https://doi.org/10.1038/s41928-024-01335-5

The quantum limits of contact resistance and ballistic transport in 2D transistors

Deji Akinwande, Chandan Biswas & Debdeep Jena

The development of transistors based on two-dimensional semiconductors requires a consistent approach to calculating and evaluating quantum contact resistances.

The development of silicon transistor technology via dimensional downscaling is rapidly approaching its physical limits. As a result, two-dimensional (2D) semiconductors are of increasing interest in the development of next-generation energy-efficient electronics. Such atomically thin semiconductors can be used to reduce the channel length of field-effect transistors (FETs), maximizing current while reducing operational voltage and power. In highly engineered 2D FETs, channel lengths are now approaching the electron mean free path — around 10 nm for monolayer transitional metal dichalcogenides. Accordingly, charge transport is likely to approach dissipation-less ballistic transport.

In the ballistic limit, it is known that transport is limited by the contact resistance (R_c) (ref. 1), which represents the interfacial resistance from the metallic contacts (considered reservoirs of electrons) to the semiconductor channel (which offers a limited density of states). As with most physical phenomena in solid-state physics that experience dimensional reduction, as the channel is scaled below the mean free path, the transistor resistance does not vanish but rather approaches a minimum or quantum resistance contact (R_{cq}) limit[2]. It is thus becoming common in the community to use R_{cq} as a benchmark to judge the quality of contacts to scaled 2D semiconductors[3–6]. Eventually, for fully ballistic transport, the overall transistor device conductance under low-field operation will be around $1/R_{cq}$, as demonstrated for graphene nanoribbons[7].

There are, however, at least three different formulas for R_{cq} for 2D FETs made from the prototypical 2D semiconductor molybdenum disulfide (MoS_2). The general width-normalized expression in all the formulas takes the form of the Landauer–Buttiker formalism with a unity transmission (T) factor (that is, no Schottky barrier)[8]:

$$R_{cq}W = \frac{h}{2e^2}\left(\frac{1}{M_{\text{eff}}}\right) = \frac{h}{2e^2}\left(\frac{\alpha}{\sqrt{n_{2D}}}\right) \quad (1)$$

where the prefix is the quantum resistance derived from Planck's constant, h, and the electron charge, e, including spin degeneracy ($g_s = 2$). W is the channel width, M_{eff} is the (effective) number of modes, n_{2D} is the carrier density in the 2D semiconductor and α is a unitless parameter. To be precise, equation (1) represents ground-state conditions with an ideal parabolic dispersion and is often used to benchmark the contact resistance in scaled devices.

In some work[2,6], $\alpha = 2/\pi$ is used, resulting in $R_{cq}W \approx 26/\sqrt{n_{2D}}$; in other work[4], $\alpha = \sqrt{\pi}/2$ is used, resulting in $R_{cq}W \approx 36/\sqrt{n_{2D}}$; and in other work[1], $\alpha = \sqrt{\pi/2}$ is used, resulting in $R_{cq}W \approx 51/\sqrt{n_{2D}}$. Perhaps owing to this ambiguity, some work[3] has used $\alpha = \sqrt{\pi/2}$ in one place (the discussion) and $\alpha = \sqrt{\pi}/2$ in another (the benchmarking plot). (n_{2D} and $R_{cq}W$ are in units of 10^{13} cm^{-2} and Ω µm, respectively.)

To help advance the study of contacts and ballistic transport in 2D semiconductors — and prevent erroneous conclusions — it is essential that we have a precise and consistent formula.

Exact derivation of quantum contact resistance

An exact ground-state expression for R_{cq} can be derived via several methods. These include the semiclassical analysis of ballistic transport, which was used by Sharvin to study the ballistic conductance of a classical point contact in metal physics in the 1960s[9], and the later Landauer–Buttiker formalism, which is used in mesoscopic physics to study quantized transport such as quantum point contacts[10]. Here we use a continuum analysis that derives from band transport with no assumptions except for ground-state temperature.

