Precision control of thermal transport in cryogenic single-crystal silicon devices

Citation: Journal of Applied Physics 115, 124508 (2014); doi: 10.1063/1.4869737
View online: http://dx.doi.org/10.1063/1.4869737
View Table of Contents: http://scitation.aip.org/content/aip/journal/jap/115/12?ver=pdfcov
Published by the AIP Publishing

Articles you may be interested in
Anomalous diameter dependence of thermal transport in ultra-narrow Si nanowires

Atomistic simulations of heat transport in real-scale silicon nanowire devices
Appl. Phys. Lett. 100, 223107 (2012); 10.1063/1.4723632

Thermal conductivity of bulk and thin-film silicon: A Landauer approach

Dissipative quantum transport in silicon nanowires based on Wigner transport equation
J. Appl. Phys. 110, 093710 (2011); 10.1063/1.3654143

Single-crystal silicon/silicon dioxide multilayer heterostructures based on nanomembrane transfer
Appl. Phys. Lett. 90, 183107 (2007); 10.1063/1.2734367
I. INTRODUCTION

The management of heat in dielectric beams and membranes is an important part of the design in micro-electromechanical systems (MEMS). While the physics of thermal transport in dielectric materials is well understood, the effect of fabrication processes on the surface physics is less clear and widely treated as a hidden variable in the evaluation of the device performance. This uncertainty can prolong the design-test cycle, where a parameter such as beam geometry is iterated until the target thermal conductance $G$ is obtained.

In our application, the MEMS is a cryogenic detector known as the Transition-Edge Sensor (TES). Figure 1 shows a TES designed for measurements of the polarization state of the Cosmic Microwave Background. The TES is comprised of a superconducting MoAu bilayer deposited on a single-crystal silicon substrate. The silicon is etched around the bilayer and other electrical components to form a membrane that is supported by dielectric beams. The total thermal conductance $G(T)$ of the beams is a critical parameter that determines the sensitivity ($\propto \sqrt{G}$), response time ($\propto 1/G$), and saturation power of the TES detector ($\propto \int GdT$) (saturation energy in the case of a TES micro-calorimeter). The response time is also a function of the total heat capacity $C(T)$ of the detector, which is the sum of the heat capacity of the silicon membrane and normal metal films.

At temperatures far below the Debye temperature of the lattice ($\Theta_D = 645 K$), anharmonic phonon collisions in the bulk of the dielectric are negligible, and the phonon-mean-free path $\ell$ is restricted by the cross sectional dimensions of the beam. In state-of-the-art far-infrared and millimeter wave cryogenic detectors operating at sub-Kelvin temperatures, the beam length $L$ is much larger than $\ell$, and the thermal conductance is often described by the Fourier law for phonon diffusion, $G(T) = \kappa(T) w t/L$, where $w$ and $t$ are the width and thickness of the beam, and $\kappa(T)$ is the bulk thermal conductivity of the dielectric. $\kappa(T)$ is determined empirically with the fabrication and testing of numerous beams of varying width and length, and the aspect ratio $w/L$ is tuned to achieve the desired conductance. However, this approach rarely provides a complete view of the thermal physics, and higher order effects such as surface roughness of the beams have been identified as the likely cause of the reported non-uniformity in $G$ across detector arrays.

A simple analysis can readily show the sensitivity of phonon propagation in long beams ($L \gg \ell$) to the details of the surface physics. For a random Gaussian surface, the roughness can be characterized by the ratio of the standard deviation $\sigma$ of the surface features to the peak thermal wavelength $\lambda_{th}(T)$ of the incident phonon field. When $\sigma/\lambda_{th} \ll 1$, phonons are reflected specularly from the surface. As the wavelength becomes much smaller than the scale size of the surface, $\sigma/\lambda_{th} \gg 1$, phonons are scattered over all angles, or diffused. Assuming isotropic diffusion, it can be shown that for a rectangular beam, the phonon mean-free-path $\ell$ scales logarithmically with $r = w/t$ ($\ell \approx 0.75 w r^{-1} \ln 2 r$ for $r > 1$).
In the presence of surface roughness, the thermal conductance of the phonon transport in a short beam is highly biased by the fabrication and is likely non-Gaussian. Dielectric beams can be roughened in various forms depending on the compatibility and properties of the etchants with the dielectric material. Preliminary tests can shed light on the roughening steps, however, fabrication processes can be difficult to control in practice. Details such as base pressure and process chamber conditioning during dry-etch steps can vary between batches of devices, especially in shared facilities. As a result, the characteristics of the roughness of a single surface can change significantly over spatial scales shorter than the beam length, and between wafers. Furthermore, not all surfaces of a beam are rough to the same degree. These effects add to the unpredictable variability in $N$ and therefore $G$ in beams dominated by diffusive thermal transport.

The work in this paper is focused on understanding the effects of fabrication on the thermal conductance in single-crystal silicon beams. The beam geometries explored span several orders of magnitude in $L/wt$. The results demonstrate that $G$ in short ($L \leq 10 \mu m$, $L/\ell \leq 1$) uniform beams with ballistic-dominated phonon transport is insensitive to the fabrication conditions and can be realized with precision in devices on different wafers. This approach is suitable for TES detector arrays in ground-based and air-borne telescopes, where the detector noise equivalent power requirement is greater than $10^{-18} W/\sqrt{Hz}$.

Throughout this paper, the terms diffuse and ballistic are used to describe the scale size of the phonon transport in a silicon beam. In the presence of surface roughness, the thermal transport in a long beam with $L/wt > 1$ is dominated by diffuse reflections off the beam walls, and $L/\ell \gg 1$. For a short beam, ballistic transport dominates and $L/\ell \leq 1$. We emphasize that these limits bound the diffusive-ballistic conductance observed in practice.

