Bistable States of Quantum Dot Array Junctions for High-Density Memory

David M.-T. Kuo and Yia-Chung Chang

Department of Electrical Engineering, National Central University, Chungli 320, Taiwan
Department of Physics, Academia Sinica, Taipei 115, Taiwan

Received June 5, 2009; accepted August 3, 2009; published online October 20, 2009

We demonstrate that two-dimensional (2D) arrays of coupled quantum dots (QDs) with six-fold degenerate p orbitals (including spin degeneracy) can display bistable states, suitable for application in high-density memory device with low power consumption. Due to the inter-dot overlap of p_\perp orbitals in these QD arrays, two dimensional conduction bands can be formed in the x–y plane, while the p_\parallel orbitals remain localized in the x–y plane such that the coupling between p_\parallel orbitals located at different dots can be neglected. We model such systems by taking into account the on-site repulsive Coulomb interactions (U) between electrons in any of the three p orbitals, which also lead to a coupling between the localized p_\parallel orbitals with the 2D conduction bands formed by p_\perp/p_\parallel orbitals. The Green’s function method within an extended Anderson model is used to calculate the tunneling current through the QDs. We find that bistable tunneling current can exist for such systems due to the interplay of the on-site Coulomb interactions between the p_\parallel orbitals and the delocalized nature of conduction band states derived from the hybridization of p_\perp/p_\parallel orbitals. This bistable current depends critically on the strength of U, the band width, and the ratio of the left and right tunneling rates. The behavior of the electrical bistability can be sustained when the 2D QD array reduces to a one-dimensional (1D) QD array, indicating the feasibility for high-density packing of these bistable nanoscale structures.

© 2009 The Japan Society of Applied Physics

1. Introduction

Intrinsic hysteresis in DC current–voltage (I–V) characteristics is one of the most intriguing problems for resonant tunneling diodes (RTDs). Such a bistability has an important application in memory devices. Whether this phenomenon exists in nanoscale devices such as single-electron transistors (SETs) and single molecular transistors (SMTs) has been theoretically investigated in refs. 5–7. Alexandrov and coworkers pointed out that the tunneling current through a highly degenerate states of a single quantum dot (QD) (molecular) can lead to a switching effect only in the case of attractive electron–electron interactions, which is mediated by the electron–phonon interaction. On the basis of Hartree approximation and polaron effect Galperin et al. proposed that the hysteresis of I–V characteristics can be observed in a single molecular junction with effective attractive electron Coulomb interaction, although they considered only one energy level. Recently, Magna and Deretzis showed hysteresis feature of tunneling current in a polaron model beyond the Hartree approximation.

Although previous theoretical studies predicted the existence of hysteresis in a QD (or molecular) junction, such a phenomenon still lacks conclusive experimental support. Moon et al. have experimentally examined the tunneling current through a carbon nanotube QD, which exhibits a periodic oscillatory behavior with respect to the applied gate voltage, arising from an eightfold degenerate state. In addition, Liljeroth et al. have reported a periodic oscillatory differential conductance as a result of tunneling current through a single spherical PbSe QD with a sixfold degenerate state. These two experiments did not exhibit the bistable tunneling current. Their results indicate that electron–phonon interactions in nanotube QDs or PbS QDs are not sufficient to yield the strongly attractive electron–electron interaction needed for observing the bistability. Thus it remains unclear whether a single QD junction can display the bistable memory effect.

Recently, it was demonstrated that semiconductor quantum dot arrays (QDAs) can be chemically fabricated to form a superlattice. Via nanoscale manipulation, experimentalists can now control the lattice constant and QD size to tune charges of QDA in the Coulomb blockade regime or semiconducting regime. Consequently, QDA is not only a good physical system for investigating strongly correlated problem but also a promising integrated electronic device. Although many theoretical efforts have been devoted to the charge transport through a single QD, many studies are on the tunneling current through a QDA junction. In this article we illustrate that the on-site repulsive Coulomb interactions present in a QDA junction involving degenerate p-like orbitals can lead to bistable tunneling current, making it a good candidate for high density storage device. Such a mechanism only appears in an array of QDs (in which a band is formed), and the effect will diminish when the band width reduces and eventually disappear when the QDs become isolated.