For a 2D semiconductor connected to ideal source/drain contacts ($T = 1$), the application of an external energy stimulus (Δ) perturbs the Fermi–Dirac step function about the Fermi energy (E_F). Employing an n-type semiconductor as a model and requiring degenerate doping ($E_F >$ conduction band minimum (CBM)) to ensure band transport, the resulting near-equilibrium net current (I) is due to states within Δ of the Fermi surface (Fig. 1a):

$$I = 2e\frac{A}{L}g_v \int_{E_F-\Delta}^{E_F+\Delta} \frac{g(E)}{2}\langle v(E)\cos\theta\rangle\, F(E-E_F)\,dE, \quad (2)$$

where the factor of 2 in the numerator is for spin degeneracy, g_v is for valley degeneracy, $F(\cdot)$ is the Fermi–Dirac statistics and $v(E)$ is the electron group velocity. A is the area — width (W) × length (L) — of the semiconductor channel. The area-normalized density of states per spin is $g(E)$, where the factor of 1/2 indicates that only half of the states can contribute to the net current in the channel, say, positive velocity (right moving) states from source to drain (Fig. 1b). The average velocity, $\langle\bullet\rangle$, is needed to account for the angular distribution of velocities at any constant energy surface in the band structure. In 2D space, this angular average is taken over a semi-circle representing the right moving momentum states (k) and is given by $\int_{-\pi/2}^{\pi/2} v(E)\cos\theta\, d\theta = (2/\pi)v(E)$.

Equation (2) can be reduced to a more fundamental expression by replacing band properties, $g(E)$ and $v(E)$, with their basic definitions:

$$I = 2eW\frac{g_v}{\pi}\int_{E_F-\Delta}^{E_F+\Delta}\left(\frac{dN}{dE}\right)\left(\frac{1}{\hbar}\frac{dE}{dk}\right)F(E-E_F)\,dE, \quad (3)$$

where $\hbar$ is the reduced Planck's constant and N is the area-normalized number of states, $N = k^2/4\pi$. Afterwards, R_{cq} can be expressed in terms of the derivative of $F(\cdot)$, which becomes a Dirac-delta function (δ) owing to the step profile of the ground-state Fermi–Dirac statistics:

$$R_{cq}^{-1} = \frac{dI}{dV} = \frac{2e^2}{h}\frac{Wg_v}{\pi}\int_{E_F-\Delta}^{E_F+\Delta} k(E)\,\delta(E-E_F)\,dE = \frac{2e^2}{h}\left(\frac{Wg_v}{\pi}k_F\right). \quad (4)$$

The item in parenthesis can be considered M_{eff}. This equation reveals that transport is determined by the Fermi momentum (k_F), channel width and the valley degeneracy. Indeed, the case for $g_v = 1$ is the Sharvin conductance of a ballistic point contact in two dimensions[9,10].

It is also helpful to express k_F in terms of the carrier density, a more controllable parameter in experimental devices. Considering a parabolic dispersion, $k_F = \sqrt{2\pi n_{2D}/g_v}$, this leads to:

$$R_{cq}W = \frac{h}{2e^2}\left(\sqrt{\frac{\pi}{2}}\frac{1}{\sqrt{g_v n_{2D}}}\right) \approx \frac{51}{\sqrt{g_v n_{2D}}}\ \Omega\,\mu m \quad (5)$$

where n_{2D} is in units of 10^{13} cm^{-2}. This result is the particular $d = 2$ case of a generalized d-dimensional quantum contact resistivity given by

$$\rho_{cq} \approx \frac{h}{e^2}\left(\frac{1}{d^{1/6}}\frac{1}{(g_s g_v)^{1/d}}\frac{1}{(n_d)^{(d-1)/d}}\right) \quad (6)$$

The derivation of this general result will be presented in a separate work.

R_{cq} metric for contact resistance and ballistic transport

In the case of semiconductors with a single valley ($g_v = 1$), $R_{cq}W \approx 51/\sqrt{n_{2D}}$, in agreement with (equation (6.3.8) in ref. 11). Examples include 2D heterostructure quantum wells made from semiconductors with a minimum at the Γ point such as GaAs or some p-type monolayer transitional metal dichalcogenides with strong spin–orbit coupling, which lifts the valence band valley degeneracy per spin at the K points of the Brillouin zone. However, for monolayer MoS_2 and similar transitional metal dichalcogenides, the CBM is at the K point of the Brillouin zone and has a valley degeneracy of 2. Hence, $R_{cq}W \approx 36/\sqrt{n_{2D}}$ is the appropriate formula. Inversion layers in silicon due to quantization also have double degeneracy for the CBM[11].

It is also worth reflecting on the meaning of R_{cq} and its application to transistor devices, whose total resistance (R_{tot}) consists of the semiconductor channel ($R_{channel}$) plus its source/drain contacts (R_c): $R_{tot} = 2R_c + R_{channel}$. A common assumption is that the two contacts to the channel are symmetric, yielding a total contact resistance of $2R_c$.

