Time-resolved imaging of negative differential resistance on the atomic scale

Rashidi, Mohammad; Taucer, Marco; Ozfidan, Isil; Lloyd, Erika; Labidi, Hatem; Pitters, Jason L.; Maciejko, Joseph; Wolkow, Robert A.

Publisher's version / Version de l'éditeur:

Condensed Matter: Mesoscale and Nanoscale Physics, 2016-08-23

Time-Resolved Imaging of Negative Differential Resistance on the Atomic Scale

Mohammad Rashidi[1,2*], Marco Taucer[1,2], Isil Ozfidan[1], Erika Lloyd[1], Hatem Labidi[1,2], Jason L. Pitters[2], Joseph Maciejko[1,3], and Robert A. Wolkow[1,2]

[1]Department of Physics, University of Alberta, Edmonton, Alberta, T6G 2J1, Canada.

[2]National Institute for Nanotechnology, National Research Council of Canada, Edmonton, Alberta, T6G 2M9, Canada.

[3]Canadian Institute for Advanced Research, Toronto, Ontario M5G 1Z8, Canada

*Correspondence to: rashidi@ualberta.ca

Negative differential resistance remains an attractive but elusive functionality, so far only finding niche applications. Atom scale entities have shown promising properties, but viability of device fabrication requires fuller understanding of electron dynamics than has been possible to date. Using an all-electronic time-resolved scanning tunneling microscopy technique and a Green's function transport model, we study an isolated dangling bond on a hydrogen terminated silicon surface. A robust negative differential resistance feature is identified as a many body phenomenon related to occupation dependent electron capture by a single atomic level. We measure all the time constants involved in this process and present atomically resolved, nanosecond timescale images to simultaneously capture the spatial and temporal variation of the observed feature.

Half a century ago, William Shockley, upon hearing of Esaki's new diode and the strange negative differential resistance (NDR) it displayed (*1*), declared the Esaki diode, or tunnel diode, would play a dominant role in the then rapidly emerging field of semiconductor electronics. However, challenges related to CMOS fabrication compatibility have persistently blocked the broad application of tunnel diodes. Because of the continued attraction of reduced device count and improved performance that NDR offers (*2*), researchers have continued to this day to seek physical mechanisms that yield NDR (*3–8*). Indications of NDR at the nanoscale have been enticing (*9*, *10*) but resisted reduction to practical implementation. Molecular-scale candidate assemblies have likewise fallen short of practical requirements and in some cases it has been revealed that molecular degradation occurring during measurement results in a current voltage trace that transiently masquerades as NDR only to fail shortly thereafter (*11*, *12*). Because of the false starts, a robust and well understood mechanism for NDR is desirable.

Well characterized atomscale ensembles, probed with the tip of a scanning tunneling microscope (STM) have recently led to renewed hope. One particular embodiment involves a silicon dangling bond (DB) on an otherwise boron terminated silicon (111) surface (*13*, *14*). Crucially, the observation of NDR is reproducible and does not require a chance unknown contamination of the metal scanned probe as had been the case in an early report (*10*, *15*). While a practical scalable fabrication process has not been reported, it is at least conceivable that one could be found. But before that practical step is taken, two more basic challenges must be overcome to create a robust basis for NDR. One issue is that DBs on the boron silicon surface occur at random during sample preparation and cannot be deliberately fabricated. Practical implementation requires control over placement of individual DBs and density of multiple DBs.

Secondly, while a logical description of factors playing a role has been offered, no predictive theory has yet been established.

In this work, we address both of these issues. We show that DBs on a hydrogen-terminated Si(100) surface (H:Si), which can be patterned with atomic precision (*16–18*), exhibit the desired NDR. In addition to scanned probe-based current-voltage spectroscopy as has been used before, we apply a new all-electronic time-resolved STM technique (*19–25*) for the first time to isolate single atomic state carrier capture events, including measurement of an absolute capture rate on the nanosecond scale. New variants of that technique yield nanosecond scale temporal resolution without loss of atom-scale spatial imaging to reveal the lateral variation of the NDR effect. A Green's function transport model gives an account of the various rates participating in the process at different bias and tip height regimes. This model convincingly reproduces the measured data and thereby establishes a new mechanism to explain the physical process underlying the atom scale NDR process.

