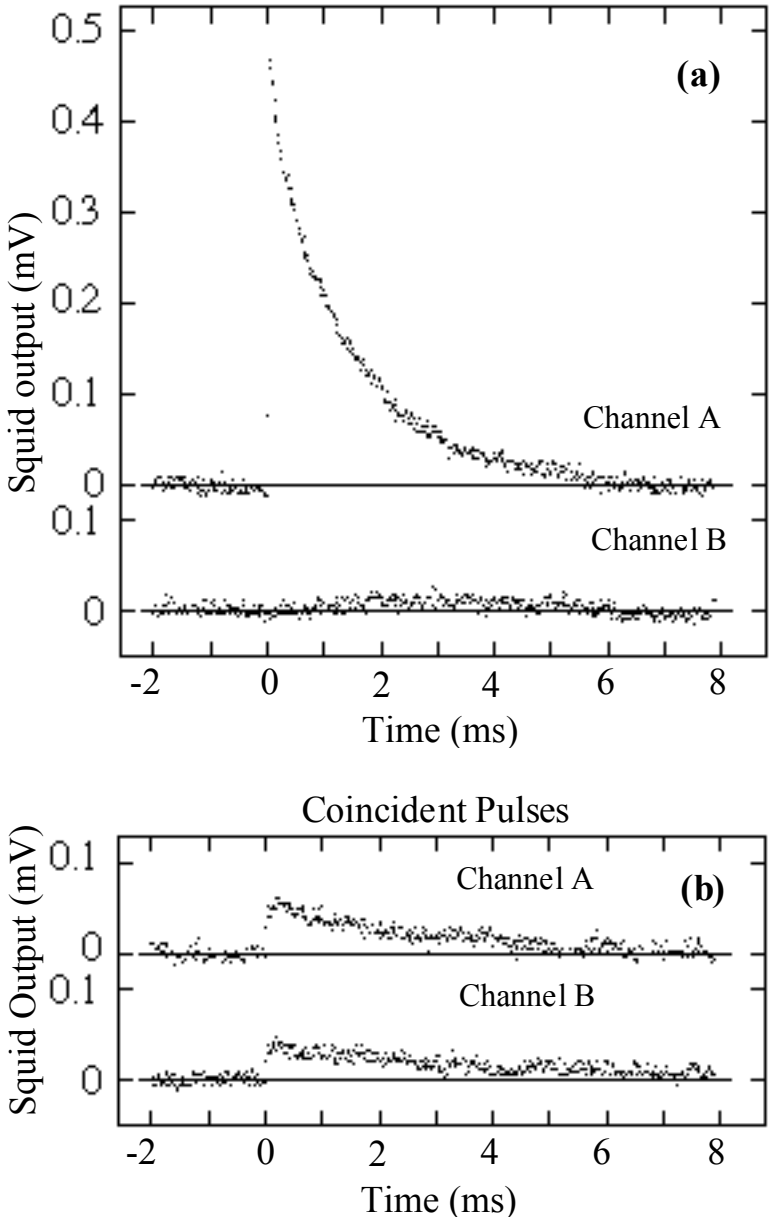


Fig 9.8 Small coincident pulses in our first two channel detector. (a) An event occurring beneath one sensor. (b) An event occurring between two sensors.

When the event occurred directly under one W sensor (fig. 9.8a), it caused a large signal in that sensor, with a fast rise time (~ 5 μs). The coincident signal was very small, and had a long (~ 1 ms) rise time. This is consistent with very diffusively propagating energy. When the event occurred between the two sensors (fig 9.8b), small signals were seen in both channels, with rise times between those of the large and small signals.

In fig. 9.9, the pulse height in one channel is plotted against the pulse height in the other. The events clustered around zero pulse height in both channels are noise triggers. The dimensions of this cluster set the scale of our noise. It is seen that the real events are clearly separated from the noise triggers, indicating that we are seeing events occurring across the entire crystal. The FWHM of noise triggers on the A vs. B plot is seen to be ~2% of the size of the largest signals (the signals directly under the pads).

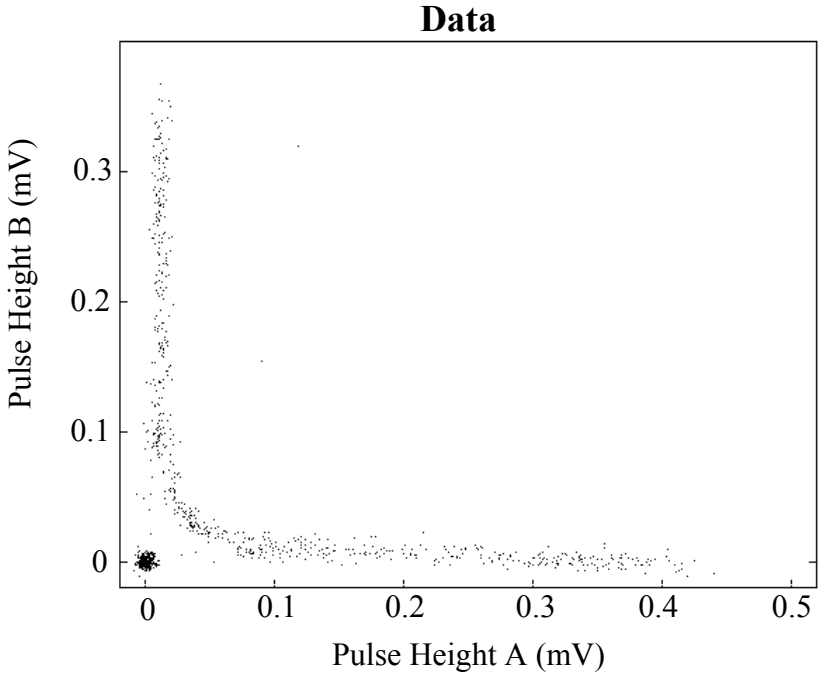


Fig 9.9 Coincident pulses when the substrate is biased close to the transition. Pulse height in channel A is plotted against pulse height in B.

As will be seen in section 9.8, the small coincident signal size in fig 9.9 is consistent with a probability of athermal phonon absorption at the surface of the crystal. Surface phonon absorption may also account for the long coincident rise times; if surface

absorption is large, the coincident signal may be due to the thermal phonons emitted over a longer time period by the channel nearest the event.

It is also possible that the low purity silicon used in these devices may have caused the phonons to propagate very diffusively, accounting for the long rise times. This ETF-TES was fabricated on very low purity off-the-shelf silicon, with a resistivity of ~ 1 Ωcm. All of our subsequent sensors have been fabricated with high purity magnetic Czochralski silicon.

9.8. Monte Carlo Models of Diffusive Phonon Propagation

To model our experimental data, we have returned to the phonon propagation Monte Carlos we used to interpret our W ZTD-TES results (see section 5.6 and Appendix A). In these models, phonons are assumed to propagate ballistically through the crystal. At surfaces, the direction of phonons is randomized. The geometry of the detector is modeled. When a phonon impinges on the W film or a heat sink, it is absorbed with a probability of 0.3. The phonon continues to scatter until it is absorbed either in a heat sink or a TES. The signal in each channel is assumed to be proportional to the total number of phonons it absorbs.

In an effort to reproduce the pulse height A vs B data that we observed, the possibility of phonon absorption at the bare surface of the crystal was introduced. In fig 9.10, a Monte Carlo model for an x-ray interaction in the center of the chip is shown. 10,000 phonons are modeled, and a dot is placed at the site of the absorption of each phonon. The full trajectory of several phonons is traced for illustration. In fig. 9.10a, a model with no surface absorption is shown. In fig 9.10b, a model with a 15% probability of absorption at the surface is shown.

To model our pulse height A vs. B results, 2,000 x-ray events were modeled. The x-ray events were modeled as being evenly distributed throughout the collimation slit. The y-position of the event was taken to be zero, and the x-position was a random number across the 10 mm wide dimension of the crystal. For each x-ray event, the trajectories of 1,000 phonons were modeled. The pulse height in each channel was assumed to be proportional to the number of phonons absorbed.

In fig 9.11a the results of a Monte Carlo simulation with no surface absorption are plotted. Note that the width of the line in the simulation is determined by the number of phonons modeled, while in the data the width is determined by the measurement noise.

As can be seen, the zero absorption Monte Carlo predicts that events with equal pulse height in each channel should have pulse heights of about one fifth of the maximum pulse height (fig 9.11a). In the data, events with equal pulse heights in each channel are seen to have pulse heights closer to one tenth of the maximum pulse height (fig 9.11c). When a 20% probability of phonon absorption at the surface is introduced into the model, the Monte Carlo matches the data much more closely (fig 9.11b).

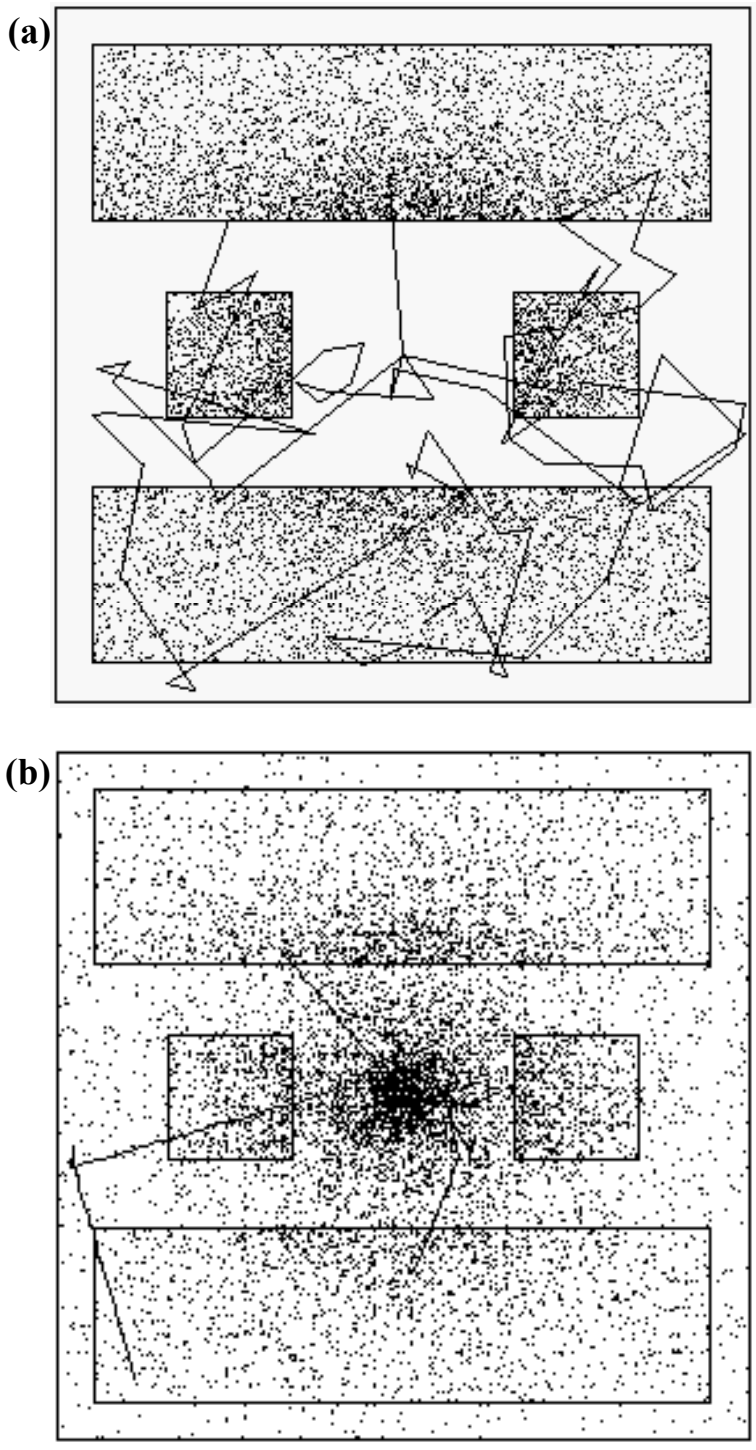


Fig 9.10 Phonon scattering Monte Carlos. Several trajectories are shown for illustration. (a) No surface absorption. (b) 15% probability of surface absorption.

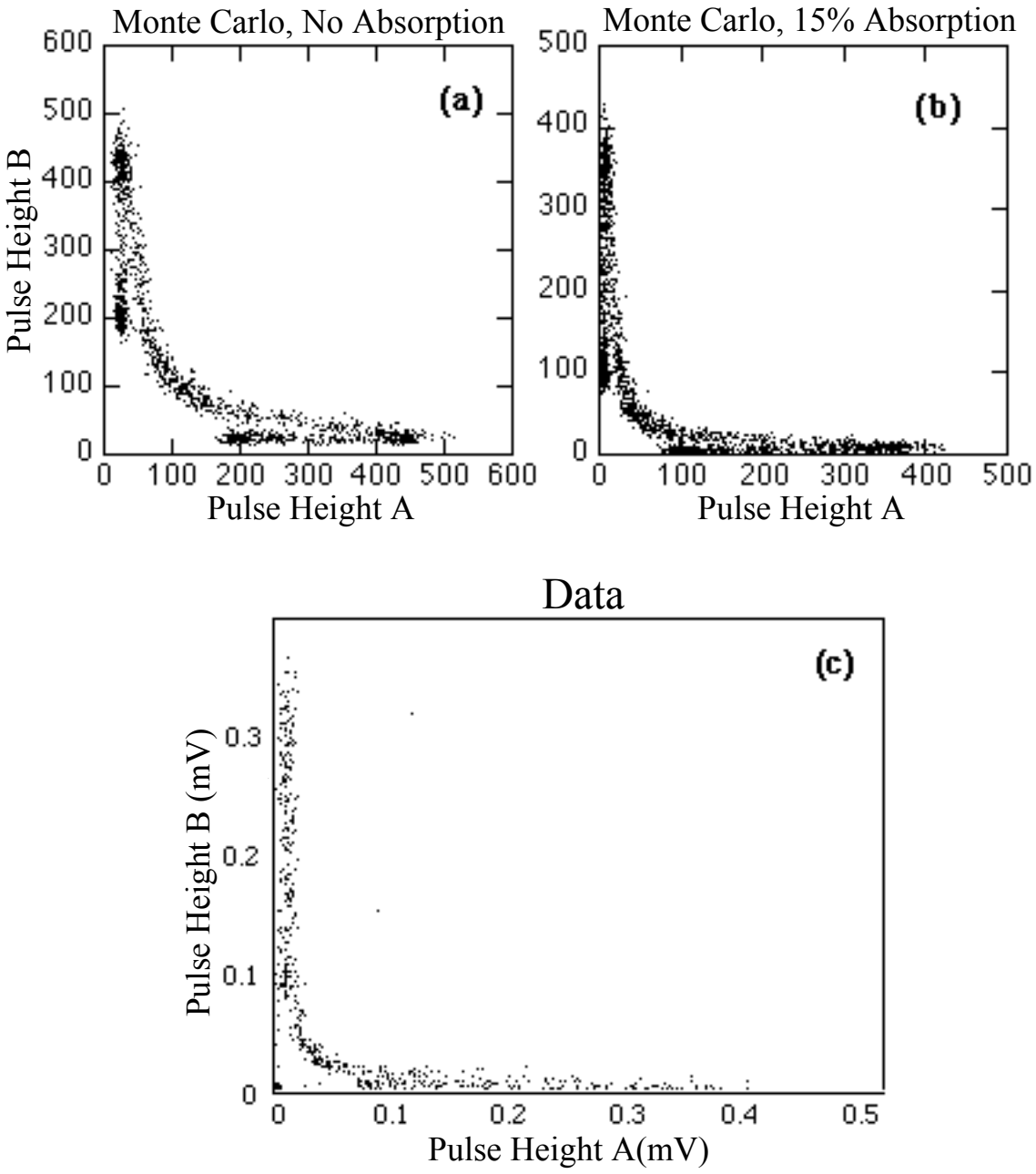


Fig 9.11 Pulse height A vs. B plots. (a) Monte Carlo phonon scattering simulation with no surface absorption. (b) Monte Carlo with 15 % surface absorption. (c) Experimental data is in closest agreement with 15 % surface absorption Monte Carlo.

