

On-chip quantum confinement refrigeration overcoming electron-phonon heat leaks

S. Autti ^{1,*}, J. R. Prance ¹ and M. Prunnila ²¹*Department of Physics, Lancaster University, Lancaster LA1 4YB, United Kingdom*²*VTT Technical Research Centre of Finland Ltd., P.O. Box 1000, 02044 VTT Espoo, Finland*

(Received 20 August 2024; revised 6 December 2024; accepted 26 March 2025; published 14 April 2025)

Circuit-based quantum devices rely on keeping electrons at millikelvin temperatures. Improved coherence and sensitivity as well as discovering new physical phenomena motivate pursuing ever lower electron temperatures, accessible using on-chip cooling techniques. Here we show that a two-dimensional electron gas (2DEG), with the sub-band populations manipulated using gate voltages, works as an on-chip cooler only limited by a fundamental phonon heat leak. The 2DEG can, for example, be realized in a silicon-based double-gate complementary metal oxide semiconductor (CMOS) transistor. A single-shot 2DEG cooler can reduce the electron temperature by a factor of 2 with a hold time up to a second, achieved by expanding the electron gas into an additional sub-band. Integrating an array of such coolers—using, e.g., CMOS fabrication techniques—to obtain continuous cooldown may allow reaching down to microkelvin device temperatures.

DOI: [10.1103/PhysRevB.111.L161404](https://doi.org/10.1103/PhysRevB.111.L161404)

Solid-state quantum devices are typically operated at the lowest accessible temperatures, down to around 10 mK, because the physical processes that these devices are based on vanish or weaken at higher temperatures. Lowering the temperature further is an obvious strategy for new discoveries and improved device performance in applications such as quantum computing, but lowering the temperature of electrons inside the device faces a fundamental barrier due to thermal decoupling from the refrigerator. The lowest device electron temperatures that can be reached by external heat sinking are typically ≈ 4 mK [1–4]. Below 4 mK the thermal link from electrons via phonons to the refrigerator sample holder becomes vanishingly small. Here, the tiniest heat leak entering the system, for example via measurement wires or due to external radiation, can prevent further cooldown even if the heat sink temperature approaches absolute zero. Immersing the device and the wiring in a superfluid [5] allows lowering the electron temperature to below one millikelvin, but this only moves the decoupling to a somewhat lower temperature.

The low-temperature thermal bottleneck can be circumvented by building an on-chip cooldown cycle. The first generation devices were made by placing miniaturized demagnetization refrigerators directly on chip [6–9], providing access deep into the microkelvin regime. This technique can only be applied to devices that tolerate magnetic fields ~ 100 mT and above, and it relies on a custom refrigerator platform and special device design and fabrication, such as electroplated refrigerant on the device. In this Letter we

study theoretically an electric field equivalent of on-chip adiabatic demagnetization. The occupation of sub-bands in a two-dimensional electron gas can be controlled by manipulating the profile of the confinement in one spatial dimension by the potentials of electrostatic gate electrodes (Fig. 1). The heat capacity in this system only depends on the number of occupied bands, which means that an adiabatic expansion of the electron population to a larger number of bands leads to a temperature decrease, providing an integrated and scalable on-chip electronic refrigeration method. The concept of this adiabatic expansion process has been discussed in Ref. [10]. However, in a real physical system or a device, a cooldown process is useful only if the coolant can absorb and surmount the unavoidable heat leaks from the environment and stay cold. Here we study the sub-band cooling effect in a Si-based two-dimensional electron gas (2DEG) against the only fundamental heat leak present [6] originating from electron-phonon coupling. By using an electron-phonon coupling model for disordered Si-based 2DEG [11], we show that significant cooling is possible below 20 mK, potentially allowing devices to reach even microkelvin electron temperatures.

The 2DEG heat capacity per unit area is $C = \pi^2 k_B^2 T D_0(E_F)/3$ [12,13], where k_B is the Boltzmann constant, E_F is the Fermi energy, and $D_0(E_F)$ is the density of states for one 2DEG sub-band. In the absence of changes in $D_0(E_F)$, the 2DEG electron temperature follows

$$\frac{dT_e}{dt} = (P_0 - P_{ep})/C, \quad (1)$$

where P_{ep} is the heat flow from 2DEG electrons to phonons in the underlying material and P_0 contains spurious heat leaks to the 2DEG.

