INSTITUTE OF PHYSICS PUBLISHING SUPERCONDUCTOR SCIENCE AND TECHNOLOGY

Supercond. Sci. Technol. **15** (2002) R1–R16 PII: S0953-2048(02)08326-4

TOPICAL REVIEW

Hot-electron effect in superconductors and its applications for radiation sensors

Alexei D Semenov[1,2], **Gregory N Gol'tsman**[2,3] **and Roman Sobolewski**[3,4]

[1] DLR Institute of Space Sensor Technology and Planetary Exploration, 12489 Berlin, Germany
[2] Physics Department, State Pedagogical University of Moscow, 119891 Moscow, Russia
[3] Department of Electrical and Computer Engineering and Laboratory for Laser Energetics, University of Rochester, Rochester, NY 14627-0231, USA
[4] Institute of Physics, Polish Academy of Sciences, PL-02904 Warszawa, Poland

E-mail: Alexei.Semenov@dlr.de

Received 20 June 2001, in final form 4 January 2002
Published 11 March 2002
Online at stacks.iop.org/SUST/15/R1

Abstract
The paper reviews the main aspects of nonequilibrium hot-electron phenomena in superconductors and various theoretical models developed to describe the hot-electron effect. We discuss implementation of the hot-electron avalanche mechanism in superconducting radiation sensors and present the most successful practical devices, such as terahertz mixers and direct intensity detectors, for far-infrared radiation. Our presentation also includes the novel approach to hot-electron quantum detection implemented in superconducting x-ray to optical photon counters.

1. Historical prelude and introduction

The term 'hot electrons' was originally introduced to describe nonequilibrium electrons (or holes) in semiconductors (for a review see, e.g. [1]). The term encompasses electron distributions that could be formally described by the Fermi function but with an effective elevated temperature. The concept is very fruitful for semiconductors, where the mobility of carriers can be shown to depend on their effective temperature. In metals, however, electrons do not exhibit any pronounced variation of mobility with their energy. As a result, heating of electrons in a metal does not affect the resistance [2], unless the change in the effective temperature is comparable with the Fermi temperature.

Schklovski [3] was the first to discuss the idea of combining steady-state electron heating with the strong dependence of the resistance on the effective electron temperature in a metal film undergoing the superconducting transition. In the steady-state regime, however, electron heating is always masked by the conventional bolometric effect; therefore, experimental results on the heating of electrons by the dc current were not very convincing. The regime of dynamic electron heating by external radiation was studied in a series of experimental and theoretical papers [4–6]. It was immediately realized that the very short relaxation time of electron excitations would make it feasible to design extremely fast radiation sensors with sensitivity much better than that of conventional bolometers.

During the last decade, a new generation of hot-electron superconducting sensors have been developed. These include submillimetre and terahertz mixers, direct detectors and photon counters for the broad spectral range from microwaves to optical radiation and x-rays. Activity in the field of hot-electron superconducting sensors is growing rapidly. These sensors have already demonstrated a performance that makes them devices-of-choice for many far-infrared, infrared and optical wavelength applications, such as plasma diagnostics, laser studies, ground-based and airborne heterodyne astronomy, and single-photon detection and quantum communications. Parallel development of compact cryocoolers and terahertz radiation sources opens hot-electron sensors for satellite astronomy and communication applications. This paper reviews the physical background of the hot-electron phenomenon in superconducting films

0953-2048/02/040001+16$30.00

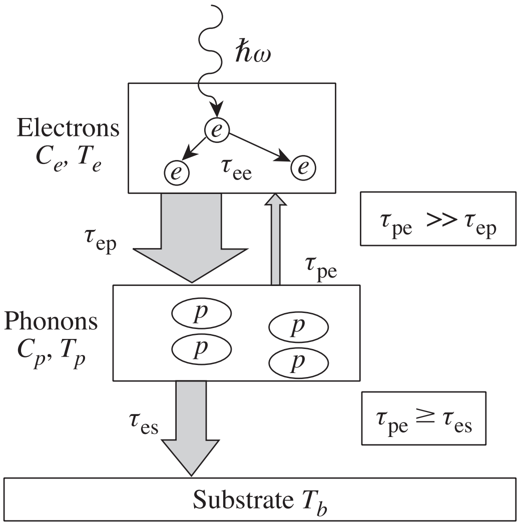

Figure 1. Thermalization scheme showing subsequent channels of the energy transfer in a hot-electron device that relaxes towards global equilibrium.

and discusses various technical realizations of hot-electron radiation sensors.

2. Physics of hot electrons

Thermal dynamics in a superconducting film on a dielectric substrate can be described in terms of four coexisting subsystems: Cooper pairs, quasiparticles (electrons from broken Cooper pairs), phonons in the film and phonons in the substrate. Thermal equilibrium exists when all of these can be described by equilibrium distribution functions with the same temperature. If any distribution does not satisfy these conditions, the situation is considered nonequilibrium. General treatment of a nonequilibrium state requires a solution of the integral kinetic equations for space- and time-dependent distribution functions. To avoid the above complexity, various simplifying assumptions are used to reduce the general problem to analytically solvable rate equations.

2.1. Hot-electron cooling and diffusion

The hot-electron model is most relevant for nonequilibrium superconductors maintained at temperature T near the superconducting transition temperature T_c, where quasiparticles and phonons can be described by thermal, normal-state distribution functions, each with its own effective temperature. The electron and phonon effective temperatures (T_e and T_p) are assumed to be established instantly and uniformly throughout the whole specimen. This assumption implies that a rapid thermalization mechanism exists inside each subsystem.

The main steps of the hot-electron phenomenon that lead to the global equilibrium are depicted in figure 1. Introducing characteristic times of the energy exchange between subsystems reduces the problem of the global equilibrium recovery to a pair of coupled heat-balance equations for T_e and T_p. The intrinsic thermalization time τ_T should be short compared to energy exchange times. This two-temperature (2-T) approach was used for the first time by Kaganov *et al* [2] to describe steady-state electron heating in metals. Below T_c, the electron specific heat exhibits an exponential temperature dependence that makes equations nonlinear for even small deviations from equilibrium. The description can, however, be simplified in the vicinity of T_c. At this temperature, the superconducting energy gap is strongly suppressed, concentration of Cooper pairs is very small and unpaired electrons exhibit no significant superconducting peculiarities: they are regarded as normal electrons having the ordinary Fermi distribution function. In the normal state, the specific heat of electrons has a much weaker temperature dependence, which can be neglected for small deviations of T_e from equilibrium. With these assumptions, the equations describing the hot-electron effect in superconductors become linear and can be written as

$$\frac{dT_e}{dt} = -\frac{T_e - T_p}{\tau_{ep}} + \frac{1}{C_e} W(t) \qquad (1a)$$

$$\frac{dT_p}{dt} = \frac{C_e}{C_p} \frac{T_e - T_p}{\tau_{ep}} - \frac{T_p - T_0}{\tau_{es}} \qquad (1b)$$

where $W(t)$ represents the external perturbation (i.e. the power per unit volume absorbed by the electron subsystem), τ_{ep} and τ_{es} are the electron energy relaxation time via electron–phonon interaction and the time of phonon escape into the substrate, respectively, C_e and C_p are the electron and phonon specific heats, respectively and T_0 is the ambient (substrate) temperature. To derive the 2-T equations we used the condition of the energy-flow balance in equilibrium $\tau_{pe} = \tau_{ep}(C_p/C_e)$, where τ_{pe} is the phonon–electron energy relaxation time.

The first implementation of the electron-heating model to superconductors was made by Schklovski [3], who used a more general, nonlinear form of the heat-balance equations to describe hysteresis of the critical current in a thin lead film. An analytical solution of equation (1) was first obtained by Perrin and Vanneste [4] for sinusoidal perturbations and by Semenov *et al* [5] for an optical pulse excitation. In the latter case, thermalization of electrons was interpreted as an increase of T_e. The increase was assumed to occur during a time interval that depended on both the duration of the optical pulse and the intrinsic thermalization time. The model was used to describe the response of superconducting NbN and $YBa_2Cu_3O_{7-\delta}$ (YBCO) films in the resistive state to near-infrared and visible radiation [5, 7]. Figures 2 and 3 show a good agreement between experimental signals and the theoretical simulation.

Figure 4 presents the detailed thermalization diagrams for both YBCO (figure 4(*a*)) and NbN (figure 4(*b*)) thin films excited by 100 fs wide optical pulses. The diagrams depict the process in the same manner as in figure 1 but now include the actual values of the characteristic time constants for both materials. The values were obtained from the 2-T model via the fit of equations (1) to the experimental photoresponse data. The measurements were performed using the electro-optic sampling system, which allowed us to obtain the intrinsic, time-resolved dynamics of the electron thermalization process in 3.5 nm thick NbN [8] and 100 nm thick YBCO films [9]. We note that, in general, the dynamic of the YBCO thermalization is roughly one order of magnitude faster than that of NbN. In both cases, the energy flow from electrons to phonons dominates the energy backflow due to reabsorption

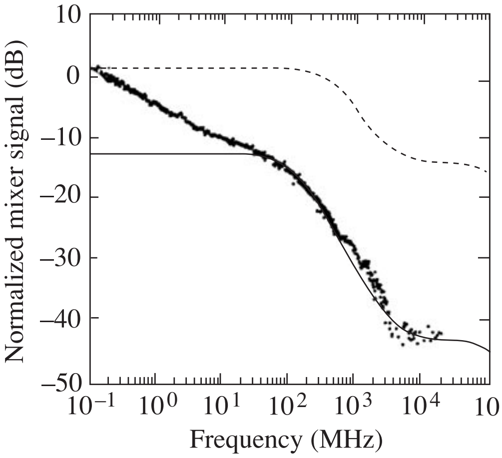

Figure 2. Response of a YBCO HEP to optical radiation (dots) versus modulation frequency [7]. The solid line was calculated using equations (1). The discrepancy at low frequencies is due to phonon diffusion in the substrate that was not accounted for in the model. (Reproduced by permission).

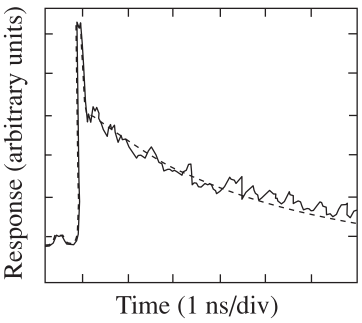

Figure 3. Response of a YBCO HEP to a femtosecond infrared pulse: experimental data (solid line) [5] and simulations (dashed line) based on the 2-T model. (Reproduced by permission).

of nonequilibrium phonons by electrons; however, while the energy backflow in YBCO can be neglected because of the very large ratio $C_p/C_e = 38$, in NbN it constitutes a non-negligible 15% ($C_p/C_e = 6.5$) of direct electron–phonon energy relaxation. Consequently, in a YBCO film excited on the femtosecond timescale, the nonthermal (hot-electron) and thermal, bolometric (phonon) processes are practically decoupled, with the former totally dominating the early stages of electron relaxation. On the other hand, the response of NbN devices is determined by the 'average' electron-cooling time τ_e, which is given by $\tau_e = \tau_{ep} + (1 + C_e/C_p)\,\tau_{es}$ [4, 5] and corresponds to the time that elapses from the peak response until the magnitude of the response declines to $1/e$ of the maximum value. If the external perturbation is substantially longer than τ_{pe} (that is, >100 ps for YBCO films), the YBCO response is dominated by the bolometric process, as was shown by the bulk of the early photoresponse measurements [10]. The very large difference in the τ_{es} values for YBCO and NbN is mainly due to the drastic difference in thickness of the tested films. Additionally, NbN films are, in general, better acoustically matched to the substrate. This significantly reduces τ_{es}.

