

The role of many-particle excitations in Coulomb blockaded transport

This article has been downloaded from IOPscience. Please scroll down to see the full text article.

2008 J. Phys.: Condens. Matter 20 374109

(<http://iopscience.iop.org/0953-8984/20/37/374109>)

View [the table of contents for this issue](#), or go to the [journal homepage](#) for more

Download details:

IP Address: 129.252.86.83

The article was downloaded on 29/05/2010 at 15:04

Please note that [terms and conditions apply](#).

The role of many-particle excitations in Coulomb blockaded transport

B Muralidharan^{1,3}, L Siddiqui¹ and A W Ghosh²

¹ School of Electrical and Computer Engineering and the Network for Computational Nanotechnology, Purdue University, West Lafayette, IN 47907, USA

² Department of Electrical and Computer Engineering, University of Virginia, Charlottesville, VA 22904, USA

³ Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

Received 22 May 2008

Published 26 August 2008

Online at stacks.iop.org/JPhysCM/20/374109

Abstract

We discuss the role of electron–electron and electron–phonon correlations in current flow in the Coulomb blockade regime, focusing specifically on non-trivial signatures arising from the breakdown of mean-field theory. By solving transport equations directly in Fock space, we show that electron–electron interactions manifest as gateable excitations experimentally observed in the current–voltage characteristic. While these excitations might merge into an incoherent sum that allows occasional simplifications, a clear separation of excitations into slow ‘traps’ and fast ‘channels’ can lead to further novelties such as negative differential resistance, hysteresis and random telegraph signals. Analogous novelties for electron–phonon correlation include the breakdown of commonly anticipated Stokes–anti-Stokes intensities, and an anomalous decrease in phonon population upon heating due to reabsorption of emitted phonons.

(Some figures in this article are in colour only in the electronic version)

1. Introduction

The experimental study of electron flow through nanostructures has been a dynamic field of activity, with an eye on extending and complementing present day transistor technologies, as well as generating entirely new applications. Nanoscale electronic transport spans a broad range of natural and artificially fabricated nanostructures, from carbon nanotubes, graphene nanoribbons and silicon nanowires [1] to spintronics and organic molecular electronics [2, 3]. In particular, there has been enormous interest in quantum dot structures for exploring novel transport phenomena and device applications beyond the transistor switching paradigm, such as the exploration of double quantum dot structures [4] for spin-based qubit manipulation and detection [5]. Electron transport through natural and artificial molecules forms a key research topic, especially for switching, sensing [6] and quantum-computation-based applications [7].

In typical transport simulations, for instance in the widely implemented non-equilibrium Green’s function (NEGF) formalism, it is common to include the effect of electron–electron or electron–vibronic interactions approximately through an effective one-electron potential or self-consistent field (SCF) that needs to be computed self-consistently. There

are many exceptions, however, where such an approximation may break down, especially when interaction energies dominate other energy scales of interest such as the level broadening and the device temperature. One such regime, well known as Coulomb blockade (CB) [8], occurs when the device or channel capacitance is low enough that an active electron inside the channel can prevent a subsequent one from entering. Such a single-particle quantization of charge transfer is frequently observed in chemical reactions [9], but is a relative newcomer in electronic transport measurements [8]. The sequential addition of electrons in integer amounts disallows mean-field treatments which tend to smear out charges and interactions by treating all electrons on an equal footing. In contrast, solutions involve products of electronic occupancy and atomic displacement operators including an exclusion principle term that requires keeping track of every possible electronic or vibrational configuration through the employment of the many-particle Hilbert (Fock) space. Under these strong correlation conditions, energy levels must be calculated not through a simple band theory or an effective potential, but as differences between *total energies* of the neutral and the cationic/anionic/vibronically excited species. This is extremely difficult since it requires enumeration of all many-particle configurations (2^N of them, for N basis sets involving

electronic and phononic coordinates!). However, such a complexity is necessary, as exclusion in Fock space creates a rich spectrum of excitations as well as universal scaling rules for the current plateau heights that are hard to capture *a priori* using a modified one-electron potential³. The alternate Fock space viewpoint (in general, a many-body density matrix theory) has been somewhat restricted for describing quantum dot transport [8, 10] and has been relatively unexploited in molecular electronics. The focus of this paper is on illustrating the Fock space viewpoint of transport, and its experimental ramifications as far as observable signatures of many-particle excitations go.

Our recent work in the area of molecular transport [11, 12, 14, 13], as well as compelling recent experiments triggered towards spin-based quantum computation [5, 15–18], both argue for increased activity in this area. This paper focuses on the role of both electron–electron and electron–phonon correlations in non-equilibrium transport. The paper is organized into three broad sections. In the first section we introduce the Fock space viewpoint. The second section runs through the formalism of Fock space transport. Section 3 discusses Coulomb blockade signatures created by electron–electron interactions. We elaborate on the non-trivial role of electronic excitations [12] in the interpretation of frequently observed I – V characteristics [20, 21]. It is further shown that electronic excitations can also result in intrinsic asymmetries within the channel that can provide an elegant approach to understanding negative differential resistance (NDR), hysteresis effects [14], and random telegraph noise [23]. The fourth section discusses the additional subtleties imposed by strong electron–phonon interactions. We show that the inclusion of Fock space excitations within the electron–phonon manifold not only explains anomalous scaling of phonon conductance side bands [25, 24], but also predicts an anomalous temperature distribution of the phonon population.

2. Theoretical background

The Fock space approach to Coulomb blocked transport was proposed originally by Beenakker [10] in order to explain the CB conductance peak spacings and heights observed in semiconducting quantum dots. Figure 1 explains the difference between the one-particle and many-particle Fock space pictures. A set of N single-particle energy levels generates 2^N Fock states corresponding to emptying or filling each of these levels with one electron. In the one-electron picture, transport involves the addition and removal of electrons between a set of channel levels and two macroscopic electrodes (figure 1(a)). The levels themselves are computed by solving the one-electron Schrödinger equation, including electronic interactions approximately through a mean-field potential that modifies these levels dynamically. In the Fock space approach, however, the addition and removal of

³ In principle, a non-local time dependent density functional theory (TDDFT) could capture these excitations even using an $N \times N$ one-electron potential matrix, provided the time dependence of the functional is rich enough to capture the additional excitations of the $2^N \times 2^N$ space through its poles. In practice, however, this has not been done, even for the simplest system of two spin-degenerate levels.

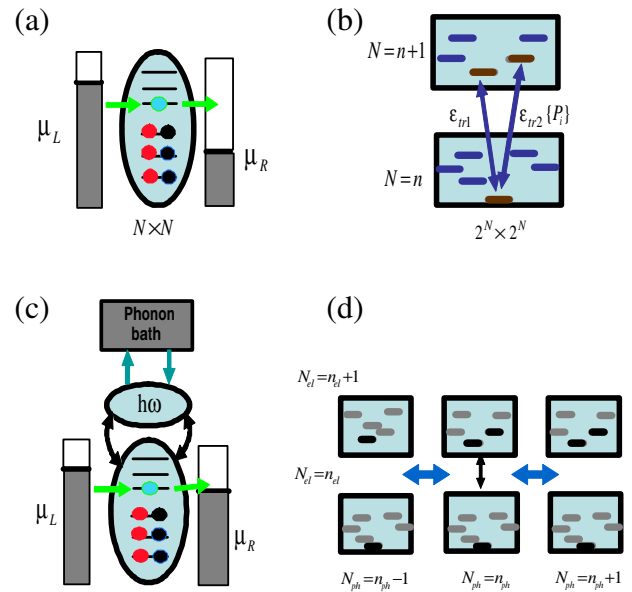


Figure 1. Introducing Fock space transport. (a) A schematic of electronic transport in the one-particle picture. The channel is coupled to two electrodes which add and remove electrons to and from the channel, resulting in current flow. Given a set of N basis functions the transport problem computationally scales as $N \times N$. (b) In the Fock space picture such addition and removal processes can be understood in terms of transitions between states that differ by a single electron number. In such a scheme one needs to keep track of all possible configurations in the channel, thus obtaining an exponential $2^N \times 2^N$ scaling of the transport problem. (c) In the case of strong electron–phonon interactions, the phonon system is coupled to a phonon bath which is maintained at equilibrium. (d) The inclusion of phonons calls for using copies of the electronic Fock space, each corresponding to a phonon number. Electronic transport results in vertical transitions while phonon coupling results in horizontal ones.

electrons leads to a transition between two entirely different multielectron configurations, in this case, configurations that differ by a single electron (figure 1(b)). In other words, electronic transport processes show up as *vertical* transitions between total energies of electronic Fock states differing by a single electron. The situation changes a bit when transport involves the emission or absorption of phonons. In the one-particle picture, an electron jumps between two one-electron levels differing by the phonon energy (figure 1(c)), the phonon system itself driven towards equilibrium through a separate coupling to a thermal bath. In the Fock space approach figure 1(d), phonon-assisted transport shows up as *horizontal* transitions between multiple copies of the electronic Fock space that each correspond to a different phonon number.

In the following, we will progressively build complexity into our Fock space approach and attempt to identify the experimental ramifications of doing so.