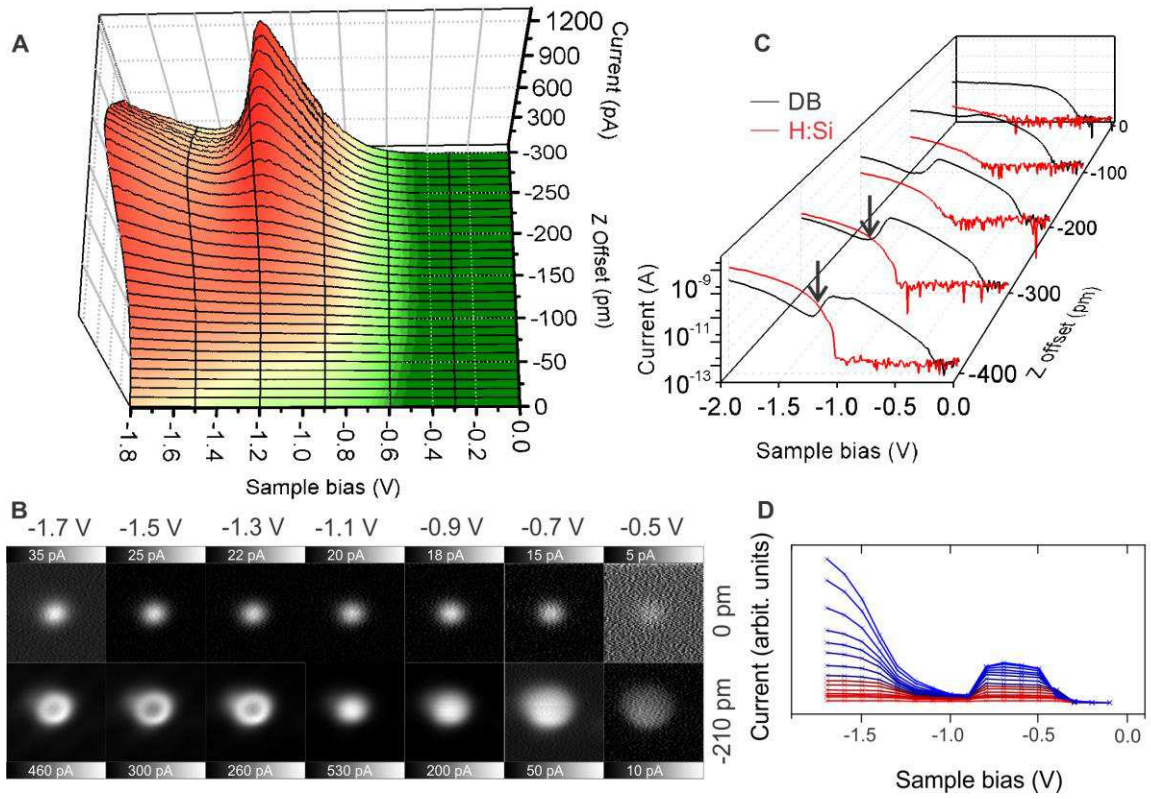

Fig. 1. Tunneling spectroscopy of DBs on degenerately arsenic-doped H:Si(100) surface. (A) Typical *I* (*V*) spectra of a single DB at different tip-sample distances measured at 4.5 K. The initial tip height (zero tip offset) is set with the tunneling conditions of -1.80 V and 30.0 pA on top of the DB. The tip is moved subsequently 310 pm closer to the DB in 10 pm increments. (B) Constant height STM images of the DB in (A) at different tip-height offsets (indicated on the right side) and sample bias voltages (indicated on top of each column). The tunneling current range for each STM image is indicated inside the color-bars attached to their frame. The size of the images is 3×3 nm^2. (C) Semi-log plots of *I*(*V*) spectra at different tip-height offsets on a DB (different DB than the one shown in A and B) and H:Si at the same tip-height. The zero tip height is set under the tunneling condition of -2.0 V and 50 pA on the DB. (D) Calculated *I*(*V*) spectra of a single silicon dangling bond at different tip-sample distances using NEGF.

A DB has a single localized level in the bulk band gap that can hold up to 2 electrons. A DB with no electrons will be in a positive charge state (DB^+), a single electron puts the DB in a neutral charge state (DB^0), and the addition of a second electron brings the DB to a negative charge state (DB^-). DB^+, DB^0, and DB^- should be regarded as three occupations of a single state with three distinct energy levels. Peaks in differential STM $I(V)$ spectroscopy do not directly report the charge states (*26*), DB^+, DB^0, and DB^-, rather, charge transition levels denoted (+/0) and (0/-) are observed. These levels correspond to the energies that must be imparted to an electron for it to drive a transition between charge states. The charge transition energies are due to polarization, relaxation, and in the case of (0/-) the on-site interaction energy, or Hubbard U, associated with the addition of an electron.

The measurements are performed on degenerately arsenic-doped H:Si heated to 1050ºC for oxide desorption (*27*). This preparation temperature leads to negligible depletion of dopants near the surface (28). Figure 1A shows $I(V,Z)$ spectra of a single DB recorded with an STM probe at the DB center and at different tip-height offsets. Unlike the results on samples with substantial dopant depletion region at the surface (*25*, *29*), the current onset arises in the bulk band gap. This indicates that the DB is being supplied by the bulk conduction band. For close tip-surface distances, the tunneling spectra exhibit a current drop (NDR) at approximately -1.20 V. For bias voltages more negative than the NDR onset, $I(Z)$ spectra exhibit a peak, while for bias voltages less negative than the NDR onset, the tunneling current is exponentially dependent on the tip-height. Corresponding constant height STM images for various bias voltages and two different tip-height offsets are shown in Fig. 1B. For closer tip-surface distances (-210 pm tip-height offset is shown as an example) the DB exhibits a dark center at bias voltages more

negative than the NDR onset, indicating the spatial dependence of the NDR feature. The slight asymmetry in the images is due to an asymmetric tip shape.