2. Formalism

Figure 1 illustrates the system of a QD array embedded in an insulator connected with metallic electrodes. The system can be described by the extended Anderson Hamiltonian, $H = H_0 + H_T + H_d$. $H_0 = \sum_{k,\sigma,\beta} \epsilon_k c_{k,\sigma,\beta}^\dagger c_{k,\sigma,\beta}$ describes the electronic states in the metallic leads. Here $a_{k,\sigma,\beta}^\dagger$ (or $a_{k,\sigma,\beta}$) creates (destroys) an electron of wave vector $k$ and spin $\sigma$ with energy $\epsilon_k$ in the $\beta$ metallic electrode. The $H_T$ term describes the coupling between the electrodes and the $p$ orbitals of the QD array:

$$H_T = \sum_{k,\sigma,\beta} V_{k,\beta} a_{k,\sigma,\beta}^\dagger d_{l,\sigma} + \sum_{k,\sigma,\beta} V_{k,\beta}^* d_{l,\sigma}^\dagger a_{k,\sigma,\beta} + \sum_{k,\sigma,\beta} V_{k,\beta} c_{k,\sigma,\beta}^\dagger c_{k,\sigma,\beta},$$

(1)

where $(d_{l,\sigma})$ creates (destroys) an electron in the $p_l$ orbital of the QD at site $l$. $V_{k,\beta}$ describes the coupling between the continuous states in the electrodes and the localized $p_l$ states. $c_{k,\sigma,\beta}^\dagger$ (or $c_{k,\sigma,\beta}$) creates (destroys) an electron in the two-dimensional (2D) conduction bands of QD array formed by the $p_{x,y}$ orbitals. $p_l$ labels the 2D wave vectors. $l$ labels the

---

\*E-mail address: yichang@gate.sinica.edu.tw

DOI: 10.1143/JJAP.48.104504
conduction bands (including spin). $V_{k,p,l}$ describes the coupling between the electrodes and the $p_{xy}$ orbitals of the QD array, which is much smaller than $V_{k,p,l}$ since the $p_{xy}$ orbitals are much more localized along the $z$-axis than the $p_z$ orbitals. At last, the $H_d$ term describes electronic states and their interactions in the QD array

$$H_d = \sum_{l,\sigma} E_p d_{l,\sigma}^d d_{l,\sigma}^d + \sum_{l,\sigma} U d_{l,\sigma}^d d_{l,-\sigma}^d d_{l,-\sigma}^d d_{l,\sigma}^d + \sum_{p,l} [\varepsilon_{p,l} + U(N_c - N_0)]c_{p,l}^d c_{p,l}^d + U_d \sum_{l,p,l',\sigma} c_{p',l'^y}^d c_{p',l'^z}^d e^{i(p' - p) \cdot R} d_{l,\sigma}^d d_{l,\sigma}^d, \tag{2}$$

where $E_p$ and $U_d$ denote, respectively, the energy level and electron Coulomb interaction of the $p_z$ orbital of QD at site $l$. $\varepsilon_{p,l}$ and $U$ denote, respectively, the 2D band energy and the on-site Coulomb interaction between two electrons in the $p_{xy}$ orbitals. $U_d$ denotes on-site repulsive Coulomb interactions between the $p_z$ and $p_{xy}$ orbitals. $N_{QD}$ denotes the number of QDs in the matrix, $R_l$ is the positions of QDs in the array. Note that if we ignore the quadrupole and higher-order terms in the expansion of 1/$R_l^2$, then the Coulomb repulsion integrals between two electrons in any of the three degenerate $p$-like orbitals are the same. $N_l$ is the occupation number per unit cell for the $l$-th conduction band, and $N_c = \sum_l N_l$ is the total occupation number per unit cell in the $p_{xy}$ orbitals which form the conduction bands. A mean-field theory (which is justified for extended states) has been applied to the 2D conduction bands ($\varepsilon_{p,l}$) to obtain the band normalization effect from the self energy of $U(N_c - N_l)$. Here, we focus on the $p_z$ orbital rather than the ground state orbital, since in the range of applied bias considered, the QD ground state energy level is deeply below the Fermi levels of both electrodes and the electron tunneling through the QD ground state is blocked. Consequently, carriers in the QD ground states only lead to a constant-shift to all the $p$ orbitals. It is worth noting that eq. (2) is the same as the Falicov–Kimball (FK) model if we turn off the self energy term, $U(N_c - N_l)$. The FK model has been used extensively to study the semiconductor–metal transition in a transition metal as well as rare-earth oxides consisting of localized orbitals and delocalized orbitals.\(^\text{18,21}\) The total number of carriers in localized orbitals and delocalized orbitals is conserved in the FK model, but not in our case, where carriers can be injected from the electrodes and allowed to tunnel out. In this study, we employ the extended Anderson model (the combination of the FK model and the Anderson model) for QD array junctions to study the hysteresis phenomena of tunneling current.