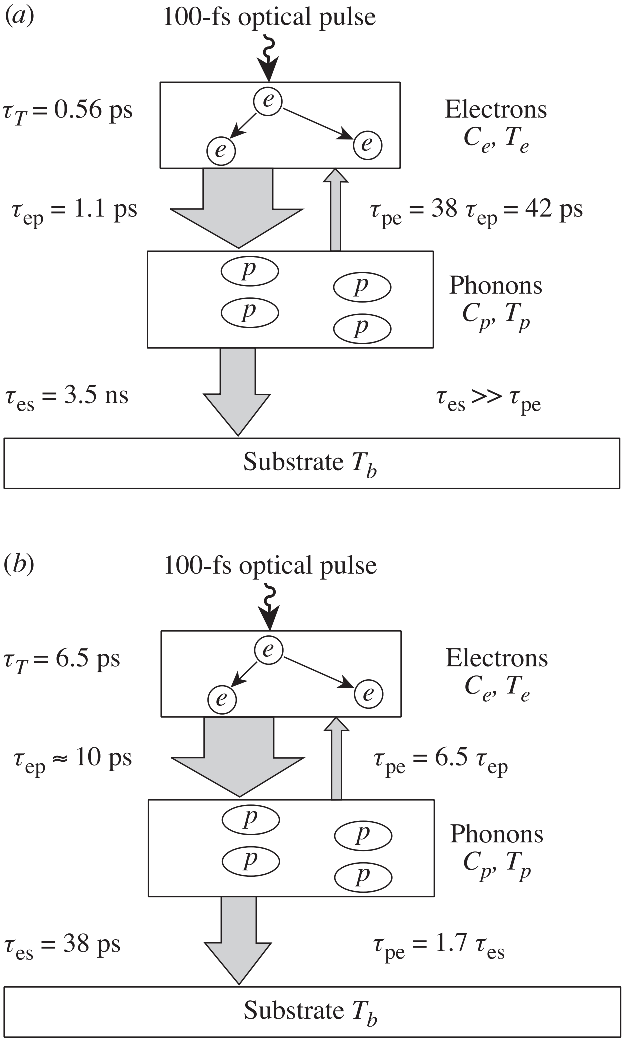

Figure 4. Hot-electron relaxation diagrams and characteristic time constants for (*a*) thin-film YBCO [9] and (*b*) ultrathin NbN film [8]. (Reproduced by permission).

Electron heating in the limiting case of a very short phonon escape time, $\tau_{es} \ll \tau_{ep}$, τ_{pe}, was first studied by Gershenzon *et al* [6] for Nb films. Although for this material [11] $C_p/C_e \approx 0.25$ and, consequently, $\tau_{ep} > \tau_{pe}$, the effective escape of phonons to the substrate prevents energy backflow to electrons. As a result, τ_{ep} alone controls the response of ultrathin (<10 nm) Nb films. Typical electron relaxation time in Nb is ≈ 1 ns at 4.2 K, which is over an order of magnitude larger than in NbN.

The 2-T model represented by equations (1) is essentially the small-signal model. Deviations of the effective temperatures from the equilibrium due to both the joule power dissipated by the bias current and the absorbed radiation power are assumed to be small compared to their equilibrium values. The theory of operation of a hot-electron photodetector (HEP) was developed on the basis of this model by Gershenzon *et al* [12], and a novel hot-electron mixer (HEM) was proposed [13].

The 2-T approach neglects, however, diffusion of electrons and assumes that the effective temperatures remain uniform within the whole device. A different approach was proposed by Prober [14], who considered diffusion of hot electrons out of the active area, rather than the energy transfer to phonons, as the main mechanism of the electron cooling.

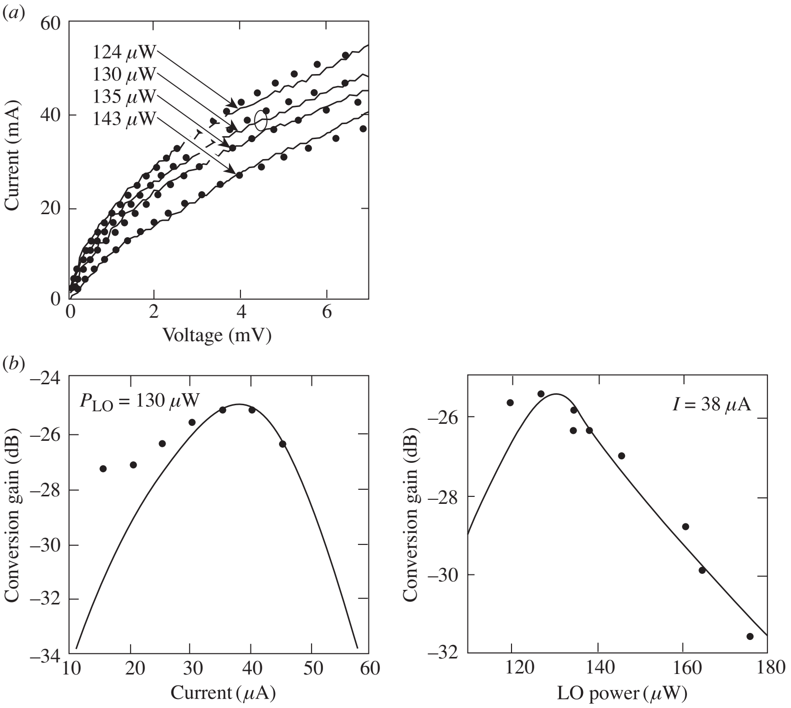

Figure 5. (*a*) Current–voltage characteristics for different LO power values. (*b*) Conversion gain curves for an NbN HEM compared with results of the uniform model (solid lines) [15]. (Reproduced by permission).

If the device length L is short compared to the thermal diffusion length $L_{\mathrm{th}} = (D\tau_{\mathrm{e}})^{1/2}$, where τ_{e} is the electron-cooling time and D is the electron diffusivity, relaxation of T_{e} is controlled by the electron out-diffusion time $\tau_{\mathrm{d}} = (L^2/\pi^2 D)$. In the limiting case $L \ll L_{\mathrm{th}}$, T_{e} remains almost uniform throughout the device length. The device can then be described by equation (1*a*), in which τ_{ep} and T_{p} should be substituted for τ_{d} and T_0, respectively. For longer devices, both the actual distribution of T_{e} along the device length and the phonon contribution to the electron relaxation should be taken into account.

2.2. *Large-signal models*

The common disadvantage of the small-signal model described above is that the optimal values of the bias current (for HEPs) and power of the local oscillator (for the HEM theory) are not derived in the framework of the model, but rather taken from the experiment or independently estimated. To include the bias current and the local oscillator (LO) power in a consistent manner, one should specify the structure of the resistive state and account for the dependence of the electron-cooling rate on the deviation from equilibrium. For large deviations from equilibrium, the heat-balance equations become nonlinear.

The large-signal mixer theory was developed by Nebosis *et al* [15] for the uniform resistive state (which is, of course, a very crude approximation). The authors assumed a finite value of τ_{es} and introduced the superconducting critical current. Reasonable quantitative agreement (see figure 5) was found between the experimental data for NbN mixers and the theoretical results. Karasik *et al* [16] implemented a similar approach for modelling a bolometric mixer fabricated from a high-temperature superconducting material. Floet *et al* [17] considered the nonuniform resistive state of a hot-electron bolometer in the small-signal regime for $\tau_{\mathrm{es}} = 0$, while Merkel *et al* [18] developed the numerical large-signal nonlinear model for a finite, nonzero value of τ_{es}. Both models described the resistive state of the mixer at optimal operation conditions in terms of a normal hot spot, maintained by self-heating. The hot spot occupies only a portion of the device length, thus assuring a mixer resistance between zero and the normal-state value. In this approach, the LO power is assumed to be uniformly absorbed in the mixer, whereas the joule power dissipation due to the bias current appears in the hot-spot region only. Since the diffusion of electrons is introduced in the basic equations, this model naturally covers all intermediate cases between the extreme diffusion cooling ($L \ll L_{\mathrm{th}}$) and phonon-cooling ($L \gg L_{\mathrm{th}}$) regimes. Neglecting phonons ($\tau_{\mathrm{es}} = 0$) and simultaneously assuming $\tau_{\mathrm{T}} = 0$, one can reduce the problem to the following system of equations [17] for T_{e}:

$$
\begin{aligned}
&-K\frac{\mathrm{d}^2 T_{\mathrm{e}}}{\mathrm{d}x^2} + \frac{C_{\mathrm{e}}}{\tau_{\mathrm{e}}}(T_{\mathrm{e}} - T_0) = j^2\rho_{\mathrm{n}} + P_{\mathrm{RF}} \quad &\text{(inside hot spot)}\\
&-K\frac{\mathrm{d}^2 T_{\mathrm{e}}}{\mathrm{d}x^2} + \frac{C_{\mathrm{e}}}{\tau_{\mathrm{e}}}(T_{\mathrm{e}} - T_0) = P_{\mathrm{RF}} \quad &\text{(outside hot spot)}
\end{aligned}
\tag{2}
$$

where K is the thermal conductivity, j is the bias current density, ρ_{n} is the resistivity of the mixer in the normal state and P_{RF} is the LO power absorbed per unit volume. This description allows for an analytical solution, which returns the bias current as a function of the hot-spot length and, thus, a voltage drop across the device. Results of simulations [17] are in good agreement with the experimental current–voltage (I–V) characteristics, especially for large P_{RF} values, which drive the mixer almost into the normal state. Surprisingly,

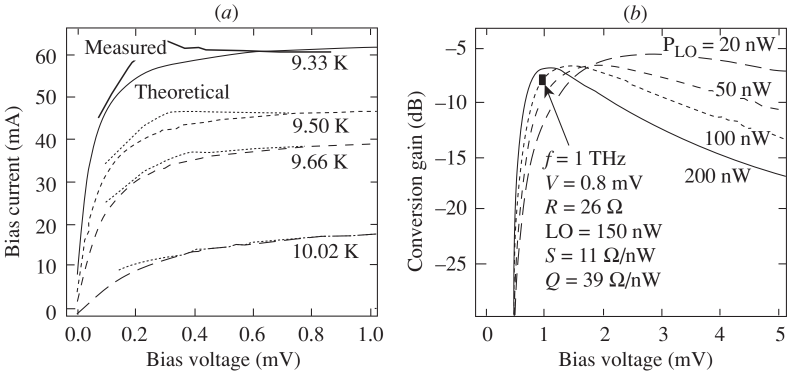

Figure 6. (*a*) Current–voltage characteristics and (*b*) conversion gain of an NbN HEM simulated in the framework of the hot-spot model [18]. Experimental characteristics are shown for comparison. (Reprinted, with permission from Merkel *et al*, *IEEE Trans. Appl. Supercond.* **9** 4201©1999 IEEE).

results based not only on equation (2), but also on the more accurate numerical model [18] shown in figure 6, do not differ much from simulations based on the uniform 2-T model (figure 5). With the appropriate set of fitting parameters, both approaches describe the *I–V* characteristics of the HEM fairly well and predict reasonable values of the conversion efficiency and noise temperature.