2.1. Single-level system, SCF analysis

Let us start with the smallest interacting system, namely, a real or artificial molecule with a single energy level capable of accommodating two spins (figure 2(a)). The onsite energy of the level is ϵ_0 and the Coulomb charging energy U . The

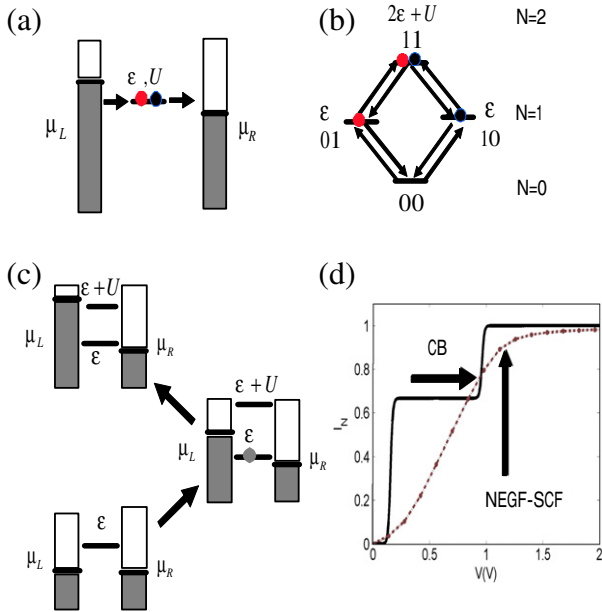


Figure 2. Fock space transport through a singly degenerate energy level. (a) The singly degenerate energy level can carry two electrons, one of each type. (b) The Fock space thus comprises four states. Transitions between these result in (c) two transport channels ϵ and $\epsilon + U$. Successive accessing of these transport channels results in two current jumps at different bias points separated by a plateau. A NEGF-SCF calculation is shown dotted for comparison.

level is coupled to contacts which are held separately at thermal equilibrium at their respective bias-separated electrochemical potentials $\mu_{L,R}$ (L: left, R: right). A starting point for our analysis is the Hubbard Hamiltonian for the molecule

$$\hat{H} = \epsilon \hat{n} + U \hat{n}_\uparrow \hat{n}_\downarrow \quad (1)$$

where the operators $\hat{n}_{\uparrow,\downarrow}$ have eigenvalues 0 and 1, while $\hat{n} = \hat{n}_\uparrow + \hat{n}_\downarrow$. Exactly diagonalizing this Hamiltonian leads to a Fock space consisting of four many-electron states (figure 2(b)), an empty zero-electron state $|00\rangle$ with energy 0, two one-electron states $|01\rangle$ and $|10\rangle$ corresponding to an up or a down spin with energy ϵ , and a doubly occupied up-down spin electron state $|11\rangle$ with energy $2\epsilon_0 + U$. Equilibrium occupancies of these many-electron states are given by the Boltzmann distribution $P_N = e^{-(E_N - \mu N)/k_B T} / \Omega$, where $k_B T$ is the thermal energy, $\mu = E_F$ is the equilibrium contact Fermi energy or electrochemical potential, and $\Omega = \sum_N e^{-(E_N - \mu N)/k_B T}$ is the grand partition function. The average electron occupancy is then given by $\langle N \rangle = \sum_N N P_N$.

One can bypass the many-electron Fock space treatment by employing a suitable self-consistent (SCF) potential acting in the one-electron subspace, modifying the energies accordingly. In the *spin restricted* approach that treats the two spins equally,

$$U_{\text{RSCF}} = \langle \partial \hat{H} / \partial N \rangle. \quad (2)$$

where $\langle \dots \rangle$ denotes a quantum mechanical average. The interacting term can be written as

$$\begin{aligned} H_{\text{int}} &= U \hat{n}_\uparrow \hat{n}_\downarrow \\ &= (U/2) \sum_\sigma \hat{n}_\sigma \hat{n}_{\bar{\sigma}} \end{aligned}$$

$$\begin{aligned} &= (U/2) \sum_\sigma \hat{n}_\sigma (N - \hat{n}_\sigma) \\ &= (U/2) N \sum_\sigma \hat{n}_\sigma - (U/2) \sum_\sigma \hat{n}_\sigma^2 \\ &= U N (N - 1) / 2, \end{aligned} \quad (3)$$

where we have used the fact that $N = \sum_\sigma \hat{n}_\sigma$, and $\hat{n}_\sigma^2 = \hat{n}_\sigma$, since \hat{n} can only take values of zero or one. The spin restricted SCF potential is then given by

$$U_{\text{RSCF}} = \langle \partial H_{\text{int}} / \partial N \rangle = U (\langle N \rangle - 1/2). \quad (4)$$

For a given electrochemical potential, one guesses the value of the average occupancy $\langle N \rangle$, uses it to calculate the SCF potential, and then calculates in turn the mean-field occupancy $\langle N \rangle$ of the level $\tilde{\epsilon} = \epsilon + U_{\text{SCF}}$ using the Fermi-Dirac distribution $f(\tilde{\epsilon}) = 1/[1 + e^{(\tilde{\epsilon} - \mu)/k_B T}]$, proceeding along these lines until self-consistent convergence is achieved.

It is easy to see that the equilibrium occupancy N [12] with respect to the chemical potential μ (we drop the angular term indicating average here) should qualitatively differ between the SCF and many-body results. In the former, the electron occupancy is a fractional amount, adiabatically changing from zero to two. In the many-body result, however, one does not simply multiply the results for one electron by two, but the electron occupancy changes abruptly between zero to one, followed by a plateau of width U over which the electrons are blocked by the Coulomb interaction, after which the electron number reaches two abruptly.

One could capture this blocked effect using an *unrestricted* self-consistent potential (USCF) by dictating that the up and down spins do not feel potentials due to themselves. By eliminating this self-interaction, the unrestricted potential for a particular spin is then given by

$$\begin{aligned} U_{\text{USCF}} &= \langle \partial H_{\text{int}} / \partial \hat{n}_\sigma \rangle \\ &= U n_{\bar{\sigma}} \\ &= U (N - n_\sigma) \end{aligned} \quad (5)$$

with $n = \langle \hat{n} \rangle$ and $\bar{\sigma}$ represents the spin opposite to σ . This spin dependent unrestricted potential eliminates the self-interaction of the level to which charge is being added. A self-consistent solution of the occupancy yields an $N - \mu$ [12] plot very similar to the exact result, showing that an unrestricted calculation can capture *equilibrium* Coulomb blockade effects.

2.2. Single-level system, Fock space transport

Non-equilibrium turns out to be hard to mimic with any SCF theory, even with considerable latitude in our choice of the SCF potential. Let us assume that the contact injection rates are given by $\gamma_{1,2}/\hbar$, ignoring level broadening for the moment. Here 1, 2 also refers to the left (L) and right (R) contact. Both conventions are used in this paper to be consistent with NEGF literature. One can write down a master equation for a transition between the many-electron levels driven by the contacts

$$\dot{P}_i = - \sum_j R_{ij} P_i + \sum_j R_{ji} P_j \quad (6)$$

where i, j represent the many-body states. The master equation, intuitively quite transparent from figure 2(b), can be formally derived by decoupling the contact and molecular density matrix equations in the steady state limit and then invoking a Markov approximation that ignores memory effects such as energy dependences in the contact broadening. For our simple example of the dot with two spin levels, the rates of transition between the four Fock states, numbered as $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$, are given by

$$\begin{aligned} R_{00 \rightarrow 01} &= R_{00 \rightarrow 10} = (\gamma_1 f_1 + \gamma_2 f_2)/\hbar \\ R_{01 \rightarrow 00} &= R_{10 \rightarrow 00} = (\gamma_1 \bar{f}_1 + \gamma_2 \bar{f}_2)/\hbar \\ R_{01 \rightarrow 11} &= R_{10 \rightarrow 11} = (\gamma_1 f'_1 + \gamma_2 f'_2)/\hbar \\ R_{11 \rightarrow 01} &= R_{11 \rightarrow 10} = (\gamma_1 \bar{f}'_1 + \gamma_2 \bar{f}'_2)/\hbar \end{aligned} \quad (7)$$

where $f_{1,2} = 1/[1 + e^{(\epsilon_0 - \mu_{1,2})/k_B T}]$, $\bar{f}_{1,2} = 1 - f_{1,2}$, $f'_{1,2} = 1/[1 + e^{(\epsilon_0 + U - \mu_{1,2})/k_B T}]$, and $\bar{f}'_{1,2} = 1 - f'_{1,2}$. In short, electron addition processes are governed by probability of occupancy f at the corresponding transition energies ϵ or $\epsilon + U$ by each contact electrochemical potential, while electron removal processes are governed by the probability of vacancy $1 - f$. Thus the two current onset points occur at the bias situations shown schematically in figure 2(c).

At steady state, it is straightforward to solve these equations (only three of which are independent), along with the normalization $\sum_i P_i = 1$. The probabilities are then used to calculate the current injected by one contact (say the left one, 'L') as

$$I^L = \sum_i (\pm e/\hbar) \left[- \sum_j R_{ij}^L P_i + \sum_j R_{ji}^L P_j \right] \quad (8)$$

where the rates R^L are obtained by only considering the individual left contact contributions to the corresponding rate, for example, $R_{00 \rightarrow 10}^L = \gamma_1 f_1/\hbar$. The \pm signs correspond to addition/removal of electrons by the left contact. The resulting I - V characteristic is shown in figure 2(d).

In an SCF treatment, the current, shown dotted in figure 2(d), is obtained by solving the rate equations in the one-electron subspace. The electron occupancy is given by $N = (\gamma_1 f_1 + \gamma_2 f_2)/(\gamma_1 + \gamma_2)$, where the Fermi functions are evaluated at the energy $\tilde{\epsilon} = \epsilon + U_{\text{SCF}}$. The SCF potential in turn depends on N as described earlier, so the calculation is done self-consistently. The converged energy is then used to calculate the current as $I = (2e/\hbar)\gamma_1\gamma_2/(\gamma_1 + \gamma_2)[f_1 - f_2]$.