Figure 1C compares $I(V)$ spectra at different tip heights over a DB and H:Si. The spectra with -200 to -400 pm tip-height offset show that the current onset is step-like at the voltage of roughly -0.2 to -0.3 V. This is an indication that the tip Fermi level becomes resonant with a charge transition level of the DB. The width of the step corresponds to the broadening of the DB state. The tunneling current subsequently rises exponentially up to the NDR onset. The spectra with -300 pm and -400 pm tip-height offset exhibit a crossing point between H:Si and DB (indicated by arrows in Fig. 1C). This is possible only under the condition of the DB electron occupation falling below one, making it positively charged. The resulting downward band bending reduces the overlap between the tip's empty energy levels and the sample's valence band, thereby decreasing conductance over the DB compared to H:Si. This indicates that for sample biases more negative than the NDR onset, the tip Fermi level must be below both charge transition levels of the DB.

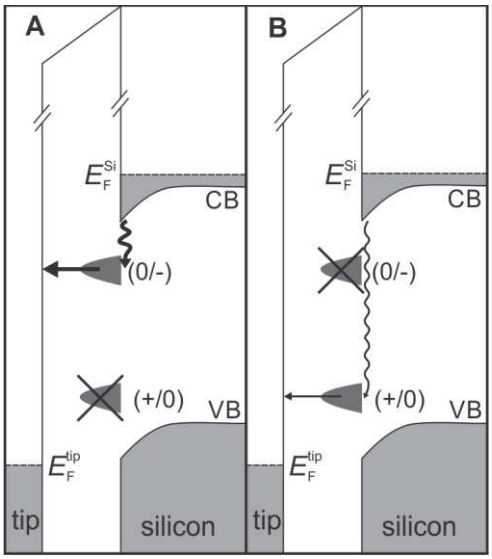

Fig. 2. The underlying mechanism of NDR. (A, B) Energy diagrams of the system of study in the NDR energy regime displaying silicon Fermi energy, conduction band, valence band, DB charge transition levels and all electron transport channels from the bulk to the tip through the DB. The tip Fermi level is below both the DB charge transition levels. When the occupation of the DB is ≥1, transport occurs through the (0/-) level (A). Occasionally, the tip fully empties the DB making its electron occupation zero (B). This switches off the dominant current flow through (0/-) level until the DB gets filled again by a single electron from the conduction band.

We implemented a non-equilibrium Green's function (NEGF) method (supplementary online text) to interpret our experimental results (Fig. 1D). These calculations corroborate our experimental finding that the current onset of $I(V)$ spectra occurs when the tip Fermi level becomes resonant with the (0/-) level of the DB. The electrons start flowing from the conduction band through the (0/-) level of the DB and into the tip. As the sample bias becomes more negative, in agreement with the experiment, the NEGF calculations predict NDR. At the NDR onset, the tip becomes resonant with the (+/0) level, and the DB can occasionally be fully emptied by the tip. During the time that the DB is fully empty, the supply of electrons from the

conduction band to the DB has to flow through the (+/0) level. Since this level is energetically much lower than the (0/-) level, the process is more inelastic and hence slow in comparison (*30*). This results in NDR. A schematic of the underlying mechanism is presented in Fig. 2 and Movie S1. The NDR persists until the (+/0) level becomes resonant with the bulk valence band due to tip-induced band bending. As the sample bias becomes more negative, greater overlap between the DB and the bulk valence band allows for increased current flow. The separation between the higher and lower levels in a non-equilibrium setting is a many-body phenomenon. At a single particle level, ignoring the interactions, NDR would not have been present in the system.