Using Keldysh’s Green’s function technique,\(^\text{22,23}\) we obtain the tunneling current $J = J_1 + J_{2D}$ consisted of two components, where $J_1 = \sum_{l,\sigma} J_{l,\sigma}$ and $J_{2D}$ denote, respectively, the tunneling current through localized $p_z$ orbitals and 2D conduction bands. The expression of $J_{l,\sigma}$ is given by

$$J_{l,\sigma} = -\frac{e}{\hbar} \int \frac{d\varepsilon}{\pi} \left[ f_\varepsilon(l) - f_{\varepsilon}(l) \right] \times \frac{\Gamma_{l,\sigma}(\varepsilon) \Gamma_{l,\sigma}(\varepsilon)}{\Gamma_{l,\sigma}(\varepsilon) + \Gamma_{R,l,\sigma}(\varepsilon)} \text{Im} G_{l,\sigma}(\varepsilon), \tag{3}$$

where $f_\varepsilon(l) = f(\varepsilon - \mu_L)$ and $f_{\varepsilon}(l) = f(\varepsilon - \mu_R)$ are the Fermi distribution functions for the left and right electrodes, respectively. The chemical potential difference between these two electrodes is related to the applied bias by $\mu_L - \mu_R = eV_b$. $\Gamma_{l,\sigma}(\varepsilon)$ and $\Gamma_{R,l,\sigma}(\varepsilon) = 2\pi N_0 |V|_{b,k}|^{2} \delta(\varepsilon - \varepsilon_k)$ denote the tunneling rates from the $p_z$ orbitals to the electrodes. Notations $e$ and $\hbar$ denote the electron charge and Plank’s constant. In the wide-band limit, these tunneling rates are approximately energy-independent. Therefore, the calculation of tunneling current is entirely determined by the spectral function $A = \text{Im} G_{l,\sigma}(\varepsilon)$, which is the imaginary part of the retarded Green’s function $G_{l,\sigma}(\varepsilon)$. As for $J_{2D}$, its expression is derived as

$$J_{2D} = -\frac{e}{\hbar} \sum_{p,l} \int \frac{d\varepsilon}{\pi} \left[ f_\varepsilon(l) - f_{\varepsilon}(l) \right] \times \frac{\Gamma_{p,l,\sigma}(\varepsilon) \Gamma_{R,p,l,\sigma}(\varepsilon)}{\Gamma_{p,l,\sigma}(\varepsilon) + \Gamma_{R,p,l,\sigma}(\varepsilon)} \text{Im} G_{p,l,\sigma}(\varepsilon), \tag{4}$$

where $\Gamma_{p,l,\sigma}(\varepsilon)$ and $\Gamma_{R,p,l,\sigma}(\varepsilon) = 2\pi N_0 |V|_{b,k}|^{2} \delta(\varepsilon - \varepsilon_k)$ denote the tunneling rates from the 2D conduction bands to the electrodes. The imaginary part of retarded Green’s function $G_{p,l,\sigma}(\varepsilon)$ provides the density of states (DOS) of 2D conduction band which is determined by the detail QD array structure. Due to $\Gamma_{p,l,\sigma}/\Gamma_{R,p,l,\sigma} \ll 1$, we take the following approximation $J \approx J_{l,\sigma} = \sum_{l,\sigma} J_{l,\sigma}$. The current expression of eq. (4) was used to describe the bistability in double barrier resonant tunneling structures by ref. 24 employing the nonequilibrium Schwinger–Keldysh formalism. Although the calculation of $G_{l,\sigma}(\varepsilon)$ and $G_{p,l,\sigma}(\varepsilon)$ can be determined by the Dyson equations,\(^\text{5,25,26}\) $G_{l,\sigma}(\varepsilon)$ and $G_{p,l,\sigma}(\varepsilon)$ are calculated by the equation of motion method,\(^\text{27,28}\) which is one of convenient methods for studying nanodevices in the nonequilibrium condition.