In Sec. II, we briefly describe the device fabrication and readout hardware employed for the low-temperature measurements of $G$ and $C$. In Sec. III, we describe the theory of radiative phonon transport in diffusive-ballistic beams. In Sec. IV, we present results that illustrate the effect of fabrication on the diffusive-ballistic nature of the thermal conductance. In Sec. V, we describe the superior performance of ballistic beams for the control of $G$ of TES detectors. Finally, we demonstrate that because of the low specific heat capacity of single-crystal silicon, the heat capacity of the detectors fabricated on this substrate can be effectively controlled using a normal metal film with high specific heat.

II. METHODOLOGY

A. Fabrication

The fabrication of the TES detectors reported in this paper is described in detail in Ref. 13, but for completeness, we provide a brief description here. Fabrication starts with a float-zone-refined p-type (100) boron doped silicon-on-insulator wafer. Heat flow is primarily in the [110] direction. The single-crystal silicon layer is 5 $\mu m$ thick, with a room temperature resistivity of 5 k$\Omega$ cm and a carrier density of $5 \times 10^{11} cm^{-3}$. A MoAu bilayer is deposited by a combination of dc plasma sputter deposition and electron beam evaporation. The Mo and Au are photolithographically patterned to define the superconducting sensor. The Au layer is etched by ion milling and the Mo layer is etched by reactive ion etching (RIE). Further metallic components are deposited on the device layer to enable the required operation of the pixel at rf frequencies, and to enable on-chip diagnostic capabilities such Johnson noise (JN) thermometry as needed. Au resistors are deposited by electron beam evaporation and patterned by lift-off. Following an in situ rf Ar plasma cleaning of the wafer, a Nb microstrip layer is sputter deposited, photolithographically patterned, and etched in fluorine-based RIE chemistry. The silicon membrane is defined and reactive ion etched down to the buried oxide layer. The silicon wafer is then bonded to a pyrex handle wafer using a sacrificial wax. The backside of the silicon wafer is patterned by photolithography and the bulk silicon is etched by inductively coupled plasma deep reactive-ion etching (DRIE). The wax is then dissolved in solvent to release the devices.

B. Experimental hardware

Measurements were performed in a mechanically cooled system that reaches 3 K at the coldest stage. Further cooling of the detector stage down to 100 mK was achieved using an adiabatic demagnetization refrigerator. The devices were mounted and thermally anchored to a gold-plated copper package, and wirebonded to a Superconducting Quantum Interference Device (SQUID) cold stage amplifier for Johnson noise current measurements. Further amplification was achieved with a SQUID series array mounted at the 3 K stage. The temperature of the 100 mK cold stage was read
with a calibrated RuO$_2$ thermometer\textsuperscript{14} with calibration traceable to the National Institute of Standards and Technology. The precision of the temperature control was 32 $\mu K$ rms at 150 mK.

The room temperature electronics comprised of a custom low-noise pre-amplifier, feedback electronics for flux-locked-loop operation, commercial low-noise high-gain amplifier and low-pass filter (SRS-650), and commercial 16 Bit digital-to-analog electronics (NI PCI-6120) for digitization of the amplified feedback signal.

C. Johnson noise thermometry

The method of JN thermometry is frequently utilized for the characterization of the thermal properties of thin-films below 1 K.\textsuperscript{15} At audio frequencies (0.5–20 kHz) and sub-Kelvin temperatures, the current noise spectral density of a resistor $R$ in series with the SQUID input coil takes the form $S_I(f) = 4k_BT_e \left[1 + (f/f_c)^2\right]^{-1}/R$, where $T_e$ is the electron temperature, $f_c = R/(2\pi L)$ is the characteristic frequency of the circuit, and $L$ is the total inductance of the input coil and stray inductance in the input circuit loop. This equation directly relates measurements of $S_I$ to $T_e$. For the test devices that are described in Ref. 13, the JN thermometer and heater are $240 \times 230 \times 0.27 \mu m^3$ Au resistors deposited on top of interdigitated Nb leads. The Kapitza conductance at the Au-silicon interface and the electron-phonon conductance in the Au are approximately 10 nW/K each at 100 mK.\textsuperscript{16,17} These conductances are a factor of 50 greater than the highest silicon beam conductance measured in our devices. We, therefore, assume the electron system in the thermometer is in thermal equilibrium with the phonons in the silicon membrane. The maximum expected temperature gradient across the membranes is less 0.5 mK across the 0.1–1 K temperature range.

The random error in a temperature measurement from the JN current can be estimated from the radiometer equation $\sigma_T(T) = T/\sqrt{B \tau_m}$,\textsuperscript{15,18} where $B = 5$ kHz is the bandwidth over which the current noise power spectral density was measured and $\tau_m = 102$ s is the measurement time. Hence, the fractional error throughout the temperature range of our measurements is less than 0.5%. Measurements of the JN current as a function of bath temperature were used to fully calibrate the cryogenic and warm readout amplifiers. Typical readout noise of 15 pA/$\sqrt{Hz}$ referenced to the input circuit was achieved. The JN signal was 30 – 100 pA/$\sqrt{Hz}$ in the range of 0.1–1 K.

To measure $G$ of the silicon beams, a small signal dc heater power was applied to the membrane and the increase in temperature was recorded. The relationship between heater power and membrane temperature at a given bath temperature $T_B$ can be parameterized by the power law function $P = P_f (T_B^n - T_B^m)$, where $P$ is the heater power, $T$ is the membrane temperature, $K$ is a constant that is proportional to the number of phonon modes, and $n$ is the temperature exponent that is indicative of the temperature dependence of the phonon mean-free-path. The first order derivative yields the thermal conductance $G(T) = nKT_B^{n-1}$. In our measurements, the change in temperature was small, $\sim 10\%$, and the approximation $G(T) \approx \Delta P(T)/\Delta T$ was valid.