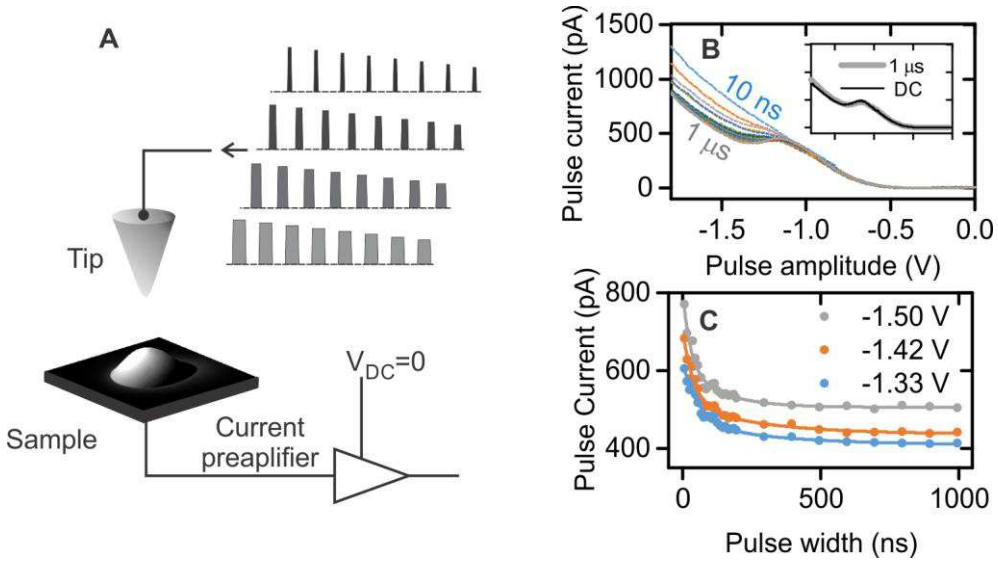

Fig. 3. Time-resolved *I*(*V*) spectroscopy. (A) Time-resolved *I*(*V*) spectroscopy measurement scheme. The DC bias is set to zero and the tunneling current is measured from a series of varying-amplitude nanosecond-regime pulses. (B) Time-resolved *I*(*V*) spectra recorded with tip-height offset of -210 pm and different pulse widths (10 ns to 1 µs). The inset compares the DC *I*(*V*) and the time-resolved *I*(*V*) with1 µs pulse width. The duty cycle for each curve is 0.05. The initial tip height (zero tip offset) is set with the tunneling conditions of -1.80 V and 30.0 pA on top of the DB. (C) Pulse current as a function of pulse width for 3 different pulse amplitudes are shown as examples. Solid lines are exponential fits.

We perform all-electronic time-resolved STM measurements (*19–25*) to capture all the relevant rates in the NDR energy regime (*27*). A schematic of time-resolved *I*(*V*) spectroscopy measurement is shown in Fig. 3A. Time-resolved *I*(*V*) measurements for a given tip height and different pulse widths are shown in Fig. 3B and compared with DC *I*(*V*). The NDR feature is extinguished for short pulses. The time-resolved *I*(*V*) spectra approach the DC *I*(*V*) for longer

pulses (inset in Fig. 3B). The measured tunneling current for a given bias voltage exhibits an exponential decay with increasing pulse width (Fig. 3C).

From the results in Fig. 3B and 3C, we are able to capture all the time constants related to NDR (Fig. 4). The time constants of the exponential decays in Fig. 3C correspond to the time needed to fully empty the singly occupied DB by the tip (T_{2t}). Figure 4B plots T_{2t} for different bias voltages in the NDR regime. During this time, the current flows only through the (-/0) level. Since the electron supply time constant from the bulk to the (0/-) level (T_{1b}) is much longer than its emptying time by the tip (T_{1t}), the measured current with the shortest pulse (10 ns) inversely corresponds to T_{1b}. Therefore, we are able to extract T_{1b} from the spectrum of 10 ns width pulse (Fig. 4C). The time constant of electron supply from the bulk to the (+/0) level (T_{2b}) can be calculated by (supplementary online text):

$$T_{2b}(V) = \frac{\left[e - I_{DC}(V)T_{1b}(V)\right]T_{2t}(V)}{I_{DC}(V)T_{1b}(V)},$$

where e is the elementary charge. Figure 4D shows T_{2b} for different bias voltages in the NDR regime. This is the direct measurement of electron capture time by a mid-gap state through an inelastic process.

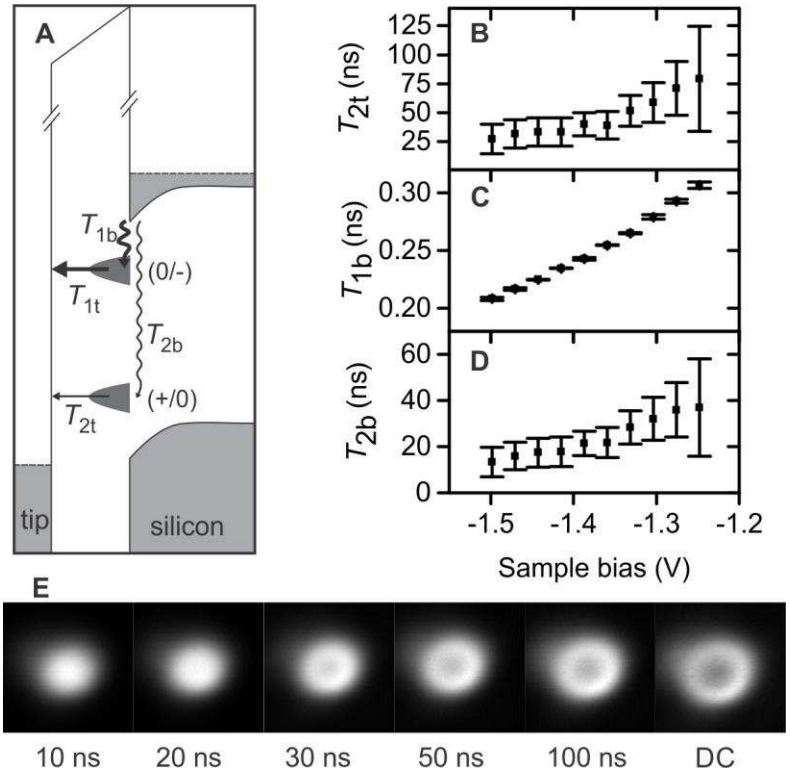

Fig. 4. Measurement of single electron dynamics related to NDR and time-resolved real-space imaging of NDR. (A) Energy diagram of the system of study in the NDR regime displaying all the relevant time constants. (B-D) The time constants displayed in (A) are extracted from the time-resolved $I(V)$ spectra. (E) Time-resolved STM topographs. DC bias is kept at zero, pulse amplitude is -1.5 V and the measurement duty cycle for each pulse width is 0.05. Image size is 2×2 nm^2 and the tip-height offset is -210 pm. DC image at -1.5 V is added for comparison. The zero tip-offset is set with the tunneling conditions of -1.80 V and 30.0 pA on top of the DB.