Historically, Sharvin's analysis of ballistic transport through a narrow constriction with semi-infinite electrodes[9] showed that the point contact conductance was the total conductance of the constriction determined by the number of available momentum states at the Fermi level and width. Subsequently, quantization of the conductance of the quantum point contact made from a GaAs–AlGaAs semiconductor mesoscale constriction was reported[10]; this quantum conductance was also understood to be the total conductance of the semiconductor constriction[12]. Similarly, this has been the case for two-point conductance measurements of ballistic graphene nanoribbons[7].

In this regard, for scaled semiconductors, R_{cq} should be interpreted as the total resistance in the limit of ballistic transport with ideal Schottky-barrier-free interfaces. In this limit, $R_{cq} = 2R_c$ as the channel becomes dissipation free. Furthermore, recent reports[3,4,6] have demonstrated that semimetal contacts (such as Bi and Sb) can avoid metal-induced gap states caused by metal–transition metal dichalcogenide contact. This approach could be suitable for realizing dissipation-free ballistic transport.

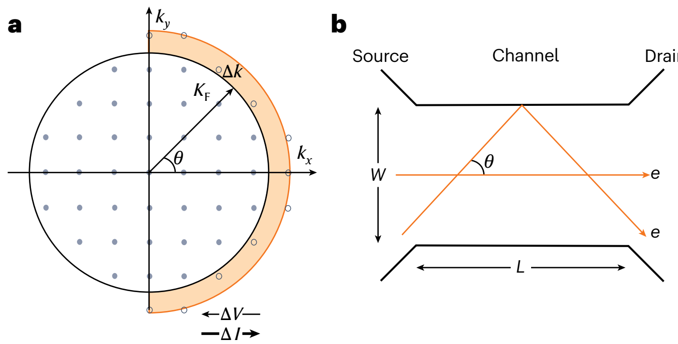

Fig. 1 | Fermi surface and transport in a ballistic channel. a, The Fermi surface in 2D momentum space. The ballistic current in the x direction is from the angular average of the momentum states in the shaded area activated by the external electric potential. The Fermi wavevector and electric potential difference are denoted as K_F and ΔV, respectively. **b**, Transport in a ballistic channel of width, W, and length, L, connected to source and drain contacts serving as ideal electron reservoirs. The illustration depicts two forward trajectories: one in a straight path, and the other at an angle (θ) that undergoes specular reflection.

In reports on advanced 2D FETs, there is currently a practice to use R_{cq} as a metric to benchmark individual contacts[1–6], which is a desirable trend towards extracting the maximum current density in scaled devices. However, R_c is compared with R_{cq} in these reports featuring otherwise diffusive channel lengths. We argue that the precise comparison is to contrast $2R_c$ with R_{cq}, reflecting the underlying quantum physics. Precise comparison also avoids the eventual scenario of a truly ballistic experimental 2D FET, whose total resistance will then be around $2R_c$; the current practice of using R_c will then result in an error of about 2.

Using $2R_c$ as the appropriate experimental metric, we benchmark it to R_{cq} for recent reports of advanced 2D FETs (Fig. 2), which indicates a wider gap to the theoretical quantum limit than currently believed. Ultimately, to advance 2D FETs, the total resistance should be reported and compared with R_{cq} for near-ballistic or ballistic devices[5]. Indeed, the ballistic ratio or ballisticity parameter (β) – a metric to quantify the degree to which carrier-density dependent transport is ballistic – could be redefined as $\beta = 2R_c/R_{tot}$, which has a maximum of unity. This ballisticity parameter based on contact resistance measurements could be helpful to experimentalists in assessing transport in scaled 2D transistor devices.

Outlook

This article makes two key points, that the analysis and development of 2D transistors requires, (1) consistent use of the appropriate formula and value of the quantum limit of the ballistic contact resistance, and (2) the benchmarking of the total contact resistance ($2R_c$) – rather than R_c – to the quantum limit. On the basis of R_{cq}, the maximum available current density in scaled ballistic 2D channel devices is estimated to exceed 3 mA μm^{-1}, indicating considerable room for improvement compared with contemporary 2D FETs. Evidence of ballistic transport can be established with temperature-independent or length-invariant current measurements, which can be supplemented with a total resistance profile that should feature an inverse dependence on the square root