D. TES current-voltage thermometry

We have also utilized TES current-voltage curves to determine the thermal conductance of detectors in which normal metal films for JN thermometry were omitted. When a TES is voltage-biased on the normal-to-superconducting phase transition, the silicon membrane temperature is fixed at the critical temperature of the MoAu bilayer, $T_C$ $\cong$ 150 mK. The thermal conductance is then determined from a fit to the Joule power dissipated in the TES as a function of bath temperature. In this case, $P = P_f (T_B^n - T_B^m)$ and $G(T_B) = nK T_B^{n-1}$, where the parameters $K$ and $n$ are equivalent to those derived from JN thermometry in the same temperature range.

III. MODELING HEAT TRANSPORT

In building a model of heat flow in a silicon beam, it is important to first identify the thermal baths that define the boundary conditions in the model. A thermal bath is ideally a blackbody that provides a matched termination for phonons at all wavelengths. In a detector such as that shown in Fig. 1, each silicon beam is connected to two thermal baths, the frame and the membrane. The electron system in the Pd film is strongly coupled to the phonon system in the membrane. At 150 mK, the electron-phonon conductance in the Pd is 30 nW/K,\textsuperscript{15} and the Kapitza conductance between the Pd and the silicon membrane is $\sim 130$ nW/K.\textsuperscript{16} For comparison, the highest beam conductance measured is 240 pW/K at 150 mK. The silicon frame is effectively a blackbody since its lateral dimensions are much larger than the peak thermal wavelength. At 100 mK, the peak thermal wavelength is $\lambda_{th} \sim 4hc/k_BT = 0.6 \mu m$ in silicon, where $h$ and $k_B$ have the usual definitions and $c = 5070$ m s$^{-1}$ is the speed of sound for transverse waves. A phonon scattered from a beam into the silicon frame has a low probability of coherently backscattering into the beam. Given the 5 $\mu m$ thickness of the silicon layer, and the thermal wavelengths above 100 mK ($\lambda_{th} \approx 0.6$ (0.1$/T$) $\mu m$), the membrane and the frame can be modeled as three-dimensional multi-moded thermal baths, where the number of modes is given by the Planck distribution and is much larger than unity.

The beams behave as three-dimensional multi-moded acoustic waveguides since the widths ($w > 13$ $\mu m$) and 5 $\mu m$ thickness are much larger than $\lambda_{th}$ at 0.1–1 K. Phonons emitted by the baths couple radiatively to the silicon beams, propagate unimpeded in the bulk of the beams, and scatter at the beam surfaces. When the surfaces of the beams are perfectly smooth, phonons coupled to the beams are scattered elastically without any loss of coherence in the beams. The conductance in this case is determined by the beam areas and is referred to as ballistic. In the case of rough surfaces, phonons scatter diffusively with a large probability of backscatter. The scattering is elastic but the phonon wavevectors are randomized at the surface. If the surface features form cavities much larger than $\lambda_{th}$, the surface can be thought of as a blackbody. The conductance is thus boundary-limited and was first described theoretically by Casimir.\textsuperscript{20} An important assumption in Casimir’s formalism is the isotropic (Lambertian) re-emission of phonons by the beam surfaces.
In practice, the scattered pattern of phonons from a rough surface is strongly dependent on the surface characteristics. When the roughness can be described statistically, the reflection pattern is averaged over the ensemble of the surface realizations. For a surface with a Gaussian correlation function, the roughness can be characterized by the correlation length $\xi$ and standard deviation $\sigma$ of the surface features above a mean height. Figure 2 illustrates the scattered pattern as a function of $\sigma/\lambda$. For a nearly smooth surface, $\sigma/\lambda \ll 1$, phonons are largely reflected in the specular direction. When $\sigma/\lambda \gg 1$, phonons incident on the surface are diffused in all directions, however, the flux is not Lambertian. Nonetheless, in numerous theoretical studies of boundary-limited scattering, the reflected pattern is decomposed into a diffuse Lambertian fraction $f$ and a specular fraction $1 - f$, as illustrated in Fig. 2. The diffuse reflection is modeled as inelastic scattering that redistributes the phonon wavevector, although strictly speaking all reflections are elastic at the surface facets that define the roughness. The surface can then be thought of as a blackbody with emittance $f$. Hence, a completely lossless but rough surface can be modeled as a lossy surface. This approach is numerically less intensive than calculations of the reflection pattern based on the exact surface profile, or the statistics of the surface.

Our model of heat flow in a silicon beam has the following assumptions: (i) The ends of the beam constitute thermal baths with constant temperature and unity diffuse fraction. The baths continue to distances much larger than the beam length as shown in Fig. 3(a). This construct is numerically convenient from the perspective of the heat transport model presented, and is similar to that adopted in Ref. 22. (ii) The surfaces of the beam reflect phonons with a diffuse fraction $f$ that is isotropic, and a specular fraction given by $1 - f$. (iii) The gradient of the temperature profile across a beam is linearly dependent on $f$ and constant as a function of length. This assumption is valid in multimoded beams, where the number of modes is much larger than unity and the diffuse phonon reflection can be treated as diffuse emission. Models that iteratively solve for the temperature profile in multimoded beams show negligible deviations from a constant temperature gradient. When the phonon mean-free-path is much smaller than the beam length, or $f$ is close to unity, phonon diffusion is the dominant term of the total heat flux, and the constant temperature gradient is an excellent approximation. The approximation also holds when $f$ approaches zero, or the mean-free-path is comparable to the beam length, since the temperature is constant at the thermal baths, which are represented by diffuse surfaces that extend to distances much larger than $L$. (b) Each surface element of the beam radiates phonons into all solid angles. The radiated flux from the infinitesimal area $dA_i$ that reaches the shaded beam port is reduced depending on the diffuse fraction of the beam surfaces. (c) Using the method of images, the path of specularly scattered phonons can be reimagined to one vector. The components of the vector map the number of scattering events at each surface of the beam.