The entropy of an electron gas can be obtained by integrating the ratio of its heat capacity and temperature, $C(T)/T$,

*Contact author: s.autti@lancaster.ac.uk

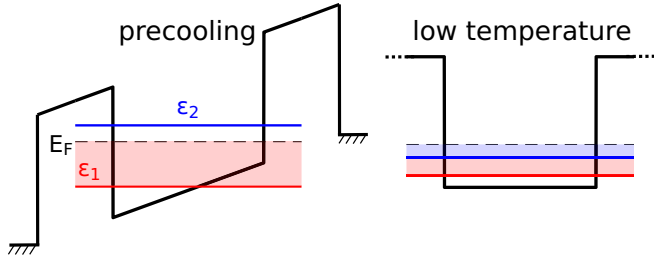


FIG. 1. Schematic illustration of quantum confinement cooling using a 2DEG. During precooling, a gate voltage is applied across the confined dimension of the 2DEG to create a triangular potential (thick black line). This increases the spacing between bands corresponding to ϵ_1 and ϵ_2 so that only ϵ_1 is below the Fermi energy. During the cooldown, the gate voltage is reduced to zero so that both $\epsilon_1, \epsilon_2 < E_F$. The (partially overlapping) red and blue shaded areas are proportional to the population of each sub-band. The tilt of the potential and the level spacing are exaggerated for illustrational purposes.

yielding

$$S(T) = \pi^2 k_B^2 T D(E_F) / 3, \quad (2)$$

where $D(E_F)$ is the sum of the densities of state of the occupied sub-bands. Therefore, $T \propto 1/D(E_F)$ in any isentropic process [10].

Let us consider the lowest two sub-bands of an isolated 2DEG at temperature $T_1 \ll E_F/k_B$. The sub-bands have initial energies ϵ_0 and ϵ_1 such that $\epsilon_0 < E_F < \epsilon_1$; only the lowest sub-band is populated. We assume that $\epsilon_1 - \epsilon_0$ can be manipulated adiabatically [14,15] by changing top and back gate voltages that control the confinement profile (Fig. 1). If we change $\epsilon_1 - \epsilon_0$ so that $\epsilon_0 < \epsilon_1 < E_F$, the total density of available states becomes $D = 2D_0$ and the 2DEG temperature is halved. Note that E_F will decrease during this process if the electron density is held constant. In a system with N sub-bands, with $N - 1$ initially above E_F , the final temperature after bringing all of them inside the Fermi surface is $T = T_1/N$. It has been shown in experiments that at least three sub-bands ($N = 3$) can be manipulated this way [16]. Continuous cooling can be achieved by operating multiple 2DEGs in turn to pump heat from a common bath. This is possible as long as each individual 2DEG is able to hold a reduced temperature for longer than the time taken to cool it, and to switch a thermal connections to the common bath.

In experiments, the 2DEG temperature evolves adiabatically provided the changes in the sub-band population are slow as compared with the rates of inter- and intraband inelastic scattering, typically reported to be $\sim$ ps or below at $T_e = 4.2$ K [17–19]. In theory, the inelastic electron-electron scattering rate within a subband scales approximately as T_e^2 [20,21] and weaker scaling is observed in experiment [18]. Assuming a scaling of T_e^2 yields ~ 100 ns at $T_e = 10$ mK, so the condition of adiabaticity should be satisfied for gate operations that last a μ s or more even in this pessimistic case. That is, the cooldown performance is primarily hindered by heat leaks from the environment. We consider a 2DEG in silicon placed in a commercial dilution refrigerator that cools the phonons in the underlying semiconductor to $T_p = 10$ mK

[6]. Where relevant, the 2DEG surface area is taken to be $A = 100 \mu\text{m} \times 100 \mu\text{m}$ and the thickness $d = 100$ nm.

Phonons can carry heat across several material interfaces and remain at the refrigerator temperature throughout the on-chip cooldown process [6]. The electron-phonon coupling decreases rapidly at low millikelvin temperatures, allowing the on-chip electrons to reach temperatures below the phonon temperature. We write the electron-phonon heat flow per unit area as $P_{ep} = \Omega(T_e^6 - T_p^6)$, where the coupling for a single sub-band can be written as [11]

$$\Omega = \frac{\hbar D_0}{2\pi^2 \rho \kappa} B_5 \frac{1}{\kappa l_s} \left\langle \frac{\sin^2 \theta \Xi_s^2}{(l_s \kappa)^{-2} + \sin^2 \theta} \right\rangle \left(\frac{k_B}{\hbar v_s} \right)^6. \quad (3)$$