The spatial evolution of a single DB at nanosecond time scale can be detected by performing time-resolved topography measurements (*27*). As expected from the time-resolved $I(V)$ measurements, for short pulses the DB appears as a simple protrusion (Fig. 4E), indicating an absence of NDR. As the pulse becomes longer, the conductivity at the center of the DB

decreases and DB appearance approaches that of the DC measurements. These measurements further corroborate our interpretation.

NDR continues to be a desirable but elusive functionality in electronic circuity. Here, we come to a detailed understanding of a structural and material configuration yielding robust NDR on silicon and which is therefore compatible with CMOS technology. As dangling bonds can be made en masse without need for scanned probe type tools, there appears to be the prospect for transitioning to a real technology that offers improved devices in the near future.

References and Notes:

1. L. Esaki, New phenomenon in narrow germanium p-n junctions. *Phys. Rev.* **109**, 603–604 (1958).

2. A. Seabaugh *et al.*, Transistors and tunnel diodes for analog/mixed-signal circuits and\nembedded memory. *Int. Electron Devices Meet. 1998. Tech. Dig. (Cat. No.98CH36217)*, 429–432 (1998).

3. W. S. Choi, S. A. Lee, J. H. You, S. Lee, H. N. Lee, Resonant tunnelling in a quantum oxide superlattice. *Nat. Commun.* **6**, 7424 (2015).

4. M. L. Perrin *et al.*, Large negative differential conductance in single-molecule break junctions. *Nat. Nanotechnol.* **9**, 830–4 (2014).

5. Y. Wu *et al.*, Three-terminal graphene negative differential resistance devices. *ACS Nano*. **6**, 2610–2616 (2012).

6. Y. Du *et al.*, Symmetrical negative differential resistance behavior of a resistive switching device. *ACS Nano*. **6**, 2517–2523 (2012).

7. M. Galperin, M. A. Ratner, A. Nitzan, A. Troisi, Nuclear coupling and polarization in molecular transport junctions: beyond tunneling to function. *Science (80-.).* **319**, 1056–60 (2008).

8. Y.-C. Lin *et al.*, Atomically thin resonant tunnel diodes built from synthetic van der Waals heterostructures. *Nat. Commun.* **6**, 7311 (2015).

9. P. Bedrossian, D. M. Chen, K. Mortensen, J. A. Golovchenko, Demonstration of the tunnel-diode effect on an atomic scale. *Nature*. **342**, 258–260 (1989).

10. I. W. Lyo, P. Avouris, Negative differential resistance on the atomic scale: implications for atomic scale devices. *Science*. **245**, 1369–1371 (1989).

11. J. L. Pitters, R. A. Wolkow, Detailed studies of molecular conductance using atomic resolution scanning tunneling microscopy. *Nano Lett.* **6**, 390–397 (2006).

12. N. J. Tao, Electron transport in molecular junctions. *Nat. Nanotechnol.* **1**, 173–181 (2006).

13. M. Berthe *et al.*, Probing the carrier capture rate of a single quantum level. *Science*. **319**, 436–438 (2008).

14. T. H. Nguyen *et al.*, Coulomb energy determination of a single Si dangling bond. *Phys. Rev. Lett.* **105**, 1–4 (2010).

15. Y. Xue *et al.*, Negative differential resistance in the scanning-tunneling spectroscopy of organic molecules. *Phys. Rev. B*. **59**, R7852–R7855 (1999).

16. M. B. Haider *et al.*, Controlled coupling and occupation of silicon atomic quantum dots at room temperature. *Phys. Rev. Lett.* **102**, 2–5 (2009).

17. S. R. Schofield *et al.*, Quantum engineering at the silicon surface using dangling bonds.

Nat. Commun. **4**, 1649+ (2013).

18. J. L. Pitters, L. Livadaru, M. B. Haider, R. A. Wolkow, Tunnel coupled dangling bond structures on hydrogen terminated silicon surfaces. *J. Chem. Phys.* **134** (2011), doi:10.1063/1.3514896.

19. G. Nunes, M. R. Freeman, Picosecond resolution in scanning tunneling microscopy. *Science*. **262**, 1029–1032 (1993).

20. S. Loth, M. Etzkorn, C. P. Lutz, D. M. Eigler, A. J. Heinrich, Measurement of fast electron spin relaxation times with atomic resolution. *Science*. **329**, 1628–30 (2010).

21. S. Loth, S. Baumann, C. P. Lutz, D. M. Eigler, A. J. Heinrich, Bistability in atomic-scale antiferromagnets. *Science*. **335**, 196–9 (2012).

22. C. Grosse, M. Etzkorn, K. Kuhnke, S. Loth, K. Kern, Quantitative mapping of fast voltage pulses in tunnel junctions by plasmonic luminescence. *Appl. Phys. Lett.* **103**, 183108 (2013).

23. C. Saunus, J. Raphael Bindel, M. Pratzer, M. Morgenstern, Versatile scanning tunneling microscopy with 120 ps time resolution. *Appl. Phys. Lett.* **102**, 051601 (2013).