A nonthermal regime of the diffusion-cooled HEM was described by Semenov and Gol'tsman [19]. The authors considered a small device made from a clean material, in which τ_T is larger than τ_d. The device was operated in the nonthermal regime and had the advantage of a short response time (or, equivalently, a large bandwidth) in the heterodyne mode. On the other hand, incomplete thermalization hampered the responsivity and increased the relative contribution of the Johnson noise to the total electric noise of the device. Compared to HEMs operated in the thermal regime, the nonthermal mixer required more power from LO. At low temperatures, however, the nonthermal regime of operation provided almost quantum-limited sensitivity.

The electric noise of a hot-electron sensor comprises of the same components as the noise of any conventional bolometer: shot noise, Johnson noise, thermal noise and flicker noise. To our knowledge, there is no consistent theory for flicker noise, so its contribution may be determined only experimentally. Unless the sensitivity of the bolometer reaches the quantum limit, the noise due to fluctuations in the background radiation can be neglected. The typical length of hot-electron devices studied so far has been much larger than the diffusion length associated with the electron–electron scattering. In this limiting case, the superposition of the Johnson noise and shot noise reduces to the Nyquist form, i.e. the spectral density of the voltage noise is $S_V = 4k_BTR$, where R is the resistance of the device. This noise has a 'white' spectrum up to very high frequencies. The corresponding contribution to the system-noise temperature in the heterodyne regime increases rapidly when the conversion efficiency rolls off at intermediate frequencies (IFs) larger than $1/\tau_e$.

Thermal noise contributes to the total spectral density the amount $4k_BT^2I^2\tau_e(\partial R/\partial T_e)^2/(C_ev)$, where I is the bias current and v is the volume of the sensor. Since the conversion efficiency is proportional to $I^{-2}\tau_e^2P_{RF}(\partial R/\partial T_e)^2/(C_ev)^2$ and has the same roll-off frequency, the noise temperature of the mixer due to thermal fluctuations is given by $T_N \approx T_e^2C_ev/(\alpha\tau_eP_{RF})$, where α is the optical coupling efficiency. The contribution to the noise temperature due to thermal fluctuations does not depend on IF, neither does the corresponding noise-equivalent power (NEP) in the direct-detection mode, NEP $\approx (T_e/\alpha)\,(k_BC_ev/\tau_e)^{1/2}$. In contrast, the contribution due to the Nyquist term increases rapidly at IFs larger than $1/\tau_e$ and usually limits the IF noise bandwidth of the mixer.

Though the above simple treatment of the bolometer noise explains the main features, it does not provide an appropriate tool for computations. To obtain exact results, one should take into account the positive feedback via the load resistor and self-heating by the bias current. The former enhances the system output noise because the bolometer rectifies part of its own noise voltage drop across intrinsic resistance. The latter effect typically increases the IF bandwidth in the heterodyne regime and decreases the response time in the direct-detection mode. It is of little practical use, however, because operation near the thermal roll-off requires very precise stabilization of the ambient temperature. For a HEM with dc resistance R at the operation point and connected to the IF load with impedance R_L, the dependence of the conversion efficiency $\eta(\omega)$ and single-sideband noise temperature $T_{SSB}(\omega)$ on the IF was derived in the framework of the uniform model [15]

$$\eta(\omega) = \frac{2\alpha}{I^2}\frac{R_L}{(R_L + R_\infty)^2}\frac{C^2P_{RF}}{\left(C\frac{R-R_L}{R_L+R_\infty} + \xi\right)^2 + \varphi^2} \tag{3}$$

$$T_{SSB}(\omega) = \frac{2\,T_e\,R_\infty I^2}{\alpha\,C^2P_{RF}}(\xi^2 + \varphi^2) + \frac{2T_e^2\,\tau_e}{C_eV\alpha\,P_{RF}} \tag{4}$$

where

$$C = I^2\tau_e\frac{\partial R/\partial T_e}{C_eV}$$

$$\xi(\omega) = \frac{1 + \omega^2\,(\tau_1\tau_3 + \tau_2\tau_3 - \tau_1\tau_2)}{1 + (\omega\,\tau_3)^2}$$

$$\varphi(\omega) = \frac{\omega\,(\tau_1 + \tau_2 - \tau_3) + \omega^3\tau_1\tau_2\tau_3}{1 + (\omega\,\tau_3)^2}$$

$$\tau_{1,2}^{-1} = \frac{1}{2}\left(\frac{1}{\tau_{ep}} + \frac{C_e}{C_p\tau_{ep}} + \frac{1}{\tau_{es}}\right)\times\left[1 \pm \sqrt{\frac{\frac{4}{\tau_{ep}\tau_{es}}}{\left(\frac{1}{\tau_{ep}} + \frac{C_e}{C_p\tau_{ep}} + \frac{1}{\tau_{es}}\right)^2}}\right]$$

and

$$\tau_3 = \frac{\tau_{ep}\tau_{es}}{\tau_e}.$$

In the above equations, R_∞ is the impedance of the bolometer at very high IF and $\partial R/\partial T_e$ is the slope of the superconducting transition at the operation point on the scale of the electron temperature. The slope of the transition cannot be derived from first principles in the framework of the uniform model. Its temperature dependence should be calculated in a phenomenological manner (see, e.g. [15]), or the value at the specific operation regime should be concluded from the experiment. Ekström *et al* [20] showed that the magnitude of the parameter C in equations (3) and (4) can be determined from the experimental dc *I*–*V* characteristics as

$$C = \frac{\frac{dV}{dI} - R}{\frac{dV}{dR} + R} \tag{5}$$

where dV/dI is the differential resistance of the HEM at the operation point. The advantage of the hot-spot model [18] is that it allows for numerical computation of the superconducting transition slope for arbitrary values of the LO power, bias current and ambient temperature.

3. Cooper-pair, kinetic-inductive photoresponse

Although, as mentioned in section 2.1, in the discussion of equations (1), the response of a superconductor that is kept well below T_c to external radiation cannot be adequately treated in the framework of the hot-electron approximation, we decided to include superconducting detectors operating at $T \ll T_c$ in our review. Rothwarf and Taylor [21] were the first to successfully develop the phenomenological description for nonequilibrium Cooper-pair recombination and breaking processes (the so-called RT model). At low temperatures, when energies of nonequilibrium quasiparticles after thermalization are spread over a narrow interval above the superconducting energy gap 2Δ, the appropriate parameters to characterize this nonequilibrium state are the number, Δn_q, of excess quasiparticles and the number, Δn_p, of excess, the so-called, 2Δ phonons. The 2Δ phonons are emitted in the Cooper-pair recombination process and, since they have the energy of at least 2Δ, they are responsible for secondary breaking of Cooper pairs. For small perturbations, concentrations of Δn_q and Δn_p are given by the following linearized RT rate equations:

$$\frac{d}{dt}\Delta n_q = -\frac{\Delta n_q}{\tau_R} + \frac{2\Delta n_p}{\tau_B} \tag{6a}$$

$$\frac{d}{dt}\Delta n_p = -\frac{\Delta n_p}{\tau_B} - \frac{\Delta n_p}{\tau_{es}} + \frac{\Delta n_q}{2\tau_R} \tag{6b}$$

where τ_R and τ_B are the quasiparticle recombination time and the time of breaking Cooper pairs by 2Δ phonons, respectively. We note that equations (6) are mathematically analogous to the 2-T model (equations (1)). Like the 2-T model, the RT approach assumes that there is a quick, intrinsic thermalization mechanism inside both the quasiparticle and phonon subsystems.

When photons with energy typically much larger than 2Δ are absorbed by a superconducting film maintained at $T \ll T_c$, they produce a time-dependent population $\Delta n_q(t)$ of nonequilibrium quasiparticles, leading to a temporary decrease in the superconducting fraction of electrons, $f_{sc} = (n_0 - n_q)/n_0$, where $n_q = n_q(0) + \Delta n_q(t)$ is the instant concentration of the quasiparticles, $n_q(0)$ is their equilibrium concentration and n_0 is the total concentration of electrons. Because the pairs are characterized by nonzero inertia, this process can be modelled as time-varying kinetic inductance [22, 23]:

$$L_{kin}(t) = \frac{L_{kin}(0)}{f_{sc}} \tag{7}$$

where $L_{kin}(0) = \mu_0(\lambda_L)^2/d$ is the equilibrium value per unit area of the film, λ_L is the magnetic penetration depth and d is the film thickness. The L_{kin} of a superconducting film makes it possible to monitor the concentration of Cooper pairs. In a current-biased superconducting film, after the destruction of a certain number of Cooper pairs, the remaining pairs accelerate to carry the same bias current. Because of nonzero inertia of pairs, acceleration requires an electric field. This intrinsically generated electric field is seen from the exterior as a voltage pulse developing across the film. Mathematically, this voltage transient is given by

$$V_{kin} = I\frac{dL_{kin}}{dt}. \tag{8}$$

For the limiting case of very fast thermalization, i.e. when τ_T is small compared to both τ_R and τ_B, the kinetic-inductive response was described by Semenov *et al* [24] as the product of the analytical solution of equations (6) and a fitting factor exponentially growing in time. The latter parameter corresponds to the multiplication cascade of quasiparticles during thermalization. This approach describes well the experimental results obtained with pulsed and modulated cw excitations for NbN films (figure 7 and [24]). Figure 8 presents the V_{kin} transient, recorded for a YBCO microbridge at $T = 60$ K, biased in the superconducting state and excited by 100 fs optical pulses. The measurements were performed using the electro-optic sampling system, [9, 27] while numerical fits were based on equations (7) and (8), with $f_{sc}(t)$ obtained using either equations (1) or (6). The positive part of the transient represents the Cooper-pair breaking process, while the negative part represents pair recombination. We note that in YBCO, contrary to metallic superconductors, both processes are of the same, 1 ps, duration and the RT model fits better experimental data.

4. Single-photon-detection mechanisms

So far, this discussion has been limited to integrating detectors in which the energy of a large number of absorbed photons is distributed among an even larger number of elementary thermal excitations in the detector. That is, individual photons cannot be distinguished, and only the average radiation power absorbed by the detector is measured. In the particular case of a thermal detector, e.g. a bolometer or a hot-electron detector near T_c, this average absorbed radiation power corresponds to enhanced effective temperatures of phonons and electrons, respectively. In a quantum (photon) detector, a single photon creates excitations that are collected and counted before they relax and before another photon is absorbed. Thus, the detector

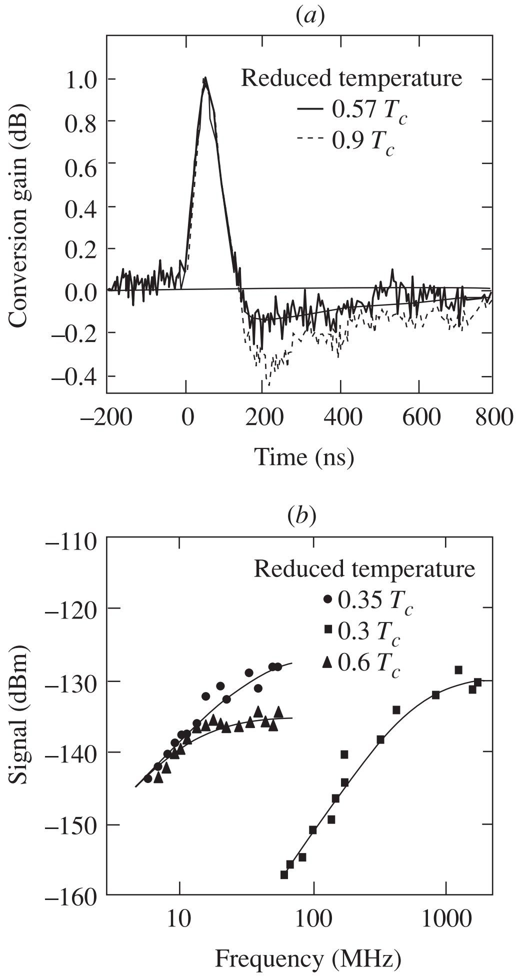

Figure 7. (*a*) Conversion gain and (*b*) signal response of an NbN HEP to pulsed and modulated cw optical radiation in comparison with model simulations based on equations (6) [24]. (Reprinted, with permission from Semenov *et al IEEE Trans. Appl. Supercond.* **7** 3083 ©1997 IEEE).

separately registers each absorbed photon, while the number of collected excitations measures the energy of absorbed photons.