In this expression v_s is the sound velocity, ρ is the mass density of the material, v_F is the Fermi velocity, $B_5 \approx 122$, l_e is the electron mean free path, the length scale $l_s = l_e v_F / (2v_s)$, κ is the Thomas-Fermi screening wave vector, and Ξ_s is the angle-dependent deformation potential. Electron mean free path is defined by electron mobility μ through $l_e = v_F \mu m^* / e$, where m^* is the effective mass. Values and expressions needed for evaluating Eq. (3) are given in Table A1 in the Supplemental Material [22]. Equation (3) applies in the diffusive limit of electron-phonon coupling, where thermal phonon wavelength $q_T = \frac{k_B T_p}{\hbar v_s}$ exceeds l_e . The prefactor with average over a solid angle that appears in the equation can be approximated as

$$\left\langle \frac{\sin^2 \theta \Xi_s^2}{(l_s \kappa)^{-2} + \sin^2 \theta} \right\rangle \approx \Xi^2 \left(1 - \frac{1}{\sqrt{(l_s \kappa)^2 + 1}} \right), \quad (4)$$

where Ξ is a “scalar” deformation potential constant in the range 1–10 eV. We use $\Xi = 10$ eV in the simulation below as a pessimistic estimate [23]. Note that thanks to the solid-angle average pre of $F(T)$ of Ref. [11] there is no unphysical divergence of $F(T)$ at $n_e \rightarrow 0$, unlike in the formulas of Ref. [24].

The 2DEG cooldown simulation shown in Fig. 2 starts from an elevated T_e , reaching a stable value exponentially. The population of the lowest band is initially $n_1 = 10^{13} \text{ cm}^{-2}$ and that of the second sub-band $n_2 = 0$. The 2DEG needs to be well isolated electrically to prevent the Wiedemann-Franz heat leak from rapidly overwhelming the electron heat capacity [see Fig. 3(c)]. For simplicity, $P_0 = 0$ in this calculation, and therefore $T_p = T_e = 10$ mK in equilibrium.

At $t = 0$, we move a third of the electron population to the upper band in a linear sweep that lasts $10 \mu\text{s}$. We label this operation a “transient.” In experiments, the timescale of the transient is limited from below by the charging time of the top gate whose capacitance is $\lesssim 10$ pF and the charging time can be as low as ~ 1 ns with a low-resistance connection ($\sim 10 \Omega$).

With two bands populated, the total electron phonon coupling is given by the sum of the couplings calculated separately for the two sub-bands using Eq. (3). We assume that the density of states is doubled immediately as the electrons start to move (this assumption is justified below). The imposed temperature reduction is much faster than the dynamics described by Eq. (1), causing a discrete jump in the simulated T_e . Consequently, the electron temperature is initially halved to $T_e = 5$ mK and it then returns exponentially to the equilibrium value with the time constant $\tau \approx 0.1$ s.