24. S. Yan, D.-J. Choi, J. A. J. Burgess, S. Rolf-Pissarczyk, S. Loth, Control of quantum magnets by atomic exchange bias. *Nat. Nanotechnol.* **10**, 40–45 (2014).

25. M. Rashidi *et al.*, Time-Resolved Single Dopant Charge Dynamics in Silicon. *arXiv*, 1–17 (2015).

26. X. de la Broïse *et al.*, Theory of scanning tunneling microscopy of defects on semiconductor surfaces. *Phys. Rev. B.* **61**, 2138–2145 (2000).

27. Materials and methods are available as supplementary materials on Science Online.

28. J. L. Pitters, P. G. Piva, R. A. Wolkow, Dopant depletion in the near surface region of thermally prepared silicon (100) in UHV. *J. Vac. Sci. Technol. B*. **30**, 021806 (2012).

29. H. Labidi *et al.*, Scanning tunneling spectroscopy reveals a silicon dangling bond charge state transition. *New J. Phys*. **17**, 073023 (2015).

30. S. Datta, *Quantum Transport: Atom to Transistor* (Cambridge University Press, 2005).

31. M. Rezeq, J. Pitters, R. Wolkow, Tungsten nanotip fabrication by spatially controlled field-assisted reaction with nitrogen. *J. Chem. Phys*. **124**, 204716 (2006).

32. J. A. Wood *et al.*, Iridium single atom tips fabricated by field assisted reactive gas etching. *Appl. Surf. Sci.* **367**, 277–280 (2016).

33. R. M. Feenstra, SEMITIP SOFTWARE (2011), (available at http://www.andrew.cmu.edu/user/feenstra/semitip_v6/).

Acknowledgments:

We would like to thank Martin Cloutier and Mark Salomons for their technical expertise, and Jacob Burgess and Sebastian Loth for stimulating discussions. We also thank NRC, NSERC, AITF, CRC, CIFAR and Compute Canada for support.

Supplementary Materials For:

Time-Resolved Imaging of Negative Differential Resistance on the Atomic Scale

Mohammad Rashidi*, Marco Taucer, Isil Ozfidan, Erika Lloyd, Hatem Labidi, Jason L. Pitters, Joseph Maciejko, and Robert A. Wolkow

*Correspondence to: rashidi@ualberta.ca

Materials and Methods

Poly-crystalline and single crystal tungsten (111), as well as Iridium are used as tip materials to confirm that the observed effect is tip independent. Tungsten and iridium are electrochemically etched, cleaned, and sharpened by field ion microscopy (*31*, *32*)
The samples are cleaved from a 3–4 mΩ.cm n-type arsenic doped Si(100) wafer (Virginia Semiconductor Inc.) and are degassed for about 12 hours at 600°C. Samples are then flash annealed at 1050°C several times to remove the oxide layer and other surface contaminates. Hydrogen termination occurs by exposure to H_2 gas at 1×10^{-6} Torr with a heated tungsten filament to crack the hydrogen. With the cracked hydrogen in the chamber the sample is left to etch at room temperature for one minute, then flashed again to 1050 °C, and then brought to 330°C for 30 seconds to hydrogen terminate.

To confirm the temperature independence of the NDR feature, DC $I(V)$ measurements were performed at 4.5K, 77K and room temperature in UHV using Omicron low and room temperature STMs. All subsequent data shown in this paper was collected at 4.5 K with a base pressure of 5×10^{-11} Torr. Nanonis SPM controller and software are used for data acquisition. Single DBs can be found natively on the surface or created by placing the tip over a hydrogen atom and applying a positive sample voltage pulse (2.0 to 2.4 V) to desorb the hydrogen atom.

Different crystals, different tip materials, entirely different chambers and STM instruments, different measurement temperatures, and different operators within our labs confirmed the results. The variability of spectra taken at one DB with different tip atomic structure is essentially the same variability between curves taken at different DBs. Therefore we conclude that within our measurement capabilities, all DBs give essentially the same NDR behavior.

$I(V)$ spectra were measured in constant-height mode and a reference spectroscopy over H:Si was used to check for tip changes. STM images are used to judge tip quality. In-situ tip shaping techniques include voltage pulses and gentle controlled contacting of tip to sample to subtly change and improve its structure. All DC and Fast images shown are taken in constant height mode with the tunneling set-point of 30 pA and -1.8 V on top of a DB. The feedback loop is shut off while the tip is rastered over the area of interest and the corresponding current is collected. The resulting image is a current map at the specified height. Consecutive images can be taken at different heights relative to the initial set point.

The Omicron STM is equipped commercially with RF wiring that has a 500 MHz bandwidth. Fig. 3A shows schematically the experimental setup. An arbitrary function generator (Tektronix AFG3252C) was used to generate a cycle of voltage pulses that are fed to the tip. An RF switch (Mini-Circuits ZX80-DR230-S+) is used to change the tip between ground and the output of the arbitrary function generator. DC bias voltages are applied to the sample side where the tunnel current is also collected through an Omicron preamplifier.

Time resolved measurements are distinct from their DC counterparts in that the tunneling current is induced by a series of short voltage pulses rather than a continuous DC bias. The voltage pulses can be tuned significantly with the following parameters: Pulse width, pulse

amplitude, repetition rate, duty cycle. An appropriate measuring time can be chosen based on the conditions of the experiment.

