Thermoelectric conversion at 30 K in InAs/InP nanowire quantum dots

Domenic Prete,1 Paolo Andrea Erdman,1 Valeria Demontis,1 Valentina Zannier,1 Daniele Ercolani,1 Lucia Sorba,1 Fabio Beltram,1 Francesco Rossella,1 Fabio Taddei,1 and Stefano Roddaro1,2

1 NEST, Scuola Normale Superiore and Istituto Nanoscienze-CNR, Piazza S. Silvestro 12, I-56127 Pisa, Italy
2 Dipartimento di Fisica “E. Fermi”, Università di Pisa, Largo Pontecorvo 3, I-56127 Pisa, Italy

We demonstrate high-temperature thermoelectric conversion in InAs/InP nanowire quantum dots by taking advantage of their strong electronic confinement. The electrical conductance $G$ and the thermopower $S$ are obtained from charge transport measurements and accurately reproduced with a theoretical model accounting for the multi-level structure of the quantum dot. Notably, our analysis does not rely on the estimate of co-tunnelling contributions since electronic thermal transport is dominated by multi-level heat transport. By taking into account two spin-degenerate energy levels we are able to evaluate the electronic thermal conductance $K$ and investigate the evolution of the electronic figure of merit $ZT$ as a function of the quantum dot configuration and demonstrate $ZT \approx 35$ at 30 K, corresponding to an electronic efficiency at maximum power close to the Curzon-Ahlborn limit.

The progress in the fabrication and control of nanostructured systems has opened new prospects for thermoelectric (TE) research [1, 2] and has provided new ways to create improved TE devices [3–7]. In particular, quantum dots (QDs) were soon identified as ideal systems for the implementation of efficient heat engines [2] and for the creation of nanoscale thermometers [8]. The key property is their discrete density of states that yields a strong energy selectivity in their transmission profile thus opening the way to the realization of TE systems with an optimized performance [2]. As a consequence, TE effects in QDs have been extensively investigated both theoretically [9–32] and experimentally [33–43]. The TE behavior of electronic devices is characterized by the so-called figure of merit $ZT$ that is larger in more efficient devices. Indeed, within linear-response regime, both the maximum efficiency and the efficiency at maximum power are solely determined by the figure of merit, and they are growing functions of $ZT$ [7]. The figure of merit is defined as $ZT = GS^2T/K$, where $G$ is the electrical conductance, $S$ is the Seebeck coefficient, $K$ is the thermal conductance and $T$ is the temperature. In bulk materials the maximization of $ZT$ has been proved to be highly non-trivial since $G$, $S$ and $K$ are intrinsically related. On the contrary, heat and charge transport in QD-based TE devices can be disentangled and larger values of $ZT$ can be achieved. The performance of thermal machines based on QDs has been theoretically studied by many authors [22, 25, 44–48], while only one experimental investigation is available up to now [40]. We also note that QDs can provide a key building block for the fundamental investigation of quantum and stochastic thermodynamic effects, as shown for small-sized systems, where thermal fluctuations are of much relevance [49–51].

The investigation of TE effects in single-electron systems is performed in temperature regimes for which the thermal energy $k_BT$ is smaller than the Coulomb gap. In addition, it is often desirable to implement devices where the individual energy levels are well-resolved with respect to the thermal energy so that only few levels contribute to the heat transport. Here, we exploit the strong confinement of QDs realized in InAs/InP heterostructured nanowires (NWs) [52–54] to demonstrate field-effect control on the thermopower $S$ up to temperatures of the order of 30 K. The high temperature regime has been little explored in similar systems [39], usually operated at much lower temperatures [38, 40]. Our devices allow for the application of a temperature gradient along the NW, the measurement of charge current and the Seebeck coefficient, as well as the direct measurement of local temperatures. The strong confinement of the electrons in our QDs allows us to obtain charging energies $E_C$ exceeding 10 meV, with inter-level spacings $\Delta\varepsilon$ of the order of 5 meV and thus to operate our devices at relatively high temperatures and close to the QD depletion. Using a theoretical model accounting for sequential tunnelling mediated by two spin-degenerate levels [18, 44, 55] we can accurately reproduce the experimental conductance and thermopower. This allows us to estimate the electronic thermal conductance $K$, the power factor $GS^2$ and the electronic figure of merit $ZT$.