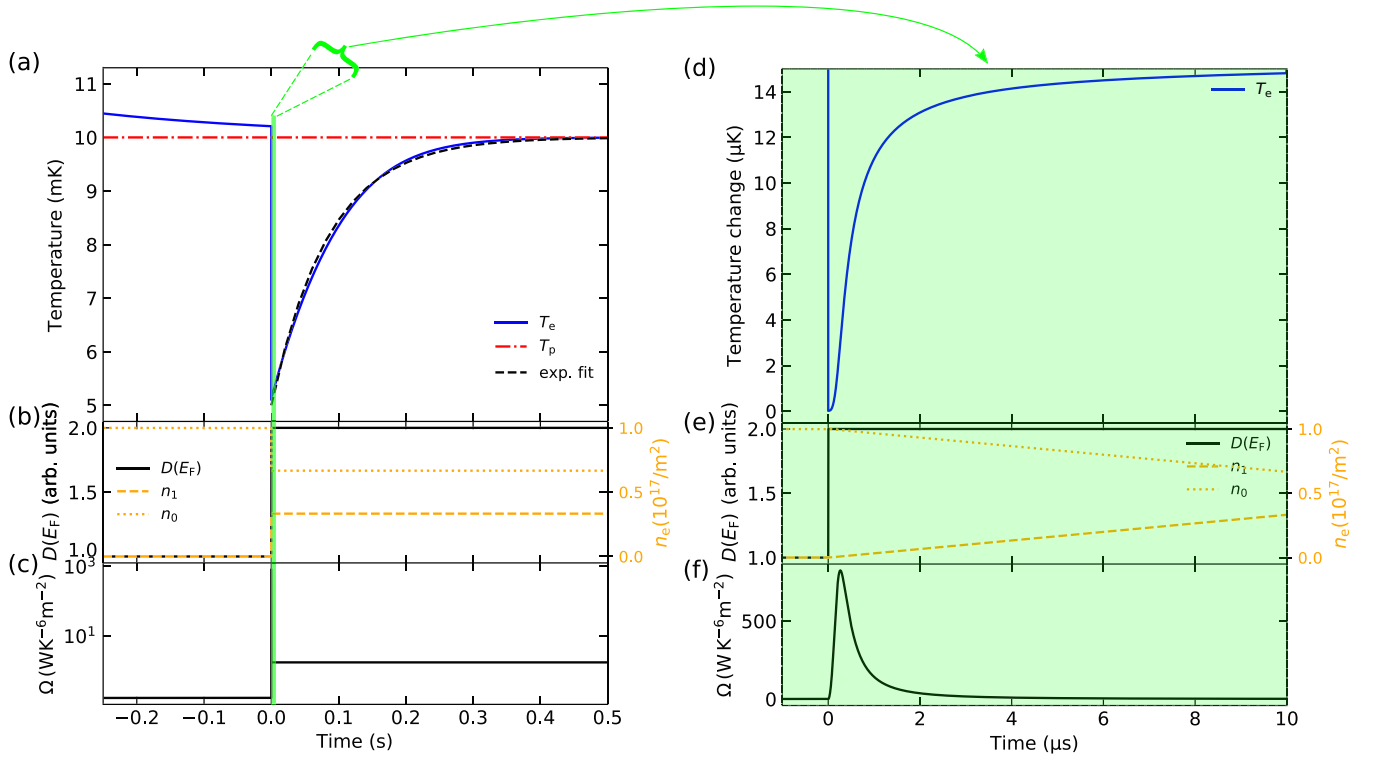


FIG. 2. Simulated 2DEG cooldown performance: At $t < 0$ s only the lowest-energy band is populated ($n_1 = 10^{13} \text{ cm}^{-2}$, $n_2 = 0$). The spurious heat leak is set to $P_0 = 0$ for simplicity. Thus, the simulated electron temperature shown in panel (a) (blue line) comes to equilibrium at $T_e = T_p$. At $t = 0$, one-third of the electrons are moved to the second-lowest band in a linear ramp that lasts $10 \mu\text{s}$ [right y axis in panel (b)]. During this process, the heat leaks are negligible and the density of states $D(E_F)$, shown on the left y axis in panel (b), is doubled, corresponding to bringing a second 2DEG sub-band inside the Fermi sphere. Thus, T_e is halved, after which it returns to an equilibrium value approximately exponentially (black dash line) with the time constant $\tau = 90 \text{ ms}$. The electron-phonon coupling increases in this process [panel (c)]. A zoomed-in version of the population ramp showing the transient in detail is in panels (d)–(f). When the electrons start to move from n_0 to n_1 at $t = 0$ s, the coupling becomes momentarily very large [panel (f)]. The warm-up rate just after $t = 0$ is large but it quickly decreases as the population ramp continues, and the temperature increase remains negligible during the transient. Before $t = 0$ the electron-phonon coupling via the upper band is taken to be zero. We have assumed here that the density of states [panels (b) and (d) left y axis] increase takes place at $t = 0$, meaning that the heat capacity increase and thus the temperature decrease take occur instantaneously at $t = 0$. In a realistic experiment, the cooldown may be smoothed for example due to spatial inhomogeneities of the 2DEG electron density, but for an adiabatic process this does not change the cooldown performance, only the details during the transient.

The electron-phonon coupling in Eq. (3) depends on sub-band electron density. For typical parameter values it has a maximum between $n_e = 0$ and $n_e = 10^{12} \text{ cm}^{-2}$. Most of the parameters in Eq. (3) have simple dependencies on electron density provided in Table A1 of the Supplemental Material [22], but how the electron mean-free-path depends on the sub-band electron density is more subtle. For a single occupied sub-band, as the total electron density decreases, the charge carrier elastic scattering time and mobility—and, thereby, the mean-free-path—typically decrease below the so-called peak mobility defining the crossover between Coulomb and surface roughness scattering [25]. To capture the essential features of the mean-free-path and mobility density dependence, we assume that the mobility for the upper band decreases linearly to zero as a function of electron density below the final value of n_2 , and above this value the mobility is the same for both bands ($\mu = 10^4 \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1}$).

We now concentrate on the transient shown in the right panels of Fig. 2. The coupling increases by an order of magnitude during the transient, but the electron temperature only increases by a few μK before the transient ends. This

conclusion holds even if we fix the mobility (mean free path) to the value that corresponds to the maximum coupling, reached at some point during the transient. Therefore, the transient can safely be ignored, and the cooldown performance remains largely the same regardless of the precise timing of the increase in the density of states, changes in electron mobility during the transient, and also regardless of inhomogeneities in the system that mean the population transfer between the 2DEG bands does not take place uniformly within the 2DEG area.

In addition to the heat flow between electrons and phonons, experiments at microkelvin and millikelvin temperatures typically observe a spurious heat leak P_0 , originating from noise entering via measurement contacts, a slowly decaying intrinsic relaxation in dielectrics [6,26,27], or possibly from external radiation such as cosmic rays [28,29]. It is known that that P_0 , scaled to the thickness of the 2DEG system considered here, can be $1 \text{ aW}/(10^4 \mu\text{m})^2$ [6] in structures in a magnetic field. We consider this to be the worst case scenario. With this large heat leak, the dynamic steady state temperature of electrons is $T_e \approx 30 \text{ mK}$ (when electron-phonon coupling is

absorbing P_0) for $T_p = 10$ mK. The cooldown time constant is reduced from ~ 0.1 s to ~ 10 ms.