It should be noted that this data was not initially very reproducible. When we tried to take another set like it later, we were able to get large pulses in each channel, but were not able to see coincidences. We now know that this was due to a change in the surface phonon absorption due to variations in an adsorbed helium layer from residual exchange gas. This effect will be discussed further in section 9.11.

9.9. Deconvoluting Position and Energy

One of our goals in this research was to create a sensor that measured both event position and energy. It is useful to measure the position of the event in order to veto events occurring near the surface of the crystal, as a background suppression measure. Since our pulse height A vs B plots are extremely dependent on event position, we have the potential for very sensitive x-y position resolution. As will be discussed in Chapter 11, timing comparisons with a charge collection device on the back side of the crystal should allow sensitive z-position resolution.

Unfortunately, in the results we have so far presented (fig 9.11c), it is difficult to deconvolute position and energy. If all energy were absorbed in the W pads, and none was lost, pulse height A + B would be constant, and the data would form a straight line. Instead, for events between the two sensors, most of the energy is lost, and the data becomes extremely hyperbolic. In comparing two points in one of the arms of the hyperbola, it is difficult to tell whether the larger event is just closer to the sensor, or has a higher energy.

The less energy that is absorbed in the heat sinks, the less hyperbolic is the data, and the easier the deconvolution of position and energy. In order to explore this possibility, we fabricated the next round of detectors with smaller gold heat sinks that were placed further from the center of the chip.

9.10. The second round of W ETF-TES

Two improvements were made in fabricating our second round of W ETF-TES. First, in an effort to reduce the energy arrival times, the new devices were fabricated on high purity magnetic Czochralski silicon (2 kΩ cm) instead of the low purity, off-the shelf silicon we used previously (~ 1 Ωcm). Note that these substrates were also 1 mm thick, instead of the previous 500 μm thick substrates.

Secondly, in an effort to increase the fraction of energy absorbed in the channels, a different mask was used in the gold evaporation step. This mask deposited four small gold pads in the corners of the chip, instead of two large stripes nearer the middle.

When these devices were cooled, substantially larger coincident pulses were observed. Rise times were fast, and coincidences were easily observed with the refrigerator operating at base temperature.

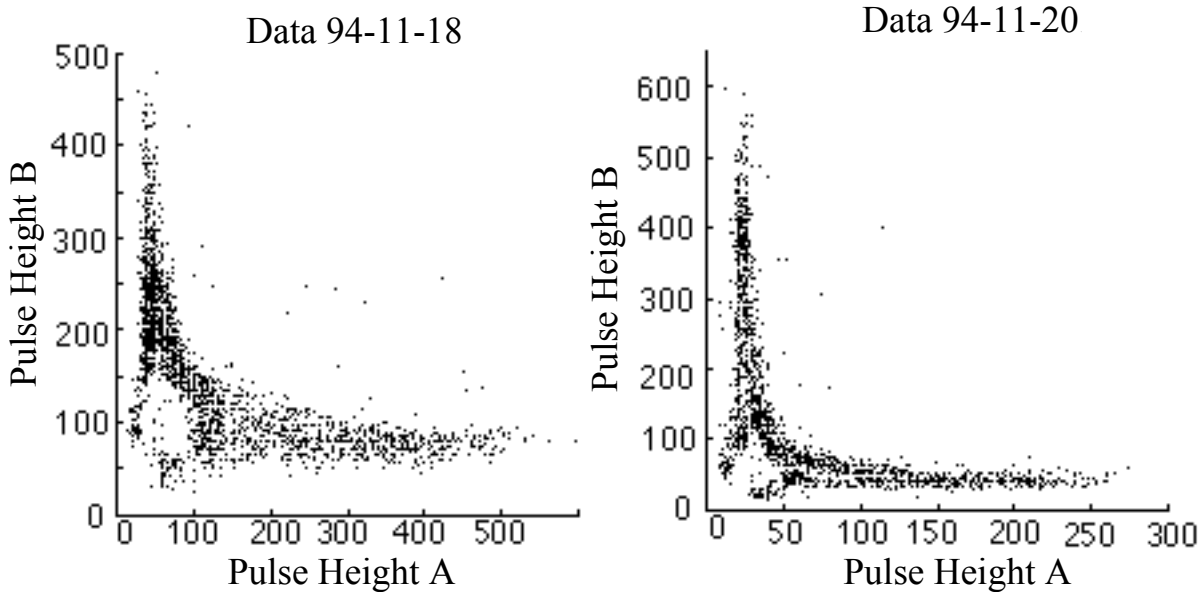


Fig 9.12 Two data sets taken with high purity silicon and small heat sinks. (a) A data set taken shortly after cooling down, with substantial distortion due to SQUID nonlinearity. (b) A data set taken 2 days later, with more surface phonon absorption.

In fig 9.12a, pulse height A is plotted vs pulse height B for one data set. This data set is asymmetric, due to the fact that at this point our SQUID arrays were running in an open loop mode. Any nonlinearities in the SQUID I-V plot thus showed up in the data set. We are now running the SQUID in a feedback mode where these nonlinearities are no longer present.

Note that in fig 9.12a and b, the events clustered on the two "spurs" that are seen near the origin correspond to x-rays incident on the area covered with gold heat sinks.

9.11. Avoiding an adsorbed helium layer

As can be seen, the data is much less hyperbolic than our previous set. As time went on, however, our data sets showed the coincident pulses getting smaller, as more energy was absorbed at the bare silicon surface (fig 9.12b). We now understand that this was due to variations in an adsorbed helium layer on the surface of the silicon crystal.

When we initially cool down the dilution refrigerator, helium exchange gas is added into the inner vacuum can (IVC). When liquid helium is transferred into the dewar outside of the IVC, the exchange gas allows the components inside the IVC to cool down to near 4K. The helium exchange gas is then pumped out for several hours with a leak checker. When the 1K pot is first turned on, most remaining exchange gas adsorbs onto it. When the 1K pot is turned on, the sample tends to be warmer than the IVC wall or the 1K pot, so very little helium adsorbs onto the crystal, and we initially see good coincidences.

As time goes on, the liquid helium level drops, and the top portions of the IVC warm above 4K. Any helium adsorbed to these surfaces comes off. Most of it adsorbs to the 1K pot or other cold portions of the refrigerator, but some of it finds its way to the sensor, which is now quite cold (~ 50 mK). An adsorbed helium layer is a fairly efficient absorber of phonons, so the size of coincident pulses begins to deteriorate. If the 1K pot is ever allowed to run dry, it releases enough helium to make coincidences almost disappear.

We have developed a new refrigerator cooling procedure to eliminate any residual exchange gas. In this technique, the exchange gas is added, and the contents of the IVC are allowed to cool to 4 K. Then the leak checker is set to pump on the IVC. After five minutes, most of the exchange gas has been removed, and the liquid helium is siphoned out of the dewar, allowing the IVC to warm up. This warming dramatically increases the rate at which exchange gas is removed. The sample stage itself is heated as needed to bring it to a temperature somewhere between 4K and 77 K (a reading of 1.1kΩ on the RuO thermometer resistor.) The leak checker is left pumping for about 15 minutes, until it drops to the lowest scale. At this point, any residual exchange gas has been reduced by more than three orders of magnitude over the amounts left in our old technique. Liquid helium is transferred once again into the dewar, and the 1K pot is turned on. Since there is no exchange gas, the sample stage remains well above 4K. It is cooled by slowly circulating the "mash" (or the $^3He/^4He$ coolant mixture) through the dilution unit for about an hour, until enough heat has been removed that the mash begins to condense. At

this point the refrigerator may be cooled normally. Because the contents of the IVC get quite warm in this procedure, and there is no exchange gas, it can take a day of operation before the refrigerator settles down stably to its base temperature.

We have found that using this cooling technique, and running the SQUID with feedback, makes our results extremely reproducible.

CHAPTER 10: THE W/AL QUASIPARTICLE TRAP ASSISTED ETF-TES

In order to scale the W ETF-TES up to a practical dark matter or neutrino detector, we must use a much larger absorber mass. To scale up to a large absorber while maintaining good position and energy resolution, it is necessary to instrument a much larger detector surface area. Instrumenting large areas cannot be done simply by increasing the W pad area without making the W heat capacity quite large.

Fortunately, it is possible to collect phonon energy over large surface areas by using a technique called quasiparticle trapping, first suggested by N. Booth for use with tunnel junction detectors [27]. In this technique, superconductors with two different gaps are used. The high gap superconductor, which is chosen to have a gap smaller than the incident phonon energy, is deposited as a large area pad. The lower gap superconductor is placed in intimate electrical contact with this pad. When phonons impinge on the high gap superconductor, they break Cooper pairs, forming quasiparticles. The quasiparticles relax to the gap edge, releasing phonons. If these phonons are above the gap, they create more quasiparticles, while any sub-gap phonons escape back into the substrate and are lost. The device geometry is chosen so that the time for quasiparticles to diffuse to the low gap superconductor is short as compared to the quasiparticle recombination time and the desired risetime. When a quasiparticle diffuses into the low gap superconductor, it relaxes to the new gap edge, releasing phonons, and is trapped in the potential well of the lower T_c material. The phonons released in this process create new quasiparticles, or escape into the substrate if they are sub gap. In Booth's devices, the quasiparticles are then measured when they cross a superconducting tunnel junction [27].

A similar technique may be used with transition edge sensors. The gap of Al is smaller than the athermal phonon energy, so it may be used as the high gap superconductor. Large Al pads are placed in electrical contact with the TES. Incident athermal phonons create quasiparticles, which diffuse to the TES on times short compared to the quasiparticle recombination time. Since the TES is on or near its transition edge, it has many thermally excited quasiparticles, and behaves much like a normal metal. When a quasiparticle diffuses into the W film, instead of just being "trapped," it interacts strongly with the electrons in the W, and its energy is rapidly thermalized. Note that this process is substantially more efficient than quasiparticle trapping with tunnel junctions. In the tunnel junction devices, the quasiparticles require many more attempts before crossing the interface, resulting in a long tunneling time. Also, in the tunnel junction devices, the subgap phonons that are shed when the

quasiparticles relax to the gap edge of the low T_c superconductor are lost into the substrate. In the quasiparticle trap assisted ETF-TES, the entire quasiparticle energy is thermalized when it enters the W film. Note, however, that in both cases about 40% of the energy is lost to sub gap phonons when the quasiparticles are initially created in the high gap film [28].

10.1. Quasiparticle trap assisted titanium TES

A. T. Lee has previously demonstrated the technique of using quasiparticle traps to improve the energy collection of our titanium transition-edge sensors [12]. In this experiment, short (7 μm long) Ti links were used to connect large Al pads. Some energy collection enhancement was demonstrated, although not as much as we expected.

One problem with this technology was that the bias temperature of our Ti films (~380 mK) was not sufficiently below the Al transition (~600 mK) for the Al films to be completely thermally insulating. As such, in addition to providing quasiparticle trapping, the Al films acted as heat sinks, and accelerated the rate of heat flow into the substrate. This effect, combined with the short intrinsic time constant of detectors operating at this higher temperature, meant that some portion of the energy deposited by the quasiparticle trapping was lost into the substrate before the collection of energy was complete. The loss of energy was more severe for larger area pads, as the quasiparticle trapping time goes as the square of the pad dimension.

For perfect quasiparticle trapping, we would expect about 60% energy collection, with the rest of the energy escaping as sub-gap phonons from the high gap material. For 20 μm x 20 μm pads, energy collection of ~ 30 % was demonstrated. For 40 μm x 40 μm pads, energy collection of ~ 14 % was demonstrated.

10.2. A parallel array W ETF-TES

In order to apply quasiparticle trap energy collection to a W ETF-TES, we must instrument it in a widely spaced parallel array. This spacing will allow the collection of energy over a large area, with many quasiparticle traps. The fabrication of a parallel array W ETF-TES will be discussed before the implementation of quasiparticle trapping.

In the parallel array W ETF-TES, a large number of W lines (typically between 100 and 1000) are connected in parallel by two narrow aluminum bars. When the aluminum bars are held at a constant potential difference, each W link individually self-biases on its transition, and act as an ETF-TES. When a particle event deposits energy

into the W lines, the current through each line is reduced as that energy is removed by electrothermal feedback. Since the lines are connected in parallel, the current through all of the lines is added before being measured by the SQUID. The total energy deposited in all of the lines is then equal to the voltage bias multiplied by the integral of the current pulse.

One advantage of instrumenting a large array of such parallel lines, instead of a pad, is that the effect of any gradient can be reduced. Consider, for example, a device of 100 parallel elements. In order to create a device with a total impedance of $\sim 1\ \Omega$, each individual segment must be $\sim 100\ \Omega$. This 100 Ω line may, however, be packed into a tight meander in order to minimize the variation in T_c across it. As each element individually self-biases, only the total T_c variation across an individual element is important. For a 100 element device, the distance across one such meander is ~ 20 μm, instead of the ~ 2 mm length scale of a pad, and the effect of a gradient is reduced by two orders of magnitude. The improvement is even more dramatic when compared to the possible variation in T_c across an entire 3" wafer.

There were initially some concerns about the stability of the parallel array configuration. Unlike the pad-format W ETF-TES, each individual element of a 100 element parallel array is a W line with impedance ~ 100 Ω. Since the sheet resistance of our 30 nm thick W films is about 1 Ω/square, a 100 Ω, 2 μm wide line would be 200 μm long. Each element can itself be considered to be made up of 100 2 μm x 2 μm sub-elements connected in series. The ETF-TES is stable in a parallel sense, where each individual element has the same bias voltage, and any element that cools will draw enough joule power to keep it at the point of stability. When wired in series, however, this Joule power is distributed in the wrong places for self-biasing. Since the current is constant through one of these parallel lines, the sections of the line that are hottest, and have the highest resistance, receive the most Joule power.

The first worry is that this condition will lead to oscillations. A moment of deeper reflection reveals, however, that the relevant heat flow equation has no "momentum" term, which is necessary for oscillations to take place. The second concern is that one of several quasi-stable states might occur, where, for instance, one half of the line goes superconducting, and the other goes completely normal. Further modeling indicates that, as long as the thermal conductivity down the line is constant, and the heat flows down the W line is much faster than into the substrate (as is true for operation at 70 mK), the device should return to a unique equilibrium state on the timescale for heat to flow the length of the W line. Note that the stability is not as certain if the gap in intermediate sections of the film is elevated (by the proximity effect), reducing the

thermal coupling between sections of a W line. This possibility will be discussed further in section 10.8.1.

10.3. The quasiparticle trap assisted Al/W ETF-TES

The second major advantage of using a parallel array, and one of the original reasons for the proposal of the ETF-TES, is that it interfaces well with an Al quasiparticle trapping scheme. The characteristic time for quasiparticles in an Al pad to diffuse to a W segment increases as the square of the length scale of the Al pad. The quasiparticle diffusion time thus sets a limit on the size of the Al pad, since, for optimal performance, it should be no longer than the effective electrothermal time constant. Thus, to instrument a large area, a number of Al pads must be used, each in electrical contact with a W ETF-TES segment. The ETF-TES allows individual parallel segments to be spaced widely, with each segment self-biasing in the middle of its own transition.