The hot-electron quantum detector was first proposed by Kadin and Johnson [26]. In this model, a photon absorbed somewhere in the resistive film initiates a growing hot spot. The resistance inside the hot spot is larger than that in the surrounding area. Even if the size of the hot spot is much smaller than that of the film, the voltage drop across the current-biased film ‘feels’ the presence of the hot spot. The disadvantage of this approach for practical devices stems from the fact that the film has to be operated near its T_c and can withstand only a very small current density without being driven into the normal state. Since the detector response is proportional to the bias current, the small operating current requires a complicated, SQUID-based read-out scheme [27].

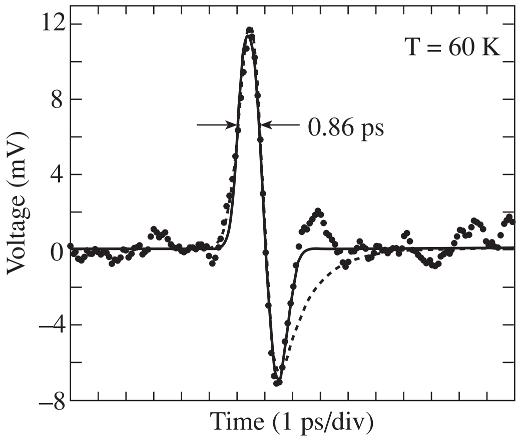

Figure 8. Kinetic-inductive response of a YBCO HEP to 100 fs wide optical pulses, operated at low temperatures in the superconducting state. Dots correspond to the experimental data, while simulated transients were obtained using the RT model (equations (6)) (solid line) and the 2-T model (equations (1)) (dashed line) [9]. (Reproduced by permission).

4.1. Supercurrent enhanced hot-spot optical single-photon detector

Semenov *et al* [28] proposed a different quantum detection regime in a superconducting stripe that is operated well below T_c and carries a bias current only slightly smaller than the critical value at the operating temperature. Generation of a hot spot at the position where the photon has been absorbed creates a local region with suppressed superconductivity. The supercurrent is forced to flow around this normal (resistive) spot, through the parts of the film that remain superconducting. If the diameter of the resistive spot is such that the current density in the superconducting portion of the film reaches the critical value, a resistive barrier is formed across the entire width of the stripe, giving rise to a voltage pulse with a magnitude proportional to the bias current.

The physical difference of the quantum detection proposed in [28], as compared to [26], is that the resistive state and, thus, the response appear to be caused by the collaborative effect of the bias current and the hot-spot formation. In the hot spot, the nonequilibrium quasiparticle concentration increases due to hot-electron thermalization (multiplication) and decreases due to electron out-diffusion. The normal spot at the absorption site occurs when the concentration of nonequilibrium electrons exceeds the critical value corresponding to the local normal state. If the film thickness d is small compared to L_{th}, the concentration of nonequilibrium-thermalized quasiparticles is given by

$$\frac{\partial}{\partial t}\Delta n_q = D\,\nabla^2\Delta n_q + \frac{\Delta n_q}{\tau_e} + \frac{d}{dt}M(t) \qquad (9)$$

where $M(t)$ is the multiplication factor and D is the normal-state electron diffusivity. The maximum value that $M(t)$ reaches during the avalanche multiplication process is called quantum yield or quantum gain; it is proportional to the energy of the absorbed quantum. Under the assumption that the $M(t)$ rate is much larger than the $1/\tau_e$ rate and that the photon is absorbed at $t = 0$ and $r = 0$, the solution for the time-dependent quasiparticle concentration profile takes the form

$$\Delta n_q(r,t) = \frac{M(t)}{4\pi Dd}\frac{1}{t}\,e^{-t/\tau_e}e^{-r^2/4Dt}. \qquad (10)$$

The diameter of the normal spot is determined from the condition $n_q(0) + \Delta n_q(r,t) > n_0$. The maximum diameter of the normal spot increases with the quantum energy. The model [28] predicts an almost-Gaussian response pulse with

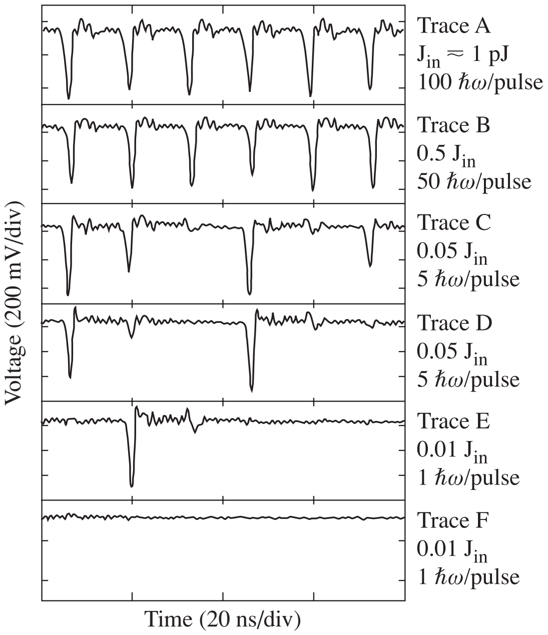

Figure 9. Response of an NbN quantum detector to trains of 100 fs optical pulses with a different number of photons per pulse (see text for details).

a magnitude that, up to a certain extent, does not depend on the photon energy. On the other hand, the pulse duration is a function of the maximum spot size, providing the basis for spectral sensitivity of the device. Finally, the single-quantum detection regime should have a cut-off wavelength that depends on operating conditions (bias current and temperature) and the detector size. Since such a detector counts individual photons, it should have ultimate background-limited sensitivity through the whole range of operation conditions.

Gol'tsman *et al* [29] experimentally demonstrated the supercurrent-assisted, hot-spot-detection mechanism for single optical (790 μm wavelength) photons. Figure 9 shows a collection of 'snapshots' recorded by a 1 GHz bandwidth oscilloscope for different energies per laser pulse, incident on the NbN quantum HEP. Each snapshot presents an 80 ns long record of the response to six successive 100 fs wide pulses and was randomly selected out of a real-time detector output data stream. Trace A in figure 9 corresponds to an average of 100 photons per pulse hitting the detector. In this case, the HEP responded to each optical pulse in the laser train. The same 100%-efficient response was observed (trace B) when there were approximately 50 photons per pulse. As the incident laser intensity was further decreased (with other experimental conditions unchanged), the quantum nature of the detector response emerged. Instead of the linear decrease of the signal amplitude with incident light intensity, which is characteristic of the classical detector, the response amplitude of the single-photon HEP remained nominally the same. In addition, some of the response pulses were missing because of the limited quantum efficiency of the device as well as fluctuations in the number of photons incident on the detector. The quantum voltage response of the HEP is most apparent in the bottom two pairs of traces: C and D (five photons/pulse) and E and F (one photon/pulse). Each pair corresponds to two different randomly selected records obtained under exactly the same experimental conditions. Note that in each case the detector response is very different.

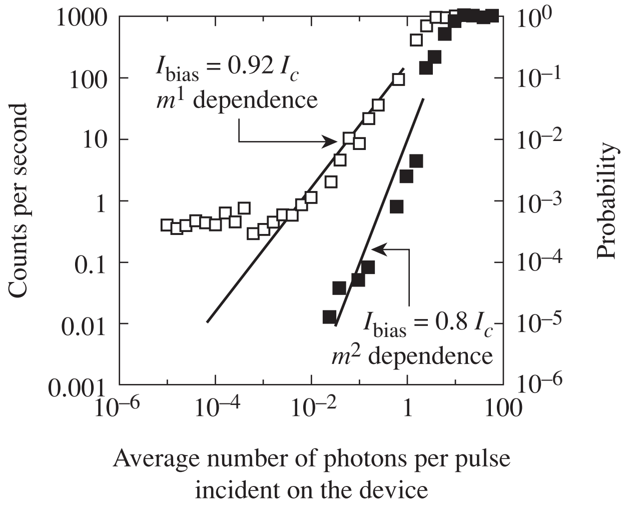

Figure 10. Count rates and the corresponding counting probability for an NbN quantum detector as a function of the radiation intensity. Depending on bias current, the detector can count single-photon (open squares) or two-photon (solid squares) events [29]. (Reproduced by permission).

Averaging over a long observation time, however, showed that both the average number of captured pulses and their magnitude remained constant if the pulse energy was unchanged. This unambiguously demonstrated the single-photon operation of the device.

For a mean number of photons per pulse (m), the probability $P(n)$ of absorbing n photons from a given pulse is proportional to

$$P(n) \sim \frac{e^{-m}(m)^n}{n!}. \tag{11}$$

When the mean number of photons $m \ll 1$ (achieved, e.g. by attenuating the radiation fluence to reduce the total number of photons incident on the detector to an average of much less than one photon per pulse)

$$P(n) \sim \frac{m^n}{n!}. \tag{12}$$

Consequently, for very weak photon fluxes, the probability of detecting one photon, two photons, three photons, etc is

$$P(1) \sim m,\ P(2) \sim \frac{m^2}{2},\ P(3) \sim \frac{m^3}{6}, \text{etc.} \tag{13}$$

Figure 10 plots the probability of the detector producing an output voltage pulse as a function of the number of photons per pulse incident on the device area, for two different values of the bias current. The left vertical axis indicates the mean number of detector counts per second. The right vertical axis corresponds to the probability of detecting an optical pulse. Open squares correspond to the bias current $0.92I_c$, where I_c is the critical current at the operation temperature. Saturation occurs at high incident photon fluxes. For smaller fluxes, as predicted by equation (12), the experimental data show the linear decrease of detection probability with the average number of incident photons over four orders of magnitude, clearly demonstrating the single-photon detection. At very low photon doses, experimental data points saturate at the level of 0.4 s^{-1} counts (probability 4×10^{-4}) since the experiment was performed in an optically unshielded environment. This level is regarded as the laboratory photon background.

The solid squares in figure 10 correspond to the same device, operated under the same conditions as those for the open-square data, but biased with $0.8I_c$. Experimental data points now follow a quadratic dependence of detection probability (see equation (13)), showing the two-photon detection. As expected for a two-photon process, the quantum efficiency is significantly lower than that for the single-photon detection. At the same time, the photon background is no longer observed since the probability of two uncorrelated, stray photons hitting the device within its response duration is negligibly small.

4.2. X-ray superconducting single-photon detectors

A nonequilibrium model of a single quantum x-ray detector with the read-out via the superconducting tunnelling junction was developed by Twerenbold [30]. Typically, a tunnel-junction detector consists of a relatively thick absorber film with an underlying thinner trapping layer, which forms one junction electrode. A photon captured in the absorber generates a high-energy photoelectron that relaxes via hot-electron multiplication into the energy gap of the absorber. Nonequilibrium quasiparticles excited during the cascade diffuse to the adjacent trapping layer, which has a smaller energy gap. There, quasiparticles scatter inelastically, reaching an energy level corresponding to the trapping-layer energy gap. The latter process is called ‘trapping’ because it confines the charge to the region close to the tunnel barrier. The tunnel junction is externally biased in such a way that trapped quasiparticles can tunnel directly to the electrode characterized by the lower energy gap. The same potential barrier prevents them from returning. They can, however, break Cooper pairs in the low-gap electrode and then form new pairs with unpaired electrons in their own electrode. Thus, the process returns unpaired electrons to the initial electrode, increasing the number of tunnelling events per quasiparticle and providing intrinsic charge amplification. The time integral of the current transient gives, with no free parameters, the charge that has been transferred through the tunnel junction. This latter value is proportional to the number of quasiparticles created in the cascade and, consequently, to the x-ray quantum energy.