In the absence of a magnetic field, P_0 can be much lower. Transport measurements across an interface between two dielectric materials in the zero-temperature limit yield $P_0 \sim zW/(10^4 \text{ } \mu\text{m})^2$ [26,27]. Thus, we also consider a well isolated 2DEG where we find P_0 by fitting it so the initial temperature is $T_e = 15$ mK, yielding $P_0 \approx 20$ zW (still pessimistic as compared with the best results reported in literature [5]). With this heat leak the cooldown time constant is already similar to that obtained with $P_0 = 0$. We emphasise that in a realistic best-case scenario [5] the electrons would be cooled close to 1 mK before starting the on-chip cooldown cycle, with the phonon temperature yet lower than this. Here, τ becomes orders of magnitude longer than a second.

We can also try to simulate the 2DEG cooldown performance based on experimental measurements of electron-phonon coupling in silicon and GaAs 2DEGs [30–34]. This effort is hindered by the unavailability of measurements of the electron-phonon coupling below 100 mK, but in the absence of more applicable experimental data we extrapolate them to 10 mK. These measurements have also been carried out at electron densities around $n_e = 10^{11} \text{ cm}^{-2}$, where the electron-phonon coupling is two orders of magnitude stronger than at $n_e = 10^{13} \text{ cm}^{-2}$. Keeping these shortcomings in mind, and without reducing the coupling at higher electron densities in the simulation, the obtained cooldown time constants [30–34] range between ~ 10 ms and 1 s, in good agreement with the purely theoretical values. These calculation outcomes are listed in Table A2 in the Supplemental Material [22]. That table also shows that the theoretical values obtained with $P_0 > 0$ are similar to the ones obtained neglecting the spurious heat leaks.

We propose two schemes for measuring the performance of the 2DEG refrigeration scheme [Figs. 3(a) and 3(b)].

Manipulation of the 2DEG sub-band occupation requires gates above and below the electron gas. Both devices shown in Fig. 3 are based on a 2DEG in an intrinsic silicon-on-insulator (SOI) structure [14,15], with a bulk back gate and a polysilicon top gate to adjust the sub-bands as in Fig. 1. The cooled region of 2DEG is isolated from external electrical systems by the in-plane shape of the polysilicon gates. The temperature of the cooled region could be measured using either quantum dots [Fig. 3(a)] or superconducting junctions [Fig. 3(b)]. Both types of device allow the temperature of the isolated region to be measured without introducing additional heat leaks; indeed they could be operated to provide even some background cooling [35–37].

Figure 3(c) shows how multiple cooled 2DEG regions could be operated in parallel and in series to reach lower temperatures and achieve continuous cooling. Such schemes are possible because the thermal connection between different 2DEG regions can be easily controlled using additional gates to electrically connect or isolate the regions. We note that below ~ 1 mK electron thermalization via the interband coupling becomes so slow that the gate operations should be made longer than $10 \text{ } \mu\text{s}$. In this temperature regime the phonon heat leak is so strongly suppressed that this will not harm the cooldown performance.

The schemes proposed in Fig. 3 could serve to validate and characterize the 2DEG refrigeration as described here. They also show how to integrate 2DEG cooler with quantum dot based spin qubits. The transient between one and two populated sub-bands is experimentally unexplored, and measuring the electron-phonon coupling during this process can be a challenging task. The theoretical assumptions we have made to describe the transient seem well justified, but characterizing the proposed devices would serve as a way to explore this regime experimentally.