The RSCF model that treats spins equally tends to give an adiabatically increasing current that reaches its maximum contact-dominated value $2e/\hbar \times \gamma_1\gamma_2/(\gamma_1 + \gamma_2)$ when the contact electrochemical potential fully crosses the level. It is important to note that charging alone can smear out this current, leading to a low conductance value spread out over a wide voltage range comparable to U . The RSCF potential $U(N - 1/2)$ causes a continuous shift in levels with charge addition, which is accomplished in fractions. The unrestricted approach USCF gives an intervening Coulomb blockade plateau of width U that separates the first spin addition (or removal) event from the second. The intervening 'open-shell'

plateau is at half the maximum value for complete level filling, which is understandable because the two spins are treated on an equal footing chemically, and therefore carry equal current.

Compared with the SCF results above, the exact solution of the many-body rate equations reveals an interesting surprise that is actually quite illuminating. Of all many-electron configurations, the one-electron states (and only those) are doubly degenerate, giving us a normalization condition that differs from a simpler version that ignores spins and simply multiplies all results by two. As a consequence of this sum rule, which takes Pauli exclusion into account (preventing double up or down spin states, for example), the exact value of the open-shell current plateau depends on the Fermi functions as $\gamma_1\gamma_2/[\gamma_1(1 + f_1) + \gamma_2(1 + f_2)]$ in the large charging ($U \rightarrow \infty$) limit. For sharp levels at positive bias ($f_1 = 1, f_2 = 0$), this reaches $\gamma_1\gamma_2/[2\gamma_1 + \gamma_2]$. For the strongly non-equilibrium situation corresponding to equal resistive couplings ($\gamma_1 = \gamma_2$), the Coulomb plateau carries two-thirds of the maximum closed-shell current, in contrast with the USCF result that gives a factor of half. This implies an interesting history dependence, in that the first spin added to the empty level carries more current than the second! If we keep track of the entire many-electron configuration space, it is easy to see that this counter-intuitive result arises because there are two ways of adding the first spin and only one way of adding the second (the other channel eliminated by exclusion). This subtlety is completely washed away when we choose to work in a reduced $N \times N$ (or $2N \times 2N$ for unrestricted) subspace instead of the full $2^{2N} \times 2^{2N}$ configuration space.

The SCF potential $\langle \partial H / \partial n \rangle$ was calculated by writing the electron operators $\hat{n} = \langle \hat{n} \rangle + \delta \hat{n}$, expanding the Coulomb term $U \hat{n}_\uparrow \hat{n}_\downarrow$ and dropping the correlation terms $\delta \hat{n}_\uparrow \delta \hat{n}_\downarrow$ completely, i.e., the Hartree-Fock approximation. One could include parts of the correlation term phenomenologically, by dictating that $\hat{n}_i \hat{n}_j \approx (1 - g_{ij})(n_i \langle n_j \rangle)$, with g_{ij} representing the exchange-correlation hole. This is in the spirit of Kohn-Sham theory, where the potential is calculated by various approximate means. However, the effect of g is simply to renormalize the charging energy U that it adjoins, influencing at best the *width*, but not the *height* of the current plateaus. Thus, even for the simplest quantum dot, unrestricted potentials in the one-electron subspace cannot capture the non-equilibrium properties correctly.

It is further illuminating when a comparison between the USCF and exact model is performed. Recall that in the USCF approach, we introduce the aspect of 'self-interaction correction' as described in equation (5). Here, the presence of an electron of a particular spin adds a Coulomb cost to the other, but not to itself. Figure 3(b) shows the discrepancy between the USCF and the many-body transport calculations for a dot with two spin levels, coupled equally to two contacts (driving it far from equilibrium). As is clear from the results, the discrepancy is in the onset voltages, widths and heights of the various plateaus. While the plateau widths could be adjusted to fit the exact results by renormalizing the charging energies parametrically to account for correlation effects, the heights are independent of these values, and depend only on universal factors arising from Pauli exclusion, and cannot be

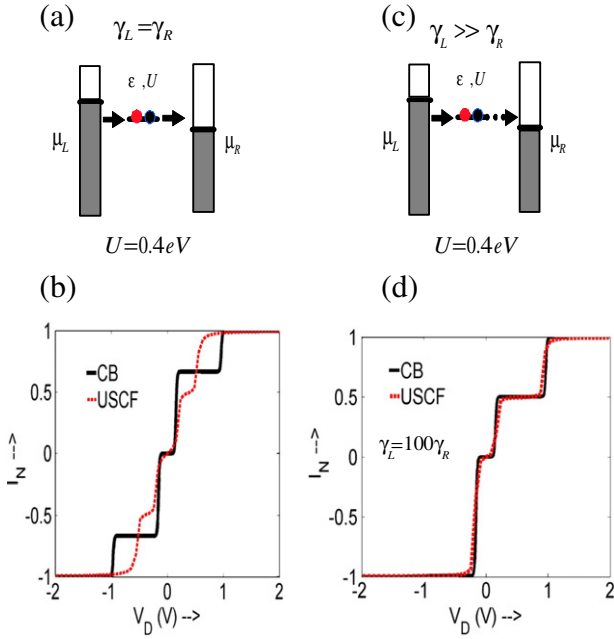


Figure 3. Comparison between USCF and exact transport results for (a) extreme non-equilibrium ($\gamma_1 = \gamma_2$), (b) in which the I - V s show significant discrepancy between USCF (dotted red) and exact (bold black) results. (c) The case of strong asymmetry implies a near equilibrium with the first contact ($\gamma_1 = 100\gamma_2$). (d) The discrepancies in plateau onset, width and height are resolved in this equilibrium limit, but are hard to resolve once we go far from equilibrium.

fixed in such a straightforward way, or by the usual DFT approach of progressively improving on correlations in the electronic structure. The issue is further underscored by the fact that in the strongly asymmetric limit ($\gamma_1 = 100\gamma_2$), where the system is essentially driven into equilibrium with the left contact, the agreement between USCF and many-body results is substantially improved. It is worth clarifying at this time though that in the multilevel generalization, this correspondence becomes a lot harder to establish *even near equilibrium*, since even the number of current plateaus generated by a $2^N \times 2^N$ Fock space approach differs substantially from its $2N \times 2N$ USCF counterpart (unless the broadening functions bear enough poles through their energy dependences to precisely account for those missing conductance peaks).

The situation gets simpler if we have asymmetric contacts $\gamma_1 \gg \gamma_2$, so that the system approaches equilibrium with the left contact. As figure 3(d) shows, the USCF agrees with the many-body limit in this asymmetric coupling case, as both approaches are dealing with a near-equilibrium problem. For positive bias on the weaker contact, the stronger contact keeps the level filled (which it can do in two ways, adding an up or a down spin, assuming the level was empty to begin with). For opposite bias, the stronger contact empties this level, which can now be done in only one way (up OR down depending on what occupied the level). The competition between ‘shell tunneling’ and ‘shell filling’ makes the I - V strongly asymmetric, with the first plateau half the second for positive bias on the weaker contact, and merging with the second for opposite bias. This

ratio of one to two, observed experimentally [26], arises in a straightforward way from our analyses since the ratio of the first and second plateau currents is given for positive bias by $(\gamma_1 + \gamma_2)/(2\gamma_1 + \gamma_2) \approx 1/2$ for $\gamma_1 \gg \gamma_2$, and by $(\gamma_1 + \gamma_2)/(\gamma_1 + 2\gamma_2) \approx 1$ for negative bias. The asymmetry arises from the difference in the number of spin addition and removal channels for positive and negative bias, and leads to an asymmetry in the current levels.

We will next show how this model is extended for a larger molecule, and how additional physics due to correlations and excitations start to arise.

2.3. General approach for multilevel systems

In the case of a larger molecule, one begins with the model Hamiltonian in second-quantized notation:

$$\hat{H} = \sum_{\alpha} \epsilon_{\alpha} \hat{n}_{\alpha} + \sum_{\alpha \neq \beta} t_{\alpha\beta} c_{\alpha}^{\dagger} c_{\beta} + \sum_{\alpha, \sigma} U_{\alpha\alpha} \hat{n}_{\alpha\sigma} \hat{n}_{\alpha\bar{\sigma}} + \frac{1}{2} \sum_{\alpha \neq \beta} U_{\alpha\beta} \hat{n}_{\alpha} \hat{n}_{\beta}, \quad (9)$$

where $\hat{n}_{\alpha} = c_{\alpha}^{\dagger} c_{\alpha}$, α, β correspond to the orbital indices of the orbitals for various sites on the molecule, and $\sigma, \bar{\sigma}$ represent a particular spin and its reverse. Exactly diagonalizing this Hamiltonian yields a large spectrum of closely spaced excitations in every charged molecular configuration. Using the equation of motion of the density matrix of the composite molecule and leads and assuming no molecule–lead correlations, one can derive [28, 27] a simple master equation for the density matrix of the system. Ignoring off-diagonal coherences, we are left with a master equation [27] in terms of the occupation probabilities P_i^N of each N -electron many-body state $|N, i\rangle$ with total energy E_i^N . The master equation then involves rates of transition $R_{(N,i) \rightarrow (N\pm 1, j)}$ between states differing by a single electron, leading to a set of independent equations defined by the size of the Fock space [10]