10.4. Al/W ETF-TES time constants

As shown in section 9.5, the optimal resolution is achieved when the time for energy to be thermalized in the detector is of the same order as the effective electrothermal feedback time constant. It is also desirable to make the effective time constant short, so as to operate as far into the extreme feedback regime as possible, and reduce the $1/f$ noise incorporated into the measurement. Thus, in order to optimize the performance of a QTA ETF-TES, it is necessary to understand the various time constants associated with detector operation, and how the device geometry determine these time constants.

Two of the time constants associated with the Al/W ETF-TES, the intrinsic time constant for heat to escape into the substrate, and the effective time constant due to electrothermal feedback, have already been discussed extensively. We have also mentioned the electronic time constant $\tau = L/R$ associated with the SQUID input coil inductance and the bias resistance of the W film.

When a particle interaction occurs in the crystal, athermal phonons propagate to the surface. The phonons that are incident on the sensor cause an initial "prompt" signal that arrives faster than the electronics time constant. Over longer timescales, energy is absorbed from multiply scattered phonons. If the silicon is pure and the surface is clean, the length of this multiply scattered phonon time constant is determined by the size of the

crystal, the sensor area coverage, and the efficiency of phonon absorption. If the crystal is larger, this time constant will increase, as it will take longer for phonons to traverse the absorber. If the sensor area coverage and absorption efficiency is higher, this time constant will get shorter, since the phonons must scatter fewer times before being absorbed. This time constant was observed to be ~ 300 μs for our W pad ETF-TES on very low purity silicon (~1 Ωcm). It should be shorter for the Al/W ETF-TES, which is fabricated on higher purity silicon (~2 kΩcm), has superior area coverage (because of Al quasiparticle trapping), and has higher absorption efficiency (because Al films have a better acoustic impedance match to Si than W does). If necessary, a Monte Carlo phonon scattering model with the device geometry might be used to analyze this time constant.

Once the phonon energy is absorbed in the quasiparticle traps, there is a characteristic time that it takes for this energy to enter the W film (the quasiparticle diffusion time). The mean free path of electrons in the Al films used in our first round of W/Al ETF-TES is limited by our film thickness to $l \approx 60nm$. The Fermi velocity of Al is $v_f = 2.02 \times 10^8 \, cm/s$ [18]. The diffusivity of our Al films is thus $D = v_f l/3 \approx .04 m^2/s$ [29]. In the W/Al ETF-TES we tested, our Al pads are ~100 μm long. The mean time for a quasiparticle to diffuse a distance L is $t = L^2/2D$ [29]. The mean quasiparticle diffusion time we expect is thus ~100 ns. Quasiparticle recombination times in Al at 100 mK are > 60 μs [30], so we expect that most of the quasiparticles are trapped in the W.

Finally, there is a characteristic time constant for the heat to diffuse down the W lines, and come into equilibrium. We will refer to this time constant as the "thermalization time constant." It is desirable to keep the normal resistance of the sensor to ~ 1 Ω, in order to minimize pickup from the environment, and reduce the heat dissipation in the voltage bias resistor. As mentioned previously, if 100 - 1000 line segments are instrumented, the impedance of each individual segment must be ~ 100 Ω - 1 kΩ. If 2 μm lines are used, and considering that the sheet resistance of our 30 nm W films is ~ 1 Ω/square, each segment will be ~ 200 μm - 2 mm long. The first W/Al ETF-TES we tested was fabricated with W lines 800 μm long. The Fermi velocity of electrons in W is $v_f = 1.35 \times 10^8 \, cm/s$ [18]. The density of states in W is $N(0) = 0.16 \times 10^{23} \, eV^{-1} cm^{-3}$ [18]. The effective density of electrons in W is $n = N(0) m_e v_f^2 / 3 = 5.49 \times 10^{22} \, cm^{-3}$, where m_e is the mass of the electron. The mean free path of an electrons gas is

$$l = \frac{m_e v_f}{n e^2 \rho} \tag{10.1}$$

where ρ = 6 μΩcm is the measured resistivity of our W films, and e is the charge of the electron. Eqn 10.1 yields a mean free path in our W films of ~ 15 nm, which is half of the thickness of our W films. Since we are not yet film thickness limited, thinner W films could be used without increasing diffusion times. The Diffusivity in our W films is thus $D = v_f l/3 \approx .007 m^2/s$. The characteristic time for a quasiparticle to diffuse 800 μm is $t = L^2/2D \approx 50\mu s$.

The thermalization time constant may, however, be reduced by using more than one quasiparticle trap per parallel W segment. To understand this point, assume that the flux of phonon energy on the quasiparticle traps is approximately the same over several hundred micrometers. If several Al quasiparticle traps are used to make thermal contact with one W link at approximately equal intervals, and equal heat is dumped into the W film from each trap, then to first order the thermalization time constant will be determined by the distance between the links, rather than by the complete line length. This will lead to a shorter "effective" thermalization time constant. It should be noted, however, that there may be a longer time constant associated with unequal distribution of energy in the quasiparticle traps, and edge effects associated with the ends of the W line. This effect should, however, be second order. In the first W/Al ETF-TES we tested, quasiparticles are injected at 8 points along the 800 μm. Thus, to first order thermalization will occur across 100 μm length scales, or $t = L^2/2D \approx 0.7\mu s$.

10.5. Design of a W/Al ETF-TES

We have fabricated W/Al ETF-TES with several different geometries. In each case, several hundred W links are wired in parallel. The links are deposited in a tight meander in order to minimize the variation in T_c across the link. Depending on the design, a number of different Al quasiparticle traps are placed around each link, and electrically connected to it at approximately equal intervals. The links are patterned into a maze-like configuration that both preserves a small area coverage and an equal spacing between the quasiparticle traps. The connection to the quasiparticle traps is made via a spur of W that protrudes from the meander and runs some distance under the Al film. It is undesirable for the Al film to be in direct contact with the meander itself, as the proximity effect might in that case raise the T_c of a small section of the W film sufficiently to split the W line into a series of thermally isolated segments. This effect would prevent the whole W segment from coming into thermal equilibrium, and might lead to an unstable configuration, where one segment might become completely normal while another is driven completely superconducting. In fig. 10.1, a mask with four

parallel W segments is shown (the masks we actually use contain hundreds of segments). The Al bars that maintain the bias voltage across the parallel segments are interleaved in a comb-like configuration.

In most of our first masks, and in fig. 10.1, the W spur that connects to the quasiparticle traps was made long enough to protrude well into the Al film. This spur was intended to assist the quasiparticle trapping process by creating a two step trap. If the effective coherence length in the W films is long compared to the film thickness, the T_c of the W films will be raised by the proximity effect. This effect might cause an efficient trapping of quasiparticles into the W spurs, which would then diffuse into the W meander, where they would be thermalized. In retrospect, using a long W spur may instead have increased our thermalization time constant. Since the diffusion times in the Al are much shorter than the diffusion times in the W lines, a long spur would have the effect of increasing the time constant over which the heat is injected into the W meanders.

In order to test the effect of the long W spur, we have fabricated several four channel devices in which, on the same chip, two channels have short W spurs, and two channels have long W spurs. By comparing the signals from each case under the same operating conditions, we expect to be able to determine the effect of the length of the W spur.

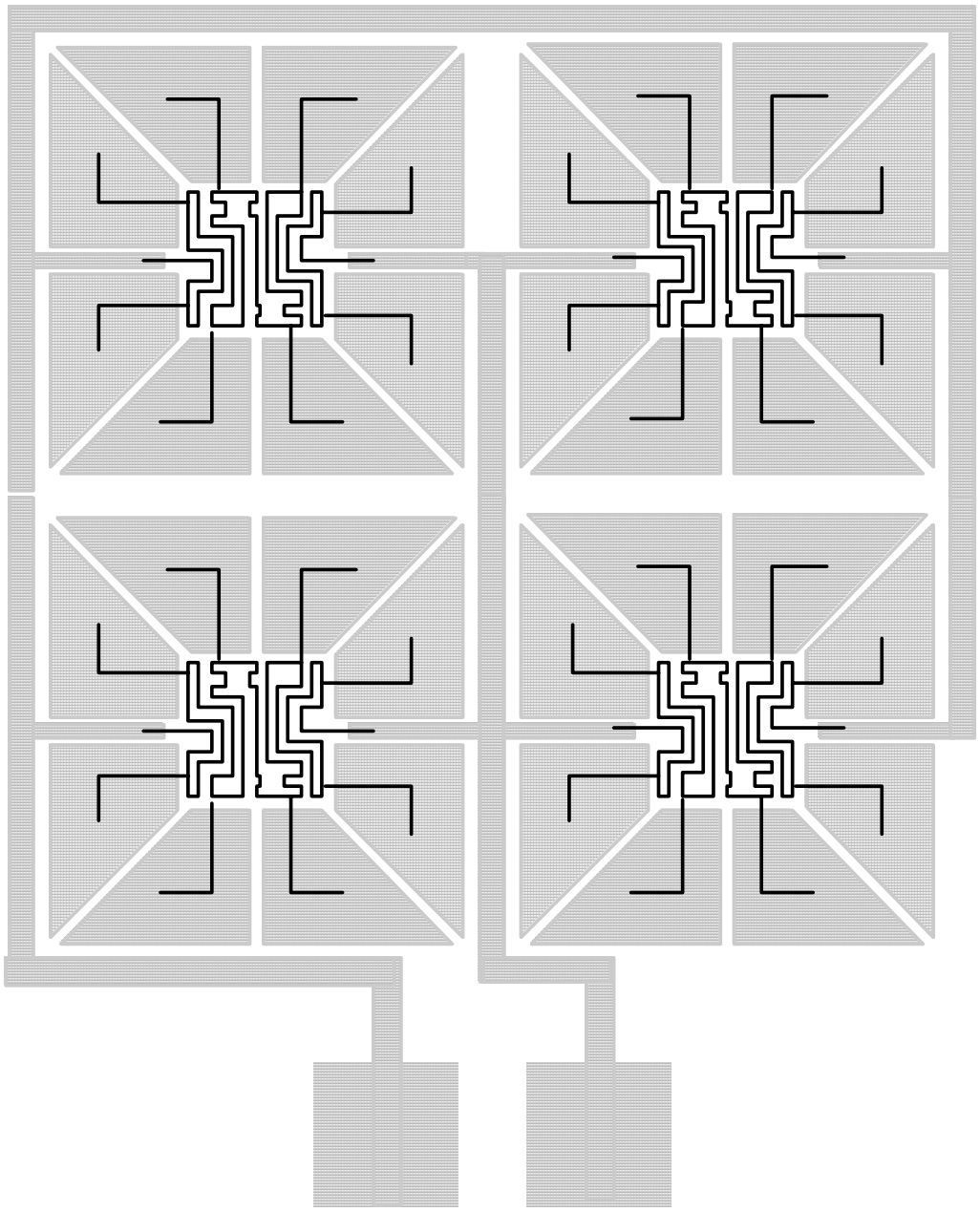


Fig 10.1 The layout of a 4-segment Al/W parallel ETF-TES. The actual masks contain several hundred segments.

10.6. Fabrication of the W/Al ETF-TES

Fabricating a W/Al ETF-TES requires a two layer process in which the top layer is etched without affecting the bottom layer. Fortunately, the chemistries of W and Al are sufficiently different that our W and Al wet etches are both fairly selective. W does not oxidize strongly in air, but has a fairly weak oxide. The W etch is a simple oxidizer (20% H_20_2 solution) which creates and carries away a tungsten oxide in solution. Al, on the other hand, oxidizes strongly in air, and its oxide is very robust. The principal characteristic of the Al etch is its ability to dissolve Al_20_3. The Al etch rate is highly dependent on processing temperature, but at one temperature it was found to etch Al films at several hundred nm/min, while it etched W at only ~ 5 nm /min. The W etch was even more selective. Because the Al oxide is not soluble in water, Al seems to be completely stable in the W etch.

Small levels of contaminants in Al films tend to cause an elevation of T_c, which does not hurt our quasiparticle trapping process, as long as the gap is still smaller than the characteristic athermal phonon energies, and film inhomogeneities do not cause trapping of quasiparticles away from the W interface. Thus, we are principally concerned with the quality of our W films, which are known to be strongly dependent on deposition conditions. There is also reason to believe that the properties of the W film may be dependent on the condition of the substrate. Our first candidate W/Al ETF-TES sensors were thus fabricated with the W film on the bottom, so that it could be deposited on clean silicon wafers. After the patterning of the W film, the wafers were put back in the Balzers, back-sputtered in situ to remove any residual oxide, and an Al film was deposited on top. When the second layer was patterned, all areas not covered by the mask were etched down to either bare Si or the W film.

When these first candidate sensors were tested, it was discovered that the T_c of the W film was elevated (about 150 mK) and extremely broad (about 30 mK width.) If a residual layer of higher T_c was left on the surface of a W film, the proximity effect might then create an extremely broad, elevated transition. Apparently, when the Al film was etched down to the W film, a residual layer containing some Al was left which exhibited a higher T_c. Since the wafers were overetched in the Al etch, and Al films are consumed very quickly under these conditions, it is clear that this residual layer of W and Al must be somewhat resistant to the standard Al wet etch.

Rather than attempting to develop processing steps to remove this residual layer, in our current Al/W ETF-TES sensors we have reversed the fabrication steps, and deposited the Al first. Thus, the Al film is deposited and patterned, and then the W film

is deposited. The sections of W film that will be patterned into a meander are deposited onto bare silicon, and never come into contact with Al.

When the second deposition occurs, the wafer must be backsputtered to remove any oxide before the deposition of the W layer. In order to prevent the backsputtering of Al into the bare Si areas where W is to be deposited, a thin protective layer of W is deposited on the Al film during the first deposition. As will be described, this technique prevents any Al from being backsputtered onto the bare Si areas before the second W deposition.

The first, failed W/Al ETF-TES were fabricated on 500 μm thick off-the-shelf silicon. Our current W/Al sensors are fabricated on 1 mm thick, 3" diameter [100] silicon wafers made of high purity, magnetic Czochralski silicon. The wafers are pre-cleaned with a dirty sulfuric dip, followed by HF, and then a clean sulfuric dip. In our previous sensors, these steps were followed by another HF dip to remove all the remaining oxide, but this step was intentionally omitted to leave a residual native oxide beneath the deposited Al films. An Al film 60 nm thick is then deposited in a Balzers 450 DC magnetron sputtering system. Without breaking vacuum, a 30 nm thick W film is deposited over the Al with a -200 V DC bias.

The wafers are then baked in a singe furnace at 150 C for 30 minutes before applying the standard resist in the Silicon Valley Group coater.

The masks for patterning these two layer devices was designed using Magic. Since this is a two layer mask, it is important to include the alignment information used by the Ultratech aligner on this mask.

The alignment is done on an Ultratech 1000 projection aligner in CIS that was modified to allow processing of thick Si wafers. The resist is developed in the Silicon Valley Group developer.

During the first etch, both the W and the Al layer must be removed where not covered by resist. This requires two etch steps. Both are done in the metal wet bench in CIS. The W etch is done with a fresh solution of 30% hydrogen peroxide. Etch times are typically 200 seconds for 30 nm thick films, but surface residues vary, and hence the activation time can cause variations. Fortunately, this step is fairly forgiving to overetches. The Al film is etched in the standard aluminum etch. Etch rates are typically hundreds of nm per minute, so etch times of 15-30s are typically used. The resist is stripped only with the clean PRS 1000 dip. It has been found that the EMT resist strip etches tungsten, so this strip is NOT used.

After the wafers are loaded in the Balzers for the second deposition, the system is baked out, and then an RF backsputter is done to remove all of the surface oxide. Since

the Al film is covered by a thin layer of W, it is expected that little Al will be backsputtered onto the Si surface. The system is then pumped down again, typically to ~ 10^{-8} torr.