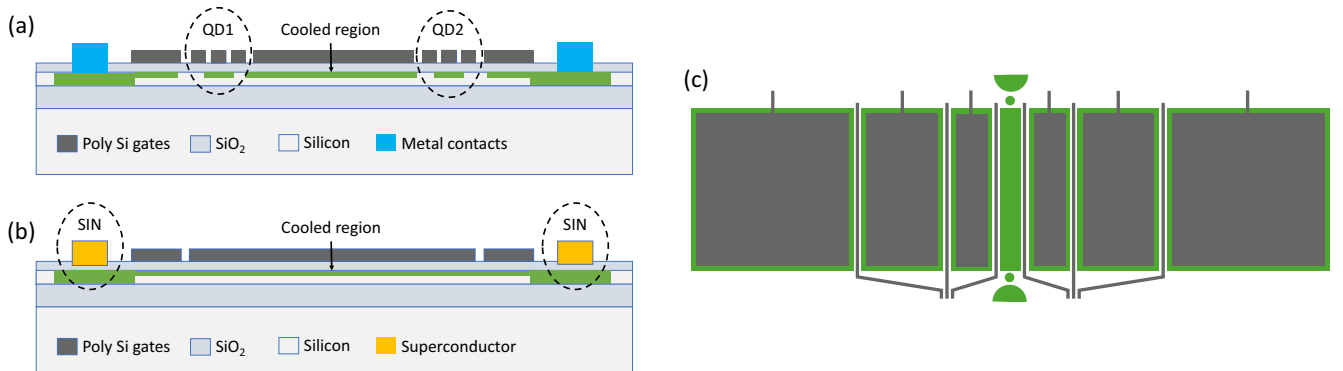


FIG. 3. Practical schemes for implementing 2DEG cooling and temperature measurement using complementary metal oxide semiconductor (CMOS) devices. The 2DEG is created in a silicon layer in silicon oxide using patterned polysilicon gates. Conducting regions are shown in green. In design (a), shown in cross section, two quantum dots (QD1 and QD2) are connected to the 2DEG via tunnel barriers. The 2DEG is cooled by changing the voltages of the polysilicon gate above it and a bulk silicon back gate (or patterned back gates) below. The temperature of the cooled 2DEG can be found from transport measurements through the 2DEG via the quantum dots as in Ref. [35]. In the design (b) two superconductor-insulator-normal metal (SIN) junctions are connected to the 2DEG. Here the N region below the S electrode is formed by degenerately doped silicon in the spirit of Ref. [36]. The electron temperature is inferred from transport measurements through the SINIS structure. In design (c), shown from above, two chains of cooled 2DEGs (green), each containing three 2DEG regions in series, are connected in parallel to a central cooled region. Electrical and thermal connection between the 2DEG regions is controlled by narrow polysilicon gates (gray) between them. Sequential cooling of the regions connected in series can be used to reach a lower final temperature. Alternating operation of the parallel chains can be used to achieve continuous cooling of the central region.

We have chosen pessimistic parameter values and operational conditions in simulating the 2DEG cooldown performance. For example, decreasing the phonon temperature to 5 mK, which has been reported in multiple 2DEG experiments and can be achieved on commercial refrigerator platforms, decreases the minimum electron temperature in single-stage operation to 2.5 mK and increases the time constant to $\tau = 1.5$ s. We note that it is possible to control the population of three bands [16], not just two. This would allow reducing the lowest temperature reached to one-third of the starting temperature. Also, elastic scattering increases the electron-phonon coupling, and thus using high-mobility systems could improve the cooldown performance. Devices of interest could be based, e.g., on SiGe [38,39] or III-V quantum wells. Finally, we note that while there can be other heat leaks to a quantum device on top of the fundamental phonon heat leak studied here, a range of experiments have shown that

these can be reduced to a level where they become irrelevant for experiments at mK temperatures [5–9].

This research is supported by the U.K. EPSRC (EP/W015730/1), the European Union's Horizon 2020 research and innovation program (European Microkelvin Platform 824109 and EFINEF 766853) and EIC Transition program (SoCool 101113086), by the Academy of Finland through the Centre of Excellence program (Projects No. 336817 and No. 312294), and by Business Finland through QuTI project (40562/31/2020) and Technology Industries of Finland Centennial Foundation.

All the simulation data in this Letter are available online [42], including descriptions of the data sets. The simulation codes can be obtained from the corresponding author upon reasonable request.