Using the equation of motion for $G_{l,\sigma}$, we obtain

\begin{align}
(\varepsilon - E_p + i\Gamma_l)G_{l,\sigma}(\varepsilon) & = \delta_{l,\sigma} + \sum_{p',p,\sigma} g_{p',p} \left( c_{p',\sigma}^d c_{p,\sigma}^d + c_{p',\sigma}^d c_{p,\sigma}^d \right) + U \sum_{p',p,\sigma} g_{p',p} \left( c_{p',\sigma}^d c_{p,\sigma}^d d_{p,\sigma}^d + c_{p,\sigma}^d c_{p',\sigma}^d d_{p',\sigma}^d \right), \tag{5}\\
(\varepsilon - E_p + i\Gamma_l)G_{p,l,\sigma}(\varepsilon) & = U(n_{p_z,-\sigma} d_{p,\sigma}^d + n_{p_z,\sigma} d_{p_z,\sigma}^d) + \sum_{p',p,\sigma} g_{p',p} \left( c_{p',\sigma}^d c_{p,\sigma}^d d_{p,\sigma}^d + c_{p,\sigma}^d c_{p',\sigma}^d d_{p',\sigma}^d \right), \tag{6}
\end{align}
The range of applied bias considered here would not be of ref. 24 with Dyson equation method. Equation (10) is the orbital in each unit cell. Worth noting that if we turn off the Coulomb interaction between the QD array and electrodes. Because of very small coupling strength, the real part of self energies is ignored in eqs. (5)–(8). In eqs. (5)–(8), we have introduced four one-particle Green’s functions $G_{\sigma}(\epsilon) = \langle d_{\sigma}d_{\sigma}^\dagger \rangle$, $G_{\sigma}(\epsilon) = \langle d_{\sigma}^\dagger c_{\sigma} \rangle$, $G_{\sigma}(\epsilon) = \langle c_{\sigma}d_{\sigma} \rangle$, and $G_{\bar{\sigma}}(\epsilon) = \langle c_{\bar{\sigma}}d_{\bar{\sigma}} \rangle$. These four single-particle Green’s function are coupled with two-particle Green’s functions via $U$ and $U_{dc}$. The equation of motion for the two-particle Green’s function (defined as $(n_{\sigma\bar{\sigma}}d_{\sigma}d_{\bar{\sigma}}^\dagger), \ (n_{\sigma\bar{\sigma}}d_{\sigma}^\dagger d_{\bar{\sigma}}), \ (c_{\sigma\bar{\sigma}}c_{\sigma}d_{\bar{\sigma}}^\dagger), \ (c_{\sigma\bar{\sigma}}d_{\sigma}^\dagger d_{\bar{\sigma}}), \ (n_{\sigma\bar{\sigma}}c_{\sigma}d_{\bar{\sigma}}), \ (n_{\sigma\bar{\sigma}}d_{\sigma}^\dagger d_{\bar{\sigma}}), \ (n_{\sigma\bar{\sigma}}^\dagger c_{\sigma}d_{\bar{\sigma}}), \ and \ (n_{\sigma\bar{\sigma}}^\dagger d_{\sigma}^\dagger d_{\bar{\sigma}})$) are coupled to the three-particle Green’s functions. In order to terminate the hierarchy of the equation of motions, we use the Hartree–Fock approximation method to decouple terms involving the $U_{dc}$ factor. Meanwhile in the calculation for $(n_{\sigma\bar{\sigma}}d_{\sigma}d_{\bar{\sigma}}^\dagger)$ and $(n_{\sigma\bar{\sigma}}^\dagger d_{\sigma}d_{\bar{\sigma}})$, the coupling terms between localized states and the electrodes is employed in the scheme considered in our previous method, which is valid for the Coulomb blockade regime. Solving eqs. (5)–(8), we obtain

$$
G_{\sigma\bar{\sigma}}(\epsilon) = \frac{\delta_{\sigma\bar{\sigma}}}{\epsilon - \epsilon_0 - \Delta_\sigma + i\Gamma_\sigma},
$$