The second W deposition typically occurs at a power of 2.5 kW, and an argon pressure of 5 mbar. A DC bias voltage of -200V is applied to the substrate during deposition. Typical drum rotation speed is 8 Hz. We found that the dc bias power slowed down the deposition, such that deposition times are 30% longer with the dc bias applied.

The second patterning proceeds much as the first, with the addition of an alignment step in the Ultratech. The only etch step is the W etch. As has been noted before, the Al film is not removed by the W etch. During the etch step, wherever there is no resist, the W film is removed down to either bare Si, or bare Al (the quasiparticle trapping pads.) Any residual layer left on the Al pads is not as critical to device performance as a residual layer left on the W would be.

After processing, the wafers are diced into 1 cm x 1 cm dies using a Kulicke & Soffa wafersaw. Blue sticky tape is used to protect the wafer surfaces (both back and front) during the dicing. In our recent experiments, we have noticed some residue left by the blue sticky tapes, so in the future the tape may be applied over a layer of photoresist.

10.7. W/Al ETF-TES Experimental Results

The results presented here are from 2 channel W/Al ETF-TES devices fabricated on 1 cm x 1 cm x 1 mm silicon chips. Each channel covers a 2 mm x 4 mm area. The channels are spaced 5 mm apart center to center, and are centered on the chip. Each channel is made of 200 parallel W lines. Each line is 2 μm wide, 800 μm long, and 30 nm thick. The device was designed to have a normal impedance of 2 Ω just above the W transition temperature. The lines are packed into a tight meander so as to minimize any gradient in T_c across the line. Eight Al quasiparticle trap pads are connected to each W line at equal intervals along its length. The W lines were wired in parallel with Al bars that were connected to Al bonding pads. Al wire bonds were used to connect the device to the SQUID and voltage bias circuit.

In the first run, the W/Al ETF-TES was fabricated with no gold heat sinks. Heat sinking was accomplished instead with four beryllium copper springs pressed against the back side of the crystal. Some later runs had gold heat sinks deposited around the edge of the crystal, as in the W ETF-TES pad detectors.

X-rays from the ^{55}Fe x-ray source were collimated in a slit across the two device channels, and were incident on the back side of the crystal. The source geometry was the same as in our W pad experiments.

When the probe was cooled, the exchange gas was fully pumped out as described in section 9.11. When base temperature was reached, the device was found to stably self bias. Pulse signal to noise was expected to be much better than in the W pads, as the heat capacity of the W lines in the W/Al ETF-TES is about 20 times smaller than the heat capacity of the W pads in our W ETF-TES. Observed pulses were large and coincident. The signal to noise was much better than that observed with the simple W pads (fig 10.2).

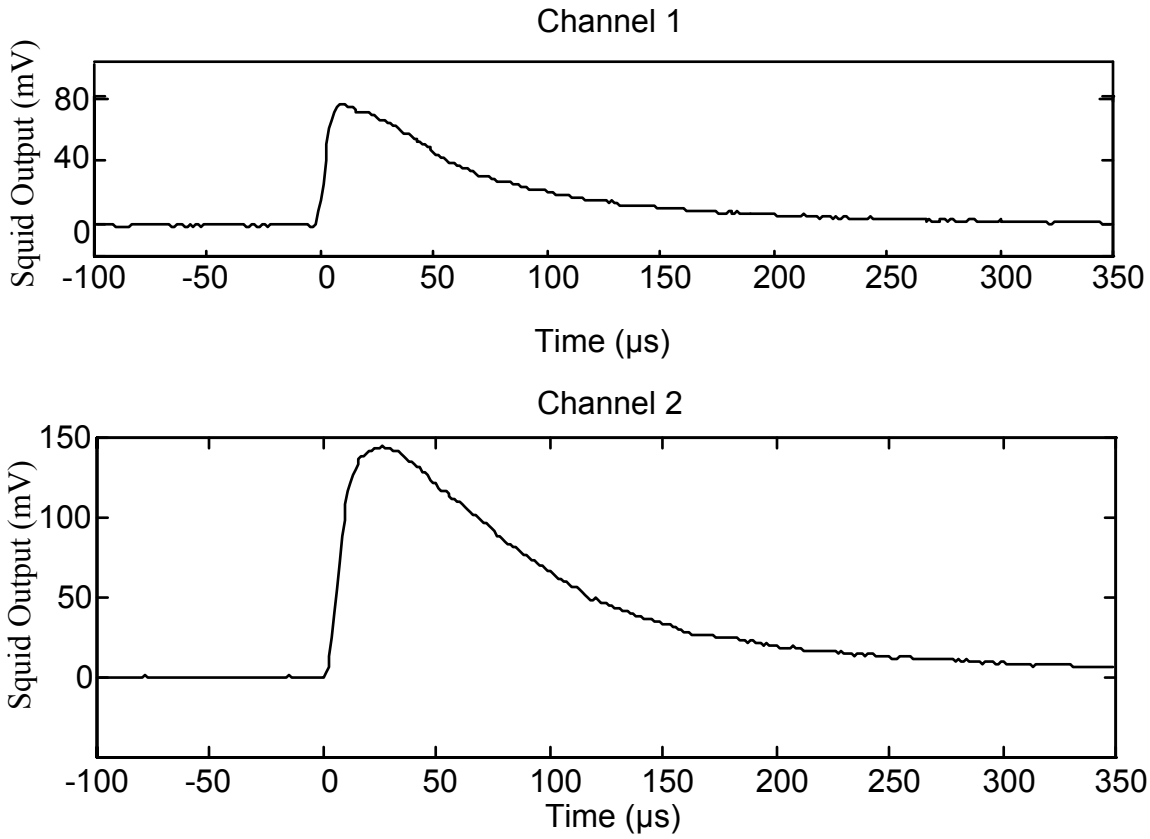


Fig 10.2 A set of coincident pulses in the W/Al ETF-TES. The event is a phonon signal from a 6 keV x-ray interacting near the back of a 1 mm thick Si crystal. 350 kHz low pass filters were used in both channels.

The total energy removed by electrothermal feedback is simply the voltage bias multiplied by the change in the SQUID current. If the entire pulse is integrated, however,

substantial low frequency noise is introduced. In an effort to estimate the pulse integral while minimizing low frequency noise, the pulse was first integrated, and then run through a high pass filter with a knee at $\omega = 1/\tau_{pulse}$, where τ_{pulse} is the observed time constant of the pulse (neglecting a long, low amplitude pulse tail that may be due to the absorption of subgap phonons directly into the W film). The height of the filtered pulse integral was taken to be an estimate of the pulse integral. This estimate will be somewhat low, since the tail of the pulse is not fully measured, but the resolution of the resulting pulse is improved.

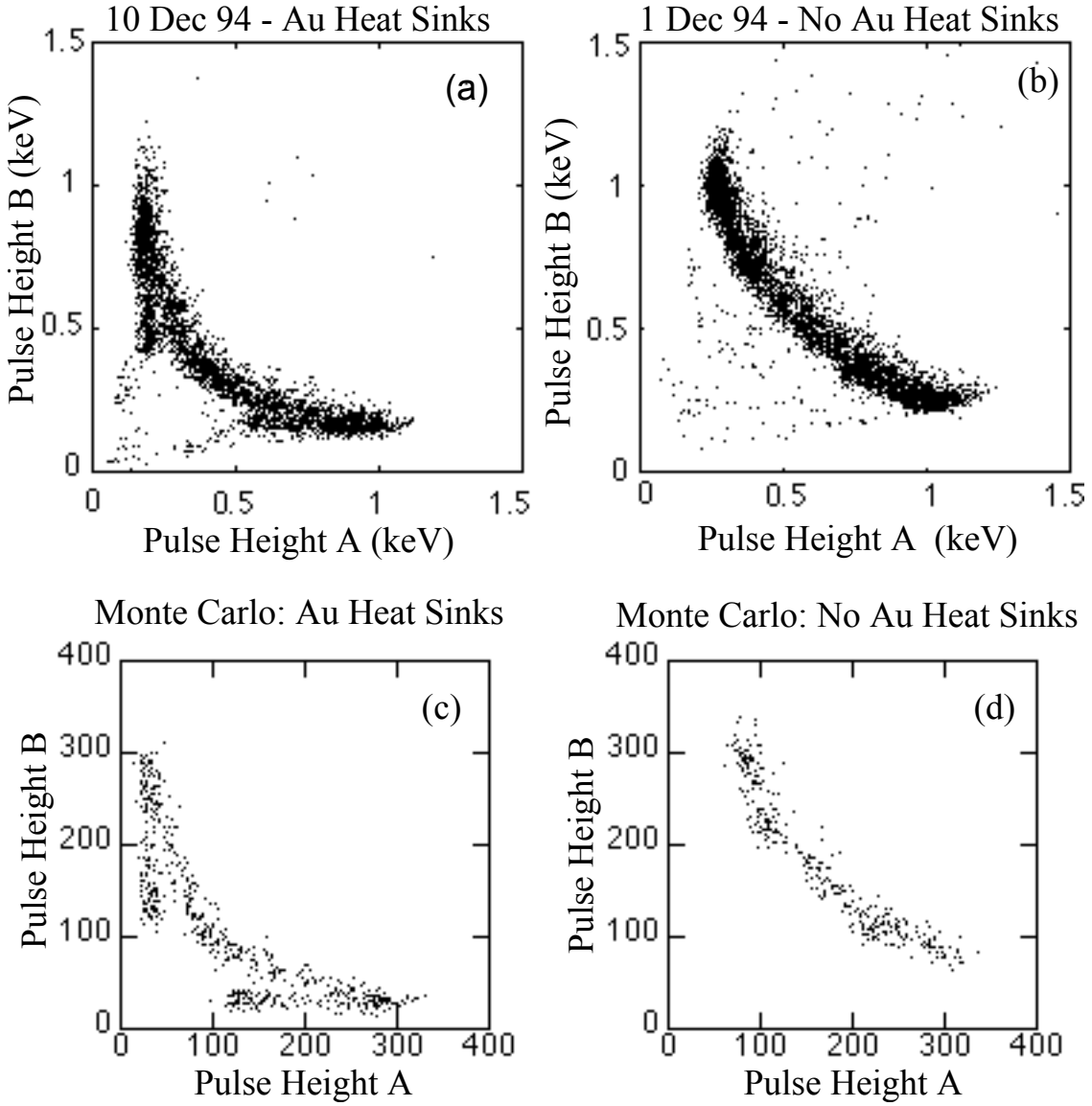


Fig 10.3 Pulse Integral A vs B plots for the W/Al ETF-TES. (a) Data with Au heat sinks (b) Data with no Au heat sinks. (c) Monte Carlo with Au heat sinks. (d) Monte Carlo with no Au heat sinks

In figure 10.3b the total feedback energy in channel A is plotted against the total feedback energy in channel B. Since no gold heat sinks were present, and little exchange gas remained to adsorb on the surface of the crystal, the data is almost a straight line. A Monte Carlo simulation models reproduces the qualitative features of the data with the device geometry, and assuming 50% probability of absorption at an Al interface (fig 10.3d).

In a later run, a device was cooled with a gold heat sink on the edge of the back side. The feedback energy A vs B is plotted in fig 10.3a. A Monte Carlo with this device geometry is shown in fig 10.3c.

10.8. Charge trapping effects

When a particle interaction occurs in a silicon crystal, both phonons and electron-hole pairs are generated. The electrons and holes rapidly relax to the band edge, shedding phonons which can be detected by the sensor. If electron-hole pairs recombine, they are converted into phonons which can also be detected. If electrons and holes persist beyond the pulse duration, their energy is not measured. Finally, if electrons or holes are trapped, they will give up a small fraction of their energy to phonons. Any statistical variation in these processes, or spatial variation which is not deconvoluted, worsens the energy measurement resolution.

When the crystal is initially cooled down, the minority impurity sites are ionized, along with an equal number of majority impurity sites [31]. We refer to this condition as "mode 1". Electrons and holes can be trapped at these sites. After sufficient exposure to a radiation source, however, all impurity sites are filled, and charge carriers can propagate all the way across the crystal ("mode 2"). At operational temperatures, thermal processes are insufficient to reionize the impurity sites.

10.9. Energy Resolution of the W/Al ETF-TES

Our group has found that titanium transition-edge sensors collect more energy when run in mode 2. In an attempt to improve the resolution of the W ETF-TES sensor, the crystal was irradiated with a photodiode for several hours to drive it into mode 2

before taking a data set . A sensor with no gold heat sink was used. The resolution was notably improved. In fig. 10.4a, the pulse integral in channel A is plotted against the pulse integral in channel B. The pulses are normalized so that the sum of the integrals in A and B is centered on 6 keV. The 5.9 and 6.5 keV x-ray lines are clearly separated.

In this data set, S. W. Nam implemented several other improvements to the data acquisition system. In all our previous sets, pulses were digitized by an 8 bit oscilloscope. In this run, the pulses were digitized by a 12 bit digitizer on a VXI bus. Also, for the first time, the SQUID arrays were run with feedback electronics.

In this set, deconvolution of the position and energy is complicated as the data folds back on itself. When a pulse occurs in the center of the crystal, the pulse integrals contain the same energy in both channels. Events that occur directly under one pad lead to maximum pulse integral in one channel, and minimum pulse integral in the other. When events occur further towards the edge of the crystal, the energy is once again shared more equally between the two channels, and the data folds back on itself. To use data from the regions of this degeneracy requires more complicated algorithms taking into account pulse timing. In order to simply estimate the resolution, the events near the edge of the crystal are cut (fig 10.4b). The integral in channel A is added to the integral in channel B to derive the total energy. A histogram of the total energy is plotted in fig. 10.5. Two separate peaks are clearly seen. Both the separation of the peaks and the ratio of the number of events in the two peaks are consistent with the identification of the peaks as the k-α (5.9 keV) and k-β (6.5 keV) x-rays from the ^{55}Fe source. The energy resolution is measured to be < 400 eV FWHM.