The theoretical energy resolution of the tunnel-junction detector is given by $2.4[h\nu\Delta(F + 1 + 1/n)]^{1/2}$, where $h\nu$ is the quantum energy, n is the number of tunnelling events per one quasiparticle and F is the Fano factor that describes the statistical fluctuations of the charge-generation process. The Twerenbold model incorporates the two-dimensional diffusion equation for Δn_q and the general nonlinear form of the RT equations.

A more general approach, including time evolution of nonequilibrium distribution functions of quasiparticles and phonons, was developed by Nussbaumer *et al* [31]. The authors solved the Chang–Scalapino equations numerically for the quasiparticle and phonon distribution functions in a spatially homogeneous situation and supplemented the solution by one-dimensional diffusion. The full theory includes the parameters that are important for the real detector, such as back tunnelling and losses of quasiparticles at the edges of the device, resulting in good agreement between the calculated transient response signals and the experimentally measured pulse shapes.

5. Hot-electron detectors

A minor, but physically very important, difference exists between a superconducting HEP and a conventional superconducting bolometer when they are operated in the transition-edge regime. In the bolometer, thermal equilibrium between electrons and phonons is established instantly, whereas in the hot-electron detector these two systems are not in equilibrium. In this review, we restrict ourselves to publications where the nonequilibrium state between the electron and phonon subsystems was clearly observed. Basically, there are two ways to decouple electrons from phonons: nonequilibrium phonons should leave the detector at a timescale that is short compared to τ_{pe}, or the intensity of external radiation should vary faster than $1/\tau_{pe}$. Depending on the superconductor and experimental arrangement, a real hot-electron detector falls somewhere between these two extremes.

5.1. Transition-edge superconducting detectors

Historically, the first HEPs were developed and studied in the early 1980s by Gershenzon *et al* [32], using ultrathin Nb films as the detector body. Niobium is characterized by relatively long τ_{pe}, typically a few hundred picoseconds at liquid helium temperature, so that $\tau_{es} < \tau_{ep}$ for films thinner than 10 nm [11]. Therefore, detectors based on thin Nb films belong to the first limiting case and their response time is approximately equal to τ_{ep}. The best performance that the Nb HEPs can achieve [33] is NEP $= 3 \times 10^{-13}$ W $Hz^{-1/2}$, detectivity $D^* = 4 \times 10^{11}$ cm $s^{1/2}$ J^{-1} and the response time of 4.5 ns. Thus, these devices are less sensitive, although much faster, than semiconductor bolometers. When the detector area was adjusted properly, Nb HEPs demonstrated a constant value of sensitivity in the range from microwaves (150 GHz) to ultraviolet (10^{15} Hz). This is actually their greatest advantage when compared to semiconductor counterparts. A Nb-based HEP was implemented to study the emission of a cyclotron *p*-germanium laser [34]. The combination of large sensitivity and short response time made it possible to identify the Landau levels responsible for lasing.

In the late 1990s, the Gershenzon group developed HEPs based on NbN superconducting films [35]. NbN has much shorter τ_{ep} and τ_{pe} than Nb; thus, even for 3 nm thick films, NbN HEPs operate in the mixed regime (i.e. τ_{ep} and τ_{es} jointly determine the response time of the detector). Detectors made from ultrathin NbN films are much faster than Nb-based devices. The intrinsic $\tau_{ep} \approx 10$ ps, while the overall response time is about 30 ps near Tc [8]. The best-demonstrated NEP $\approx 10^{-12}$ W $Hz^{-1/2}$ [36]. In spite of a rather complicated electronic band structure [37], the quantum yield in NbN reaches above 300 for near-infrared photons [38], which corresponds to one-third of the upper theoretical limit.

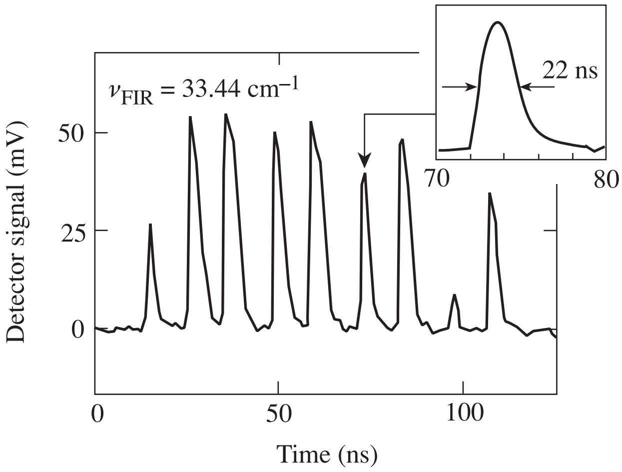

Figure 11. Pulses from a single-shot, optically pumped, far-infrared gas laser recorded with an NbN HEP [39]. The inset shows one of the pulses on an expended timescale. (Reprinted, with permission from Lang *et al Appl. Phys.* B **53** 207 ©1991 Springer-Verlag).

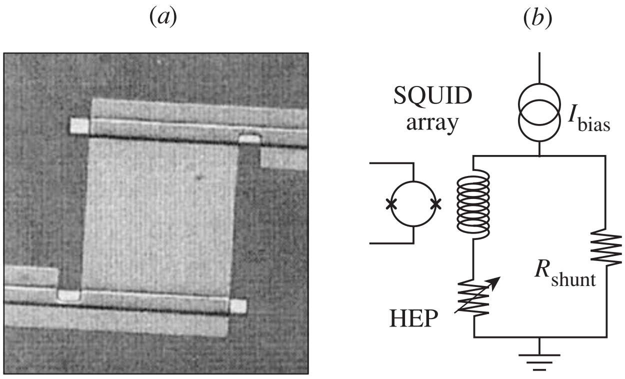

Figure 12. (*a*) Microphotograph of a transition-edge, hot-electron quantum detector and (*b*) the corresponding equivalent circuit [40]. (Reprinted, with permission from Miller *et al IEEE Trans. Appl. Supercond.* **9** 4205 ©1999 IEEE).

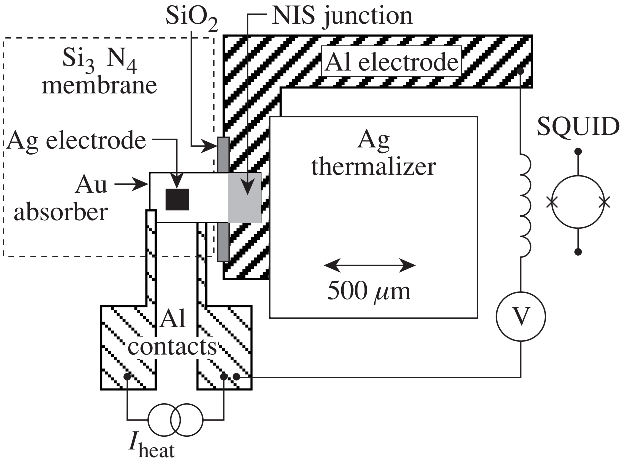

Figure 13. Detailed schematic of the hot-electron microcalorimeter developed by Nahum and Martinis [42] (see text for explanation). (Reproduced by permission).

Detectors fabricated from NbN were used to study the emission of optically pumped infrared gas lasers, in particular, pulsed lasers [39]. Figure 11 shows far-infrared laser pulses recorded with an NbN hot-electron detector. The unique combination of response time and sensitivity made it possible to detect and identify very weak emission lines.

Miller *et al* [40] have demonstrated a photon counter based on the transition-edge, hot-electron, direct detector. The device was a 20 × 20 μm^2 square of 40 nm thick tungsten film (figure 12) having $T_c = 80$ mK with a transition width of 1 mK. The device was operated at a bath temperature of 40 mK in a voltage-bias regime that maintained the sensor within the transition region via negative electrothermal feedback [41]. This mode of operation was shown [41] to increase the transition-edge sensor sensitivity and to decrease its time constant to $\tau_0/(1 + \alpha/n)$. Here τ_0 is the intrinsic time constant of the sensor, n is the power of the temperature dependence of the thermal conductance between the film and the substrate and α is the dimensionless sharpness parameter of the superconducting transition. A photon absorbed in the sensor heats the electron system above its equilibrium temperature, leading to an increase of the sensor's resistance and, consequently, to a decrease of the bias current and dissipated joule power. The integral of the drop in current (read out by an array of dc SQUIDs) gives the energy absorbed by the sensor with no free parameters. The detector described in [40] exhibited a time constant of about 60 μs and was able to register 0.3 eV (4 μm wavelength) single photons with an energy resolution of 0.15 eV. To test the detector, the authors performed an observation of the planetary nebula NGC 6572, using the 8 inch telescope. The energy resolution was somewhat lower than in the laboratory, although it was high enough to detect the strong emission lines.

A hot-electron microcalorimeter was developed by Nahum and Martinis [42]. In this type of device, photon absorption gives rise to T_e in a metal absorber and is measured using the *I–V* characteristics of a normal-insulator-superconductor tunnel junction, in which a part of the absorber forms the normal electrode. Figure 13 shows a schematic of the tested device. The current through the junction was measured with a low-noise dc SQUID. The absorber had an area of 100 × 100 μm^2 and was deposited on a silicon nitride membrane. In this configuration, the phonons that escaped from the absorber were reflected back from the membrane and were further available for the energy exchange. Thus, the Si_3N_4 membrane prevented energy loss from the electron subsystem in the absorber. The microcalorimeter was operated at 80 mK with a time constant of 15 μs and demonstrated an energy resolution of 22 eV for 6 keV photons.

In another version, Nahum and Martinis [43] proposed a microbolometer that consisted of a normal metal stripe connected to superconducting electrodes (figure 14). The device relied on Andreev reflections of low-energy, thermal quasiparticles at the edges of the stripe and on the weak electron–phonon coupling at low temperatures. Both effects confined the energy delivered by the photons, providing a large rise of T_e. This was subsequently read out by the superconductor-insulator-normal metal junction, for which the metal strip formed the normal electrode. Projected responsivity and NEP of the device with the Cu absorber operated at 100 mK were about 10^9 V W^{-1} and 3 × 10^{-18} W Hz$^{-1/2}$, respectively, which is, at least by a factor

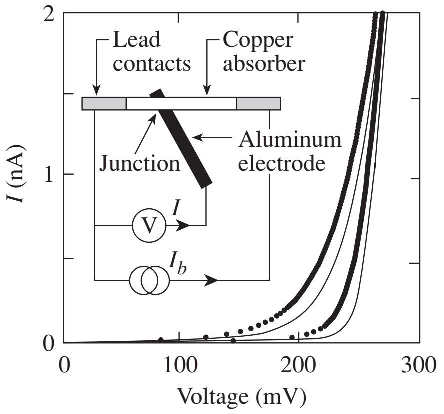

Figure 14. A hot-electron microbolometer using Andreev reflections of quasiparticles from superconducting contacts and the corresponding *I*–*V* characteristics [43]. (Reproduced by permission).

of 10, better than the performance of any currently available detectors. The time constant of the microcalorimeter is determined by the rate of energy transfer from electrons to phonons that corresponds to $\tau_{\rm ep}$ at the Fermi level. For the device under consideration in [43], the computed response time was $\tau = 20\,\mu$s. Since, for a bolometer, NEP scales as $\tau^{-1/2}$, the device performance can be further improved by increasing the response time up to a value only slightly less than that required by a specific application.