(9)

where $\Delta_\sigma = U_{dc}(N_{\sigma\bar{\sigma}} + N_{\bar{\sigma}\sigma}) + U(N_{\sigma} - N_{\bar{\sigma}})$ and

$$
G_{\sigma}(\epsilon) = \frac{1 - N_{\sigma\bar{\sigma}}}{\epsilon - E_\sigma - \Delta_\sigma + i\Gamma_\sigma} + \frac{N_{\sigma\bar{\sigma}}}{\epsilon - E_\sigma - U - \Delta_\sigma + i\Gamma_\sigma},
$$

(10)

where $\Delta_\sigma = U_{dc}N_{\sigma}$ denotes the self-energy due to the interaction of the local orbital with the band states via Coulomb interaction. $N_{\sigma}$ is the occupation number of $p_{\sigma}$ orbital in each unit cell. Worth noting that if we turn off $U_{dc}$ in eqs. (9) and (10), and then eq. (9) is the same as eq. (5a) of ref. 24 with Dyson equation method. Equation (10) is the well known Green’s function expression to describe the tunneling current through a single energy level in the Coulomb blockade regime.

To reveal the tunneling current behavior, the occupation numbers $N_{\sigma\bar{\sigma}}$ and $N_{\sigma}$ are solved self-consistently by the following equations:

$$
N_{\sigma\bar{\sigma}} = -\int \frac{d\epsilon}{\pi} \frac{G_{\sigma\bar{\sigma}}(\epsilon)}{\Gamma_\sigma + \Gamma_R} \text{Im} \ G_{\sigma\bar{\sigma}}(\epsilon),
$$

(11)

$$
N_{\sigma} = \frac{1}{2} \sum_{\sigma,\bar{\sigma}} \int \frac{d\epsilon}{\pi} \frac{G_{\sigma\bar{\sigma}}(\epsilon) + G_{\sigma\bar{\sigma}}(\epsilon)}{\Gamma_\sigma + \Gamma_R} \frac{N_{\sigma\bar{\sigma}}}{\Gamma_\sigma + \Gamma_R} \text{Im} \ G_{\sigma\bar{\sigma}}(\epsilon).
$$

(12)

As mentioned above, the coupling between the electrodes and the $p_{\sigma}$ orbitals of $QD$ array is much smaller. Therefore, we consider $\text{Im} \ G_{\sigma\bar{\sigma}}(\epsilon) \approx \pi\delta(\epsilon - \epsilon_{\sigma\bar{\sigma}} - U(N_{\sigma} - N_{\bar{\sigma}}) - U_{dc}(N_{\sigma\bar{\sigma}} + N_{\bar{\sigma}\sigma}))$ in the following numerical calculation. The range of applied bias considered here would not be enough to overcome the charging energy of $U + \Delta_\sigma$, therefore, the second term in eq. (10) can be ignored and we have $G_{\sigma\bar{\sigma}}(\epsilon) = (1 - N_{\sigma\bar{\sigma}})/(\epsilon - E_\sigma - \Delta_\sigma + i\Gamma_\sigma)$ in which we set $\Gamma_\sigma = \Gamma$. The occupation numbers at zero temperature are calculated by

$$
N_{\sigma\bar{\sigma}} = \frac{1 - N_{\sigma\bar{\sigma}}}{\pi} \frac{\Gamma_\sigma}{\Gamma_\sigma + \Gamma_R} \int_{-\infty}^{\infty} \frac{d\epsilon}{(\epsilon - E_\sigma - \epsilon_0 V_a - \Delta_\sigma)^2 + \Gamma^2},
$$

(13)

or

$$
\frac{\Gamma_\sigma}{\Gamma_L} \frac{N_{\sigma\bar{\sigma}}}{1 - N_{\sigma\bar{\sigma}}} = \frac{1}{\pi} \cot \left( \frac{\epsilon_0 V_a - E_\sigma - \epsilon_0 V_a - U_{dc} N_{\sigma\bar{\sigma}}}{\Gamma} \right),
$$

(14)

in which $\epsilon_0 V_a$ term arises from the applied bias crossing QDA and $\alpha$ is a dimensionless scaling factor determined by the QDA location, and

$$
N_{\sigma} = \frac{\Gamma_{cL}}{\Gamma_{cL} + \Gamma_{cR}} \int_{-\infty}^{\infty} \frac{d\epsilon}{\pi} \text{Im} \ G_{\sigma\bar{\sigma}}(\epsilon),
$$