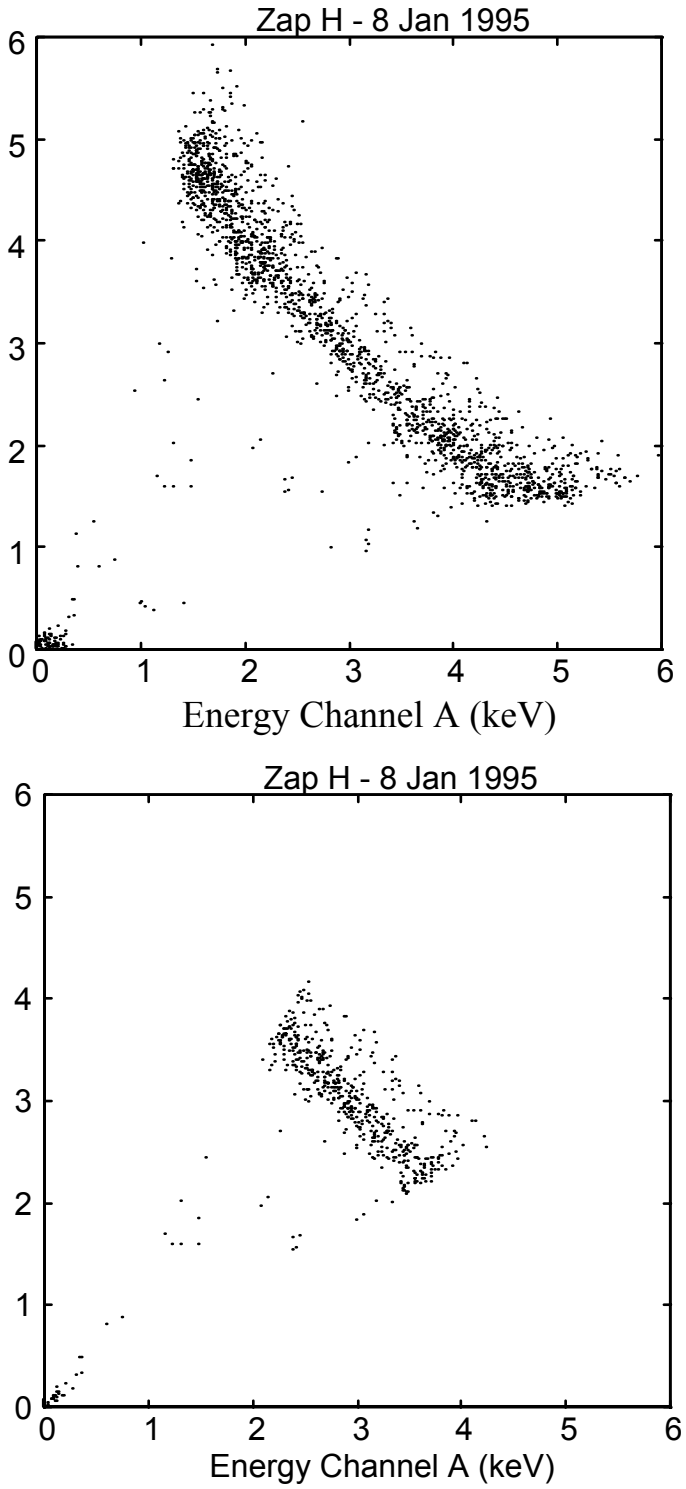


Fig 10.4 Mode 2 pulse integral A vs B plots for the W/Al ETF-TES. The events are normalized to a peak at 6 keV in PHA + PHB. (a) All the data. (b) The data that remains after a cut removes events near the edge of the crystal.

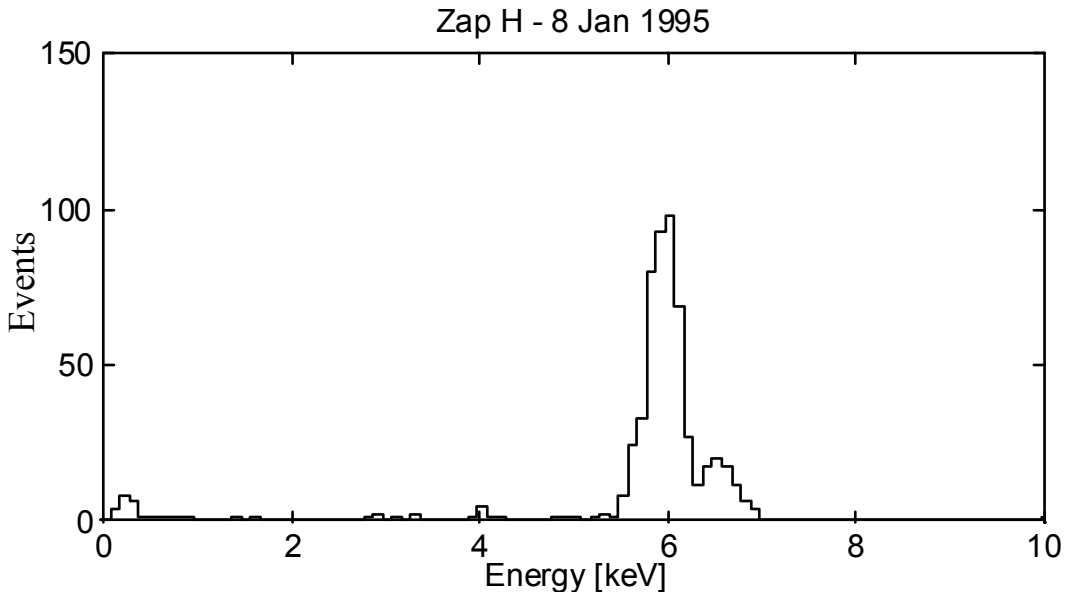


Fig 10.5 The energy histogram for the data in fig 10.4b. The energy peak is normalized to 6 keV. The 5.9 keV k-alpha and 6.5 keV k-beta peaks are clearly distinguishable. The energy resolution is < 400 eV FWHM.

10.10. Signal to noise ratio and other resolution limits

In our high impedance W ZTD-TES, the biggest pulses have a signal to noise ratio of about 10:1. Noise triggers have a FWHM resolution of about 1 keV. The FWHM resolution of the 6 keV band is ~ 1 keV. The resolution is limited by the signal to noise ratio of the signal.

In the W ETF-TES, the biggest pulses have a signal to noise ratio of about 50:1. Noise triggers have a FWHM resolution of about 120 eV. The FWHM resolution of the 6 keV band is difficult to measure, as the data is extremely hyperbolic, and position and energy information are degenerate.

In the W/Al ETF-TES results, the biggest pulses have a signal to noise ratio of about 120:1. Noise triggers have a FWHM of about 20 eV. The FWHM resolution of the 6 keV band is about 0.4 keV.

In the high impedance W ZTD-TES, the resolution is limited by the signal to noise ratio to ~ 1 keV. In the W/Al ETF-TES, the signal to noise ratio has been improved by over an order of magnitude, while the FWHM resolution is improved by a factor of only ~ 2.5. Clearly, the resolution of this device is limited at present by the energy collection from the crystal or by long timescale instabilities in the sensor, rather than by noise in the bandwidth of the pulse.

Statistical fluctuation in the number of quanta in the silicon will place additional limits on the resolution. When a 6 keV x-ray interacts in the silicon, about 2/3 of its energy goes into phonons, and the other 1/3 into electron-hole pairs. Since phonons have mean energies of ~ 4 meV, each 6 keV x-ray will produce ~ 10^6 phonons. Statistical fluctuations in this number will be small (~0.1%). Since electron hole pairs have mean energies of ~ 1 eV, each event will produce ~1,000 pairs, and one would expect ~ 3% fluctuations in the electron-hole pair energy. Since ~ 1/3 of the energy is partitioned into ionization, this effect will cause a fluctuation of about 2% in the energy available to phonons, limiting the resolution to ~ 120 eV, still well below our observed resolution. Note that electron-hole recombination at surfaces, and consequent shedding of phonons, can further reduce this limit, especially if a field is applied across the crystal.

Another possible limitation on our resolution is inadequate collimation of the source in the y direction. Since our two channel detector only contains enough information to deconvolute position from energy in the x dimension, any variation in interaction site in the y dimension will cause a smearing of the resolution. The instrumentation of a four channel detector, which will be done soon, will allow the additional deconvolution of the y dimension.

Edge effects might impose an additional limitation. The penetration depth of 5.9 keV x-rays in a Si crystal is ~ 29 μm, while 6.5 keV x-rays penetrate ~ 38 μm. This distance is small when compared to the 1mm crystal thickness. If, however, phonon or electron-hole pair interaction with the back surface of the crystals causes absorption or down conversion, variations in the interaction depth might cause substantial smearing of the resolution. This effect might be avoided by using a more penetrating source (such as a neutron source), deconvoluting the z position from the energy, and vetoing events near the surface. Phonon scattering Monte Carlos indicate that measuring the z position may be possible using simply the distribution of event energy in the various phonon channels. Additional information will come from timing information available when an ionization sensor is instrumented on the rear side of the crystal. If the beginning of the ionization signal is compared to the beginning of the phonon signal, strong z dimension information should be available. The final dark matter detector will use both phonon and ionization sensors. Finally, phonon signal risetimes may provide additional z-position resolution.

10.10.1. DC bias current fluctuations and the proximity effect

DC bias current fluctuations may place a limit on the resolution of the W/Al ETF-TES. When a W/Al ETF-TES is run, fairly large fluctuations in the DC current due to

the bias voltage are observed. When a pulse occurs, it often does not return to the same DC current baseline. These jumps always seem to correlate with large events. This sort of phenomenon was not observed with the W ETF-TES, so it seems to be related to the W/Al ETF-TES sensor technology itself, rather than the SQUID or other parts of the biasing circuitry.

Two different types of DC bias jumps seem to occur. In one kind, the baseline jumps discontinuously, but returns to its former value over several seconds (still a very long time period as compared to a pulse duration). In the second kind, the DC bias value jumps, but remains at the new value until another jump occurs.

The phenomenon may make the resolution of the sensor worse, but this point is not clear. Short runs of pulses between jumps (~100) often seem to have very good resolution, but if a long data set is cut to include only data within a range of DC bias current, there does not seem to be any improvement of resolution. This question could be answered with a more careful statistical study of the resolution of short data runs.

While the DC bias current shifts, the DC bias voltage remains constant, so the actual detector bias power changes. Since the film must remain in steady state, the power flow from the film to the substrate must also have changed, implying that either the substrate temperature has changed, or sections of the W film have gone completely superconducting or completely normal.

One possible explanation for this effect is the presence of series instabilities in the W line. We have calculated that the line should be stable if there is no variation in thermal conductivity in the film. If, however, the proximity effect raises the gap of sections of line, and hence the thermal impedance, there may be metastable states where some segments of the line are superconducting, and other segments are completely normal. Such a state may persist until an event occurs that is sufficiently energetic to drive the regions of elevated T_c normal. Recall that the Al quasiparticle traps are in electrical contact with the line at eight equally spaced positions.

To analyze this possibility, it is important to calculate the length scale of the proximity effect. The Fermi velocity of electrons in W is $v_f = 1.35 \times 10^8 \, cm/s$ [18]. The coherence length for a normal metal in the dirty limit is [19]

$$\xi_n = \sqrt{\frac{\hbar v_F l}{6 \pi k_B T}} \tag{6.3}$$

At 70 mK, the coherence length in normal W metal is calculated to be $\xi_n \approx 0.34\mu m$. This coherence length is still very short as compared to the 4 μm long W spur that connects that main W line to the Al pad.

Unfortunately, when the normal metal involved in the proximity effect is actually a superconductor slightly above its transition, the electron-electron interaction potential is nonzero and the coherence length no longer sets the length scale. Then, the order parameter scales as $e^{-k_n|x|}$ for $|x| >> k_n^{-1}$, where $k_n^{-1} = \xi_n\sqrt{1+2/\ln(T/T_{cn})}$ and T_{cn} is the transition temperature of the normal metal [19]. Since our W films have a superconducting transition at 70 mK and about 1 mK wide, T/T_{cn} will characteristically be of the order of $71/70 \approx 1.014$. Using this number, and $\xi_n \approx 0.34\mu m$, we arrive at $k_n^{-1} \sim 4\mu m$.

The spur connecting the W line to the Al pad is 4 μm long, which is about the length scale of the proximity effect, so an elevation of the gap in some regions of the W line is to be expected.

This hypothesis will be tested by running a detector without Al pads. If the proximity effect is indeed responsible for our DC bias instability, using devices with longer W spurs between the main W line and the Al may fix the problem. Such devices may indeed have resolutions closer to the statistical limits.

10.10.2. Resolution limits due to charge trapping effects

The fact that the resolution is substantially improved when the crystal is driven into mode 2 implies that when charge trapping sites are present (mode 1), an additional statistical process broadens the energy peak. It is possible that spacial inhomogeneities in the ionized site density, and insufficient deconvolution of event position and energy, are responsible for this broadening. When the crystal is driven into mode 2, these inhomogeneities are reduced, and the resolution is improved.

Although driving the crystal into mode 2 improved the resolution, it is possible that residual ionized traps may be limiting our energy resolution.

10.11. W/Al ETF-TES summary

The W/Al ETF-TES has proven extremely successful. The use of quasiparticle traps in conjunction with the ETF-TES technology allows for the instrumentation of larger detector areas. Energy resolution of < 400 eV was demonstrated. As will be

discussed in chapter 10, the W/Al ETF-TES has more than met our device goals for the construction of a dark matter detector.

CHAPTER 11: CONSTRUCTION OF A DARK MATTER DETECTOR

In order to be able to build a full scale dark matter detector, we outlined several detector goals in chapter 1.

(1) It must be possible to scale up to a ~ 1 kg absorber mass.

Using our previous titanium transition-edge sensors, it is impossible to scale up to a large absorber mass, due to the threshold phonon energy density below which they are insensitive. The incident phonon energy density is inversely proportional to the square of the absorber thickness. This threshold makes it just possible to see 60 keV x-rays through a 300 μm thick substrate. It would thus be impossible to use this technology in a larger absorber mass detector.

We have successfully constructed several detector technologies with no threshold energy density, including the W/Al ETF-TES, which will be used in the construction of a full scale dark matter detector. This technology should scale up to a large absorber mass with only minimal loss of resolution.

(2) Event threshold and energy resolution must be < 1 keV

With the W/Al ETF-TES we have detected 6 keV x-rays on the backside of a 1 mm thick Si crystal with < 400 eV FWHM resolution. We expect to be able to improve this resolution further before the construction of the dark matter detector.

(3) Event position resolution of ~ 1 mm should be possible to allow the vetoing of surface events.

In our experiments, we used a 6 keV x-ray source incident on the backside of 1mm thick silicon crystals. Since 6.5 keVs x-rays penetrate only ~ 38 μm, while 5.9 keV x-rays penetrate ~ 29 μm, they are effectively collimated in the z-axis. We have demonstrated x-axis resolutions of ~ 0.2 mm for a slit-collimated source with minimal variation in y-axis location. We will soon test a 4-channel detector that will allow x and y-axis resolution with an uncollimated source.

In a dark matter detector, events will occur at all crystal depths, so position resolution on the z-axis is also necessary. The analysis of risetime variations may

provide z-axis resolution. Using an ionization channel on the backside will, however, provide a better z-axis measurement by allowing comparison of the timing of the frontside phonon and backside ionization pulses. Using our titanium transition-edge sensors on both sides of the crystal we have demonstrated such arrival time measurements [32].

(4) The technology should lend itself well to use in a hybrid phonon/ionization measurement configuration.

In our group, M. J. Penn has successfully demonstrated simultaneous instrumentation of a phonon channel on the front side of a crystal, and an ionization channel on the backside [33]. One problem that we encountered was that the current bias through our high impedance sensors caused substantial variations in electric fields across the crystal, which caused a distortion in the ionization signal. The ETF-TES technology is much more suited to a hybrid phonon/ionization detector. Since it is a low impedance, the fields induced across the crystal by a particle event are 1000 times smaller than in the high impedance technology. We will soon try the first hybrid ETF-TES phonon/ionization detector. Using this hybrid technology with our Ti transition-edge sensors, M. J. Penn has successfully demonstrated electron/nuclear recoil discrimination due to the fact that nuclear recoils produce fewer electrons and holes than the same energy electron recoils.

We have thus achieved the detector goals necessary to build a dark matter detector. This first detector will consist of a 0.6 kg mass stack of six 1 cm thick, 3" diameter silicon wafers. B. Chugg and B. A. Young have modified the facilities in the clean rooms at the Center for Integrated Systems to allow processing of these thick wafers. The fabrication and testing of the detector will begin this year.

The experiment will take place in the Stanford Underground Facility, which has a 20 meters of water equivalent overburden of dirt. This facility is a previously unused beam dump in End Station III of the Hansen Experimental Physics Laboratory (fig 11.1). We have extended the tunnel an additional 50 feet, where a 34.5 feet of dirt overburden is available. This overburden is sufficient to remove the hadronic component of the cosmic rays.

This site was created in a collaboration under the Center for Particle Astrophysics with groups from UC Berkeley, LBL, UC Santa Barbara and Baksan. Already in place are the dilution refrigerator that will be used in the experiment, a lead shield, and moderator.