$$\frac{dP_i^N}{dt} = - \sum_{N, j} \left[R_{(N,i) \rightarrow (N\pm 1, j)} P_i^N - R_{(N\pm 1, j) \rightarrow (N, i)} P_j^{N\pm 1} \right] \quad (10)$$

along with the normalization equation $\sum_{i, N} P_i^N = 1$. For weakly coupled dispersionless contacts, parameterized using bare-electron tunneling rates γ_{α} , (α : left/right contact), we define rate constants

$$\begin{aligned} \Gamma_{ij\alpha}^{\text{Nr}} &= \gamma_{\alpha} |\langle N, i | c_{\alpha}^{\dagger} | N-1, j \rangle|^2 \\ \Gamma_{ij\alpha}^{\text{Na}} &= \gamma_{\alpha} |\langle N, i | c_{\alpha} | N+1, j \rangle|^2, \end{aligned} \quad (11)$$

where $c_{\alpha}^{\dagger}, c_{\alpha}$ are the creation/annihilation operators for an electron on the molecular end atom coupled with the corresponding electrode. The transition rates are given by

$$\begin{aligned} R_{(N,i) \rightarrow (N-1, j)} &= \sum_{\alpha=L, R} \Gamma_{ij\alpha}^{\text{Nr}} [1 - f(\epsilon_{ij}^{\text{Nr}} - \mu_{\alpha})] \\ R_{(N-1, j) \rightarrow (N, i)} &= \sum_{\alpha=L, R} \Gamma_{ij\alpha}^{\text{Na}} f(\epsilon_{ij}^{\text{Na}} - \mu_{\alpha}) \end{aligned} \quad (12)$$

for the removal levels ($N, i \rightarrow N-1, j$), and replacing using ($r \rightarrow a, f \rightarrow 1-f$) for the addition levels ($N, i \rightarrow$

$N + 1, j$). μ_α are the contact electrochemical potentials, f is the corresponding Fermi function, with single-particle removal and addition transport channels $\epsilon_{ij}^{\text{Nr}} = E_i^N - E_j^{N-1}$, and $\epsilon_{ij}^{\text{Na}} = E_j^{N+1} - E_i^N$. Finally, the steady state solution to equation (10) is used to get the left terminal current

$$I = \pm \frac{e}{\hbar} \sum_{N,i,j} \left[R_{(N,i) \rightarrow (N\pm 1,j)}^L P_i^N - R_{(N\pm 1,j) \rightarrow (N,i)}^L P_j^{N\pm 1} \right] \quad (13)$$

where states corresponding to a removal of electrons by the left electrode involve a negative sign. We will assume a break-junction configuration with equal electrostatic coupling with the leads, $\mu_{L,R} = E_F \mp eV_d/2$.

While the above equations include spectral details from the multiple excitations, a considerable simplification arises if we can incoherently sum over many of these excitations, leading to the ‘orthodox model’, where

$$I = \pm \frac{e}{\hbar} \sum_N \left[R_{N \rightarrow N\pm 1}^L - R_{N \rightarrow N\mp 1}^L \right] P^N. \quad (14)$$

The transition energies can be obtained in terms of a simple RC circuit, while the transition rates also compactify once we integrate them over the relevant energies, taking exclusion factors into account.

The implementation of the above sets of equations now sets the stage for further discussions.

3. Coulomb blockade: the case of electronic transport channels

In this section we describe how the excitation spectra of molecules may be probed using Coulomb blockade transport. While typical molecular I - V s look relatively featureless, what is not often appreciated is that in the Coulomb blockade limit (realized by engineering weak contacts or non-conductive backbones), electronic excitations do give rise to prominent features in the I - V [20, 21]. We will also show how such an excitation spectra can generate *intrinsic* I - V asymmetries, which in turn can manifest as NDR effects.

Exact diagonalization of the molecular Hamiltonian equation (9) generates a body of excitation spectra for every charged configuration. Our typical starting point is the equilibrium molecular quantum dot configuration in its ground state. Addition or removal of an electron takes the system to a new ground state corresponding to the singly charged cation or anion, marking the onset of conduction. In addition to the ground state however, each charged species bears a quasi-continuous excitation spectrum separated from its ground state energy by a gap that is determined by the energetics of the molecular system. Once the first excitation is accessed, the quasi-continuous excitation spectrum can be easily probed, as shown in figures 4(a)–(c). It is worth emphasizing that the excitations are energetically quite close as they differ in the rearrangement (but not in the number) of charges. At high bias with small broadening and in the absence of phonons, the purely electronic excitations do not have adequate time to relax to the ground state, accounting for their visibility in the current spectrum.

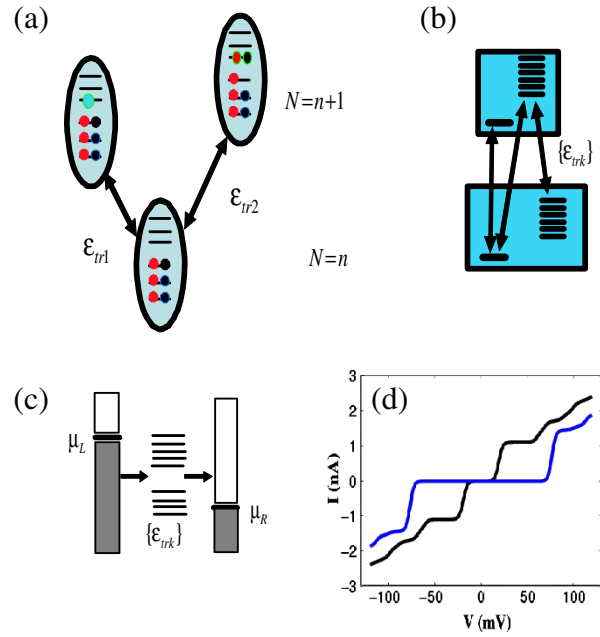


Figure 4. Electronic excitations in molecular conduction.

(a) Schematic shows three Fock states: the $N = n$ equilibrium ground state configuration and two states in the addition $N = n + 1$ spectrum. (b) Each charge configuration consists of a ground state and a set of closely spaced excitations. The accessing of such closely spaced excitations between the neutral and charged species gives rise to numerous (c) transport channels ϵ_{trk} [11]. The accessing of such transport channels gives rise to a (d) quasi-ohmic behavior in the I - V characteristics that are commonly noted in various experiments [19, 22].

The simplest and most prominent impact of Coulomb blockade on the I - V s of short molecular wires is a clear suppression of zero-bias conductance, often seen experimentally [2, 29]. However, integer charge transfer can also occur between various electronic *excitations* of the neutral and singly charged species at marginal correlation costs [30, 32]. The above fact leads to fine structure in the plateau regions [20, 19, 22, 33], specifically, a quasi-linear regime resulting from very closely spaced transport channels (ϵ_{ij}^N) via excitations. The crucial step is the accessing of the first excited state via channel $\epsilon_{10}^{\text{Nr}}$, following which transport channels involving higher excitations are accessible in a very small bias window.

The sequence of accessing of transport channels upon bias enumerated in the state transition diagrams shown in figures 4(a)–(c) determines the shape of the I - V characteristic. When the Fermi energy E_F lies closer to the threshold transport channel $\epsilon_{00}^{\text{Nr}}$ corresponding to charge transfer between two ground states, it takes an additional positive drain bias for the source to access the first excited state of the neutral system via the transition $\epsilon_{10}^{\text{Nr}}$. Under this condition, the I - V shows a *sharp rise followed by a plateau* (figure 4(d), shown in blue online), as seen in various experiments [31]. However, when transport channels that involve low lying excitations such as $\epsilon_{10}^{\text{Nr}}$ are closer to the Fermi energy E_F than $\epsilon_{00}^{\text{Nr}}$, the excitations get populated by the left contact immediately when the right contact intersects the threshold channel $\epsilon_{00}^{\text{Nr}}$, allowing for a

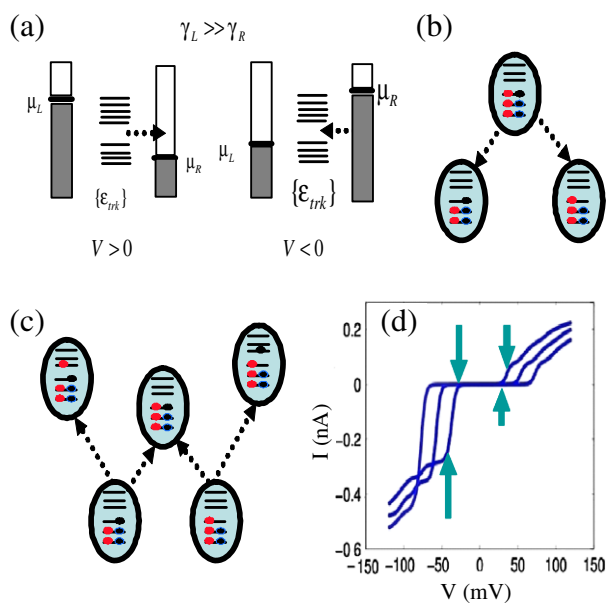


Figure 5. Current asymmetry from extrinsic ‘contact’ asymmetry. (a) In the case when the left electrode is more strongly coupled to the channel than the right one $\gamma_L \gg \gamma_R$, currents are limited by removal (addition) rates under forward (reverse) bias conditions. (b) The number of removal channels can be significantly different from the number of (c) addition channels. This gives rise to the asymmetry between (d) forward and reverse bias situations. Furthermore the application of gate potential modulates the number of transport channels available at threshold. This feature results from the Fock space excitation spectra, causing not only a change in threshold voltage with applied gate bias, but also a discernible threshold current modulation as noted here, and in various experiments [20].

simultaneous population of both the ground and first excited states via ϵ_{00}^{Nr} and ϵ_{10}^{Nr} at threshold. Under these conditions the $I-V$ shows a sharp onset followed immediately by a quasi-linear regime (figure 4(d) shown in black) with no intervening plateaus, as observed frequently in $I-V$ s of molecules weakly coupled with a backbone [19, 22, 20].