![Figure 2](image2.png)

**FIG. 2.** Beam patterns of phonons reflecting off a fairly smooth ($\sigma/\lambda \ll 1$) and a very rough surface ($\sigma/\lambda \gg 1$). The diffuse fraction of the realistic beam patterns is generally approximated as isotropic (Lambertian), which is relatively simple to integrate into models of thermal transport.
the thermal spectrum can be scattered very differently as temperature changes, and in recognizing the limitations of our model, we do not extrapolate the results to $\langle \xi \rangle$.

Figure 3(b) illustrates the radiative model. The beam is divided into infinitesimal areas $dA = dx_i dx_i$ that radiate heat into all solid angles. The phonons are scattered several times before reaching the beam exit port and thermalizing in a bath. During each scattering event, the phonon flux is reduced by a factor $f_i$ that is the diffuse fraction of the surface $i$. The nature of the problem is then to simultaneously calculate the flux that is viewed by $dA$, and the total number of scattering events from each surface as $f_i$, and the total heat flux through the beam.

For the latter, the complex path of the phonons can be unfolded, or re-imaged, to a single vector that reaches an equivalent beam exit port as shown in Fig. 3(c). The components of the vector then equate to the number of scattering events from each surface.

The details of the radiative model are described in the Appendix. We define an equivalent view factor for a rectangular waveguide

$$F(x_i) = \sum_{m=1}^{4} \left[ F_l(f_m; x_i) + \sum_{s=1}^{2} F_{l_s}(s, f_m; x_i) + F_{s_d}(f_m; x_i) \right],$$

where $F_l(f_m; x_i)$ and $F_{l_s}(s, f_m; x_i)$ are equivalent view factors between parallel and perpendicular surfaces of the beam, respectively, and $F_{s_d}(f_m; x_i)$ is the contribution of the direct (unscattered) flux from the surfaces of the beam to the cross sectional area at $x = 0$ (reference plane). $f_m$ is the diffuse fraction of the surface emitting the flux, and the index $m$ runs over all surfaces of the beam. The index $s$ runs over the surfaces perpendicular to $m$. Included in the terms $F_l(f_m; x_i)$ and $F_{l_s}(s, f_m; x_i)$ is a factor $\prod_{s=1}^{2} (1 - f_s)^{N_s}$, where $N_s$ is the total number of scattering events from each surface as described in Fig. 3(c). For the baths, $F_{b_d}(x_i)$ is obtained by setting $f_m = 1$ in Eq. (1). The total heat flux through the beam is $P = 2 F_{b_d}^{2}$.

$$P = q \left[ \int_{-L/2}^{L/2} T^4(x_i) F(x_i) dx_i - \int_{0}^{L/2} T^4(x_i) F(x_i) dx_i + T^4_b \int_{L}^{\infty} F_B(x_i) dx_i \right],$$

where $q = \pi^2 k_b^2/(120h^3\bar{v}^2)$, and $\bar{v} = \sqrt{2/\pi^2 + 1/\pi^2}$ is an average group velocity. Given the linear temperature profile $T(x_i) = x_i (dT/dx_i) + T_i$, the approximation $T_b - T_s \approx 4 (dT/dx_i) LT^3$, where $T$ is the average of the bath temperatures, and the relation $F(x_i) \equiv F(-x_i)$, Eq. (2) reduces to

$$P = 4 q T^3 \int_{-L}^{L} x_i F(x_i) dx_i + 4 q T^3 \int_{L}^{\infty} F_B(x_i) dx_i,$$

where $I_d$ is the net diffuse power from the beam surfaces, and $I_s$ is the net specular power that is exchanged between the baths. The conductance of the beam as a function of temperature when scaled by the maximum (ballistic) conductance $G_B = 4 q T^3 A$ is

$$g = \frac{G}{G_B} = \left( \frac{P}{A (dT/dx_i) L} \right) \frac{1}{G_B},$$

$$= \frac{1}{A} \left( \frac{2 I_d + I_s}{L} \right).$$

In the short beam limit, $I_d \to 0$ and $I_s \to A$. For non-zero diffuse fractions and as the beam length becomes much larger than the cross-sectional dimensions, $I_s \to 0$, and using $\kappa = C_V \bar{v} L/3$, where $C_V = 16 q T^3/\bar{v}$ is the volume specific heat of the material, the mean free path takes the form $\ell = 3 I_d/(2A)$.

When the diffuse fractions of the surfaces of the rectangular beam are different, all terms in the summation over the index $m$ in Eq. (1) must be evaluated. The numerical evaluation of the problem is reduced by a factor of two if there is symmetry in the diffuse fraction of the surfaces, $f_{Top} = f_{Bottom}$ and $f_{Left} = f_{Right}$. For the silicon beams discussed in this paper, the top surface and sidewalls are diffuse, and the bottom surface is specular. It is important to note that the diffuse fractions of the top and sidewall surfaces are uncorrelated.

The terms $F_l(f_m; x_i)$ and $F_{l_s}(s, f_m; x_i)$ in Eq. (1) are evaluated using Monte-Carlo integration with $10^5$ uniformly distributed random points averaged $10^2$ times. $F_{s_d}(f_m; x_i)$ is evaluated with an adaptive quadrature routine. Figure 4 illustrates the accuracy of the model compared to an analytic expression for the mean-free-path $\ell$ in a rectangular beam. At the diffuse limit ($L/\ell \gg 1$), the conductance scales as $1/L$, and at the ballistic limit, the conductance is independent of length. The expression $g = \ell / (L + \ell)$ is remarkably accurate over the range of beam length explored in Fig. 4. This fact highlights the importance of the phonon mean-free-path, that a single parameter can capture a significant fraction of the heat transport in the ballistic-diffusive limit.