Finally, Karasik *et al* [44] proposed the use of the dependence of the electron–phonon scattering time on the electron mean free path to control the intrinsic response time of a transition-edge detector. Increase of the intrinsic response time results in the decrease of the minimum detectable power, while at the same time, the device response time can be decreased to a reasonable value by exploiting the negative electrothermal feedback. According to estimates in [44], using this approach, a detector could be fabricated with NEP $\approx 10^{-20}$ W Hz$^{-1/2}$ and a few milliseconds response time at 100 mK bath temperature.

5.2. Superconducting kinetic-inductive detectors

The detectors described in the preceding section produce a response when the device, or at least a part of it, is in the resistive state. Kinetic-inductive integrating detectors represent their superconducting counterparts. The main advantage of superconducting kinetic-inductive detectors is their low noise power. To realize this advantage, a SQUID read-out should be used. Grossman *et al* [45] described the design of a kinetic-inductive detector/mixer with an estimated NEP = 2.5×10^{-17} W Hz$^{-1/2}$ and a bandwidth of 5.5 MHz at 100 mK. Unfortunately, a laboratory prototype showed only NEP = 4.4×10^{-11} W Hz$^{-1/2}$ [46]. Sergeev and Reizer [47] performed thorough calculations for both s-wave and d-wave superconductors, including the appropriate quasiparticle distribution function and scattering times. They found NEP and D^* close to those reported in [45]. Bluzer [23] proposed a balanced-bias scheme for a kinetic-inductive photodetector with directly coupled SQUID read-out, intended to eliminate the losses inherent in inductively coupled read-outs and increase the responsivity of the detector. The performance of the detector was simulated for a 0.1 μm thick YBCO film at 9 K, resulting in NEP = 2.5×10^{-15} W Hz$^{-1/2}$ and 10 μs response time. It is believed that the use of an LTS material should result in a two- to three-orders-of-magnitude decrease in NEP.

Table 1. Experimental performance of an NbN photodetector at 790 nm.

Response time–intrinsic/measured	10 ps/100 ps
Quantum gain factor	340
A/W responsivity	220 A/W
V/W responsivity	4×10^4 V/W
Device quantum efficiency	~20%
Operating temperature	~4 K
Dark counts per second	<0.0001
Device noise temperature	~15 K

5.3. Superconducting quantum detectors

A number of novel approaches proposed during the last decade have been aimed at the realization of detectors with ultimate quantum sensitivity. Kadin and Johnson [26] introduced the quantum detection regime in ultrathin resistive films. In the proposed mechanism, an absorbed photon induces a resistive hot spot, centred at the point where the photon hits the film. If the photon flux is sufficiently low, hot spots do not overlap until they disappear. Using material parameters of NbN, the authors estimated that a 0.1 μm^2 size sample should respond to 1 eV photons with 1 mV amplitude pulses and 10 GHz bandwidth.

A photon counter using the quantum detection regime in a current-carrying superconducting film [28] was recently demonstrated by Gol'tsman *et al* [29]. The counting element consisted of a 1.3 μm long, 0.2 μm wide microbridge, formed from a 6 nm thick NbN film deposited on a sapphire substrate. The detector was operated at 4.2 K, with a bias current of approximately 90% of $I_{\rm c}$. The voltage pulses generated by the bridge in response to absorbed photons were further amplified by a cooled, low-noise amplifier. The output pulses were time-limited by electronics and had duration of approximately 100 ps (see figure 9). The intrinsic dark count rate for the detector was measured to be below 0.001 s^{-1} (probability 10^{-6}), which corresponds to zero detected responses over 1000 s when the input was completely blocked. Table 1 presents the basic parameters of the device operated at the 790 nm wavelength. Single-photon counting was observed in the photon-wavelength range from 0.4 μm to 2.4 μm [48]. We note that the device represents a unique combination of the picosecond response time and very high responsivity. These characteristics of NbN single-photon HEPs should lead to their practical implementation in areas ranging from free-space satellite communication [49], through quantum communications and quantum cryptography [50], to ultraweak luminescence observations and semiconductor integrated circuit testing [51]. Another exciting application for this type of detector can be background-limited direct detector arrays [52] for submillimetre astronomy.

The most advanced superconducting quantum detectors are tunnel-junction detectors, which are being developed for a wide range of applications from materials science and microanalysis to particle physics and astrophysics. Only a few recent publications are mentioned here because a full review of the activities in this field is beyond the scope of this paper. Nb-based tunnel-junction detectors with Al trapping layers have reached for 70 eV photons, an energy resolution of 1.9 eV. This performance is limited by the statistics of quasiparticle multiplication [53]. A typical device had an area of 50 × 50 μm^2. The smallest detectable, 0.3 eV (4.1 μm wavelength) photon energy was achieved with Ta-based devices [54] since this material has an energy gap smaller than that of Nb. An energy resolution of 0.19 eV was demonstrated for 2.5 eV (0.5 μm wavelength) photons, using Ta-based devices with an area of 20 × 20 μm^2 and 12 μs response time.

6. Hot-electron mixers

Historically, HEMs have been divided into two large categories: lattice- or phonon-cooled [13] and diffusion-cooled [14] devices. As presented earlier, the physics for these two types of HEMs is essentially the same. Both types can be described by equations (2) using temperature-dependent parameters and proper boundary conditions. The analysis becomes easier, however, when the device is designed to be close to one of the two extremes, namely, the lattice- or the diffusion-cooling regime. Typically, lattice-cooled mixers are made from ultrathin films of NbN, whereas diffusion-cooled devices use Nb or Al.

6.1. Lattice-cooled mixers

Current state-of-the-art NbN technology is capable of routinely delivering 3.0 nm thick devices that are 500 nm^2 in size with T_c above 9 K. Near T_c, τ_{pe} is close to τ_{es}, which is about 40 ps for a 3 nm thick film (see figure 4(*b*)). The τ_{ep} at 8 K is below 20 ps, which results, with the diffusivity of 0.5 cm^2 s^{-1}, in a thermal healing length of about 30 nm. Since the device length is typically much larger, the mixer operates in the phonon-cooled regime. The mixer's intrinsic IF bandwidth is determined by the combination of τ_{ep} and τ_{es} time constants. In real devices, however, the measured bandwidth depends strongly on the bias regime. This makes it difficult to compare the published data and to reach meaningful conclusions. For HEMs on Si substrates, the best reported gain and noise bandwidths are 3.5 GHz [55] and 8 GHz [56], respectively. Further increases in the bandwidth for lattice-cooled HEMs can be achieved by using a substrate material that is better thermally coupled to the superconducting film. One promising candidate is MgO. Recent measurements have shown [57] that MgO provides, for a 3.5 nm thick bolometer, a 4.8 GHz gain bandwidth and 5.6 GHz noise bandwidth. Further progress in increasing the bandwidth may be achieved by decreasing the bolometer thickness. Recently a 9 GHz gain bandwidth was reported [58] for a 2.5 nm thick device on MgO. Unfortunately, this direction is limited because NbN films thinner than 2.5 nm become inhomogeneous and lose their superconductivity [59].

A waveguide version of the receiver with the lattice-cooled NbN HEM has been installed and operated successfully in the frequency range of 0.6 to 0.8 THz [60] and 1.04 THz [61] at the 10 m Sub-mm Telescope Facility in Mount Graham in Arizona. At this telescope, the measured noise temperature of the receiver was 560 K at 0.84 THz and 1600 K at 1.035 THz over a 1 GHz IF bandwidth centred at 1.8 GHz. The receiver was used to detect the CO molecular line emission in the Orion nebula (figure 15). It is worth noting that this was the first earth-based observation at a frequency above 1 THz. A quasi-optical version of the HEM receiver for the THz range is currently under preparation for test flights on a stratospheric airplane observatory [62]. The mixer will be incorporated into a planar logarithmic spiral antenna (figure 16), which is integrated with an extended hyperhemispherical silicon lens.

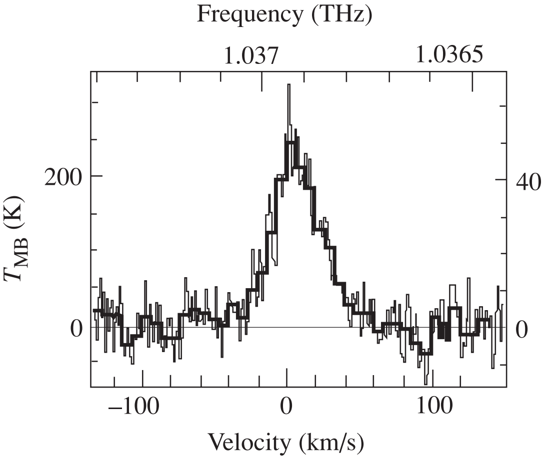

Figure 15. Terahertz CO line in the Orion IRc2 nebula recorded with an NbN hot-electron mixer at a ground-based telescope in Arizona [61]. The thick solid line shows a smoothed spectrum at a resolution of 25 MHz. The temperature scale of the spectrum is calibrated by taking into account the receiver noise temperature, the estimated atmospheric opacity and the estimated efficiency of the telescope. (Reproduced by permission).

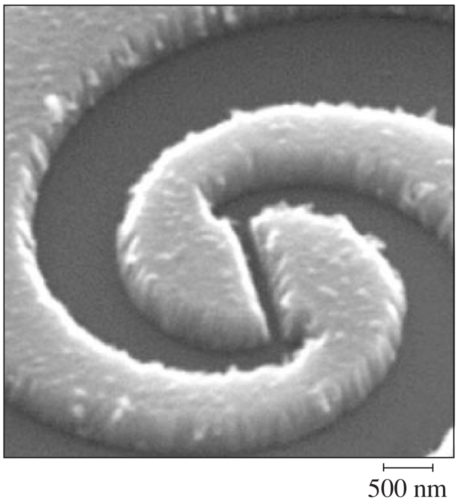

Figure 16. Micrograph of a central part of a planar logarithmic spiral antenna with the NbN hot-electron microbridge.

Major practical advantages of the lattice-cooled devices are their stability and the weak sensitivity of their noise temperature to operation parameters. Figure 17 shows that, indeed, the noise temperature of an NbN hot-electron mixer does not vary notably over a broad range of LO power and bias voltage [63].

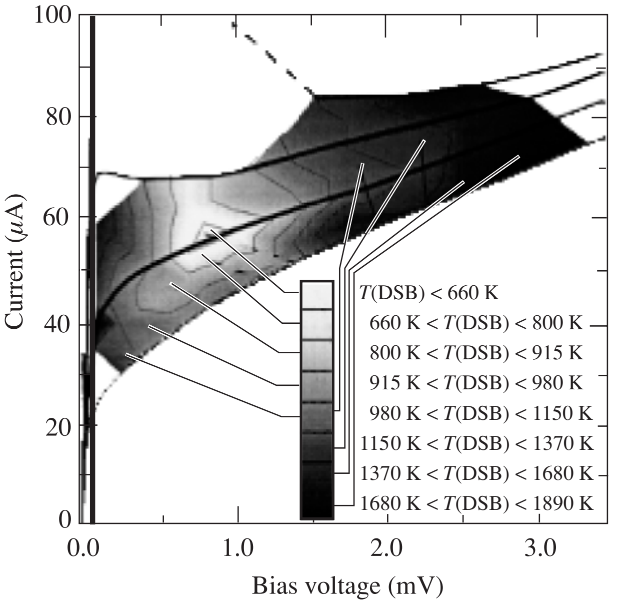

Figure 17. Double-sideband (DSB) noise temperature of a laboratory heterodyne receiver with NbN HEM at various bias regimes [63]. (Reproduced by permission).