The time-resolved topography measurements are collected similarly to the DC images. The difference is that the resulting current maps do not correspond to a response to a continuous DC bias, but to short voltage pulses. In this way the current maps are able to resolve dynamics that occur on the time scale set by the pulse width and duty cycle of the pulses. The DC bias was set to zero such that any current collected was strictly due to the pulses.

There exists an impedance mismatch between the 50 Ω of the outside circuit and the tunnel junction. In order to mitigate ringing, the pulses edges were kept to 2.5ns. Another consequence is that very short pulses experience a distortion at the junction. This creates an effective voltage at the junction that may be different than the applied voltage. A calibration technique was implemented for the affected pulses to circumvent this. A series of fast $I(V)$ curves for the pulse widths of interest were taken over H:Si, followed by a DC curve. The fast curves of the inert surface should reproduce the DC curve and an appropriate voltage shift can be extracted by comparison. This calibration reflects the difference in effective voltage and the applied voltage.

The current recorded for each $I(V)$ curve had to be normalized according to its duty cycle during measurement. The measured current is an average over a complete cycle, however we want the corresponding current when the voltage is on. The following formula was used to calculate the current induced by the voltage pulses:

$$\frac{T}{W} \times I_{\text{preamp}} = I_{eff},$$

where T is the period of the pulse series, W is the width of the pulse, I_{preamp} is the measured current from the STM preamplifier and I_{eff} is the current induced by the voltage pulses.

Non-equilibrium Green's function Calculations

We implemented a non-equilibrium Green's function (NEGF) model to aid in the interpretation of our experimental results. The current through a single DB coupled to a metallic tip and the silicon substrate can be expressed by the Landauer formula as

$$I = \frac{q}{h}\int_{-\infty}^{+\infty} \mathrm{Tr}\left[\Gamma_{\mathrm{tip}} G^R \Gamma_{\mathrm{Si}} G^A\right]\left(f_{\mathrm{tip}}(E) - f_{\mathrm{Si}}(E)\right)\mathrm{d}E,$$

where q is the electron charge, $G^{R(A)}$ is the retarded (advanced) Green's function of the coupled DB, $\Gamma_{\mathrm{tip(Si)}}$ is the coupling strength of the DB to the tip (Si) and $f_{\mathrm{tip(Si)}}$ is the Fermi function of the tip (Si). The non-equilibrium retarded Green's function satisfies the Dyson equation $G^R = G_0^R + G_0^R \Sigma^R G^R$ where G_0^R is the non-interacting Green's function of the DB. In this equation, the self-energy, Σ, entangles the interactions within the DB with that of the tip and the surface. The self-energy is not exactly solvable. Instead, we solve the problem approximately by using the exact Green's function g^R of the isolated, interacting single DB as the unperturbed Green's function and add the non-interacting tunneling self-energy Σ;

$$\left[G^R(\omega)\right]^{-1} \approx \left[g^R(\omega)\right]^{-1} - \Sigma^R(\omega).$$

The Hamiltonian for an uncoupled, interacting DB is written as an Anderson model with an onsite energy of ε_0 and Coulomb repulsion, U:

$$H = \sum_{\sigma} \varepsilon_0 c_\sigma^\dagger c_\sigma + U n_\uparrow n_\downarrow, \qquad (1)$$

where $c_\sigma^\dagger (c_\sigma)$ are operators creating (annihilating) an electron with spin-σ on the DB and $n_\sigma = c_\sigma^\dagger c_\sigma$ is the number operator. Using the equation of motion technique with the Hamiltonian (Eq. 1), we obtain the exact Green's function of the uncoupled, interacting DB,

$$g_\sigma^R(\omega) = \frac{\langle n_{\bar{\sigma}} \rangle}{\omega + i\delta - \varepsilon_\sigma - U} + \frac{1 - \langle n_{\bar{\sigma}} \rangle}{\omega + i\delta - \varepsilon_\sigma}.$$

The tunneling self-energy, Σ, is composed of an imaginary part, proportional to the coupling strength, Γ, that is responsible for the level broadening, and the real part that causes a shift in the energy levels of the DB. For simplicity, we ignore the real part of the self-energy and only include the imaginary part, $\Sigma^R = -i\Gamma/2$, in our calculation. We model the coupling strength to the tip as a constant, Γ_{tip}. For Γ_{si} we include the coupling to the Si valence/conduction bands and coupling rate to the degenerate dopant level differently, and consider the coupling to the dopants bias-dependent.