3.1. Extrinsic asymmetries—the case of rectification

The direct role of excitations in conduction becomes particularly striking under asymmetric coupling ($\gamma_L \gg \gamma_R$) with contacts [20, 34]. Due to this extrinsic asymmetry, current magnitudes are dictated by the weaker contact. This asymmetry directly affects the forward and reverse bias characteristics, leading to current rectification. This rectification is caused by the inherent asymmetry between addition and removal processes, each of which is rate limiting depending on the bias direction, as shown in figures 5(a)–(c).

In contrast to the SCF regime where unequal charging drags out a same level current over different voltage widths [35], in the CB regime we encounter clear intermediate current steps from open shells, with current heights that are themselves asymmetric at threshold (figure 5(d)). This asymmetry arises due to the difference in number of pathways for removing or adding a spin, taking in particular into account the possible excitation channels between the neutral and singly

charged species (figures 5(b) and (c)). The number of such accessible excitations at threshold can be altered with an external gate bias, leading to a prominent gate modulation of the threshold current heights, over and above the modulation of the onset voltages and the conductance gap [20] (figure 5(d)). Furthermore, it is easy to show that the asymmetry will flip between gate voltages on either side of the charge degeneracy point, as is also observed experimentally [34].

It is worth emphasizing that the sophistication arose specifically due to the presence of separate identifiable contributions from the open shells, which are normally overwhelmed by broadening if we are away from the CB limit. However, the specific identification of these excitations is not crucial to the qualitative shape of the $I-V$ s, as long as one is not looking too closely at the individual spectral features. A simpler orthodox model would then suffice to reproduce these broad features described above, such as the transition between steps and slopes in the $I-V$ s, the flipping of asymmetry and the gate modulation of the current levels [13]. However, there are notable exceptions, such as for intrinsic asymmetries, where a clear separation needs to be made between certain classes of excitations with longer lifetimes (‘traps’) and the regular excitations with shorter lifetimes (‘channels’) that are responsible for the current flow.

3.2. Intrinsic asymmetries—NDR, hysteresis and telegraph noise

The physics of NDR can be explained readily using an USCF model that is actually quite intuitive. Consider a channel and a trap, the channel being strongly coupled to the contacts and the trap weakly coupled. Accessing the channel creates a resonant onset of current. Accessing the trap subsequently would keep the conduction unaffected, as the trap is non-conducting; however, once we throw in the strong Coulomb repulsion arising from charging up the trap, it is easy to see that this repulsion can expel the conducting channel out of the conducting bias window, leading to an NDR. Self-interaction correction is crucial to this picture, as the charging should repel the channel level, but not the trap level itself. Further sophistications can arise from considering the lifetime of the trap. If the trap does not release its charge during the measurement time, then a reverse scan would keep the channel blocked and lead to a hysteresis. Such a hysteresis is scan rate dependent, as the rate determines the degree to which the trap releases its captive charge. The low current state can be reversed by going to large negative bias to expel this charge and restore initial conditions.

The realization of the NDR involves two conditions: (a) that the charging of the trap is large enough to expel much of the channel from the bias window, and (b) that the resulting current carried by the trap plus the residual tail of the channel sitting in the bias window exceeds the current carried originally by the channel alone. The first amounts to the onset condition for an NDR, while the second describes, in some sense, the effectiveness of the NDR (in terms of an inequality). With a little bit of effort, one can extract a range of parameters that satisfy both these conditions. We will instead show that

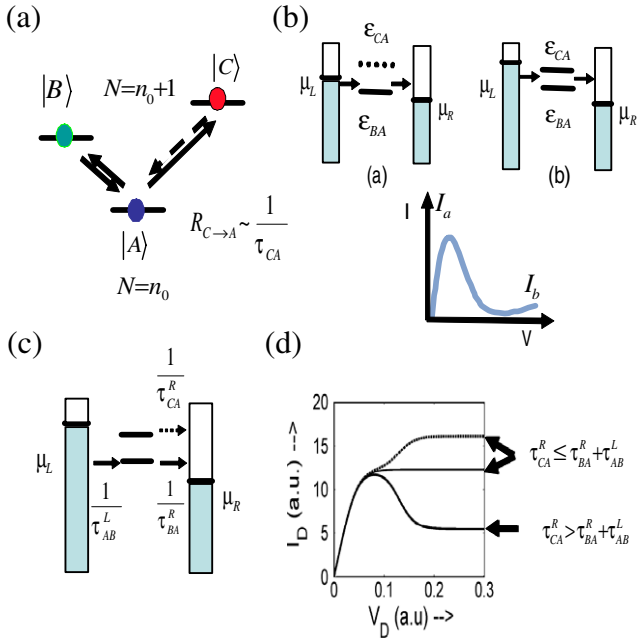


Figure 6. NDR effects from intrinsic asymmetry or ‘dark states’. (a) A generic mechanism NDR can be cast in terms of three device Fock space states, and transitions $|A\rangle \leftrightarrow |B\rangle$ and $|A\rangle \leftrightarrow |C\rangle$ between those that differ by a single electron. The dark state, $|C\rangle$ say, has a very slow removal rate in comparison to its addition rate. (b) Bias environments, (a) and (b), correspond to current rise followed by a collapse. (c) A general criterion for such a current collapse or NDR to occur with increasing voltage can be cast in terms of the dark state removal rates. (d) NDR in the I – V characteristic is contingent on the criteria derived here [14].

even in the Fock space picture (which treats the channel plus trap as a composite system), one can identify these two same conditions: the NDR starts when one encounters what we refer to as a ‘dark’ or ‘blocked’ state, and subject to this condition, the effectiveness of the NDR amounts to an inequality that we will now discuss.

A condition for NDR can be derived much more generally in terms of three Fock space states $|A\rangle$, $|B\rangle$, and $|C\rangle$ with energies $E_A < E_B < E_C$ respectively (figure 6(a)), representing three accessible states within the bias range of interest. For instance, $|A\rangle$, $|B\rangle$ could be the ground states of the $N = n_0$ and $N = n_0 + 1$ electron systems, and $|C\rangle$ the first excited state of the $N = n_0 + 1$ -electron system. Transport of electrons involves single-charge removal or addition between states $|B\rangle$, $|C\rangle$ and $|A\rangle$ that differ by an electron, via addition and removal rates $R_{A\leftrightarrow B,C} \propto \frac{1}{\tau_{A\leftrightarrow B,C}}$. Such an electron exchange is initiated when reservoir levels are in resonance with the single-electron transport channels $\epsilon_{BA} = E_B - E_A$ and $\epsilon_{CA} = E_C - E_A$ respectively. The I – V characteristic of this three-state system, shown schematically in figure 6(b), shows two plateaus with current magnitudes I_a and I_b respectively. Current collapse or NDR occurs when $I_a > I_b$.

In the above system, NDR occurs when under specific conditions, the state $|C\rangle$ can be a blocking or dark state, for which electron addition is feasible while removal is rate limiting. This can happen when there are intrinsic asymmetries within the transport problem, as in the trap–channel dichotomy

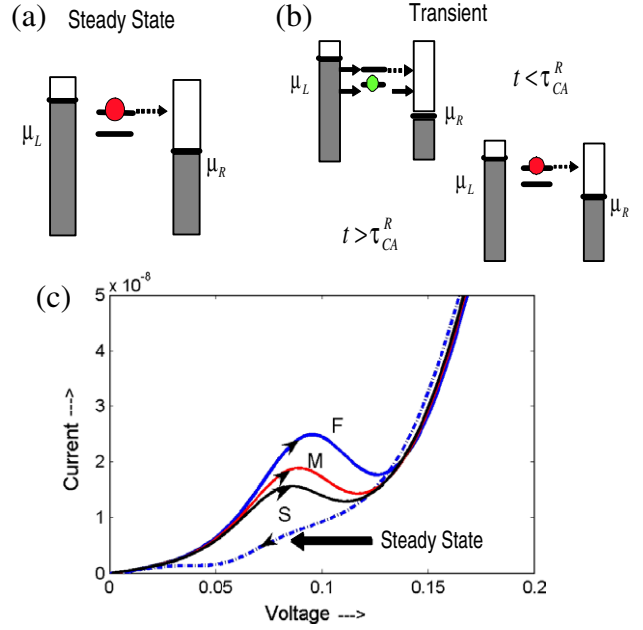


Figure 7. Hysteresis effects due to sweep rate. (a) Under steady state conditions the dark state behaves as an effective blocker. (b) A sweep rate faster than the rate determining timescale can put the transport process in a transient state. (c) Under such transient conditions sweep rate can induce disguised NDRs and often hysteresis, both depending on sweep rate [36].

described earlier. Regardless of the specific origin of blocking, a simple criterion for current collapse can be derived [14] in terms of the rate limiting process, that is the rate of electron removal from the dark state $|C\rangle$, $R_{C\leftrightarrow A}$. A better intuition is provided by thinking of rates as inverse lifetimes, as shown in figure 6(c). If the lifetime of state $|C\rangle$, τ_{CA} , exceeds the sum of the addition and removal rates of the conducting state, it becomes an effective blocker. In other words, the condition $\tau_{CA}^R > \tau_{AB}^L + \tau_{BA}^R$ determines whether NDR occurs or not. The superscripts L and R represent the left or right electrode, and under the forward bias situation add and remove electrons respectively. The consequence of the above criterion is summarized in figure 6(d), clearly indicating that NDR only occurs when $\tau_{CA}^R > \tau_{AB}^L + \tau_{BA}^R$. A similar criterion is valid for the reverse bias direction on interchanging superscripts L with R.