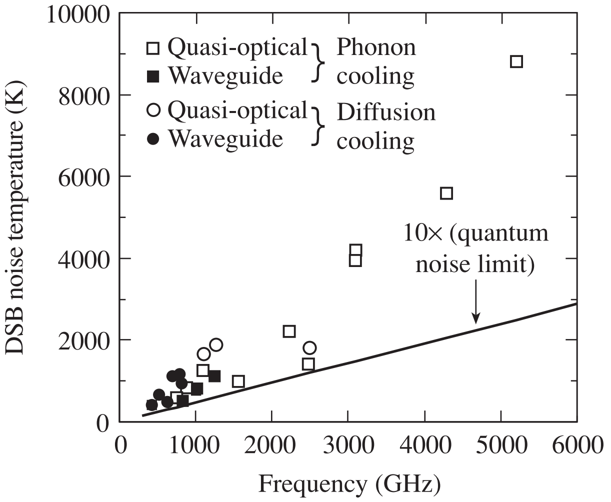

Figure 18. Best DSB noise temperatures for various types of superconducting hot-electron mixers as a function of signal frequency. The solid line is the hot-electron model prediction.

6.2. *Diffusion-cooled mixers*

The bulk of the diffusion-cooled mixers have been realized based on Nb films. At a 4.2 K bath temperature, the 10 nm thick Nb film typically has τ_{ep} of about 1 ns and a diffusivity of 2 cm^2 s^{-1} [11], which results in $L_{th} \approx 0.4$ μm. Therefore, Nb devices having a length of 0.1 μm or less operate in the pure diffusion-cooled regime. It has been shown experimentally [64] that the transition to diffusion cooling of electrons occurs at a device length $\approx$0.2 μm. The expected gain bandwidth for a 0.1 μm long device is about 7 GHz, if one assumes uniform electron heating through the length of the device. Laboratory tests at sub-THz frequencies confirmed theoretical expectations, and a 9 GHz gain bandwidth was measured for a 0.1 μm long HEM [65]. No noise bandwidth data have been reported so far for diffusion-cooled mixers. Traditionally, quasi-optical, diffusion-cooled HEMs use a twin-slot or double-dipole planar antenna and a hemispherical lens to couple the LO and signal radiations to the mixer. The best reported noise temperatures for Nb diffusion-cooled mixers are presently almost twice as large as those of lattice-cooled devices.

Another apparent difference between the two types of HEMs is the optimal bias regime, i.e. the regime resulting in the lowest noise temperature. For a lattice-cooled HEM, the optimal bias point is within the linear portion of the nonhysteretic *I–V* characteristics [63], whereas optimal operation of diffusion-cooled devices corresponds to the nonlinear portion of a hysteretic *I–V* curve [65]. The difference stems from boundary conditions imposed on the normal domain. The movement of the domain walls caused by signal radiation is not influenced by the contacts [66] if they are located far enough from the domain borders. One can envision such a domain as a freestanding domain in a stable equilibrium state. In the opposite case, when domain walls are confined near the contacts, the temperature profile at the walls slopes more steeply and the contacts restrict the movement of the domain walls thus hampering the mixer responsivity. The length of the smallest stable freestanding domain is of the order of the thermal diffusion length. Since the length of a diffusion-cooled HEM is typically smaller than the thermal diffusion length L_{th}, the conversion efficiency of the mixer is larger when the domain is 'overcooled' and is slightly shorter than the smallest freestanding domain. The actual domain length in a diffusion-cooled HEM, as seen from the resistance measurements in the normal state at the optimal operation point [65], is about 0.6 times the mixer physical length, whereas for phonon-cooled HEMs [63], the domain length is 0.2 times the device length. Since the total noise power at the HEM output is partly due to the Nyquist noise, smaller responsivity should result in a somewhat larger noise temperature. Another disadvantage of the diffusion-cooled HEM is that its hysteretic regime may cause additional instability [67] when accessed by a practical receiver.

For both types of mixers, it is common that optimal operation, aimed at the minimal noise temperature, does not provide the largest possible IF bandwidth. Both the bandwidth and the noise temperature increase with the bias current. Thus, varying the bias regime allows for a compromise between the desired bandwidth and the noise temperature acceptable for a particular application.

A diffusion-cooled Al mixer has been recently proposed [68] as an alternative to Nb devices. Measurements at 30 GHz [69] showed that a diffusion-cooled Al mixer exhibits reasonably good performance, but these data are not conclusive for the desired THz operation since the quantum energy of 30 GHz photons remains smaller than the Al energy gap. Moreover, there are concerns [19] that Al HEMs at THz frequencies would require a large LO power.

Table 2 and figure 18 summarize the current state-of-the-art noise temperatures for both the lattice-cooled and diffusion-cooled HEMs. The rapid increase in noise temperature with frequency is inconsistent with the hot-electron model. The model suggests that the noise temperature, when corrected for optical losses, should not depend on frequency unless

Table 2. Best DSB noise temperatures reported in the literature for lattice-cooled and diffusion-cooled mixers.

Lattice-cooled mixers					
Quasi-optical layout			Waveguide layout		
Frequency (GHz)	DSB noise temperature (K)	Reference	Frequency (GHz)	DSB noise temperature (K)	Reference
620	500	[70]	430	410	[73]
750	600	[65]	636	483	[73]
910	850	[65]	840	490	[61]
1100	1250	[65]	1017	750	[61]
1560	1000	[71]	1030	800	[61]
1620	700	[58]	1260	1100	[61]
2240	2200	[71]			
2500	1100	[58]			
3100	4000	[72]			
4300	5600	[72]			
5200	8800	[72]			
Diffusion-cooled mixers					
630	470	[64]	530	650	[76]
1100	1670	[74]	700	1100	[17]
1267	1880	[75]			
2500	1800	[64]			

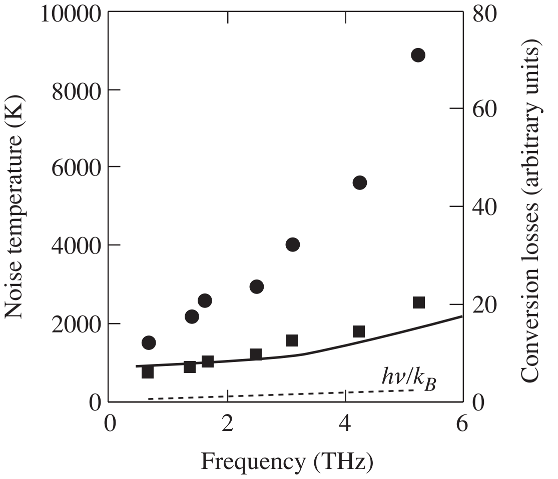

Figure 19. Frequency dependence of the noise temperature (dots) and conversion losses (squares) of an NbN HEM [67]. The solid line shows the calculated conversion losses that account for the skin effect in the device. The dashed line represents quantum-limited noise temperature $h\nu/k_B$. The scale of the right axis was adjusted to match calculated conversion losses and corrected noise temperature. (Reproduced by permission).

it approaches the quantum-limited value $h\nu/k_B$. A proper account of losses in coupling optics does not eliminate the above discrepancy; the noise temperature of the mixer alone increases with frequency, closely following the $10h\nu/k_B$ law in the frequency range from 0.6 THz to 5.2 THz. It has been shown recently [64] that the nonuniform distribution of the high-frequency current across the device may account for this effect.

In figure 19, simulated frequency dependence of the conversion efficiency is compared with the noise temperature corrected for optical losses. The good agreement between the experimental and theoretical results up to 4 THz suggests that the increase in the noise temperature should be less pronounced for narrower HEMs. However, the contact resistance increases with the decrease of the mixer width deteriorates coupling efficiency and may counterbalance the effect of the more uniform current distribution.

7. Conclusions

Superconductor hot-electron radiation sensors, operated as either THz frequency mixers or optical single-photon detectors, promise a revolutionary approach for diagnostics, radio astronomy, and quantum cryptography and communications. The unique performance of these devices in heterodyne as well as in the direct-detection regime results from a combination of the hot-electron phenomenon with the high sensitivity of a superconductor to nonequilibrium electronic states. To take full advantage of this combination, devices are routinely fabricated from ultrathin superconducting films and feature submicron lateral dimensions. They are also operated in the very-low-noise cryogenic environment.

HEMs proved their reliability and advantageous features during a 2 year test on a ground-based telescope. In the frequency range from 1 THz to 5.2 THz, HEMs outperformed Schottky diodes, making them the devices-of-choice for THz astronomy and communications.

HEPs demonstrated excellent performance in the spectral range from far-infrared wavelengths to x-rays when operated in either integrating or quantum regimes. Their future applications are expected in areas ranging from background-limited detector arrays for submillimetre astronomy and x-ray spectroscopy, through practical, high-speed quantum cryptography, to digital integrated circuit diagnostics.

Acknowledgments

The authors thank S Cherednichenko for helpful discussions and S Svechnikov for systemizing published data. This research was made possible partly by Award No RE-2227 of the US Civilian Research and Development Foundation for the Independent States of the Former Soviet Union. G N Gol'tsman acknowledges support of the INTAS (no 97-1453, no 99-569).

R Sobolewski acknowledges support of the US Office of Naval Research grant (N00014-02-1-0026).

References

[1] Conwell E M 1967 *High Field Transport in Semiconductors, Solid State Physics Supplement* **9** (New York: Academic) p 24

[2] Kaganov M L, Lifshitz I M and Tanatarov L V 1957 *Sov. Phys.–JETP* **4** 173

[3] Schklovski V A 1975 *Sov. Phys.–Solid State* **17** 3076

[4] Perrin N and Vanneste C 1983 *Phys. Rev.* B **28** 5150

[5] Semenov A D, Nebosis R S, Gousev Yu P, Heusinger M A and Renk K F 1995 *Phys. Rev.* B **52** 581

[6] Gershenzon E M, Gershenzon M E, Gol'tsman G N, Semenov A D and Sergeev A V 1984 *Sov. Phys.–JETP* **59** 442

[7] Lindgren M, Zorin M A, Trifonov V, Danerud M, Winkler D, Karasik B S, Gol'tsman G N and Gershenzon E M 1994 *Appl. Phys. Lett.* **65** 3398

[8] Il'in K S, Lindgren M, Currie M, Semenov A D, Gol'tsman G N, Sobolewski R, Cherednichenko S I and Gershenzon E M 2000 *Appl. Phys. Lett.* **76** 2752

[9] Lindgren M, Currie M, Williams C, Hsiang T Y, Fauchet P M, Sobolewski R, Moffat S H, Hughes R A, Preston J S and Hegmann F A 1999 *Appl. Phys. Lett.* **74** 853

[10] Sobolewski R 1998 *Proc. SPIE* **3481** 480

[11] Gershenzon E M, Gershenzon M E, Gol'tsman G N, Lyul'kin A M, Semenov A D and Sergeev A V 1990 *Sov. Phys.–JETP* **70** 505

[12] Gershenzon E M, Gershenzon M E, Gol'tsman G N, Lyul'kin A M, Semenov A D and Sergeev A V 1989 *Sov. Phys.–Tech. Phys.* **34** 195