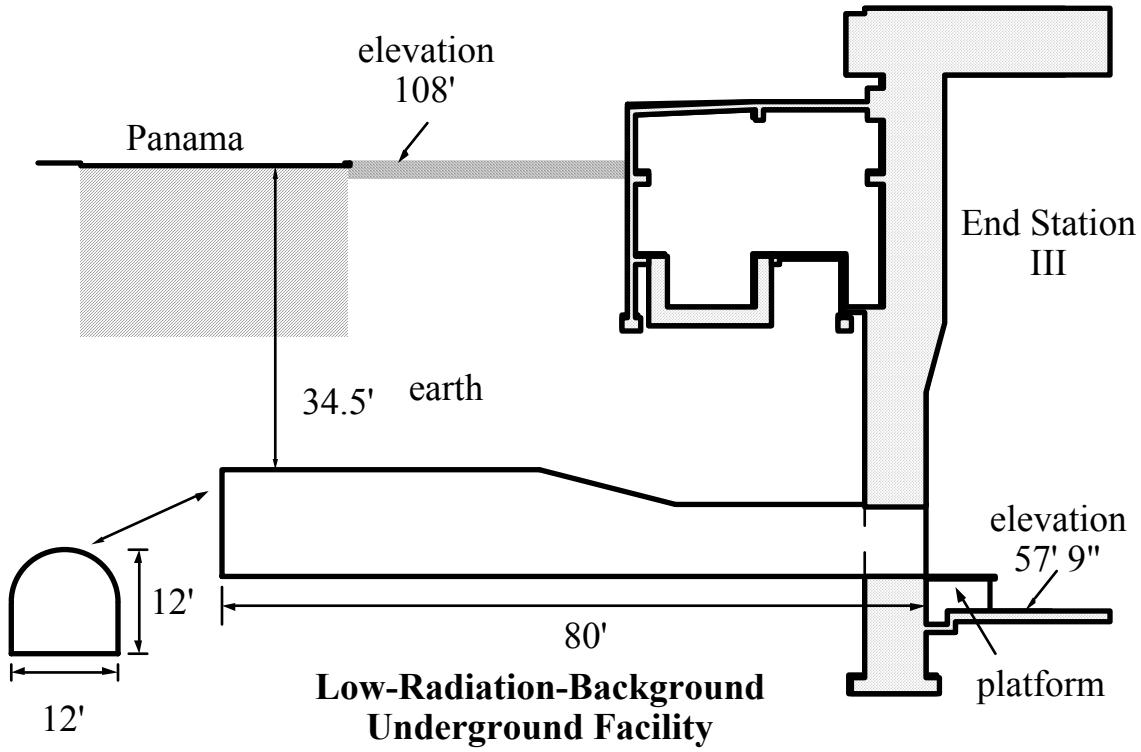


Fig 11.1 The Stanford Underground Facility, where our first dark matter search will take place. The site has a 34.5' overburden of earth.

CHAPTER 12: OTHER APPLICATIONS OF THE ETF-TES

The ETF-TES technology has a range of possible applications other than phonon-mediated particle detection. For the first time, this technology makes it practical to measure energies with a resolution below the rms thermodynamic fluctuations in the sensor.

Since α/n can be ~ 100, a resolution an order of magnitude below any existing technology of similar count rate, heat capacity, and operating temperature is theoretically achievable. Unfortunately, for larger energy events such as x-rays, saturation limitations may require that the heat capacity of the ETF-TES be made larger than that of competing technologies. The ETF-TES saturates when temperature excursions larger than the transition width occur. This width, typically ~ 1 mK, is much smaller than the sensitive range of most other technologies. As long as the pulse duration is kept smaller than the intrinsic time constant, the integral of the pulse can allow accurate measurement of the pulse energy, but better signal to noise is usually achieved by increasing the heat capacity than by running in saturation.

Nevertheless, even though heat capacities must be increased, achievable resolutions are still very competitive. It is also often advantageous to be able to manufacture a device of higher heat capacity than other technologies, but of similar resolution. For instance, a large heat capacity can allow a larger absorber.

For some applications (such as the detection of individual eV energy scale photons), the event energy is low enough that saturation is not a problem. Then the limitation on heat capacity is set by practical fabrication constraints. In this case, the heat capacity of the ETF-TES will be similar to other technologies, and the fundamental resolution limits will be an order of magnitude smaller.

In this chapter, we will discuss three possible applications of the ETF-TES. These are not meant to represent the full array of possible applications; rather, they will allow the discussion of three different regimes of application. These three technologies are an x-ray spectrometer, an infrared bolometer, and an eV-scale photon detector.

12.1. High Resolution X-ray Detection With the ETF-TES

The high resolution and high count rate of ETF-TES sensors suggests their use as high resolution x-ray detectors. In this application, it is desirable to make the film heat capacity large enough so that the highest expected incident energy drives the film

through about one tenth of a transition width. As long as $\tau_{eff} \ll \tau_0$, energy depositions exceeding this condition will still be correctly measured, but the pulse duration will begin to lengthen, and the resolution will begin to deteriorate.

As has been discussed, the optimal bias condition for an ETF-TES is near the 10% point on the transition. This point is near the bottom of the linear portion of the transition, where $dR/dT \sim R_n/\Delta T_c$, the ratio of the normal resistance to the transition width. At this point, the α parameter becomes

$$\alpha = (T/R)dR/dT \sim 10\,T_c/\Delta T_c \tag{12.1}$$

If the heat capacity is then chosen so that an x-ray of energy E drives the film 10% further up its transition, we have

$$C = 10E/\Delta T_c \tag{12.2}$$

From eqn. 8.47, the fundamental resolution for small pulses is known to be $\Delta E_{FWHM} = 2.36\sqrt{4kT_0^2C(1/\alpha)\sqrt{n/2}}$. This equation is only strictly applicable to small pulses, but we shall use it to derive an approximate resolution for the 10% transition width excursion case. Combining the above equations, we have

$$\Delta E_{FWHM} = 2.36\sqrt{4kT_0E\sqrt{n/2}} \tag{12.3}$$

If the heat capacity is such that 6 keV x-rays cause 10% excursions, and taking n = 5, we arrive at $\Delta E_{FWHM} \sim 1.1eV$.

We will now proceed to calculate the volume of W film that would correspond to this heat capacity. From eqns. 12.1 and 12.2, the heat capacity necessary to keep temperature excursions to 10% of the transition width is

$$C \sim E\alpha/T_c \tag{12.4}$$

For a 6 keV x-ray, and for $\alpha = 1000$, the required heat capacity is ~ 1.4 x 10^{-11} J/K. Since the lattice contribution to the heat capacity is proportional to T^3, while the electronic contribution is proportional to T, the electronic heat capacity will dominate at 70 mK. The electronic heat capacity of normal W metal is $C = \gamma T$. In BCS theory, the heat capacity at the superconducting transition is discontinuous. On the superconducting

side, the heat capacity is 2.43 times the value of the normal metal [34]. Assuming that at the 10% point of the transition the heat capacity is roughly the same as the superconducting value, the required film volume is

$$V = \frac{Cu}{2.43\gamma\rho T_c} \tag{12.5}$$

For W the atomic weight u= 183.85 g/mole, γ = 1.3 mJ mole^{-1}K^{-2} [35], the density ρ = 19.3 g cm^{-3}, and we arrive at V ~ 6 x 10^{-4} mm^3. This volume corresponds to an area of about 0.77 mm x 0.77 mm, and a thickness of 1 μm, enough to absorb ~ 50% of incident 6 keV x-rays.

If a larger absorber volume is desired, a low heat capacity semi-metal such as bismuth can be thermally connected to the superconducting film, and used as an absorber [36]. Bismuth dimensions of 2 μm x 1.6 cm x 1.6 cm would also absorb ~ 50% of incident 6 keV x-rays, while limiting temperature excursions to one tenth of a transition width.

In the application of x-ray spectrometry to materials analysis, Poisson statistics determine the uncertainty in x-ray emission line peak heights. The minimum detectability limit goes as the square root of the ratio of the energy resolution to the count rate [37]. Energy Dispersive Spectrometry (EDS) is typically limited to a 150 eV resolution and a 20 kHz count rate for 6 keV x-rays. Cryogenic calorimeters are theoretically limited to a 1 eV resolution and a 1 kHz count rate, yielding uncertainties 3 times better than EDS [37]. ETF-TES sensors are theoretically limited to about 1.1 eV resolution with count rates above 50 kHz, leading to a potential performance ~ 15 times better than EDS.

12.2. Detection of infrared radiation with the ETF-TES

A W ETF-TES can be used as a fast, sensitive detector of infrared and millimeter waves [38]. In such a device, a small normal metal absorber is coupled to an antenna or waveguide. This absorber is then electrically and thermally connected to a W ETF-TES. Because the electrothermal feedback will operate on time constants fast as compared to any variation in the incident power, the device will operate in a "locked in" feedback mode where temperature variations are small, and the measured power is simply the voltage bias multiplied by the change in the SQUID current. The measurement of the power is thus made with no free parameters.

The device would consist of a W pad deposited on a silicon substrate. A small spur of some normal metal such as Cu would be electrically and thermally connected to the W pad. The spur would have impedance ~ 50 Ω for a good impedance match to the antenna circuit. Al rails would be used to apply the bias voltage evenly across the W pad. The substrate would be heat sunk extremely well, perhaps through a gold film deposited on the back side.

The important figures of merit for this application are the power responsivity S (here in units of amps/watt) and the Noise Equivalent Power (NEP). We now proceed to calculate these quantities.

If the thermalization time of the W film is short as compared to any other detector time constant, the power flow equation for the device will be

$$C\frac{dT}{dt} = \frac{V^2}{R} - K\left(T^n - T_s^n\right) + P_s e^{i\omega t} \tag{12.6}$$

Where ω is the modulation frequency, and P_s is the signal power. P_s should not be confused with P_0, the bias power. Proceeding as in chapter 8, we expand eqn 12.6 for small departures from the equilibrium temperature.

$$\Delta T = \frac{P_s}{\frac{P_0\alpha}{T} + g + i\omega C} \tag{12.7}$$

Now we assume operation in the extreme feedback regime, where $P_0 = gT_0/n$, take into account that $\alpha >> n$, and find the absolute value of eqn. 12.7. We arrive at

$$\Delta T \approx \frac{P_s n}{g\alpha\sqrt{1+\omega^2\tau_{eff}^2}} \tag{12.8}$$

Where $\tau_{eff} = Cn/g\alpha$, the effective time constant due to electrothermal feedback. The current fluctuation in the ETF-TES due to a temperature fluctuation ΔT is $\Delta I = -I_0\alpha\Delta T/T_0$. Substituting in eqn. 12.8, we arrive at

$$\Delta I = -\frac{P_1}{V_0\sqrt{1+\omega^2\tau_{eff}^2}} \tag{12.9}$$

implying a power responsivity of

$$S = -\frac{1}{V_0\sqrt{1+\omega^2\tau_{eff}^2}} \tag{12.10}$$

which has units of amps/watt.

The NEP of this device will have contributions from the phonon noise, the Johnson noise, and the SQUID noise. As explained in chapter VIII, if $T_{substrate}^n << T_0^n$, the phonon noise will have power spectral density $2kT_0^2 g$. The Johnson noise current is highly suppressed below the knee. From eqn. 8.33 it has form

$$I_J^2 = \frac{4kT_0}{R_0}\frac{\left(n^2/\alpha^2 + \omega^2\tau_{eff}^2\right)}{1+\omega^2\tau_{eff}^2} \tag{12.11}$$

The NEP contributions from the Johnson and SQUID current noise are found by dividing these quantities by the power responsivity. When these three noise sources are added in quadrature, we arrive at the total NEP

$$NEP^2 = 2kT^2 g + 4kTP_0\omega^2\tau_{eff}^2 + I_{squid}^2 V_0^2(1+\omega^2\tau_{eff}^2) \tag{12.12}$$

It is desirable to bias such that the SQUID noise is smaller than the phonon noise for frequencies below the knee. For a film with T_c = 70 mK, in the extreme feedback regime, when the device is biased at ~ 1 Ω, the current equivalent noise due to the phonon noise is found to be

$$I_n = S\sqrt{2kT_0^2 g} = \frac{\sqrt{2kT_0P_0 n}}{V_0\sqrt{1+\omega^2\tau_{eff}^2}} \approx \sqrt{2kT_0 n/R} \approx 3.1\,pA/\sqrt{Hz} \tag{12.13}$$

whereas a SQUID with an acceptable input inductance can achieve 1 pA/√Hz.

Assuming the SQUID noise contribution is small, and using $P_0 = gT/n$, we find that the NEP of the ETF-TES is thus

$$NEP^2 = 4kT_0^2 g\left(1/2 + \omega^2\tau_{eff}^2/n\right) \tag{12.14}$$

In conventional bolometric infrared sensors, the time constant, $\tau = C/g$, is of the same order as the modulation frequency [39]. In the ETF-TES, the effective time constant can be made of order 10 μs, and for most applications $\omega\tau_{eff} << 1$. In this case, the Johnson noise is suppressed by the electrothermal feedback, the NEP becomes independent of the modulation frequency, and the NEP is a simple function of the thermal conductance.

Dynamic range sets the most significant limit on the NEP of an ETF-TES. The incident power must be kept smaller than the bias power to avoid saturation, so the maximum detectable power is $P_{\max} = gT_0/n$. If the thermal conductance is made lower, the saturation power drops accordingly. Thus, the NEP may be expressed as a function of the maximum incident power. Assuming that $\omega\tau_{eff} << 1$, we arrive at $NEP^2 = 2kT_0P_{\max}n$. If it is desirable to detect incident power as high as 1 pW, the NEP is found to be 3 x 10^{-18} W/√Hz. This NEP is a factor of 10 lower than that of the best conventional composite bolometers in a similar temperature range.

In summary, the ETF-TES has important advantages over conventional bolometers for the detection of infrared radiation. It operates in a lock-in mode that allows the direct measurement of the incident power, with no free parameters. The electrothermal feedback suppresses the Johnson noise, and the NEP is independent of the modulation frequency. Finally the NEP achievable in principle is a factor of 10 lower than that of conventional bolometers.

12.3. Detection of eV-scale photons

As discussed at the beginning of this chapter, for larger energy events such as x-rays, saturation limitations require that the heat capacity of the ETF-TES be made larger than that of competing technologies. Even with this limitation, the ETF-TES is quite competitive. When looking at individual lower energy photons, however, where the heat capacity is set by fabrication constraints, the ETF-TES does not suffer from the same limitations, and becomes more promising than other calorimetric technologies.

The proximity effect sets a final restriction on the minimum ETF-TES heat capacity. As discussed in section 10.8.1, for W films biased on the transition, the order parameter scales as $e^{-k_n|x|}$ for $|x| >> k_n^{-1}$, and $k_n^{-1} \sim 4\mu m$. A W pad 40 μm long, 4 μm wide, and 30 nm thick would be 10 times as long as the proximity effect length scale, and would have a 1 Ω impedance at the 10% point on the transition. From eqn. 12.5, the heat capacity associated with a W pad of this volume is

$$C = \frac{2.43\gamma\rho T_c V}{u} = 696\, eV/K \qquad (12.14)$$

where for W the atomic weight u= 183.85 g/mole, γ = 1.3 mJ mole^{-1}K^{-2}, and the density ρ = 19.3 g cm^{-3}.

From eqn. 8.47, the fundamental resolution for small pulses is known to be $\Delta E_{FWHM} = 2.36\sqrt{4kT_0^2 C(1/\alpha)\sqrt{n/2}}$. For $\alpha = 1000$, and taking n=5, the fundamental limit on the resolution is ~ 3.2 meV FWHM. From eqn. 12.4, the maximum pulse energy (the pulse energy that will drive the sensor through 10% of its transition width), is $E \sim CT_c/\alpha$, or 50 meV, the energy of a photon in the far infrared.