Since the coupling strength effects the broadening of the DB levels and the onsite energy, and the Coulomb repulsion affects the onsets of the current, we base ε_0 and U on previous calculations and fit the coupling strength to the experiment[1]. Finally, we calculate the amount of tip induced band bending (TIBB) with increased tip-sample distance and bias voltage using the SEMITIP software (*33*). Due to all approximations, the theory serves to explain the physical phenomenon leading to NDR qualitatively and not to reproduce exactly the measured $I(V)$ spectra. While general features are well reproduced, theory does not fully capture the experimental suppression of current in the very close tip regime. Specifically, while the experiment shows a reduction in the local current minimum at approximately -1.4V as the tip approaches the sample, in our theoretical calculations the minimum value of the current increases with the reduced tip-sample distance. This discrepancy may be a result of neglecting the dynamic

[1] Calculation values: $E_F^{Si} = 0$, E_{VB}=-1 eV, ε_0 = -0.8 both with respect to E_F^{Si}. U = 0.74 eV, Γ_{Si} = 0.08, Γ_{DB}=0.002 + 0.005V_{bias}, $\Gamma_{tip} = 3\times10^4 \exp(-2.4\times d)$, where d is the approximate tip-sample distance in nm.

interactions in self-energy which is expected to play a more important role as the distance between the tip and DB decreases.

Relations between single electron rates

The dynamical transfer of charge to and from the DB determines the occupation probabilities and the total tunneling current, and ultimately gives rise to NDR. In this section we describe the equations governing the single electron transfer rates. This allows us to relate our experimental measurements to physically meaningful timescales.

When the tip Fermi level is below the (0/-) level, but above the (+/0) level, the DB is always occupied by at least one electron and all current flows through the (0/-) level. In this regime (less negative than the NDR onset), the total current is

$$I_{\mathrm{T}} = e\frac{\Gamma_{1\mathrm{b}}\Gamma_{1\mathrm{t}}}{\Gamma_{1\mathrm{b}}+\Gamma_{1\mathrm{t}}}, \tag{2}$$

where all rates are related to the corresponding time constants by $\Gamma = 1/T$. In this regime, the *I(Z)* curves show a saturation, indicating that $\Gamma_{1\mathrm{b}} \ll \Gamma_{1\mathrm{t}}$ for the smallest tip-sample separations.

When the tip Fermi level is below both transition levels (more negative than the NDR onset), it becomes possible for the DB to be completely unoccupied. The probability for the DB to be positive, neutral, or negative can be described by the probability vector $\mathbf{P} = \left(P^{+}, P^{o}, P^{-}\right)^{T}$, whose elements sum to unity. The time-dependence of the probability vector is described by the differential equation $\dot{\mathbf{P}} = \mathbb{M}\mathbf{P}$, where $\mathbb{M}$ is

$$\mathbb{M} = \begin{pmatrix} -\Gamma_{2\mathrm{b}} & \Gamma_{2\mathrm{t}} & 0 \\ \Gamma_{2\mathrm{b}} & -\left(\Gamma_{1\mathrm{b}}+\Gamma_{2\mathrm{t}}\right) & \Gamma_{1\mathrm{t}} \\ 0 & \Gamma_{1\mathrm{b}} & -\Gamma_{1\mathrm{t}} \end{pmatrix}.$$

The steady state probability vector corresponds to the zero eigenvalue, and it can be expressed as

$$\mathbf{P}_{ss} = \frac{1}{\Gamma_{1t}\Gamma_{2t} + \Gamma_{1t}\Gamma_{2b} + \Gamma_{1b}\Gamma_{2b}} \begin{pmatrix} \Gamma_{1t}\Gamma_{2t} \\ \Gamma_{1t}\Gamma_{2b} \\ \Gamma_{1b}\Gamma_{2b} \end{pmatrix}.$$

The steady state applies to the DC *I(V)* measurements, and the current can in fact be decomposed into a current passing through each of the transition levels, $I_{DC} = I_{(+/0)} + I_{(0/-)}$, where

$$I_{(+/0)} = e\frac{\Gamma_{1t}\Gamma_{2t}\Gamma_{2b}}{\Gamma_{1t}\Gamma_{2t} + \Gamma_{1t}\Gamma_{2b} + \Gamma_{1b}\Gamma_{2b}},$$

and

$$I_{(0/-)} = e\frac{\Gamma_{1b}\Gamma_{1t}}{\Gamma_{1b} + \Gamma_{1t} + \frac{\Gamma_{1t}\Gamma_{2t}}{\Gamma_{2b}}}.$$

Because Γ_{2b} is smaller than all other rates, $I_{(+/0)}$ must be small. Furthermore, the current through the (0/-) level is altered by the addition of the third term in the denominator, $\Gamma_{2t}\,\Gamma_{2t}/\Gamma_{2b}$ as compared with Equation 2. It is the presence of this term whenever the tip Fermi level is below the (+/0) level which suppresses current, leading to NDR.

Finally, using the fact that $\Gamma_{1b} << \Gamma_{1t}$ for the smallest tip-sample separations, we can relate the time constant for the inelastic process, T_{2b} to the experimentally measured quantities T_{1b}, T_{2t} and I_{DC},

$$T_{2b} = \frac{e - I_{DC}T_{1b}}{I_{DC}T_{1b}}T_{2t}. \qquad (3)$$

Movie S1 caption:

A cartoon to present the electronic trasnport from bulk condcution band to the tip through charge trasnition levels of a single DB in the NDR energy regime. The tip Fermi level is below both the DB charge transition levels. When the occupation of the DB is bigger than one, the electrons flow from the conduction band to the tip through the (0/-) level. This is the main electronic path. Occasionally, the DB gets fully emptied and the dominant current flow through (0/-) level is switched off. Until the DB gets filled again by a single electron from the conduction band, there is no current flow.