-
- [1] J. S. Xia, E. D. Adams, V. Shvarts, W. Pan, H. L. Stormer, and D. C. Tsui, Ultra-low-temperature cooling of two-dimensional electron gas, *Phys. B: Condens. Matter* **280**, 491 (2000).
 - [2] N. Samkharadze, A. Kumar, M. J. Manfra, L. N. Pfeiffer, K. W. West, and G. A. Cs  thy, Integrated electronic transport and thermometry at milliKelvin temperatures and in strong magnetic fields, *Rev. Sci. Instrum.* **82**, 053902 (2011).
 - [3] D. I. Bradley, R. E. George, D. Gunnarsson, R. P. Haley, H. Heikkinen, Yu. A. Pashkin, J. Penttil  , J. R. Prance, M. Prunnila, L. Roschier, and M. Sarsby, Nanoelectronic primary thermometry below 4 mK, *Nat. Commun.* **7**, 10455 (2016).
 - [4] A. T. Jones, C. P. Scheller, J. R. Prance, Y. B. Kalyoncu, D. M. Zumb  hl, and R. P. Haley, Progress in cooling nanoelectronic devices to ultra-low temperatures, *J. Low Temp. Phys.* **201**, 772 (2020).
 - [5] L. V. Levitin, H. van der Vliet, T. Theisen, S. Dimitriadis, M. Lucas, A. D. Corcoles, J. Ny  ki, A. J. Casey, G. Creeth, I. Farrer, D. A. Ritchie, J. T. Nicholls, and J. Saunders, Cooling low-dimensional electron systems into the microkelvin regime, *Nat. Commun.* **13**, 667 (2022).
 - [6] S. Autti, F. C. Bettsworth, K. Grigoras, D. Gunnarsson, R. P. Haley, A. T. Jones, Yu. A. Pashkin, J. R. Prance, M. Prunnila, M. D. Thompson, and D. E. Zmееv, Thermal transport in nanoelectronic devices cooled by on-chip magnetic refrigeration, *Phys. Rev. Lett.* **131**, 077001 (2023).
 - [7] D. I. Bradley, A. M. Gu  nault, D. Gunnarsson, R. P. Haley, S. Holt, A. T. Jones, Yu. A. Pashkin, J. Penttil  , J. R. Prance, M. Prunnila, and L. Roschier, On-chip magnetic cooling of a nanoelectronic device, *Sci. Rep.* **7**, 45566 (2017).
 - [8] M. Palma, C. P. Scheller, D. Maradan, A. V. Feshchenko, M. Meschke, and D. M. Zumb  hl, On-and-off chip cooling of a Coulomb blockade thermometer down to 2.8 mK, *Appl. Phys. Lett.* **111**, 253105 (2017).
 - [9] N. Yurttag  l, M. Sarsby, and A. Geresdi, Indium as a high-cooling-power nuclear refrigerant for quantum nanoelectronics, *Phys. Rev. Appl.* **12**, 011005(R) (2019).
 - [10] L. G. C. Rego and G. Kirczenow, Electrostatic mechanism for cooling semiconductor heterostructures, *Appl. Phys. Lett.* **75**, 2262 (1999).
 - [11] M. Prunnila, Electron-acoustic-phonon energy-loss rate in multicomponent electron systems with symmetric and asymmetric coupling constants, *Phys. Rev. B* **75**, 165322 (2007).
 - [12] Q. Li, X. C. Xie, and S. Das Sarma, Calculated heat capacity and magnetization of two-dimensional electron systems, *Phys. Rev. B* **40**, 1381 (1989).
 - [13] E. Grivei, J.-M. Beuken, G. Mariage, V. Bayot, and M. Shayegan, Heat capacity of a 2DEG, *Phys. B: Condens. Matter* **256–258**, 90 (1998).
 - [14] M. Prunnila and J. Ahopelto, Two sub-band conductivity of Si quantum well, *Physica E* **32**, 281 (2006).
 - [15] S. Zhao, G. Yuan, D. Zhang, Y. Liu, J. Lu, W. Han, and J. Luo, Electron transport characteristics in dual gate-controlled 30 nm-thick silicon membrane, *J. Phys. D* **55**, 495105 (2022).
 - [16] L. Yang, X. Wang, T. Wang, J. Wang, W. Zhang, P. Quach, P. Wang, F. Liu, D. Li, L. Chen *et al.*, Three subband occupation of the two-dimensional electron gas in ultrathin barrier AlN/GaN heterostructures, *Adv. Funct. Mater.* **30**, 2004450 (2020).
 - [17] M. Cass  , B. Cardoso Paz, G. Ghibaudo, T. Poiroux, E. Vincent, P. Galy, A. Juge, F. Gaillard, S. de Franceschi, T. Meunier *et al.*, Evidence of 2D intersubband scattering in thin film fully depleted silicon-on-insulator transistors operating at 4.2 K, *Appl. Phys. Lett.* **116**, 243502 (2020).
 - [18] R. A. Davies and M. Pepper, Electron-electron scattering in silicon inversion layers, *J. Phys. C* **16**, L353 (1983).
 - [19] S.-C. Lee and I. Galbraith, Intersubband and intrasubband electronic scattering rates in semiconductor quantum wells, *Phys. Rev. B* **59**, 15796 (1999).
 - [20] L. Zheng and S. Das Sarma, Coulomb scattering lifetime of a two-dimensional electron gas, *Phys. Rev. B* **53**, 9964 (1996).
 - [21] G. F. Giuliani and J. J. Quinn, Lifetime of a quasiparticle in a two-dimensional electron gas, *Phys. Rev. B* **26**, 4421 (1982).
 - [22] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/PhysRevB.111.L161404> for details on the numerical simulations, which includes Refs. [40,41].
 - [23] R. Hull, *Properties of Crystalline Silicon*, Emis Series No. 20 (IET, London, 1999).
 - [24] A. Sergeev, M. Yu. Reizer, and V. Mitin, Deformation electron-phonon coupling in disordered semiconductors and nanostructures, *Phys. Rev. Lett.* **94**, 136602 (2005).