Further sophistications may arise due to ‘transient’ probing. The criterion obtained above applies for steady state, where the dark state gets occupied with certainty, resulting in an inevitable current blockade regime. Let us consider such a blocked current state, shown dotted in figure 7(c). A ‘disguised’ NDR can also be achieved if the voltage sweep rates are faster than the rate determining dark state lifetime. Such a sweep rate simply means that the dark state has not yet been occupied and thus a current determined by the addition and removal times $1/(\tau_{AB}^L + \tau_{BA}^R)$ can still flow. This situation is shown in figure 7(c), in which the onset of the dark state can be delayed provided the sweep rate is fast enough. But once the dark state forms, the current is blocked and thus remains so. This implies that on the backward sweep current remains blocked, resulting in a sweep rate dependent hysteresis. Such

a behavior has been noted by Kiehl *et al* [36] in a molecular system, which was attributed to a slow charge trapping process, which again fits into the dark state picture. Just as noted here, Kiehl *et al* also observe that the NDR effect diminishes with decreasing sweep rate and ultimately vanishes at steady state, implying the absence of a real hysteresis at steady state. This kind of hysteresis is just an artifact of sweep rates being faster than the rate determining processes.

It must be mentioned that certain hysteretic processes can occur even at steady state due to the true bistable nature of the system. A classic example is the charging induced hysteresis in resonant tunneling diode (RTD) structures [37]. Other interesting examples include bistability caused by hyperfine interactions [39] causing a hysteresis with applied magnetic field, and optical bistability created by superlattice dielectrics that show strong terahertz non-linearity due to Bloch oscillation and dynamic localization of electrons [40, 41].

While we discussed the interesting case of dark states earlier, touching upon relevant consequences such as NDR and hysteresis, we avoided particular examples. Recently, we applied our dark state model [14] to explain the NDR observed due to subtle spin correlation effects in double quantum dots [38]. Indeed this NDR results in the spin blockade regime [38] which currently forms a key concept in the area of single-spin manipulation and control [5, 15–18].

Finally, the transient probing with a high resolution probe would allow us to explore the approach to resonance with the trap states. As we discuss in [23], the stochastic blocking and unblocking of the channel by the occupation/deoccupation of the traps near resonance generates a flicker in the output current known as random telegraph noise. The ratio of capture and emission times is given by a Boltzmann factor whose argument depends on the trap energy location. Shifting this trap with a combination of gate and drain voltages leads to an associated scaling of the capture to emission time ratio, allowing one to infer the spatial and spectral location of the trap states. This provides a powerful ‘barcode’ for characterizing single-molecular defects.

Our analyses over the last few sections focused on electron–electron correlations and their treatment in a Fock space approach that captures the relevant physics (if at the expense of computational simplicity). In the next section, we will discuss how this approach can be extended to incorporate electron–phonon correlations in transport.

4. Electron–phonon interactions

So far, our primary concern was Coulomb blockade and electronic excitations. Here, we show that coupling Coulomb blockade with phonon-assisted tunneling results in non-trivial physics. This section is primarily motivated by a recent series of experiments performed on suspended carbon nanotube quantum dots [24]. In such a suspended system, the phonons are driven far out of equilibrium and couple strongly with the electronic system. Our approach then is to invoke the Fock space of the electron–phonon system as indicated in figures 1(c) and (d).

4.1. SCF treatment—IETS and phonon sidebands

The SCF approach to electron–electron interactions has been discussed at length in our earlier papers and contrasted with the Fock space approach. It is worth quickly touching upon the SCF treatment of phonons, before diving into its Fock space analog. In presence of dephasing scattering events, the current at the left contact can be written in the NEGF formalism [42] as

$$I_L = \frac{2e}{h} \int dE \text{Tr}[\Sigma_L^{\text{in}} A - \Gamma_L G^n] \quad (15)$$

where A is the spectral function, Γ_L is the broadening by the left contact, Σ_L^{in} is the in-scattering self-energy from the left contact and G^n is the correlation function describing the energy dependent occupancy of the levels, taking quantum interference into account. The influence of scattering by contacts and phonons resides in $\Sigma^{\text{in}} = \Sigma_L^{\text{in}} + \Sigma_R^{\text{in}} + \Sigma_{\text{ph}}^{\text{in}}$. The equations for the contact Σ s are well known—we will focus here on the additional contributions from the phonon scattering processes. Within the self-consistent Born approximation (self-consistency needed to conserve current), the equations connecting the phonon contributions form two groups—a set of dynamic equations, and a set of static equations. The non-equilibrium dynamics describing the filling and emptying of states is described by

$$\begin{aligned} G^{n,p}(E) &= G(E) \Sigma_{\text{ph}}^{\text{in,out}}(E) G^\dagger(E) \\ \Sigma_{\text{ph}}^{\text{in,out}}(E) &= D_0(\omega) \otimes [n_B(\omega) G^{n,p}(E \mp \hbar\omega) \\ &\quad + (n_B(\omega) + 1) G^{n,p}(E \pm \hbar\omega)] \end{aligned} \quad (16)$$

where $n_B(\omega) = [e^{\hbar\omega/k_B T} - 1]^{-1}$ is the equilibrium Bose–Einstein distribution of the phonons, and D_0 is its deformation potential, in other words, $(D_0)_{ij} = h_i h_j$, h_α being the electron–phonon coupling constant at the real space point α (in general, h is a non-diagonal matrix in an arbitrary basis set, and D_0 is a fourth-rank tensor). \otimes denotes an element-by-element (as opposed to matrix) multiplication. The equations above are for a single phonon mode at frequency ω , and will need to be integrated over a phonon density of states for a continuous distribution of phonons.

The static equations describing the states themselves are given by

$$\begin{aligned} \Gamma_{\text{ph}}(E) &= \Sigma_{\text{ph}}^{\text{in}}(E) + \Sigma_{\text{ph}}^{\text{out}}(E) \\ \Sigma_{\text{ph}}(E) &= \mathcal{H}(\Sigma_{\text{ph}}) + i \frac{\Gamma_{\text{ph}}}{2} \end{aligned} \quad (17)$$

$$G(E) = [EI - H - \Sigma_L - \Sigma_R - \Sigma_{\text{ph}}]^{-1},$$

where \mathcal{H} denotes the Hilbert transform.

The above equations capture the physics of phonon-assisted tunneling in all its various limits. In its simplest form, it contributes to dephasing that reduces the ballistic current in devices. Near resonance, the $\pm\hbar\omega$ terms in the arguments of the $G^{n,p}$ matrices contributing to $\Sigma^{\text{in,out}}(E)$ generate phonon sidebands. The scaling of these sidebands depends on the deformation potential as well as the distribution of the phonons. We have assumed this to be equilibrium Bose–Einstein n_B , although one could generalize it provided we

have a separate evolution equation for the phonon dynamics coupled to the electron transport equations that describe parametrically or otherwise how the phonons are driven away from equilibrium by the electronic subsystem, and how the other terms in the phonon evolution equation (typically involving coupling with a thermal reservoir) try to bring this back to equilibrium. In the next section, we will, in fact, include these ‘hot’ phonon equations explicitly; instead of separate coupled equations, however, we will treat the electron–phonon as a composite system, and use the couplings between the electrons, phonons, and with the contacts and the bath as processes driving the evolution of the composite system.

The NEGF formalism also allows us to get away from this phonon-assisted tunneling limit to the off-resonant limit. When the electronic levels lie far from resonance, their sidebands do not show up and we get instead the inelastic electronic tunneling spectrum (IETS) of the molecular system. For weak electron–phonon coupling (retaining leading order terms in D_0), the NEGF algebra simplifies considerably, so that the current partitions into an elastic component given by the usual Landauer formula, as well as an inelastic component that explicitly involves exclusion principle terms at different emission and absorption energies. For applied voltages larger than the phonon frequency, these terms create additional current transport channels through phonon emission, creating a slight increase in the current that only shows up as peaks in the second derivative with voltage, generating the familiar IETS spectrum. The physics resides entirely in the inelastic current—one can use sophisticated electron structure methods to extract these peak positions from the phonon frequencies, as well as the IETS peak heights from the computed deformation potentials. The results also show additional subtleties near resonance that are experimentally observed. Specifically, near resonance the elastic current also picks up phonon signatures from the phonon contributions to Σ residing in G , *generating a dip rather than a peak* that arises from the phonon sidebands described above (the equations also generate a familiar polaronic shift of the main peak). The inclusion of the equilibrium n_B distribution gives phonon emission peaks that are stronger in strength than absorption peaks by Boltzmann ratios evaluated at the electronic temperature.

While encouraging agreement between computed and experimental off-resonant behavior, specifically, IETS spectra, has been reported [43], the resonant phonon sidebands and the scaling of their conductance peak heights with current shows significantly more complex behavior arising from strong non-equilibrium electron–phonon correlations, which necessitates a Fock space approach. We will now introduce the joint electron–phonon Fock space and the experimental ramifications of electron–phonon correlation effects captured with this treatment.