Fig. 1(a) shows a scanning electron micrograph of one of the investigated devices, where different colors correspond to different elements. The device core is constituted by an InAs/InP heterostructured NW with a diameter of $52 \pm 1$ nm. As visible in the Fig. 1(b), the nanostructure embeds a 20 $\pm$ 1 nm InAs island separated from the rest of the NW by two $\approx 4$ nm-wide InP barriers. The NWs used for this study were grown by Au-seeded chemical beam epitaxy [56] and have a wurtzite crystal structure. As grown NWs were detached from the growth substrate by sonication in isopropyl alcohol (IPA) and randomly deposited on a SiO$_2$/Si p++ substrate by drop-casting and contacted by e-beam lithography followed by an evaporation of a metallic Ti/Au (10/100 nm) bilayer. The resulting device layout includes a metallic serpentine heater (red), which can be fed by a current $I_H$ to induce, thanks to Joule heating, a thermal gradient along the NW (green). Our architecture also includes a set of multiple contact electrodes (yellow) that allows to: (i) apply a voltage bias $V_{\text{bias}}$ and measure the current $I_{DS}$ flowing through the NW; (ii) perform a local mea-
FIG. 1. (a) Scanning electron micrograph of a typical device. A local heater (red) is used to establish a temperature difference $\Delta T$ between the two ends of the NW (green) embedding an InAs/InP heterostructured QD. The device is fabricated on top of a degenerately-doped SiO$_2$/Si p++ substrate (gray) and a set of Ti/Au electrodes (yellow) can be used both as electrical contacts to the NW and as local resistive thermometers. The QD electronic configuration can be controlled with a pair of side gates (SG, purple) or using the conductive substrate as a backgate electrode (BG). (b) Transmission electron micrograph of the heterostructured QDs. (c) Sketch of the energetics scheme: the QD implements a multi-level system that can mediate heat and charge transport between the a source (S) and a drain (D) electrodes, in the presence of thermal and electric biases. Two spin-degenerate levels $\varepsilon_1$ and $\varepsilon_2$ ($\Delta \varepsilon = \varepsilon_2 - \varepsilon_1$) play a relevant role in the regime studied in the experiment.

The control on the temperature of the system is obtained by setting the temperature of the bath $T_{bath}$ in which the device is immersed and then by setting a temperature gradient across the NW so that its hot end is at a temperature $T_H$ and its cold end is at a temperature $T_C$. The average temperature $T_{avg} = (T_H + T_C)/2$ is set high enough ($T_{avg} \sim 30$ K) so that we can benefit of the following advantages. (i) Since $\hbar \Gamma \ll k_B T_{avg}$ (where $\Gamma$ is the characteristic tunnelling rate through the QD barriers) and $\Delta \varepsilon$ is not much larger than $k_B T_{avg}$, sequential tunnelling processes dominate the thermoelectric transport and data analysis is relatively straightforward (see below for details); (ii) Since $\Delta T \ll T_{avg}$, non-linear response effects in thermal bias $\Delta T$ are negligible thus simplifying the theoretical analysis; (iii) charge transport can be explored for relatively large applied bias voltages (up to $V_{bias} \approx k_B T_{avg}/e$) while remaining in the linear response regime. In this situation, the current $I_{DS}$ flowing through the QD is given by