FIG. 4. Simulations of a beam with fixed cross-section compared to the analytic expression in Ref. 12. The sidewall diffuse fraction is unity. The standard error from the Monte-Carlo integration is shown for each simulation.
IV. DIFFUSIVE-BALLISTIC CONDUCTANCE

Figure 5 shows the thermal conductance of silicon beams as a function of beam area-to-length ratio. The cause of the variability in conductance (a factor of ~5 for the same beam geometry) is the roughening of the silicon beam surfaces during fabrication. The top surface of the beams is primarily roughened during a dry etch step that patterns a Nb layer. The plasma chemistry, 75% CF4, 15% O2, etches silicon at a higher etch rate than Nb. The exact processes that roughen the silicon are unclear, but the likely candidates are micro-masking due to non-volatile species in the plasma, and slower etch of native and in situ NbOx in the fluorine chemistry. The use of an etch stop such as 100 nm thick AlOx was successfully employed in early devices; however, these detectors exhibited very large uncontrolled heat capacity due to the amorphous nature of the etch stop.24 Test devices were also fabricated with a lift-off process that left the beam surfaces smooth. The conductance of these beams was higher, suggesting boundary-limited phonon propagation as the cause of the variability in conductance of the beams with rough surfaces.

We have explored the relationship between thermal conductance and surface roughness with an extensive dataset of atomic force microscopy (AFM) images. Non-destructive AFM of a silicon beam surface a millimeter in length is not practical. Instead, the silicon device layer was imaged on the handle wafer, in close proximity to each device. For a silicon surface with a Gaussian roughness, $\sigma \approx 5–20$ nm and $\zeta \approx 100–200$ nm depending on the processing conditions. However, we have found that a large fraction of the roughened surfaces did not satisfy the mathematical assumptions required to apply the theory of phonon scattering by nearly smooth Gaussian surfaces, and the measured thermal conductance differed from expectations of such a theory, even when $\sigma$ was much less than the typical phonon wavelength at the temperatures explored (0.1–1 K). We observed power spectral densities of surface height that were approximately power laws over the range of observable wavevector values. In this case, we found an empirical relationship between the power-spectral-density $S(k)$ of the surface profile and the measured magnitude and temperature dependence of the phonon mean-free-path in our silicon beams, $\ell = \pi t/(4k^2\sqrt{S(k)/(2\pi)})$, where $t$ is the silicon beam thickness and $k = 0.010 T/(\text{nm K})$ is the phonon wavenumber. As expected from Beckmann’s theoretical treatment of scattering from non-Gaussian surfaces, the observed relationship involves the correlation function of the local tilt angle of the surface facets, $\theta_{\text{rms}}(k) = k^2 \sqrt{S(k)/(2\pi)}$.

It is challenging to precisely predict the phonon mean-free-path from information contained in the AFM data due to practical sampling considerations. This occurs for several reasons: The statistics of the surface roughness are observed to vary significantly across the wafer and are sparsely sampled in $8 \times 8 \mu m^2$ patches set by the resolution and stability of the AFM. In addition, for a multi-moded beam with thickness of order the width, details such as surface roughness of the beam sidewalls are significant, as is shown in Sec. V. Non-destructive AFM imaging of the sidewall geometry is extremely difficult to perform, and such a task becomes impractical when considering the number of silicon beams that constitutes a detector imaging array. However, we have found that AFM images can provide qualitative insight to the thermal conductance in beams where the surface roughness relative to the thermal wavelength has a significant amplitude and is non-Gaussian.

Figure 6(a) shows the conductance of smooth (lift-off) and roughened (dry etch) silicon beams when compared to the boundary-limited conductance in each case.12 The phonon propagation transitions from diffuse to ballistic at 300 mK, where the peak thermal wavelength is $\lambda_{\text{th}} \approx 200$ nm. The minimum in conductance suggests the transition is resonant for the rough beam.26 The inset in Fig. 6(a) shows the inverse phonon mean free path $\ell^{-1}$. The scale height of the scatterers $\zeta$ can be estimated from the temperature at which $\ell^{-1}$ peaks. Using the Rayleigh criterion, $\zeta \approx \lambda_{\text{th}}/8\sin\gamma$,11 where $\lambda_{\text{th}}$ is the phonon wavelength at the peak temperature, and $\gamma$ is the phonon angle of incidence to the surface. Assuming an average over all angles, $(\sin\gamma) = 2/\pi$, the Rayleigh criterion suggests $\zeta \approx 40$ nm for the rough beam and $\zeta \approx 6$ nm for the smooth beams. These estimates are consistent with the roughness measured from the AFM images shown in Fig. 6(b). The rough (dry etch) surface is heavily pitted, with an average pit depth of 30 nm and average size of 250 nm that is comparable to $\lambda_{\text{th}}$ at 300 mK. The scale size is much smaller for the smooth (lift-off) surfaces, $\sigma \approx 3$ nm, which leads to the onset of resonance at a much higher temperature.

In addition to phonon scattering by rough surfaces of a beam, organic residue such as photoresist on the surfaces can enhance diffuse scattering. Thin layers of metals and non-metals hundreds of Angstroms thick have been shown to induce diffuse transport in silicon.26 In our devices, the...
reduction and variability in $G$ in a few cases were at least qualitatively attributed to residue accumulation on beams of order a millimeter in length. Care must be taken when devices are cleaned since the cleaning process may introduce further residue, or roughness (see example of silicon dipped in HF in Ref. 26).

Table I summarizes the silicon beam geometries and conductances shown in Fig. 5. The exponent $\beta \equiv n - 1$ is inversely proportional to $G$, indicating a transition from diffuse to ballistic transport. Low $G$ devices have boundary-limited transport, and $\beta \approx 3$. High $G$ devices are approaching ballistic transport, and $\beta < 3$ in the transition region.