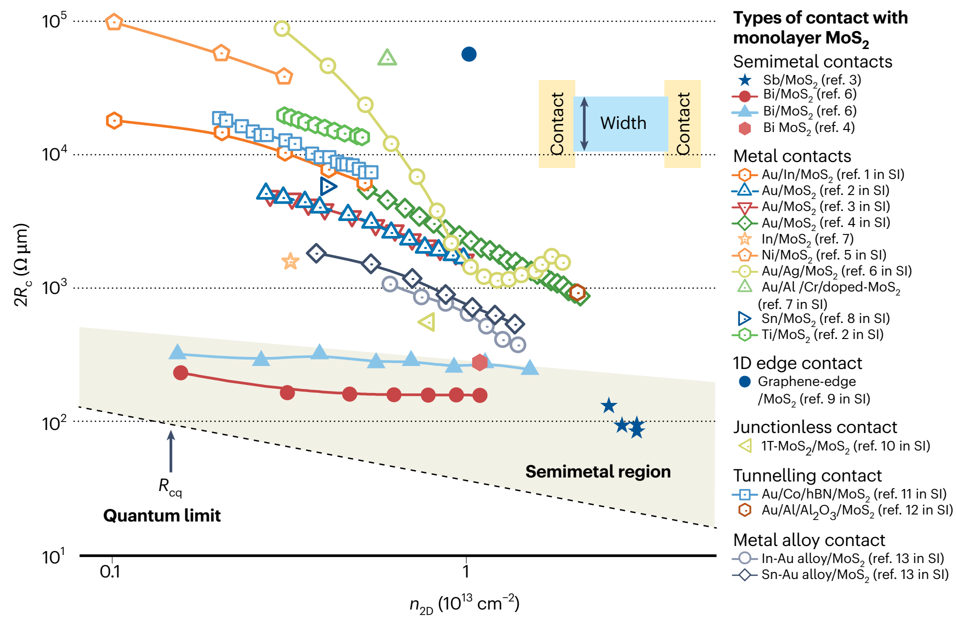

Fig. 2 | Benchmarking of total contact resistance, $2R_C$, versus n_{2D} plot of experimental MoS_2 monolayer transistors. Total contact resistance, $2R_C$, rather than R_C, is the appropriate parameter to compare with the ground-state ballistic quantum limit. Semimetal contacts currently provide the lowest contact resistances approaching the quantum limit. All measurements were reported at room temperature, except the Bi/MoS_2 data (red connected circles), which was reported at 15 K. The inset image illustrates a typical transistor configuration. See Supplementary Information (SI) for specific values and references for data points.

of the carrier density at low temperatures. Future theoretical analysis should consider deriving compact analytical formulas including the effects of band non-parabolicity and temperature on the quantum limit of ballistic contact resistance, which will better reflect practical 2D materials at operational conditions.

Deji Akinwande[1] ✉, **Chandan Biswas**[1] **& Debdeep Jena**[2]

[1]Microelectronics Research Center, The University of Texas at Austin, Austin, TX, USA. [2]Electrical and Computer Engineering, and Materials Science and Engineering, Cornell University, New York, NY, USA.
✉e-mail: deji@ece.utexas.edu

Published online: 07 February 2025

References

1. Wang, Y. & Chhowalla, M. *Nat. Rev. Phys.* **4**, 101–112 (2022).
2. Jena, D., Banerjee, K. & Xing, G. H. *Nat. Mater.* **13**, 1076–1078 (2014).
3. Li, W. et al. *Nature* **613**, 274–279 (2023).
4. Shen, P.-C. et al. *Nature* **593**, 211–217 (2021).
5. Jiang, J., Xu, L., Qiu, C. & Peng, L.-M. *Nature* **616**, 470–475 (2023).
6. Mondal, A. et al. *Nat. Nanotechnol.* **19**, 34–43 (2024).
7. Baringhaus, J. et al. *Nature* **506**, 349–354 (2014).
8. Buttiker, M., Imry, Y., Landauer, R. & Pinhas, S. *Phys. Rev. B.* **31**, 6207–6215 (1985).
9. Sharvin, Y. V. J. *Exp. Theor. Phys.* **16**, 517 (1965).
10. van Wees, B. J. et al. *Phys. Rev. Lett.* **60**, 848 (1988).
11. Datta, S. *Quantum Transport: Atom to Transistor* (Cambridge Univ. Press, 2005).
12. van Houten, H., van Wees, B. J. & Beenakker, C. W. J. in *Physics and Technology of Submicron Structures* (eds Heinrich, H. et al.) 198–207 (Springer, 1988).

Acknowledgements

This work was supported in part by a National Science Foundation (NSF) Future of Semiconductors (FuSe) grant 2329191, and by Samsung. D.A. acknowledges the Cockrell Family Regents Chair Professorship. D.J. acknowledges the support of the Joint University Microelectronics Program (JUMP) from DARPA and the Semiconductor Research Corporation (SRC).

Author contributions

D.A. led the comment analysis with technical contribution from D.J. Data representation was performed by C.B. The Comment was written by all authors.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41928-024-01335-5.

Peer review information *Nature Electronics* thanks the anonymous reviewers for their contribution to the peer review of this work.