$$I_{DS} = G(V_{bias} + S\Delta T),$$

where $G$ is the electrical conductance and $S$ is the Seebeck coefficient. The thermovoltage at open circuit conditions is thus given by $V_{th} = -S\Delta T$. It should be noted that both $G$ and $S$ are temperature dependent even iif not explicitly indicated. Typical transport data from our QDs can be seen in Fig. 2(a), reporting a colorplot of $I_{DS}$ as a function of the gate voltage $V_{gate}$ and the DS bias $V_{bias}$ at $T = 4.2$ K. Any of the three avail-
able gates (SG1, SG2 and BG) can be used to operate the device, but in the experiment reported here the QD was controlled using the lateral gate SG1 because it yielded the most stable control of the electronic configuration. Coulomb diamonds are clearly visible for various QD filling numbers $N$, and indicate a charging energy on the first spin-degenerate level $E_c = 14 \pm 2$ meV and an inter-level spacing $\Delta \varepsilon = 5 \pm 1$ meV. In Fig. 2(a) $N = 0$ corresponds to a completely depleted QD, as further discussed in the Supplementary Information and as indicated by high-temperature transport data. However, the behavior of the device is not expected to critically depend on the presence of filled levels as long as they are far away in energy with respect to the relevant energy scales, in particular $k_B T_{\text{avg}}$.

In Fig. 2(b) we report a set of $I_{DS}(V_{\text{bias}})$ curves corresponding to vertical cross-sections at fixed $V_{\text{gate}}$ taken from the region highlighted by the black rectangle in panel (a). The thermoelectric response of the device has been obtained from transport data in the linear regime. From the curve $I_{DS}(V_{\text{bias}})$ curve restricted to the linear response range (see e.g. Fig 2 (b)), we determine $G$ (slope) and $I_{\text{th}} = G S \Delta T$ (intercept) for each $V_{\text{gate}}$ value and for a fixed $\Delta T$ using Eq. (1). From the knowledge of the slope and intercept of the lines it is possible to compute $V_{\text{th}} = -S \Delta T$. Fig. 2(c)-(f) show the plots of $G(V_{\text{gate}})$ and $-V_{\text{th}}(V_{\text{gate}})$ for two different bath temperatures and heating currents $I_{\text{th}}$, such that the average temperatures of the QD are $T_{\text{avg}} = 24.9$ K in panels (c) and (e) and $T_{\text{avg}} = 31.2$ K in panels (d) and (f), respectively.

Experimental data are analyzed with a theoretical model based on a master equation accounting for sequential tunneling of electrons between the leads and the QD [9, 44, 58]. As already mentioned in the introduction, we can reasonably assume that co-tunneling processes are negligible when two conditions are met, namely $h \Gamma \ll k_B T_{\text{avg}}$ and $\Delta \varepsilon \lesssim k_B T_{\text{avg}}$. The first condition is required since small values of $h \Gamma/k_B T_{\text{avg}}$ suppress the coefficient of co-tunneling rates. The second condition comes from the fact that co-tunneling contributions decay as a power law with the energy difference $\delta E$ between the chemical potential of the leads and the nearest resonance (measured in units of $k_B T_{\text{avg}}$) [58]. This means that they may prevail over sequential contributions for large enough $\delta E$, since sequential tunneling contributions decay exponentially with $\delta E/k_B T_{\text{avg}}$ [55, 58, 59]. Remarkably, the presence of two spin-degenerate levels, with spacing equal to $\Delta \varepsilon$, sets a bound to the energy difference $\delta E$. The second condition (for which $\Delta \varepsilon \lesssim k_B T_{\text{avg}}$), therefore, ensures that the exponential suppression of sequential tunneling is limited and we can safely disregard co-tunneling effects. This is particularly relevant for values of $V_{\text{gate}}$ such that the chemical potential of the leads is between the two spin-degenerate levels, where the thermopower is maximal even though the conductance is negligible.