- [25] T. Ando, A. B. Fowler, and F. Stern, Electronic properties of two-dimensional systems, *Rev. Mod. Phys.* **54**, 437 (1982).
- [26] P. J. Heikkinen, S. Autti, V. B. Eltsov, R. P. Haley, and V. V. Zavjalov, Microkelvin thermometry with Bose–Einstein condensates of magnons and applications to studies of the AB interface in superfluid ^3He , *J. Low Temp. Phys.* **175**, 681 (2014).
- [27] J. J. Hosio, V. B. Eltsov, R. De Graaf, M. Krusius, J. Mäkinen, and D. Schmoranzner, Propagation of thermal excitations in a cluster of vortices in superfluid ^3He -B, *Phys. Rev. B* **84**, 224501 (2011).
- [28] S. Autti, A. Casey, N. Eng, N. Darvishi, P. Franchini, R. P. Haley, P. J. Heikkinen, A. Jennings, A. Kemp, E. Leason, L. V. Levitin, J. Monroe, J. March-Russel, M. T. Noble, J. R. Prance, X. Rojas, T. Salmon, J. Saunders, R. Smith, M. D. Thompson *et al.*, QUEST-DMC superfluid ^3He detector for sub-GeV dark matter, *Eur. Phys. J. C* **84**, 248 (2024).
- [29] S. Autti, A. Casey, N. Eng, N. Darvishi, P. Franchini, R. P. Haley, P. J. Heikkinen, A. Kemp, E. Leason, L. V. Levitin *et al.*, QUEST-DMC: Background modelling and resulting heat deposit for a superfluid helium-3 bolometer, *J. Low Temp. Phys.* **215**, 465 (2024).
- [30] O. Prus, M. Reznikov, U. Sivan, and V. Pudalov, Cooling of electrons in a silicon inversion layer, *Phys. Rev. Lett.* **88**, 016801 (2001).
- [31] A. Mittal, R. G. Wheeler, M. W. Keller, D. E. Prober, and R. N. Sacks, Electron-phonon scattering rates in GaAs/AlGaAs 2DEG samples below 0.5 K, *Surf. Sci.* **361**, 537 (1996).
- [32] Y. Y. Proskuryakov, J. T. Nicholls, D. I. Hadji-Ristic, A. Kristensen, and C. B. Sørensen, Energy-loss rate of a two-dimensional electron gas measured using a mesoscopic thermometer, *Phys. Rev. B* **75**, 045308 (2007).
- [33] N. J. Appleyard, J. T. Nicholls, M. Y. Simmons, W. R. Tribe, and M. Pepper, Thermometer for the 2D electron gas using 1D thermopower, *Phys. Rev. Lett.* **81**, 3491 (1998).
- [34] M. Prunnila, P. Kivinen, A. Savin, P. Törmä, and J. Ahopelto, Intervalley-scattering-induced electron-phonon energy relaxation in many-valley semiconductors at low temperatures, *Phys. Rev. Lett.* **95**, 206602 (2005).
- [35] J. R. Prance, C. G. Smith, J. P. Griffiths, S. J. Chorley, D. Anderson, G. A. C. Jones, I. Farrer, and D. A. Ritchie, Electronic refrigeration of a two-dimensional electron gas, *Phys. Rev. Lett.* **102**, 146602 (2009).
- [36] E. Mykkänen, J. S. Lehtinen, L. Grönberg, A. Shchepetov, A. V. Timofeev, D. Gunnarsson, A. Kemppinen, A. J. Manninen, and M. Prunnila, Thermionic junction devices utilizing phonon blocking, *Sci. Adv.* **6**, eaax9191 (2020).
- [37] A. V. Feshchenko, J. V. Koski, and J. P. Pekola, Experimental realization of a Coulomb blockade refrigerator, *Phys. Rev. B* **90**, 201407(R) (2014).
- [38] M. Yu. Melnikov, A. A. Shashkin, V. T. Dolgoplov, S.-H. Huang, C. W. Liu, and S. V. Kravchenko, Ultra-high mobility two-dimensional electron gas in a SiGe/Si/SiGe quantum well, *Appl. Phys. Lett.* **106**, 092102 (2015).
- [39] A. A. Shashkin and S. V. Kravchenko, Recent developments in the field of the metal-insulator transition in two dimensions, *Appl. Sci.* **9**, 1169 (2019).
- [40] J. T. Muhonen, M. J. Prest, M. Prunnila, D. Gunnarsson, V. A. Shah, A. Dobbie, M. Myronov, R. J. H. Morris, T. E. Whall, E. H. C. Parker *et al.*, Strain dependence of electron-phonon energy loss rate in many-valley semiconductors, *Appl. Phys. Lett.* **98**, 182103 (2011).
- [41] F. Cardarelli, *Materials Handbook: A Concise Desktop Reference* (Springer, London, 2008).
- [42] <https://doi.org/10.17635/lancaster/researchdata/712>.