### V. CONTROL OF $G$ IN SILICON-BASED DETECTORS

In the design of prototype TES detectors, the choice of the beam length and width was guided by the data in Fig. 5. Several strategies were adopted in order to control the conductance and produce uniform detectors for the focal plane array described in the text. To minimize the top surface area of the silicon beams roughened by the dry etch, the width of the Nb leads was extended to cover a larger fraction of this surface, and the beam widths were defined with a combination of DRIE using the standard Bosch process, followed by RIE to remove the silicon down to a Nb ground plane. The two processes are separated to reduce potential metallic contamination in the DRIE chamber, while minimizing lateral silicon undercut from the isotropic RIE. The bottom surface of the beams in contact with the Nb ground plane is smooth, and scanning electron microscopy images of the sidewalls showed no significant features from the RIE etching of the beams employed in test structures, Fig. 7(a). The beam sidewalls were predicted to be nearly specular, with conductances comparable to the lift-off devices shown in Fig. 5. The top surface of the beams was expected to be slightly rough, $f \sim 0.2$, since the area exposed during the dry etch step was reduced by a factor of two. Four beams

![FIG. 6. (a) Conductance of silicon beams scaled by the boundary-limited conductance in a rectangular beam (dashed line). Solid lines are spline fits to guide the eye. The labels refer to the top surface of the beams, which is roughened by a dry etch step during fabrication (squares), but is left smooth when the dry etch step is replaced with a lift-off process (diamonds, triangles). The inset shows the inverse phonon mean-free-path as a function of temperature. The peak in $f^{-1}$ suggests resonant phonon scattering from the rough surface. (b) AFM images of $6.4 \times 6.4 \mu m^2$ patches representative of the top surface of the beams described in (a). The scale size estimated from the peaks in $f^{-1}$ is consistent with the roughness scales measured from these images. The rough surface is heavily pitted, with an average pit depth of $15 \text{nm}$. The large distribution in pit size and shape is the likely cause of the observed broad resonance. Since the scale size is much smaller for the smooth surfaces ($\sigma \approx 3 \text{nm}$), the onset of the peaks occurs at a much higher temperature (shorter thermal wavelengths). The striping in the AFM image of the smooth surface is due to residue and accumulation of particulates on the AFM tip. Beam dimensions are $L = \{0.4, 1.5, 3\} \text{ mm}$, and $w = \{24, 20, 25\} \mu m$ for devices labelled with squares, diamonds, and triangles, respectively.

![FIG. 7. SEM images of silicon beams. The Nb lead is 200\text{nm} thick. (a) The isotropic RIE chemistry leaves smooth sidewalls with a slight taper. In Fig. 5, devices marked with crosses, squares, diamonds, and circles have similar sidewall profiles. (b) To minimize the top surface area roughened by the dry etch of the Nb lead layer, DRIE was used to define the beam width. The ridges at the top of the beam are scalloping from the DRIE cycles. Possible sources for the bright features on the sidewall are leftover passivation layer and charged silicon features.]}

<table>
<thead>
<tr>
<th>Length ($\mu m$)</th>
<th>Width ($\mu m$)</th>
<th>$z$ (pW/K$^{\frac{1}{2}}$)</th>
<th>$\beta$</th>
<th>$G$ (150 mK) (pW/K)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5000</td>
<td>35</td>
<td>$1054 \pm 55$</td>
<td>3.13</td>
<td>2.78</td>
</tr>
<tr>
<td>5000</td>
<td>45</td>
<td>$1449 \pm 50$</td>
<td>2.96</td>
<td>5.28</td>
</tr>
<tr>
<td>3000</td>
<td>25</td>
<td>$499 \pm 36$</td>
<td>2.92</td>
<td>1.96</td>
</tr>
<tr>
<td>3000</td>
<td>25</td>
<td>$1241 \pm 118$</td>
<td>3.27</td>
<td>2.51</td>
</tr>
<tr>
<td>3000</td>
<td>25</td>
<td>$1210 \pm 103$</td>
<td>2.19</td>
<td>18.99*</td>
</tr>
<tr>
<td>1500</td>
<td>15</td>
<td>$625 \pm 44$</td>
<td>2.92</td>
<td>2.46</td>
</tr>
<tr>
<td>1500</td>
<td>15</td>
<td>$583 \pm 31$</td>
<td>3.21</td>
<td>1.32</td>
</tr>
<tr>
<td>1500</td>
<td>20</td>
<td>$2000 \pm 81$</td>
<td>2.38</td>
<td>21.88*</td>
</tr>
<tr>
<td>1000</td>
<td>25</td>
<td>$2406 \pm 37$</td>
<td>2.89</td>
<td>10.00</td>
</tr>
<tr>
<td>553</td>
<td>20</td>
<td>$2521 \pm 67$</td>
<td>3.11</td>
<td>6.90</td>
</tr>
<tr>
<td>544</td>
<td>20</td>
<td>$1922 \pm 132$</td>
<td>2.80</td>
<td>9.48</td>
</tr>
<tr>
<td>400</td>
<td>24</td>
<td>$4292 \pm 386$</td>
<td>2.45</td>
<td>41.10</td>
</tr>
</tbody>
</table>
785 μm in length and 13 μm in width were integrated into the TES design.

Post fabrication, the beams had well-defined top surfaces, but the sidewalls were roughened by the DRIE steps, Fig. 7(b). The total conductance was a factor of 5 below the predicted target range. In principle, the scale length and height of the sidewall features that include the DRIE scalloping as well as passivation residue present a very diffuse surface for phonons above 100 mK. Figure 8 compares the total conductance of the four beams to simulations based on the diffusive-ballistic model described in Sec. III. The diffuse fractions of the sidewall and top surface of the beams vary across a wafer and between wafers.

The inferred diffuse fraction of the top surface of the beams is \( f \leq 0.4 \). A possible source for the moderate roughness is the 200 nm thick Nb lead sputter-deposited on each beam. The lead and beam are generally considered as decoupled systems, and the Nb conductance is estimated to be far below that of the silicon beam. However, it is the diffusion process of the phonons in the silicon by the Nb layer that is of importance and generally unknown. It has been shown that even in the presence of large acoustic mismatch, phonons in silicon can be diffused at surfaces with metallic layers tens to thousands of Angstroms thick (see Fig. 8 in Ref. 26 for examples of phonon diffusion caused by Al and Ti films below 1 K).