Two spin-degenerate QD orbitals $\varepsilon_1$ and $\varepsilon_2$ are taken into account in the QD model and tunneling rates are assumed to depend on the number of electrons $N$ in the QD (see Supplementary Information for details about the model). Further important outputs of the fitting procedure are the electronic temperatures $T_{H, \text{fit}}$ and $T_{C, \text{fit}}$ that can be used to obtain a better estimate of the net temperature bias across the QD structure $\Delta T_{\text{fit}}$. The average fit temperature $T_{\text{avg,fit}} = (T_{H, \text{fit}} + T_{C, \text{fit}})/2$ matches nicely with the one obtained experimentally sing the metallic thermometers $T_{\text{avg}}$, and confirms the good calibration of the resistive thermometers. On the contrary, $\Delta T_{\text{fit}}$ is sizably smaller than $\Delta T$. This is not very surprising since the NW has a finite heat conductance leading to a partitioning of the thermal bias as observed in experiments at lower temperatures [60]. Since $\Delta T_{\text{fit}}$ is the relevant temperature bias for the analysis of the TE response of the QD, it has been used for all the subsequent estimates on the TE parameters of the nanostructure.

Fig. 3(a) and (b) show the experimental data (red dots) for the case $T_{\text{avg}} = 31.2$ K, together with the fitted curves (solid lines) obtained using the model ($\Delta T_{\text{fit}} = 2.05$ K). Different fitting models are compared: the solid black curve accounts for two spin-degenerate energy levels in the QD (here referred to as the “four-level” model); the green dashed curve accounts for a single (spin-degenerate) energy level (referred as the “two-level” model), as often done in the literature to model single [61–65] and double quantum dot system [25, 46, 66–68]. We notice that the four-level model nicely fits the experimental data for $G$ and $S$ in the whole range of values of $V_{\text{gate}}$ under consideration, while the two-level model is accurate only for values $V_{\text{gate}} < 1.6$ V in the conductance plot (the first two peaks in Fig. 3(a)) and up to $V_{\text{gate}} \simeq 1$ V in the thermopower plot (Fig. 3(b)). The theory can now be employed to calculate the electronic thermal conductance $K$ by using the fitting parameters obtained from the experimental data $G(V_{\text{gate}})$ and $S(V_{\text{gate}})$ (see Supplementary Information for details about values obtained for the parameters). The resulting $K(V_{\text{gate}})$ curves for the two models are plotted in Fig. 3(c). Remarkably, the curve relative to the two-level model departs even in the first peak from the curve of the four-level model. This is due to the fact that $K$, by definition, is the ratio between the heat current and the temperature difference in open circuit conditions, i.e. when $I_{DS}$ is zero. In a sequential single-level model, all electrons tunnel through the QD at the same energy; thus the same amount of heat is transferred to/from a given reservoir in each tunneling event. If $I_{DS} = 0$, there is not net transfer of charge, thus there is no net transfer of heat. This implies $K = 0$, regardless of the gate voltage. Conversely, if there are two or more energy levels, electrons can tunnel at different energies, and this possibility allows for a net heat transfer at zero charge current [18, 44]. Therefore, as opposed to $G$ and $S$, the value of $K$ in the sequential regime is fully determined by the multilevel structure of the QD. It is thus crucial to employ the four-level model to estimate it properly.

Using the three transport coefficients $G$, $S$ and $K$ one
can evaluate the performance of our QD system when operated as a heat engine. Indeed, the electric response to a temperature difference can be used to drive a current against a voltage bias, effectively extracting work from the system. The efficiency of a heat engine is then defined as the ratio between the extracted work and the heat provided by the hot reservoir. At fixed $V_{\text{gate}}$ and temperature difference, we define $P_{\text{max}}$ as the maximum power that can be extracted by optimizing over the applied $V_{\text{bias}}$. The corresponding efficiency will then be the efficiency at maximum power $\eta(P_{\text{max}})$. It can be shown that $P_{\text{max}}$ also corresponds to the maximum power that can be performed on a variable load resistance in series with the QD. However, as shown in Ref. [7], within the linear response regime both $P_{\text{max}}$ and $\eta(P_{\text{max}})$ can be computed simply from the knowledge of the transport coefficients $G$, $S$ and $K$ previously shown, i.e. $P_{\text{max}} = GS^2\Delta T^2/4$ and $\eta(P_{\text{max}}) = \eta_C/2 \times ZT/(ZT + 2)$, where $ZT = GS^2T/K$ is the electronic figure of merit.