Larger photon energies require larger heat capacities. If it is desired to resolve 1 eV photons, and $\alpha = 1000$, a 30 nm thick pad of area 20 μm x 200 μm is be needed, with a fundamental resolution limit of ~ 15 meV FWHM.

These pads are probably too small to be used to directly absorb photons. Unlike the x-ray detector case, a normal metal absorber cannot be used, because of heat capacity constraints. It might be possible, however, to collect the energy into the sensor without increasing the heat capacity by using either quasiparticle trapping or athermal phonon mediation.

CHAPTER 13: CONCLUSIONS

There has been tremendous progress in the last few years in the development of our phonon sensor technology for dark matter and neutrino experiments. The sensor technology developed here also shows promise in other applications, including high resolution x-ray spectroscopy, infrared bolometry, and the detection of low energy photons.

Superconducting tungsten films with 1 mK wide transitions at 70 mK have been deposited. Using these films, high impedance W sensors have been demonstrated that do not exhibit the threshold energy density limitations suffered by our titanium sensors.

The most exciting development has been the demonstration of an electrothermal feedback transition-edge sensor, in which the temperature of the superconducting W film is held constant within its transition by an electrothermal feedback process. Using this technology, we have demonstrated substantial improvements in signal to noise, linearity, dynamic range, and count rate. The fundamental limits on the resolution of this technology have been analyzed, and found to be below the rms thermodynamic energy fluctuations in the film, and better than any existing technology operating at the same temperature, count rate, and absorber heat capacity.

The enhancement of this electrothermal feedback technology with quasiparticle trapping from aluminum films has been demonstrated. Using the quasiparticle trap enhanced electrothermal feedback sensors, an energy resolution of < 400 eV FWHM has been measured for 6 keV x-rays interacting on the backside of a 1 mm thick silicon substrate, with an electronic noise of 120 eV.

Finally, the pieces are in now in place for the construction of a dark matter experiment. Sufficient event position sensitivity, threshold, and resolution have been demonstrated. S. W. Nam has developed the SQUID electronics necessary for instrumenting many channels of the W/Al ETF-TES. B. Chugg and B. A. Young have developed the procedure necessary to fabricate these devices on thick substrates. M. J. Penn has developed the ionization collection technology necessary for timing and electron/nuclear recoil discrimination. In collaboration with the Center for Particle Astrophysics, an underground site has been created, and a dilution refrigerator with appropriate radiation shielding is now in place. Fabrication of the dark matter detector will begin this year.

APPENDIX A: PHONON SCATTERING MONTE CARLOS

A.1 PhonScatt.c

```
/*   PhonScatt.c
/*
        A Monte Carlo model of a phonon-mediated particle detector.

        This program models the geometry of a phonon-mediated particle detector.
        When a particle interaction occurs in the absorber, the random trajectory
        of a fixed number of phonons is followed.  Phonons are assumed to
        propagate ballistically through the crystal.  At surfaces, the direction
        of phonons is randomized.  The geometry of the detector is modeled.  When a
        phonon impinges on the W film or a heat sink, it has a probability of
        absorption.  The phonon continues to scatter until it is absorbed either in
        a heat sink or a TES. The signal in each channel is assumed to be
        proportional to the total number of phonons it absorbs.

        This code is intended to sit inside a Macintosh shell.  The shell should
        call InitCalculation() once, an then should alternate between calling
        Calculation() and handling Macintosh events.  After a quit is received,
        the shell should call EndCalculation().
*/


#include "Calcshell.h"              /* The Macintosh shell program #includes */
#include <Quickdraw.h>
#include <stdlib.h>
#include <math.h>
#include <stdio.h>

//#define GEOMETRY 1        /* Define this to see the geometry instead of A vs B */
//#define TRAJECTORY 100 /* If plotting the geometry, how often to plot whole phonon trajectory? */

/* Define the structure to hold information about a rectangular device feature (i.e. the
chip boundaries, the boundaries of TES channel #1, etc. */
typedef struct feature_s {
        double left, right, top, bottom;        /* Feature boundaries */
        double absorb_coeff;                            /* Phonon absorption coefficient */
        double coverage;                        /* Fraction of feature surface area covered */
        long phonon_count;                              /* Number of phonons absorbed */
} feature_t;

typedef struct point_s {
        double x,y;
} point_t;

double distribution();
void reflect(point_t *s,struct feature_s *t);
void startline(point_t *pPtr);
void endline(point_t *pPtr);
void plotpoint(point_t *pPtr);

void plotAB(int a, int b);
```

```
void PlotFeature(feature_t *fPtr);
void plotxy(double a,double b);

#define SCALE_FACTOR            30              /* Screen pixels per mm */
#define CENTER_X                200             /* Position of center of chip on screen */
#define CENTER_Y                150
#define POINTX(POINT) (CENTER_X+SCALE_FACTOR*(POINT))
#define POINTY(POINT) (CENTER_Y+SCALE_FACTOR*(POINT))
#define END_OF_ARRAY            -1
#define LAMBDA0                 1.0     /* center of diffusion distribution, in mm */

/* Phonon absorption probability at W interface     */
#define W_ABS                   0.3

/* Phonon absorption probability at Al interface    */
#define AL_ABS          0.5
#define NUM_PHON_MAX    500     /* Number of phonons modeled per event */
#define NUM_EVENTS_MAX  2000    /* Maximum number of events modeled */

/* The probability of phonon absorption at uncovered surface  */
#define DOWN_CONVERT    0

/* Now include the file containing the specific device & absorber geometry */
#include "NoGold.h"
/* The name of the file where the results should be stored */
#define FILE_NAME "NoGold0.5Al0.3W.dat"

point_t phonon,front;

double pi = 3.141592653589793;

unsigned long NumPhon;     /* The total number of phonons to date */
unsigned long NumEvents = 0;
FILE *TheFile;

InitCalculation()
{
        Rect theBox;

        TheFile = fopen(FILE_NAME,"w");

#ifdef GEOMETRY
        SetRect(&theBox,20.,20.,320.,320.);
        FrameRect(&theBox);
#endif
}

EndCalculation()
{
        fclose(TheFile);
}

Calculation ()
{
        double theta,random();
        double lambda, lambda0;
        double Stheta;
```

```
        feature_t *fPtr;
        int absorbed;
        int trajectory;         /* True if we are to plot the trajectory */
        point_t start;

        long int NumPhonMax;
        long int NumAttempts;
        void PlotMask();
        char buffer[100];
        int num1, num2;


        /* Increment counter.  If program done, return. */
        if (++NumEvents > NUM_EVENTS_MAX)
                return;

        NumPhon = 0;

        /* Randomize event position along a slit */
        start.x = 10.0*(random()-0.5);
        start.y = 0.0;

        /* Reset all the phonon counts to zero */
        for(fPtr = &feature; fPtr->phonon_count != END_OF_ARRAY; fPtr++)
                fPtr->phonon_count = 0;

        /* Choose 5.9 or 6.5 keV */
        if(random()<(2.9/24.0))
                NumPhonMax = (6.49/5.89)*NUM_PHON_MAX;
        else
                NumPhonMax = NUM_PHON_MAX;

#ifdef GEOMETRY
        PlotMask();
#endif

        while(NumPhon++ < NumPhonMax) {
                absorbed = FALSE;

                phonon.x = start.x;
                phonon.y = start.y;

                NumAttempts = 0;
#ifdef GEOMETRY
                trajectory = (!(NumPhon % TRAJECTORY));
                if(trajectory)
                        startline(&start);
#endif

                while (!absorbed) {
                        NumAttempts++;

                        theta = 2 * pi * random();
                        lambda = distribution();
```

```
/* Move phonon half of the diffusion length - from the back face to the front */
                                front.x = phonon.x + lambda * cos(theta)/2.0;
                                front.y = phonon.y + lambda * sin(theta)/2.0;

                                /* If off of the chip, reflect... */
                                reflect(&front,&chip);

                                /* Now check for absorption */
                                for(fPtr = &feature; fPtr->phonon_count != END_OF_ARRAY; fPtr++)
                                        if (absorbed = absorb_check(&front,fPtr))
                                                break;

                                /* Allow for silicon surface absorption */
                                if (random() < DOWN_CONVERT)
                                        absorbed = 1;

                                /* If not absorbed, move the phonon the whole diffusion length */
                                if(!absorbed) {
                                        phonon.x += lambda*cos(theta);
                                        phonon.y += lambda*sin(theta);
                                        reflect(&phonon,&chip);
#ifdef GEOMETRY
                                        if(trajectory) {
                                                endline(&phonon);
                                                startline(&phonon);
                                        }
#endif
                                }
#ifdef GEOMETRY
                                if(absorbed) {
                                        if(trajectory)
                                                endline(&front);
                                        else
                                                plotpoint(&front);

                                }
#endif
                        }
                }
#ifndef GEOMETRY
                plotAB(TES1->phonon_count, TES2->phonon_count);
                num1 = TES1->phonon_count;
                num2 = TES2->phonon_count;
                sprintf(buffer,"%d\t%d",num1,num2);
                fprintf(TheFile,"%s\n",buffer);

#endif
}

void plotAB(int a,int b)
{
        MoveTo(20 + 300 * ((double)a)/ ((double)NUM_PHON_MAX),
                                320-300*((double)b)/((double)NUM_PHON_MAX));
        Line(0,0);
}

double random()
```

```
{
        return(((double)rand())/((double)RAND_MAX));
}

double distribution()
{
        double Stheta,lambda;
        double randval;

        randval = random();
        if (randval < .001)
                randval = .001;
        if (randval > .999)
                randval = .999;
        Stheta = sqrt(randval);
        lambda = LAMBDA0*Stheta/sqrt(1-Stheta*Stheta);
        return(lambda);
}


/* Check to see if phonons are absorbed in a feature */
int absorb_check(point_t *loc,feature_t *fPtr)
{
        if      ((loc->x    > fPtr->left)     &&
                 (loc->x    < fPtr->right)    &&
                 (loc->y    > fPtr->bottom) &&
                 (loc->y    < fPtr->top)      &&
                 (random() < (fPtr->absorb_coeff)*(fPtr->coverage)) )
        {
                (fPtr->phonon_count)++;
                return(TRUE);
        }
        return(FALSE);
}

void reflect(point_t *front_ptr,feature_t *chip_ptr)
{
        while(outbounds(front_ptr,chip_ptr)) {

                /* Do the reflections */
                if (front_ptr->x < chip_ptr->left)
                        front_ptr->x = 2*chip_ptr->left - front_ptr->x;
                if (front_ptr->x > chip_ptr->right)
                        front_ptr->x = 2*chip_ptr->right - front_ptr->x;
                if (front_ptr->y > chip_ptr->top)
                        front_ptr->y = 2*chip_ptr->top - front_ptr->y;
                if (front_ptr->y < chip_ptr->bottom)
                        front_ptr->y = 2*chip_ptr->bottom - front_ptr->y;
        }

}
int outbounds (point_t *front_ptr,feature_t *chip_ptr)
{
        int offleft,offright,offup,offdown;

        offleft = front_ptr->x < chip_ptr->left;
        offright = front_ptr->x > chip_ptr->right;
```

```
        offup = front_ptr->y > chip_ptr->top;
        offdown = front_ptr->y < chip_ptr->bottom;

        return(offleft || offright || offup || offdown);
}


void startline(point_t *pPtr)
{
        MoveTo(POINTX(pPtr->x),POINTY(pPtr->y));
}

void endline(point_t *pPtr)
{
        float x,y;
        x = POINTX(pPtr->x); y = POINTY(pPtr->y);
        LineTo(x,y);
}

void plotpoint(point_t *pPtr)
{
        float x,y;
        x = POINTX(pPtr->x); y = POINTY(pPtr->y);
        MoveTo(x,y);
        Line(0,0);
}

void PlotFeature(feature_t *fPtr)
{
        MoveTo(POINTX(fPtr->left),POINTY(fPtr->top));
        LineTo(POINTX(fPtr->right),POINTY(fPtr->top));
        LineTo(POINTX(fPtr->right),POINTY(fPtr->bottom));
        LineTo(POINTX(fPtr->left),POINTY(fPtr->bottom));
        LineTo(POINTX(fPtr->left),POINTY(fPtr->top));
}

void PlotMask()
{
        feature_t *fPtr;
        PlotFeature(&chip);
        for(fPtr = &feature; fPtr->phonon_count != END_OF_ARRAY; fPtr++)
                        PlotFeature(fPtr);
}
```

A.2 Device Geometries

A.2.1 HighZ.h

```
/*  HighZ.h
        This file contains the geometry information for a 1 cm x 1 cm silicon crystal with
        two 2 x 4 mm W high Z channels with W pads on the edge.
*/

/* Initialize the boundaries of the chip */
```

```
feature_t chip = {
        -5.0,5.0,5.0,-5.0,          /* Boundaries: left, right, top, bottom */
        0.0,                        /* Absorption coefficient */
        0.0,                        /* Coverage */
        0                           /* Phonon Count */
};
feature_t feature[] = {
        /* Initialize Transition Edge Sensor #1 */
        {       -2.0,0.0,2.0,-2.0,  /* Coordinates */
                W_ABS,                  /* Absorption coefficient */
                0.4,
                0
        },
        /* Initialize Transition Edge Sensor #2 */
        {       0.0,2.0,2.0,-2.0,
                W_ABS,
                0.4,
                0
        },
                /* Left Metal Pad */
        {       -5.0,-4.0,5.0,-5.0,/* Coordinates */
                W_ABS,                  /* Absorption coefficient */
                1.0,
                0
        },
                /* Right Metal Pad */
        {       4.0,5.0,5.0,-5.0,       /* Coordinates */
                W_ABS,                  /* Absorption coefficient */
                1.0,
                0
        },
                /* One Liner Pad */
        {       -2.0,2.0,5.0,4.0,       /* Coordinates */
                W_ABS,                  /* Absorption coefficient */
                1.0,
                0
        },
                /* Bottom Metal Pad */
        {       -4.0,4.0,-4.0,-5.0,/* Coordinates */
                W_ABS,                  /* Absorption coefficient */
                1.0,
                0
        },
        /* End of array marker */
        {       0.,0.,0.,0.,
                0.,
                0.,
                END_OF_ARRAY
        }
};

/* Initialize pointers to each feature type */
feature_t *TES1  = feature;
feature_t *TES2  = feature + 1;
feature_t *Heatsink1        = feature + 2;
feature_t *Heatsink2        = feature + 3;
feature_t *TempPad          = feature + 4;
```

```
feature_t *BondPad          = feature + 5;
```

A.2.2 BigGold.h

```
/*  BigGold.h
        This file contains the geometry information for a 1 cm x 1 cm silicon crystal with
        two 1.8 x 1.8 mm W ETF-TES channels and big gold heat sinks.
*/

/* Initialize the boundaries of the chip */
feature_t chip = {
        -5.0,5.0,5.0,-5.0,          /* Boundaries: left, right, top, bottom */
        0.0,                        /* Absorption coefficient */
        0.0,                        /* Coverage */
        0                           /* Phonon Count */
};

feature_t feature[] = {
        /* Initialize Transition Edge Sensor #1 */
        {       -3.4,-1.6,0.9,-0.9,/* Coordinates */
                W_ABS,                      /* Absorption coefficient */
                1,
                0
        },
        /* Initialize Transition Edge Sensor #2 */
        {       1.6,3.4,0.9,-0.9,
                W_ABS,
                1,
                0
        },
        /* Initialize Heat Sink #1 */
        {       -4.445,4.445,4.445,1.905,
                W_ABS,
                1.0,
                0
        },
        /* Initialize Heat Sink #2 */
        {       -4.445,4.445,-1.905,-4.445,
                0.5,
                1.0,
                0
        },
        /* End of array marker */
        {       0.,0.,0.,0.,
                0.,
                0.,
                END_OF_ARRAY
        }
};
```

```
/* Initialize pointers to each feature type */
feature_t *TES1          = feature;
feature_t *TES2          = feature + 1;
feature_t *Heatsink1     = feature + 2;
feature_t *Heatsink2     = feature + 3;
feature_t *TempPad               = feature + 4;
feature_t *BondPad               = feature + 5;
```