It is evident from the results in Fig. 8 that a solution to the problem of conductance control is a design that is least sensitive to the effect of fabrication conditions such as surface roughening and residue. Our approach was to reduce the length \( L \) of a beam, which increases the phonon mean-free-path \( \ell \) by reducing the interaction probability of the phonons with the rough sidewalls of the beam. The conductance is less sensitive to \( \ell \) in the limit \( L/\ell \leq 1 \). Figure 9 shows the simulated and measured total conductances when a ballistic beam 10 μm long and 13 μm wide is integrated into the detector as shown in Fig. 1. Given our knowledge of the surface roughness after the dry etch step (Fig. 8), a diffuse fraction \( f \leq 0.4 \) was expected and subsequently obtained in practice.

The total conductance, which is largely set by the ballistic beam, was predicted prior to fabrication. The diffuse beams merely support microstrip and bias leads necessary for the operation of the detector. However, they contribute the largest fraction to the variance in the total conductance across a wafer and between wafers. The magnitude and variance in \( G \) of these beams were estimated from the data in Fig. 8. The fractional standard deviation in total conductance is ±8% and within the requirement of the target instrument.

The range of conductances that can be accessed with a ballistic beam is suitable for low-noise bolometric detectors operating at ground-based or air-borne telescopes where the signal power is of order 1-3 pW/K. In the ideal design, a single ballistic beam sets the conductance of the detector and carries all signal and bias leads. This approach is very well suited to the fabrication of kilo-pixel arrays, especially when the signal is coupled to the detector with free-space absorbers. For space applications where the signal power is less than 1 pW, the ballistic beam can be a phononic structure with phonon bandgaps that can be engineered to reduce the conductance.

VI. CONTROL OF C IN SILICON-BASED DETECTORS

The heat capacity of the dielectric also plays an important role in the performance of thermal detectors. For a bolometer and calorimeter, \( C \) affects the detector time constant, and in some cases anomalous heat capacity associated with the dielectric has been reported to increase thermodynamic noise of these detectors.

The heat capacity of the silicon membranes was determined from \( C = G \tau_{th} \), where the thermal time constant \( \tau_{th} \) was measured from the decay in temperature of a silicon membrane to a pulse in heater power. Compared to the
Debye specific heat, $C_V^D / C_V^\text{Meas} = 0.2$ at 200 mK and 0.7 at 1 K. The systematic difference between the measurements and the Debye approximation is due to the large contribution to the total heat capacity of the devices from the normal metal Au heater/thermometer and Pd-Au termination on the silicon membranes. Nonetheless, the measured specific heat is extremely small. For the detectors shown in Fig. 1, the estimated silicon heat capacity is $\sim 2 \Omega/K$ at 150 mK, which is advantageous since the overall detector heat capacity can be controlled with a normal metal that can be reliably deposited on the membrane. We have successfully used Pd to achieve a heat capacity of 2 pJ/K.\textsuperscript{2} Measurements indicate that the specific heat of Pd below 1 K is 5.8 mJ mol\textsuperscript{-1} K\textsuperscript{-2}, which is lower than the room temperature value of 9.4 mJ mol\textsuperscript{-1} K\textsuperscript{-2}.\textsuperscript{9} We have successfully deposited 400 nm Pd layers with a lift-off process.

It is important to emphasize that there is no evidence for anomalous specific heat in the single-crystal silicon device layers. This is in contrast to the large excess specific heat of amorphous SiN\textsubscript{x} that is widely used in the fabrication of thermal detectors.\textsuperscript{29–31}

VII. CONCLUSIONS

The thermal conductance of single-crystal silicon devices can be precisely controlled using a short beam with ballistic-dominated phonon transport. The conductance of the short beam is largely set by its area and is insensitive to the details of the fabrication conditions that can vary over time. This approach is superior to the conventional use of long beams for the control of thermal conductance. Phonon transport in long beams is diffusive-ballistic and very sensitive to the detailed surface physics of the beams. The conductance is thus boundary-limited, and has been observed in beams with modest surface roughness (5–10 nm rms in surface height). Resonant phonon scattering has also been observed at 300 mK in beams with highly pitted surfaces with pit depths of order 30 nm. Boundary-limited scattering has contributed to a large variance, up to a factor of 5, in the conductance of devices fabricated and tested during this work. In contrast, when a beam with ballistic-dominated thermal transport is integrated into the design, the uniformity in conductance is reduced to a fractional deviation of ±8% in devices fabricated across two wafers, and this variability is largely set by the conductance of diffuse beams that support electrical and microwave leads to the device. In addition, we have found no evidence of excess specific heat in single-crystal silicon membranes. Hence, the total heat capacity of the detector can be effectively controlled with a normal metal film. For the TESs described in this work, the heat capacity was determined by a Pd layer of 400 nm thick. Hence, the strategies outlined for the precision control of thermal conductance and heat capacity are well suited to the fabrication of uniform large-format arrays of TESs with sensitivities approaching $10^{-18}$ W/Hz\textsuperscript{1/2}.

ACKNOWLEDGMENTS

We gratefully acknowledge financial support from the NASA ROSES/APRA program. K. Rostem was supported by the NASA Postdoctoral Program at the Goddard Space Flight Center during the initial stages of this work. We thank Samelys Rodriguez for wirebonding and cryo-cable support. We thank our collaborators A. Ali, J. Appel, and T. A. Marriage for their contribution to the data acquisition. We thank the National Institute of Standards and Technology for providing the SQUID series array and time division multiplexer chips.