The maximum power is plotted in Fig. 4(a) as a function of the gate voltage $V_{\text{gate}}$ using the theoretical (solid curve) and experimental (red circles) values of $G$ and $S$ together with the fitted value of $\Delta T_{\text{fit}}$. The experimental data (available down to $V_{\text{gate}} \simeq -1$ V) is well reproduced by the four-level model which presents several peaks, the...
largest one corresponding to $P_{\text{max}} = 0.37 \text{fW}$ occurring at $V_{\text{gate}} \approx -0.78 \text{V}$. Remarkably, the corresponding value of the electronic figure of merit [plotted in Fig. 4(b)] is $\eta \approx 0.47\eta_C$ (where $\eta_C$ is Carnot’s efficiency) for $V_{\text{gate}} \approx -0.78 \text{V}$, implying that our QD system can be operated at an electronic efficiency at maximum power very close to the Curzon-Ahlborn’s linear response upper bound $\eta_{\text{CA}} = \eta_C/2$. It is worth observing that the electronic efficiency at maximum power is computed from the knowledge of ZT [7]. Interestingly, Fig. 4(c) shows that the electronic efficiency at maximum power roughly behaves as the maximum power [Fig. 4(a)], implying that both the maximum power and corresponding efficiency can be simultaneously maximized.

In conclusion, we have explored TE phenomena in InAs/InP NW QD-based devices at high temperature in the linear regime. Experimental data were analyzed using a multi-level model based on the resolution of a master equation which allowed us to compute the electronic thermal conductance of the system. This, in combination with the experimental data of conductance and thermopower, allowed us to estimate the electronic thermoelectric figure of merit $ZT$ and the electronic efficiency at maximum power of our thermoelectric engine. We find that the ideal Curzon-Ahlborn’s upper bound is nearly attained, and that a figure of merit $ZT \approx 35$ is reached while extracting the maximum power from the system. This study demonstrates the full electrostatic control of the heat engine features of a thermally biased NW QD operating in high temperature regimes. Our results shed light on the operation of few level thermoelectric engines, a key issue for the physics and technology addressing heat and charge transport mediated by single carrier. The electronic ZT, which neglects the contribution of the phonons to the thermal conductance, is a sound characterization of the electronic properties of the device which allows to compare the electronic performance of different thermoelectric materials. Furthermore, it is directly related to the efficiency of non-equilibrium devices, such as solars cells which aim at recovering the energy of out-of-equilibrium "hot carriers" excited by light [69,70]. Possible applications include on-chip cooling, energy harvesting on cryogenic platforms and nanoscale thermometry.

We thank R. Fazio for helpful discussions and H. Courtois for carefully reading the manuscript. This work has also been supported by the SNS-WIS joint lab QUANTRA, by the SUPERTOP project, QUANTERA ERA-NET Cofound in Quantum Technologies and by the CNR-CONICET cooperation programme Energy conversion in quantum, nanoscale, hybrid devices.

**Materials and Methods.** InAs/InP heterostructured NWs were fabricated using a chemical beam epitaxy process seeded by metallic nanoparticles obtained from thermal dewetting of a Au thin film [71]. Growth was performed at $420 \pm 10 \text{°C}$ using trimethylindium (TMIn), tert-butylarsine (TBA), and tributylphosphine (TBP). The TBA and TBP are thermally cracked at around $1000 \text{°C}$ upon entering the growth chamber, while the TMIn decomposes on the substrate surface. The metallorganic pressures were 0.3, 1.0, and 4.0 Torr for TMIn, TBA, and TBP, respectively. These growth conditions ensure to achieve InAs/InP NW heterostructures with straight morphology, constant diameter, wurtzite crystal structure and atomically sharp interfaces [56,72]. InAs/InP and InP/InAs interfaces where realized without any interruption by switching the group V precursors. The position of the dot inside the NW is determined based on its average distance from the Au nanoparticle, as measured from an ensemble of wires with STEM. This leads to a typical alignment error of ±50 nm, based on NW imaging statistics and on alignment errors during the lithographic process. Ohmic contacts are obtained by thermal evaporation of a Ti/Au (10/100 nm) bilayer, after a chemical passivation step based on a standard (NH$_4$)$_2$S$_2$ solution. The NW voltage bias was obtained using the auxiliary output of a Stanford Research lock-in amplifier SR830 and a resistor divider. Charge current was measured using a DL1211 current preamplifier and an Agilent 34401A multimeter. Each of the thermometers was excited by an AC current of 1 $\mu$A, obtained using the sinusoidal output of a SR830 amplifier connected to a 1 MΩ bias resistor. The local heater was supplied using a Keithley 2400 SourceMeter. The gate voltage was controlled using a Yokogawa 7651 DC source.