A.2.4 EdgeGold.h

```
/*  EdgeGold.h
        This file contains the geometry information for a 1 cm x 1 cm silicon crystal with
        two 1.8 x 1.8 mm W ETF-TES channels and gold heat sinks near the edge.

*/
/* Initialize the boundaries of the chip */
feature_t chip = {
        -5.0,5.0,5.0,-5.0,      /* Boundaries: left, right, top, bottom */
        0.0,                    /* Absorption coefficient */
        0.0,                    /* Coverage */
        0                       /* Phonon Count */
};

feature_t feature[] = {
        /* Initialize Transition Edge Sensor #1 */
        {       -3.4,-1.6,0.9,-0.9,/* Coordinates */
                AL_ABS,                 /* Absorption coefficient */
                1,
                0
        },
        /* Initialize Transition Edge Sensor #2 */
        {       1.6,3.4,0.9,-0.9,
                AL_ABS,
                1,
                0
        },
        /* Initialize Top Left Heat Sink */
        {       -4.7,-.5,4.7,3.7,
                W_ABS,
                1.0,
                0
        },
        /* Initialize Top Right Heat Sink */
        {       .5,4.7,4.7,3.7,
                W_ABS,
                1.0,
                0
        },
        /* Initialize Bottom Left Heat Sink */
        {       -4.7,-.5,-3.7,-4.7,
                W_ABS,
                1.0,
```

```
                0
        },
        /* Initialize Bottom Right Heat Sink */
        {       .5,4.7,-3.7,-4.7,
                W_ABS,
                1.0,
                0
        },
        /* Initialize Top Mid Left Heat Sink */
        {       -4.7,-3.7,3.7,0.5,
                W_ABS,
                1.0,
                0
        },
        /* Initialize Top Mid Right Heat Sink */
        {       3.7,4.7,3.7,0.5,
                W_ABS,
                1.0,
                0
        },
        /* Initialize Bottom Mid Left Heat Sink */
        {       -4.7,-3.7,-0.5,-3.7,
                W_ABS,
                1.0,
                0
        },
        /* Initialize Bottom Mid Right Heat Sink */
        {       3.7,4.7,-0.5,-3.7,
                W_ABS,
                1.0,
                0
        },
        /* End of array marker */
        {       0.,0.,0.,0.,
                0.,
                0.,
                END_OF_ARRAY
        }
};

/* Initialize pointers to each feature type */
feature_t *TES1         = feature;
feature_t *TES2         = feature + 1;
feature_t *Heatsink1    = feature + 2;
feature_t *Heatsink2    = feature + 3;
feature_t *TempPad              = feature + 4;
feature_t *BondPad              = feature + 5;
```

A.2.5 NoGold.h

```
/*  NoGold.h
        This file contains the geometry information for a 1 cm x 1 cm silicon crystal with
        two 2 x 4 mm Al/W ETF-TES channels and no gold heat sinks.
```

```
*/
/* Initialize the boundaries of the chip */
feature_t chip = {
        -5.0,5.0,5.0,-5.0,          /* Boundaries: left, right, top, bottom */
        0.0,                        /* Absorption coefficient */
        0.0,                        /* Coverage */
        0                           /* Phonon Count */
};

feature_t feature[] = {
        /* Initialize Transition Edge Sensor #1 */
        {       -3.4,-1.6,0.9,-0.9,         /* Coordinates */
                W_ABS,                              /* Absorption coefficient */
                1,
                0
        },
        /* Initialize Transition Edge Sensor #2 */
        {       1.6,3.4,0.9,-0.9,
                W_ABS,
                1,
                0
        },
        /* Initialize Top Left Heat Sink */
        {       -4.7,-3.7,4.7,3.7,
                W_ABS,
                1.0,
                0
        },
        /* Initialize Top Right Heat Sink */
        {       3.7,4.7,4.7,3.7,
                W_ABS,
                1.0,
                0
        },
        /* Initialize Bottom Left Heat Sink */
        {       -4.7,-3.7,-3.7,-4.7,
                W_ABS,
                1.0,
                0
        },
        /* Initialize Bottom Right Heat Sink */
        {       3.7,4.7,-3.7,-4.7,
                W_ABS,
                1.0,
                0
        },
        /* End of array marker */
        {       0.,0.,0.,0.,
                0.,
                0.,
                END_OF_ARRAY
        }
};

/* Initialize pointers to important feature types */
feature_t *TES1             = feature;
```

```
feature_t *TES2                 = feature + 1;
```

APPENDIX B: PULSE SIMULATIONS

The behavior of the ETF-TES pulses were simulated with MATLAB code before the first device was cooled. In these models, the 20 mΩ shunt resistor, a 5 mΩ parasitic resistance in series with the ETF-TES, the SQUID inductance, and a 30 Ω SQUID shunt resistor were included.

B.1 TransitionModel.m

```
%
%  The transition is modeled with a Fermi function
%  DTc is the 10-90% transition width.
%
function fraction = TransitionModel(T,Tc,DTc)

operand = 4.3944491546724387655*(T ./ Tc -1) .* Tc ./ DTc;
fraction = 1 - 1 ./ (exp(operand) + 1);
```

B.2 ETFTESDampEquil.m

```
%%%%%%% ETFTESDampEquil.m                                    %%%%%%%
%
%  Find the equilibrium bias conditions of the ETF-TES
%

clear
clg
subplot(221)
% First set up the model parameters

Tc = .070;                  % Film Tc (K)
Ts = .050;                  % Substrate temperature (K)
DTc = .0007;                % Transition width (K)
Rs = .020;                  % Shunt resistance ( ohms )
Rp = .005;                  % Parasitic resistance ( ohms )
Rn = 3;                     % Film normal resistance ( ohms )
Rd = .2;                    % SQUID damping resistor ( ohms )
L = 0.25e-6;                % SQUID input loop inductance ( henries )

% Choose the initial bias current
Ibias = 41e-6*sqrt(Rn);
Ibias = Ibias * 1.75;


% Calculate thermal conductivity from intrinsic time constant

N = 5;                      % coupling power

% Calculate SigmaV, and hence G, from the known latching current
```

```
% for our high impedance detectors
SigmaV = (6e-9)^2*(1.6e6)/(Tc^N - Ts^N);

G = N*SigmaV*Tc^(N-1);

% Now compute heat capacity
% Self-consistently scale to achieve the desired intrinsic heat capacity
tau = .0003;
C = tau*G;

% Here set tiny inductance to find the equilibrium calculation
L = .075e-6;
% Initial current is unimportant.
I = 0;
Is = 0;

% Initially, set the film normal
T = .0705;
R=Rn*TransitionModel(T,Tc,DTc);

% Set up loop
Dt = .2e-6;      % Timestep (sec)
Numstep =4000;
StepT = 1:Numstep;
StepR = 1:Numstep;
StepI = 1:Numstep;
StepIs = 1:Numstep;
for step = 1:Numstep
% Electrical step ( adjust current )
% Find derivative of SQUID current
DIs = (Ibias*Rs - Is*(Rp+Rs+R)) / ( L + (L/Rd)*(Rp + Rs + R) );

% Find new SQUID current
Is = Is + DIs*Dt;
% Find new film current
I = (L/Rd)*DIs + Is;

% Electrothermal step ( adjust power )
Power = I^2 * R - SigmaV *(T^5 - Ts^5);

% Thermal step ( adjust temperature )
T = T + Power*Dt/C;
StepIs(step) = Is;
StepT(step) = T;
R=Rn*TransitionModel(T,Tc,DTc);
StepR(step) = R;
StepI(step) = I;
end

subplot(221)
plot(1:Numstep,StepT)
title('Temperature');
subplot(222)
plot(1:Numstep,StepR)
title('Resistance');
subplot(223)
plot(1:Numstep,StepI)
```

```
title('Current');
subplot(224)
plot(1:Numstep,StepIs)
title('Squid Current');

% Set up the equilibrium variables
T0 = T;
Is0 = Is;
R0 = R
```

B.3 ETFTESDampPulse.m

```
%%%%%%% ETFTESDampPulse.m                    %%%%%%%%
%
% Model a pulse
%

clg
subplot(221)
% Keep most model parameters from the equilibrium calculation


Rd = 30.0;          % SQUID damping resistor ( ohms )
L = 0.25e-6;                % Inductance ( henries )

% Initial current
Is = Is0;


R=Rn*TransitionModel(T,Tc,DTc);


% Set up loop
Dt = .05e-6;                % Timestep (sec)
Numstep =3000;
StepT = 1:Numstep;
StepR = 1:Numstep;
StepI = 1:Numstep;
StepIs = 1:Numstep;
for step = 1:Numstep

% Introduce a 1 keV pulse at step 500
if step == 500
        T = T0 + (1000*1.6e-19)/C;
end

% Electrical step ( adjust current )
% Find derivative of SQUID current
DIs = (Ibias*Rs - Is*(Rp+Rs+R)) / ( L + (L/Rd)*(Rp + Rs + R) );
% Find new SQUID current
Is = Is + DIs*Dt;
Is;
change = DIs*Dt;
```

```
% Find new film current
I = (L/Rd)*DIs + Is;
% Electrothermal step ( adjust power )
Power = I^2 * R - SigmaV *(T^5 - Ts^5);
% Thermal step ( adjust temperature )
T = T + Power*Dt/C;
StepIs(step) = Is;
StepT(step) = T;
R=Rn*TransitionModel(T,Tc,DTc);
StepR(step) = R;
StepI(step) = I;
end
time = Dt * (1:Numstep);
subplot(221)
plot(time,StepT)
title('Film temperature')
subplot(222)
plot(time,StepR)
title('Film resistance')
subplot(223)
plot(time,StepI)
title('Film current')
subplot(224)
DeltaIs = Is0 - StepIs;
plot(time,DeltaIs,'w')
title('Squid current')
xlabel('Time (sec)')
ylabel('Squid current reduction(amps)')
```

REFERENCES

1. B. L. Dougherty, in *Particle and Nuclear Astrophysics and Cosmology in the Next Millennium*, Proceedings of Snowmass '94, eds. R. Peccei and E. Kolb, World Scientific Publishing Co.

2. V. Rubin et. al, *Ap. J.* **374**, 344 (1991).

3. M. Davis et. al., *Ap. J.* **292**, 371(1985).

4. L. Krauss, *Phys. Rev. Lett.* **64**, 999 (1990).

5. B. Cabrera, B.L. Dougherty, and K. D. Irwin in *Frontiers of Neutrino Astrophysics,* Proceedings of the 1992 Neutrino Astrophysics Symposium.

6. B. Cabrera, L. M. Krauss, F. Wilczek, *Phys. Rev. Lett.* **55**, 25 (1985).

7. P. Vogel and J. Engel, *Phys. Rev.* **D39**, 3378 (1989).

8. B. Neuhauser, B. Cabrera, C. J. Martoff, and B. A. Young, *Jap. J. of Appl. Phys.* **26**, 1671 (1987), Proc. 18th Int. Conf. on Low Temperature Physics.

9. B. A. Young, B. Cabrera, A. T. Lee, C. J. Martoff, B. Neuhauser, and J. P. McVittie, *Nucl. Inst. and Meth.* **A288**,199 (1990).

10. W. L. Bond et. al., *Phys. Rev. Lett.* **15**, 260 (1965).

11. D. C. Paine, J. C. Bravman, C. Y. Yang, *Appl. Phys. Lett.* **50**,498 (1987).

12. A. T. Lee, Ph. D. thesis, Stanford University, 1992, unpublished.

13. R. P. Welty, J. M. Martinis, *IEEE Trans. Magn.* **27**, 2924 (1991), Proc. 1990 Applied Superconductivity Conference.

14. R. P. Welty, J. M. Martinis, *IEEE Trans. Appl. Superc.* **3**, 2605 (1993), Proc. 1992 Applied Superconductivity conference.

15. M. L. Roukes et. al., *Phys. Rev. Lett.* **55**, 422 (1985).

16. S. I. Dorozhkin et. al., *Solid State Comm.* **60**, 245 (1986).

17. M. Frank, D. Dummer, S. Cooper, J. Igalson, F. Probst and W. Seidel, *Nucl. Inst. and Meth.* **A345**, 367 (1994).

18. W. A. Harrison, *Electronic Structure and the Properties of Solids*, 1980, W. H. Freeman & Company.

19. See, for example, T. Van Duzer, Principles of Superconductive Devices and Circuits.

20. K. D. Irwin, B. Cabrera, B. Tigner and S. Sethuraman in *Low Temperature Detectors for Neutrinos and Dark Matter IV*, eds. N.Booth and G. Salmon, (Editions Frontieres, France, 1992).

21. C. A. van der Jeugd et. al., *Appl. Phys. Lett.* **57**, 354 (1990).

22. W. Owens, personal communication.

23. J. C. Mather, *Appl. Opt.* **21**, 1125 (1982).

24. S. H. Moseley, J. C. Mather, and D. McCammon, *J. Appl. Phys*. **56**, 1257 (1984).

25. D. McCammon, W. Cui, M. Juda, J. Morgenthaler, J. Zhang, R. L. Kelley, S. S. Holt, G. M. Madejski, S. H. Moseley and A. E. Szymkowiak, *Nucl. Inst. and Meth.* **A326**, 157 (1993).

26. A. E. Szymkowiak, R. L. Kelley, S. H. Mosely, C. K. Stahle, *Journ. Low Temp. Phys*. **93**, 281 (1993), Proc. Fifth Int'l Workshop on Low Temperature Detectors.

27. N. E. Booth, *Appl. Phys. Lett.* **50**, 293 (1987).

28. M. Kurakado, *Nucl. Instr. Meth.* **196**, 275 (1982).

29. See, for example, F. Reif, *Fundamentals of Statistical and Thermal Physics* , McGraw-Hill, New York, 1965.

30. K. E. Gray, A. R. Long, and C. J. Adkins, *Phil. Mag.* **20**, 273 (1969).

31. B. L. Dougherty, B. Cabrera, A. T. Lee, M. J. Penn, and B. A. Young, *J. Low Temp. Phys.* **93**, 399 (1993), Proceedings of Fifth International Workshop on Low Temperature Detectors.

32. B. A. Young, B. Cabrera, A. T. Lee, *Phys. Rev. Lett.* **64**, 2795 (1990).

33. M. J. Penn, B. L. Dougherty, B. Cabrera, *J. Low Temp. Phys.* **93**, 423 (1993), Proc. of the Fifth Int'l Workshop on Low Temp. Detectors.

34. See, for example, M. Tinkham, *Introduction to Superconductivity*, McGraw Hill, 1975.

35. C. Kittel, Introduction to Solid State Physics, John Wiley & Sons, 1986.

36. J. M. Martinis, private communication.

37. L. Lesyna, D. DiMarzio, S. Gottesman, M. Kesselman, *J. Low Temp. Phys.* **93**, 779 (1993), Proc. of Fifth Int'l Workshop on Low Temp. Detectors.

38. The author is grateful for useful discussions on this subject with A. T. Lee.

39. M. Nahum, J. M. Martinis, *Appl. Phys. Lett.* **63**, 3075 (1993).