APPENDIX: RADIATIVE PHONON TRANSPORT

We start from the usual definition of the infinitesimal view factor $dF_{dA_i,dA_j}$,

$$dA_{i} dF_{dA_{i},dA_{j}} = \frac{\cos \theta_i \cos \theta_j}{\pi | \vec{R} |^2} dA_{j} dA_{i}, \quad (A1)$$

where $\vec{R}$ is the vector between the infinitesimal surfaces $dA_{i}$ and $dA_{j}$, and $\theta_i, \theta_j$ are the angles between the normal vectors to $dA_{i,j}$ and $\vec{R}$ (see Fig. 3). We define an equivalent view factor for convenience,

$$F_{dA_{i},dA_{j}} = dA_{i} dF_{dA_{i},dA_{j}}, \quad (A2)$$

and note that the heat exchanged between the infinitesimal surfaces is

$$dP_{dA_{i},dA_{j}}(x) = q [T_{i}^4(x) - T_{j}^4(x)] F_{dA_{i},dA_{j}}, \quad (A3)$$

where $T_{i,j}(x)$ are the temperatures of the surface elements. The total heat flux is the integral of $dP_{dA_{i},dA_{j}}$ over the beam length,

$$P = \int_{S} dP_{dA_{i},dA_{j}}(x), \quad (A4)$$

where $S$ represents the surfaces of the waveguide, which in our case is a rectangular beam.

The equivalent view factor between any two perpendicular surface elements of the beam is

$$F_{dA_{i},dA_{j}} = \left\{ \frac{y_{i} z_{j}}{\pi [(x_{i} - x_{j})^2 + z_{i}^2 + y_{i}^2]^2} \right\} \left[ f_{n} \prod_{k} (1 - f_{k})^{N_{k}} \right] dA_{i} dA_{j},$$

$$= F_{dA_{i},dA_{j}} \left[ (m; x_{i}, y_{i}, z_{i}) dx_{i} dy_{i} dz_{i}, \right] \quad (A5)$$

where the indices $m$ and $k$ sum over all surfaces of the beam. For parallel surface elements,

$$F_{dA_{i},dA_{j}} = \left\{ \frac{t^2}{\pi [(x_{i} - x_{j})^2 + y_{i} - y_{j})^2 + t^2]^2} \right\} \left[ f_{n} \prod_{k} (1 - f_{k})^{N_{k}} \right] dA_{i} dA_{j},$$

$$= F_{dA_{i},dA_{j}} \left[ (m; x_{i}, y_{i}, z_{i}) dx_{i} dy_{i} dx_{j} dy_{j}, \right] \quad (A6)$$

$N_k$ is the number of specular reflections experienced by phonons reflecting off the element $dA_j$ first, and then all other surfaces until a reference plane (e.g., $x = 0$ in Fig. 3(a)) is reached. $N_k$ can be determined from the geometry of the
specular scattering as described in Fig. 3(c), and is the sum of
scattering events from each surface. For example, if the
radiating element is on the top surface, $N_k$ can be found for
phonons reflecting off the top and bottom surfaces, and
reflecting off the sidewalls only. In this way, a different dif-
fuse fraction can be assigned to each surface. It is also useful
to note that the linear dependence of the equivalent view fac-
tors on $f_m$ can be translated to a linear dependence of the
temperature gradient on the diffuse fraction when Eqs. (A5)
and (A6) are finally substituted into Eq. (3).

The integral equations of the equivalent view factors
over the surfaces of the beam are

$$\int \mathcal{F}_{\parallel}(f_m; x_t) \, dx_t = dx_i \int \int \mathcal{F}_{\parallel}(f_m; x_i, y_i, z_i) \, dy_i \, dx_j \, dz_j, \quad \text{(A7)}$$

$$\int \mathcal{F}_{\perp}(f_m; x_t) \, dx_t = dx_i \int \int \mathcal{F}_{\perp}(f_m; x_i, y_i, z_i) \, dy_i \, dx_j \, dz_j, \quad \text{(A8)}$$

where the limits of the integrals depend on the surface $m$
from which the phonons were originally radiated.

Since a temperature gradient is only present in the prin-
cipal direction of heat flow ($x$-axis), all surface elements at
the plane of constant $x$ are at the same temperature and radi-
ate the same flux. Hence, the equivalent view factor for the
direct (non-scattered) heat flux $F_{\parallel}(f_m; x_i)$ from the surface to
the reference plane can be calculated analytically from Eq.
(A1) by considering the view factor of a infinitesimal strip of
width $w$ to the reference plane at $x = 0$.

3V. Yefremenko, G. Wang, V. Novosad, A. Datesman, J. E. Pearson, R.
Hubmayr, K. Irwin, J. Klein et al., IEEE Trans. Appl. Supercond. 23,
2500704 (2013).
5V. Yefremenko, P. Ade, K. Aird, J. Austermann, J. Beall, D. Becker, B.
Supercond. 23, 2100605 (2013).

K. Irwin and G. Hilton, Cryogenic Particle Detection, Topics in Applied
pp. 63–150.
2K. Rostem, C. L. Bennett, D. T. Chuss, N. Costen, E. Crowe, K. L.
3J. R. Eimer, C. L. Bennett, D. T. Chuss, T. Marriage, E. J. Wollack, and L.
4C. Kittel, Introduction to Solid State Physics (John Wiley & Sons,
2005).
5P. Khosropanah, R. Hijmering, M. Ridder, J. R. Gao, D. Morozov, P. D.
6P. Beckmann and A. Spizzichino, The Scattering of Electromagnetic
9RoJ thermometer model RX-202A-CD-0.05B, LakeShore Cryotronics,
Westerville, OH 43082.
10C. P. Lusher, J. Li, V. A. Maidanov, M. E. Digby, H. Dyball, A. Casey, J.
12, 1 (2001).
12F. Giazzotto, T. T. Heikkinen, A. Luukanen, A. M. Savin, and J. P. Pekola,
14A. Vinante, P. Falferi, R. Mezzena, and M. Muck, Phys. Rev. B 75,
104303 (2007).
15H. Casimir, Physica 5, 495 (1938).
(1953).
38, 7576 (1988).
167, 141 (2012).
22F. Laermer, P. Schilp, and R. Bosch Gmbh, “Method of anisotropically
23J. Sen, T. A. Puurtinen, T. J. Isotalo, S. Chaudhuri, and I. J. Maasilta,