[13] Gershenzon E M, Gol'tsman G N, Gogidze I G, Gusev Y P, Elant'ev A I, Karasik B S and Semenov A D 1990 *Supercond., Phys. Chem. Technol.* **3** 1582
Karasika B S and Elantiev A I 1996 *Appl. Phys. Lett.* **68** 853

[14] Prober D E 1993 *Appl. Phys. Lett.* **62** 2119

[15] Nebosis R S, Semenov A D, Gousev Yu P and Renk K F 1996 *Proc. 7th Int. Symp. on Space Terahertz Technology* (Charlottesville, VA: University of Virginia) p 601

[16] Karasik B S, McGrath W R and Gaidis M C 1997 *J. Appl. Phys.* **81** 1581

[17] Floet W D, Miedema E, Klapwijk T M and Gao J R 1999 *Appl. Phys. Lett.* **74** 433

[18] Merkel H, Khosropanah P, Yagoubov P and Kollberg E 1999 *IEEE Trans. Appl. Supercond.* **9** 4201

[19] Semenov A D and Gol'tsman G N 2000 *J. Appl. Phys.* **87** 502

[20] Ekström H, Karasik B S, Kollberg E L and Yngvesson K S 1995 *IEEE Trans. Microw. Theory Tech.* **43** 938

[21] Rothwarf A and Taylor B N 1967 *Phys. Rev. Lett.* **19** 27

[22] Grossman E N, McDonald D G and Sauvageau J E 1991 *IEEE Trans. Magn.* **27** 2677

[23] Bluzer N 1995 *J. Appl. Phys.* **78** 7340

[24] Semenov A D, Heusinger M A, Renk K F, Menschikov E, Sergeev A V, Elant'ev A I, Goghidze I G and Gol'tsman G N 1997 *IEEE Trans. Appl. Supercond.* **7** 3083

[25] Lindgren M, Currie M, Williams C A, Hsiang T Y, Fauchet P M, Sobolewski R, Moffat S H, Hughes R A, Preston J S and Hegmann F A 1996 *IEEE J. Sel. Top. Quantum Electron.* **2** 668

[26] Kadin A M and Johnson M W 1996 *Appl. Phys. Lett.* **69** 3938

[27] Gupta D and Kadin A M 1999 *IEEE Trans. Appl. Supercond.* **9** 4487

[28] Semenov A D, Gol'tsman G N and Korneev A A 2001 *Physica* C **351** 349

[29] Gol'tsman G N, Okunev O, Chulkova G, Lipatov A, Semenov A, Smirnov K, Voronov B, Dzardanov A, Williams C and Sobolewski R 2001 *Appl. Phys. Lett.* **79** 705

[30] Twerenbold D 1986 *Phys. Rev.* B **34** 7748

[31] Nussbaumer Th, Lerch P, Kirk E, Zehnder A, Füchslin R, Meier P F and Ott H R 2000 *Phys. Rev.* B **61** 9719

[32] Gershenzon E M, Gershenzon M E, Gol'tsman G N, Semenov A D and Sergeev A V 1981 *JETP Lett.* **34** 268

[33] Gershenzon E M, Gershenzon M E, Gol'tsman G N, Karasik B S, Lyul'kin A M and Semenov A D 1989 *Sov. Tech.–Phys. Lett.* **15** 118

[34] Bespalov A V, Gol'tsman G N, Semenov A D and Renk K F 1991 *Solid State Commun.* **80** 503

[35] Voronov B M, Gershenzon E M, Gol'tsman G N, Gogidze I G, Gusev Yu P, Zorin M A, Siedman L A and Semenov A D 1992 *Supercond., Phys. Chem. Technol.* **5** 960

[36] Gousev Yu P, Gol'tsman G N, Semenov A D, Gershenzon E M, Nebosis R S, Heusinger M A and Renk K F 1994 *J. Appl. Phys.* **75** 3695

[37] Vonsovskii S V, Iziumov I U A and Kurmaev E Z 1982 *Superconductivity of Transition Metals: Their Alloys and Compounds* (New York: Springer)

[38] Il'in K S, Milostnaya I I, Verevkin A A, Gol'tsman G N, Gershenzon E M and Sobolewski R 1998 *Appl. Phys. Lett.* **73** 3938

[39] Lang P T, Leipold I, Knott W J, Semenov A D, Gol'tsman G N and Renk K F 1991 *Appl. Phys.* B **53** 207
Lang P T, Schatz W, Kass T, Semenov A D and Renk K F 1992 *Opt. Lett.* **17** 502

[40] Miller A J, Cabrera B, Clarke R M, Figueroa-Feliciano E, Nam S and Romani R W 1999 *IEEE Trans. Appl. Supercond.* **9** 4205

[41] Irwin K D 1995 *Appl. Phys. Lett.* **66** 1998

[42] Nahum M and Martinis J M 1995 *Appl. Phys. Lett.* **66** 3203
Irwin K D, Hilton G C, Martinis J M, Deiker S, Bergren N, Nam S W, Redman D A and Wollman D A 2000 *Nucl. Instrum. Methods Phys. Res.* A **444** 184

[43] Nahum M and Martinis J M 1993 *Appl. Phys. Lett.* **63** 3075

[44] Karasik B S, McGrath W R, LeDuc H G and Gershenzon E M 1999 *Supercond. Sci. Technol.* **12** 745

[45] Grossman E N, McDonald D G and Sauvageau J E 1991 *IEEE Trans. Magn.* **27** 2677

[46] Sauvageau J E, McDonald D G and Grossman E N 1991 *IEEE Trans. Magn.* **27** 2757

[47] Sergeev A V and Reizer M Yu 1996 *Int. J. Mod. Phys.* B **10** 635

[48] Gol'tsman G N, Okunev O, Chulkova G, Lipatov A, Dzardanov A, Smirnov K, Semenov A, Voronov B, Williams C and Sobolewski R 2001 *IEEE Trans. Appl. Supercond.* **11** 574

[49] Ortiz G G, Sandusky J V and Biswas A 2000 *Proc. SPIE* **3932** 127

[50] Gilbert G and Hamrick M 2000 *MITRE Technical Report* MTR00W0000052 *Phys. Rep.* (at press) webpage http://xxx.lanl.gov/abs/quant-ph/0009027

[51] Tsang J C and Kash J A 1997 *Appl. Phys. Lett.* **70** 889

[52] Gol'tsman G, Semenov A, Smirnov K and Voronov B 2001 *Proc. 12th Int. Symp. on Space Terahertz Technology (Pasadena, CA, Feb. 2001)* p 469

[53] Friedrich S, Le Grand J B, Hiller L J, Kipp J, Frank M, Labov S E, Cramer S P and Barfknecht A T 1999 *IEEE Trans. Appl. Supercond.* **9** 3330

[54] Verhoeve P, Rando N, Peacock A, van Dordrecht A, Poelaert A and Goldie D J 1997 *IEEE Trans. Appl. Supercond.* **7** 3359

[55] Cherednichenko S, Yagoubov P, Il'in K, Gol'tsman G N and Gershenzon E 1997 *Proc. 8th Int. Symp. on Space Terahertz Technology* (Cambridge, MA: Harvard University Press) p 245

[56] Ekström H, Kollberg E, Yagoubov P, Gol'tsman G, Gershenzon E and Yngvesson S 1997 *Appl. Phys. Lett.* **70** 3296

[57] Cherednichenko S, Kroug M, Yagoubov P, Merkel H, Kollberg E, Yngvesson K S, Voronov B and Gol'tsman G 2000 *Proc. 11th Int. Symp. on Space Terahertz Technology* (Ann Arbor, MI: University of Michigan) p 219

[58] Kroug M, Cherednichenko S, Merkel H, Kollberg E, Voronov B, Gol'tsman G N, Hueberas H W and Richter H 2001 *IEEE Trans. Appl. Supercond.* **11** 962

[59] Voronov B M 1996 Private communication

[60] Kawamura J H, Tong C-Y E, Blundell R, Cosmo Papa D, Hunter T R, Gol'tsman G N, Cherednichenko S, Voronov B and Gershenzon E 1999 *IEEE Trans. Appl. Supercond.* **9** 3753

Kawamura J H, Hunter T R, Tong C-Y E, Blundell R, Zhang Q, Katz C A, Papa D C and Sridharan T K 1999 *Publ. Astron. Soc. Pac.* **111** 1088

[61] Tong C-Y E, Kawamura J, Hunter T R, Papa D C, Blundell R, Smith M, Patt F, Gol'tsman G and Gershenzon E 2000 *Proc. 11th Int. Symp. on Space Terahertz Technology* (Ann Arbor, MI: University of Michigan) p 49

[62] Huebers H-W, Semenov A, Schubert J, Gol'tsman G N, Voronov B N, Gershenzon E M, Krabbe A and Roeser H-P 2000 *Proc. SPIE* **4014** 195

[63] Yagoubov P, Kroug A M, Merkel H, Kollberg E, Gol'tsman G, Svechnikov S and Gershenzon E 1998 *Appl. Phys. Lett.* **73** 2814

[64] Burke P J, Schoelkopf R J, Prober D E, Skalare A, McGrath W R, Bumble B and LeDuc H G 1996 *Appl. Phys. Lett.* **68** 3344

[65] Wyss R A, Karasik B S, McGrath W R, Bumble B and LeDuc H 1999 *Proc. 10th Int. Symp. on Space Terahertz Technology* (Charlottesville, VA: University of Virginia) p 215

[66] Semenov A D and Hubers H-W 2001 *IEEE Trans. Appl. Supercond.* **11** 196

[67] Semenov A D, Hubers H-W, Schubert J, Gol'tsman G N, Elantiev A I, Voronov B M and Gershenzon E M 2000 *J. Appl. Phys.* **88** 6758

[68] Karasik B S and McGrath W R 1998 *Proc. 9th Int. Symp. on Space Terahertz Technology* (Pasadena, CA: California Institute of Technology) p 73

[69] Siddiqi I, Verevkin A, Prober D E, Skalare A, Karasik B S, McGrath W R, Echternach P and LeDuc H G 2000 *Proc. 11th Int. Symp. on Space Terahertz Technology* (Ann Arbor, MI: University of Michigan) p 82

[70] Yagoubov P, Kroug M, Merkel H, Kollberg E, Schubert J and Hubers H-W 1999 *Supercond. Sci. Technol.* **12** 989

[71] Gerecht E *et al* 1999 *Proc. 10th Int. Symp. on Space Terahertz Technology* (Charlottesville, VA: University of Virginia) p 200

[72] Schubert J, Semenov A, Gol'tsman G, Hubers H-W, Schwaab G, Voronov B and Gershenzon E 1999 *Supercond. Sci. Technol.* **12** 748

[73] Kawamura J, Blundell R, Tong C-Y E, Gol'tsman G, Gershenzon E, Voronov B and Cherednichenko S 1997 *Proc. 8th Int. Symp. on Space Terahertz Technology* (Cambridge, MA: Harvard University) p 23

[74] Skalare A, McGrath W R, Bumble B and LeDuc H G 1998 *Proc. 9th Int. Symp. on Space Terahertz Technology* (Pasadena, CA: California Institute of Technology) p 115

[75] Skalare A, McGrath W R, Bumble B and LeDuc H G 1997 *IEEE Trans. Appl. Supercond.* **7** 3568

[76] Skalare A, McGrath W R, Bumble B, LeDuc H G, Burke P J, Verheijen A A, Schoelkopf R J and Prober D E 1996 *Appl. Phys. Lett.* **68** 1558