4.2. Fock space treatment of phonon sidebands

We start from a model Hamiltonian for a quantum dot having onsite energies ϵ_i , Coulomb interaction energy $U_{ii'}$, vibronic modes at energy $\hbar\omega_j$ and electron–phonon coupling λ_{ij} . The

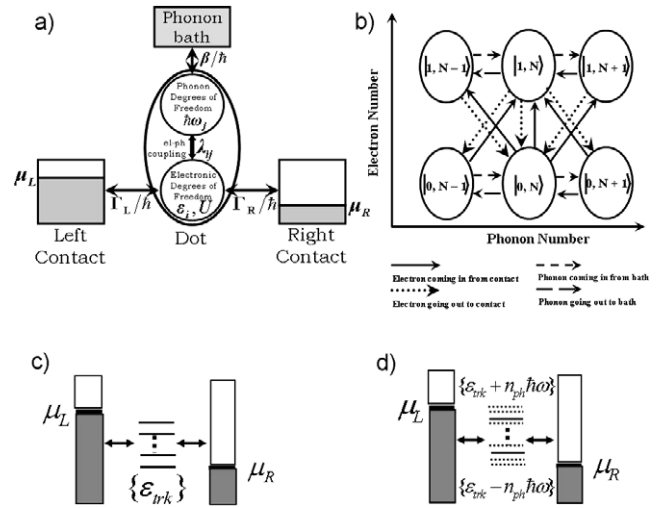


Figure 8. Fock space model to include electron–phonon interactions. (a) The dot is electrically connected to the left (right) contact (with electron tunneling rates $\Gamma_{L,R}/\hbar$) and mechanically to the phonon bath (with a phonon escape rate of β/\hbar). The dot has electronic degrees of freedom ϵ_i and phonon degrees of freedom $\hbar\omega_j$ ($i, j = 1, 2, 3, \dots$) with coupling λ_{ij} . (b) Transition between different states of a dot, having one electronic level and one phonon mode, due to coupling to the contact and the phonon bath ($|0, N\rangle$ and $|1, N\rangle$ denote the states of the system having N phonons each but zero electrons and one electron respectively). (c) Effective Coulomb blocked channel in the absence of phonons. (d) Inclusion of electron–phonon coupling adds phonon sidebands to the existing transport channels.

system is connected to electrical contacts with couplings $\Gamma_{1,2}$ and to a thermal bath with coupling β (figure 8(a)). Electronic transport due to bias applied to the contacts as well as phonon emission and absorption processes lead to transitions between many-body states $|e_{N_e}^i, k\rangle$ (k phonons and i th electronic level in the N_e electronic subspace) of the quantum dot (figure 8(b)). The rates of these transitions due to left (right) contact ($\mathcal{R}_{|e_{N_e-1}^i, k\rangle \rightarrow |e_{N_e}^i, p\rangle}^{L(R)}$) and the phonon bath ($\mathcal{R}_{|e_{N_e-1}^i, k\rangle \rightarrow |e_{N_e}^i, p\rangle}^{ph}$) are calculated by applying Fermi’s Golden Rule [25]. When an electron is added or removed via standard transport processes, it can also change the number of phonons, taking the dot from a state $|0, N\rangle$ to $|1, N \pm p\rangle$ or vice versa, where p is the number of phonons emitted or absorbed. When the quantum dot absorbs or emits a phonon, the state changes from $|X, N\rangle$ to $|X, N \pm 1\rangle$, where $X = 0$ or 1 . The consequence of phonons in transport is the addition of extra transport channels to the already existing Coulomb blockade transport channels, as shown in figures 8(c) and (d).

The rate equations in the above Fock space (8) can be solved for a finite number of phonon emission or absorption channels to extract the resulting current–voltage characteristic. Focusing on the intriguing experiments on Coulomb blocked nanotube quantum dots with prominent breathing modes [24], one sees multiple intriguing features. (a) The absorption and emission sideband heights do not scale as simply as above. This is because the number of phonons N_{ph} itself changes with current, in addition to being driven far from equilibrium in the suspended sections of the tube with small escape rate β . (b) The scaling of sidebands with phonon population

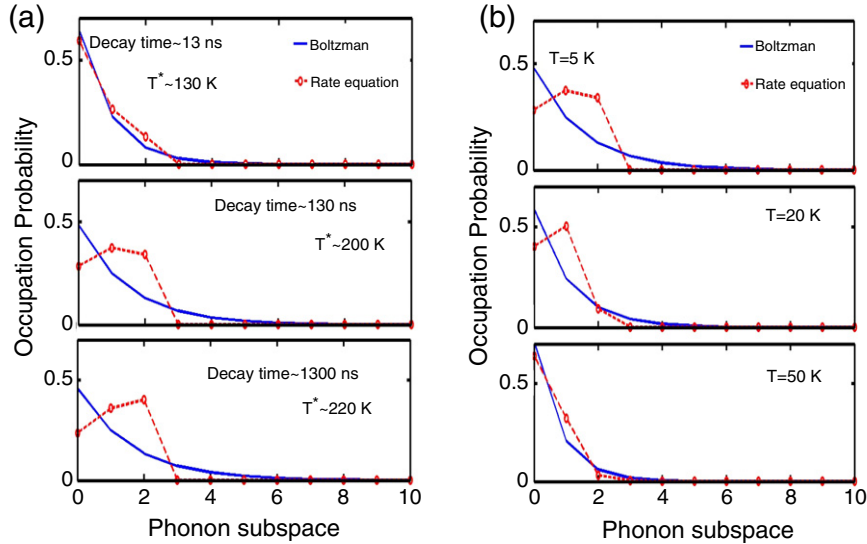


Figure 9. The phonon distribution P_k^{phB} from the Boltzmann distribution (a) (solid line) with the effective temperature T^* differs considerably from the phonon distribution P_k^{ph} calculated directly from the rate equation (circles). (b) The phonon distribution P_k^{ph} approaches the Boltzmann distribution P_k^{phB} (with proper effective temperature T^* not shown here) at elevated surrounding temperature T .

differs significantly from predictions of the analogous Tien–Gordon theory of photon-assisted tunneling (PAT) [44]. This discrepancy arises because unlike in the PAT experiments, the phonons are not coherent, and are partly correlated with the nanotube electrons [25].

4.3. Non-equilibrium phonons and effective temperature

A crucial component of the above picture is the strongly non-equilibrium distribution for the phonon population, as their rate of generation by the drive current (determined by $\Gamma_{L,R}$ and λ) exceeds their extraction rate β . From the solution to the rate equations giving us the joint electron–phonon occupation probabilities, we can calculate the probability distribution of the phonon by summing over the electronic subspaces: $P_k^{\text{ph}} = \sum_{s, N_e} P_{|e_{N_e, k}^s\rangle}$. Defining an effective temperature T^* corresponding to the non-equilibrium phonon occupation $N_{\text{ph}} = [e^{\hbar\omega/k_B T^*} - 1]^{-1}$, we find the corresponding Boltzmann distribution: $P_k^{\text{phB}} = e^{-k\hbar\omega/k_B T^*} / Z$ of phonon subspace occupation probability, where Z is the partition function (T^* is identified by fitting the higher energy tails between the two distributions). A comparison between P_k^{ph} and P_k^{phB} at different biases for different decay rates of the phonons reveals that they differ considerably as the phonon decay rate β/\hbar decreases (figure 9(a)) at some applied bias. The deviation from a Bose–Einstein like shape suggests that the phonons in the suspended nanotubes are strongly non-equilibrium, so that *temperature* is not a well-defined quantity except for describing the higher energy tails (figure 9(b)).

4.4. Effect of the surrounding temperature: anomalous phonon population

The non-equilibrium phonon population reveals a striking and somewhat counter-intuitive dependence on the substrate temperature, *the phonon population decreasing with increasing*

temperature at certain bias voltages (figure 10(a)). This peculiar behavior manifests itself as long as the surrounding temperature is smaller than the separation between the upper emission and lower absorption sidebands of two subsequent Coulomb blockade peaks. Beyond this temperature the phonon occupancy increases monotonically with temperature for all bias values, as expected. The anomalous temperature dependence arises from a trade-off between phonon generation and recombination rates in the coupled dot–lead–bath system. The behavior is observed specifically at bias values corresponding to the onset of a new phonon absorption channel (figure 10(b), bottom left). With increasing temperature from ~ 5 to 20 K, the increasing tails of the contact Fermi functions redistribute the electrons from a phonon emission sideband $\epsilon_1 + \hbar\omega$ of a lower electronic peak to an absorption sideband $\epsilon_2 - \hbar\omega$ of a higher peak (figure 10(b), bottom right). Under this specific combination of temperature and bias values, the number of electrons resonant with the phonon sidebands decreases with increasing temperature, decreasing the efficiency of the phonon-assisted transport. At higher temperatures between 20 and 50 K, the temperature is large enough to open new emission channels that eventually increase the phonon population at all bias voltages.