**SUPPLEMENTARY INFORMATION**

- Experimental methods: measurement setup, thermometers and heaters calibration, device characterization;
- Details about the theoretical models involved in this study;

[1] L. D. Hicks and M. S. Dresselhaus, *Phys. Rev. B* **47**, 12727 (1993).
[2] G. D. Mahan and J. O. Sofo, *Proceedings of the National Academy of Sciences* **93**, 7436 (1996), [http://www.pnas.org/content/93/15/7436.full.pdf](http://www.pnas.org/content/93/15/7436.full.pdf).
[3] M. Dresselhaus, G. Chen, M. Tang, R. Yang, H. Lee, D. Wang, Z. Ren, J.-P. Fleurial, and P. Gogna, *Advanced Materials* **19**, 1043 (2007).
7

F. Beltram, Nano Letters 11, 1695 (2011), pMID: 21446718, http://dx.doi.org/10.1021/nl200209m.

[58] Y. V. Nazarov and Y. M. Banter, Quantum Transport (Cambridge, New York, 2009).

[59] D. V. Averin and Y. V. Nazarov, Single Charge Tunneling, edited by H. Grabert and M. H. Devoret (Springer US, 1992).

[60] E. A. Hoffmann, H. A. Nilsson, J. E. Matthews, N. Nakpathomkun, A. I. Persson, L. Samuelson, and H. Linke, Nano Letters 9, 779 (2009), pMID: 19159269, http://dx.doi.org/10.1021/nl8034042.

[61] P. Murphy, S. Mukerjee, and J. Moore, Phys. Rev. B 78, 161406 (2008).

[62] M. Esposito, K. Lindenberg, and C. V. den Broeck, EPL (Europhysics Letters) 85, 60010 (2009).

[63] O. Entin-Wohlman, Y. Imry, and A. Aharony, Phys. Rev. B 82, 115314 (2010).

[64] N. Nakpathomkun, H. Q. Xu, and H. Linke, Phys. Rev. B 82, 235428 (2010).

[65] F. Mazza, R. Bosisio, G. Benenti, V. Giovannetti, R. Fazio, and F. Taddei, New Journal of Physics 16, 085001 (2014).

[66] P. Trocha and J. Barnaś, Phys. Rev. B 85, 085408 (2012).

[67] B. Bhandari, G. Chiriacò, P. A. Erdman, R. Fazio, and F. Taddei, Physical Review B 98 (2018), 10.1103/physrevb.98.035415.

[68] P. A. Erdman, B. Bhandari, R. Fazio, J. P. Pekola, and F. Taddei, Physical Review B 98 (2018), 10.1103/physrevb.98.045433.

[69] S. Limpert, S. Brenner, and H. Linke, New Journal of Physics 17, 095004 (2015).

[70] S. Limpert, A. Burke, I.-J. Chen, N. Anttu, S. Lehmann, S. Fahlvik, S. Brenner, G. Conibeer, C. Thelander, M.-E. Pistol, and H. Linke, Nanotechnology 28, 434001 (2017).

[71] U. P. Gomes, D. Ercolani, V. Zannier, F. Beltram, and L. Sorba, Semiconductor Science and Technology 30, 115012 (2015).

[72] V. Zannier, F. Rossi, V. G. Dubrovskii, D. Ercolani, S. Battiato, and L. Sorba, Nano Letters 18, 167 (2017).