5. Concluding remarks

For a wide variety of transport problems, a perturbative treatment of interactions coupled with a quantum kinetic theory (such as NEGF) does an admirable job of explaining and predicting experimental features and providing important physical insights. As objects scale towards nanodimensions, however, strong confinement and poor coupling with the surroundings lead to increasing degrees of correlation, especially at lower temperatures. In this limit, the Fock space approach (more generally, the many-body density matrix

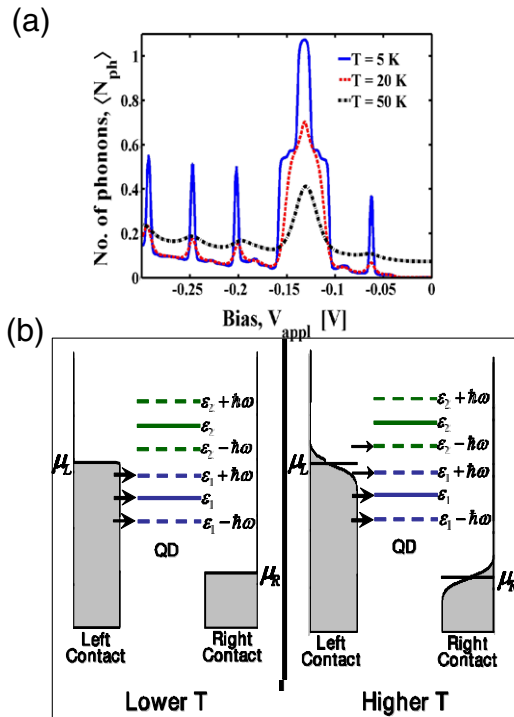


Figure 10. Anomalous temperature dependence of phonon distribution. (a) The number of phonons decreases at some bias points with increasing surrounding temperature T . (b) Bottom left: at lower surrounding temperature excitation of energy $\epsilon_1 + \hbar\omega$ associated with resonant level ϵ_1 falls inside the bias window between $\mu_L + k_B T_L$ and $\mu_R - k_B T_R$ and the excitation of energy $\epsilon_2 - \hbar\omega$ associated with resonant level ϵ_2 falls outside the bias window. Bottom right: at elevated surrounding temperature Fermi functions of the contacts broaden around respective chemical potentials and the excitation of energy $\epsilon_2 - \hbar\omega$ enters the bias window, and a lesser number of electrons appear at energy $\epsilon_1 + \hbar\omega$ in the left contact.

approach) naturally allows us to compute transport signatures, provided we have a suitable means of extracting the various correlated many-body states and the rates of transition among them. Exact diagonalization provides one option, although this quickly becomes computationally intractable. Partial configuration interaction (CI) schemes may prove to be more practical. In contrast with the NEGF-SCF limit case where one could aim for quantitative and predictive accuracy with increasing amounts of chemical sophistication, such details are hard to build into the Fock space approach and need to be replaced by simpler, model problems, with parameters that could be benchmarked with more detailed models and measurements. However, even these simple models (with a reasonable choice of parameters) show qualitatively new physics that is experimentally observable, ranging from gate-modulated current levels, gate tunable excitation spectra, scan rate dependent NDR and hysteresis, and the breakdown of our common intuition based on equilibrium phonon sideband scaling and the classical theories of photon-assisted tunneling. While these two limits (quantum wire and quantum dot) are separately well understood, at least formally, the intermediate coupling regime between the two becomes particularly challenging to model as there is no small ‘fine

structure’ parameter that would allow a convenient starting point for a perturbation expansion (e.g. a non-interacting wire or a fully interacting but isolated quantum dot). Significant progress is needed at formal, computational and experimental levels in order to probe this regime, which bears the promise of completely novel physics of non-equilibrium correlations as well as possible applications based on the interaction between conducting detector elements in the quantum wire regime and non-conducting storage elements in the quantum dot regime.

References

- [1] Lundstrom M and Guo J 2005 *Nanoscale Transistors: Device Physics, Modeling and Simulation* (New York: Springer)
- [2] Reed M A, Zhou C, Muller D J, Burgin T P and Tour J M 1997 *Science* **278** 252
- [3] Chen J, Reed M A, Rawlett A M and Tour J M 1999 *Science* **286** 1550
- [4] van der Weil W G, DeFranceschi S, Elzerman J M, Fujisawa T, Tarucha S and Kouwenhoven L P 2003 *Rev. Mod. Phys.* **75** 1 and references therein
- [5] Hanson R, Kouwenhoven L P, Petta J R, Tarucha S and Vandersypen L M K 2007 *Rev. Mod. Phys.* **79** 1217
- [6] Xiao M, Martin I, Yablonovitch E and Jiang H W 2004 *Nature* **430** 435
- [7] Loss D and DiVincenzo D P 1998 *Phys. Rev. A* **57** 120
- [8] Grabert H and Devoret M H (ed) 1992 *Single Charge Tunneling (NATO ASI Series vol 294)* (New York: Plenum)
- [9] Shriver D and Atkins P W 2006 *Inorganic Chemistry* (San Francisco, CA: Freeman)
- [10] Bonet E, Deshmukh M M and Ralph D C 2002 *Phys. Rev. B* **65** 045317
- Beenakker C W J 1991 *Phys. Rev. B* **44** 1646
- Elste F and Timm C 2005 *Phys. Rev.* **71** 155403
- [11] Muralidharan B, Ghosh A W and Datta S 2006 *Phys. Rev. B* **73** 155410
- [12] Muralidharan B, Ghosh A W and Datta S 2006 *J. Mol. Simul.* **32** 751
- [13] Miller O D, Muralidharan B, Kapur N and Ghosh A W 2008 *Phys. Rev. B* **77** 125427
- [14] Muralidharan B and Datta S 2007 *Phys. Rev. B* **76** 035432
- [15] Petta J R, Johnson A C, Taylor J M, Laird E A, Yacoby A, Lukin M D, Marcus C M, Hanson M P and Gossard A C 2005 *Science* **309** 2180
- [16] Johnson A C, Petta J R, Taylor J M, Yacoby A, Lukin M D, Marcus C M, Hanson M P and Gossard A C 2005 *Nature* **435** 925
- [17] Koppens F H L, Buizert C, Tielrooij K J, Vink I T, Nowack K C, Meunier T, Kouwenhoven L P and Vandersypen L M K 2006 *Nature* **442** 776
- [18] Nowack K C, Koppens F H L, Nazarov Yu V and Vandersypen L M K 2007 *Science* **318** 1430
- [19] Reichert J, Weber H B, Mayor M and von Lohneysen H 2003 *Appl. Phys. Lett.* **82** 4137
- [20] Park J, Pasupathy A N, Goldsmith J I, Chang C, Yaish Y, Petta J R, Rinkovski M, Sethna J P, Abruna H D, McEuen P L and Ralph D C 2002 *Nature* **417** 722
- [21] Zhitenev N B, Meng H and Bao Z 2002 *Phys. Rev. Lett.* **88** 226801
- [22] Mayor M, Weber H B, Reichert J, Elbing M, von Hanisch C, Beckmann D and Fischer M 2003 *Angew. Chem. Int. Edn* **42** 5834
- [23] Vasudevan S, Walczak K and Ghosh A W 2008 *IEEE Sens.* **8** 857
- [24] LeRoy B J, Lemay S G, Kong J and Dekker C 2004 *Nature* **432** 371

- [25] Siddiqui L, Ghosh A W and Datta S 2007 *Phys. Rev. B* **76** 085433
- [26] Deshmukh M M, Bonet E, Pasupathy A N and Ralph D C 2002 *Phys. Rev. B* **65** 073301
- [27] Braig S and Brouwer P W 2005 *Phys. Rev. B* **71** 195324
- [28] Elste F and Timm C 2005 *Phys. Rev. B* **71** 155403
- [29] Xiao X, Xu B Q and Tao N J 2001 *Nano Lett.* **4** 267
Xui B Q and Tao N J 2003 *Science* **301** 1221
- [30] Fulde P 1991 *Electron Correlations in Molecules and Solids* (*Springer Series in Solid-State Sciences* vol 100) (Berlin: Springer)
- [31] Lee J-O, Lientschnig G, Wiertz F, Struijk M, Janssen R A J, Egberink R, Reinhoudt D N, Hadley P and Dekker C 2003 *Nano Lett.* **3** 113
- [32] Narvaez G A and Kirczenow G 2003 *Phys. Rev. B* **68** 245415
- [33] Elbing M, Ochs R, Koentopp M, Fischer M, von Hanisch C, Evers F, Weber H B and Mayor M 2005 *Proc. Natl Acad. Sci.* **102** 8815
- [34] Scott G D, Chichak K S, Peters A J, Cantrill S J, Stoddart J F and Jiang H-W 2005 *Preprint* [cond-mat/0504345](#)
- [35] Zahid F, Ghosh A W, Paulsson M, Polizzi E and Datta S 2004 *Phys. Rev. B* **70** 245317
- [36] Kiehl R A, Le J D, Candra P, Hoye R C and Hoye T R 2006 *Appl. Phys. Lett.* **88** 172102
- [37] see for example Capasso F, Mohammad K and Cho A Y 1986 *IEEE J. Quantum Electron.* **22** 1853 and references therein
- [38] Ono K, Austing D G, Tokura Y and Tarucha S 2002 *Science* **297** 1313
- [39] Ono K and Tarucha S 2004 *Phys. Rev. Lett.* **92** 256803
- [40] Ghosh A W, Kuznetsov A V and Wilkins J W 1997 *Phys. Rev. Lett.* **79** 3494
- [41] Ghosh A W and Wilkins J W 2000 *Phys. Rev. B* **61** 5423
- [42] Datta S 2005 *Quantum Transport: Atom to Transistor* (Cambridge: Cambridge University Press)
- [43] Paulsson M, Frederiksen T, Ueba H, Lorente N and Brandbyge M 2007 *Preprint* [0711.3392](#) [cond-mat]
- [44] Tien P K and Gordon J P 1963 *Phys. Rev.* **129** 647