(15)

$$
\left[ \epsilon - U_{dc}(N_{\sigma\bar{\sigma}} + N_{\bar{\sigma}\sigma}) - U(N_{\sigma} - N_{\bar{\sigma}}) \right] \epsilon_i^\dagger \epsilon_i = \frac{\epsilon_0 V_a - E_\sigma - U_{dc} N_{\sigma\bar{\sigma}}}{\Gamma}.
$$

(16)

where $E_b$ denotes the bottom of the conduction band and $W$ is the band width. Such an approximation allows eq. (15) to have a simple analytic solution of the form

$$
N_{\sigma} = g - cN_{\sigma\bar{\sigma}}
$$

(17)

with $g = [E_b + (1 - \alpha)\epsilon_0 V_a - E_b]/(\gamma W + 3U)$ and $c = 2U_{dc}/(\gamma W + 3U)$, where $\gamma = (\Gamma_{cL} + \Gamma_{cR})/\Gamma_{cL}$. Substituting this into eq. (14), we obtain a simple transcendental equation, which can be solved numerically. The equation allows a maximum of three roots, out of which only two are stable roots.

We can also solve two coupled transcendental equations as given in eqs. (14) and (15) numerically for a more realistic density states, which is derived for a 2D tight-binding model. We consider a tight-binding model for $p_{\sigma\bar{\sigma}}$, orbitals arranged on a rectangular lattice with lattice constants $a$ and $b$. Figure 2 illustrates the rectangular lattice. The band structure for the $p_{\sigma}$ band is given by
\[ \epsilon_s(k) = E_p - 2v_f \cos(k,a) - 2v_i \cos(k,b). \]  
where \( v_i \) denotes the (ppr) interaction and \( v_f \) denotes the (ppr) interaction.\(^{20} \) For the \( p_x \) band, we have

\[ D_s(\epsilon) = \begin{cases} 
\frac{1}{\alpha^2} \int_0^\infty \text{d}n \left[ (2\epsilon)^2 - [2\epsilon + 2v_f(1 - \cos \eta) - \epsilon^2]^{1/2} \left( \epsilon - 2v_f(1 - \cos \eta) \right) \right] & \text{for } 0 < \epsilon < 4v_f \\
\frac{1}{\alpha^2} \int_0^\infty \text{d}n \left[ (2\epsilon)^2 - \epsilon^2 - 2v_f(1 - \cos \eta) - \epsilon^2 \right]^{1/2} \left( \epsilon - 2v_f(1 - \cos \eta) \right) & \text{for } 4v_f < \epsilon < 4v_i, \\
\frac{1}{\alpha^2} \int_0^\infty \text{d}n \left[ (2\epsilon)^2 - [2v_f + 2v_i(1 + \cos \eta) - \epsilon^2]^{1/2} \left( \epsilon - 2v_i(1 + \cos \eta) \right) \right] & \text{for } 0 < \epsilon < 4v_i 
\end{cases} \]

where \( \epsilon = \epsilon - E_p + 2v_f + 2v_i \) and \( \epsilon = E_p + 2v_f + 2v_i - \epsilon \).

If \( v_f < v_i \), then the roles of \( v_f \) and \( v_i \) should be exchanged in the above expression. The DOS described by eq. (20) contains the van Hove singularities. Similar expression \( [D_s(\epsilon)] \) holds for the \( p_x \) band with the hopping parameters \( v_f \) and \( v_i = v'_i \) replaced by \( v'_f = v_i \) and \( v'_i \). By varying these hopping parameters (for instance, fix lattice constant \( a \) and tune \( b \)), we can study the behavior of bistable tunneling current for systems between the 1D and 2D limits.

3. Results and Discussion

To numerically solve the coupled nonlinear eqs. (14) and (15), we adopt the following parameters in our calculations: \( U_{pk} = U = 50 \text{ meV}, \quad \Gamma_L = 1 \text{ meV} \quad (\Gamma_{Lc} = \Gamma_L/20), \quad \text{and} \quad \Gamma_R = 1 \text{ meV} \quad (\Gamma_{Rc} = \Gamma_R/20). \) Throughout the paper, we shall use \( T = 0 \text{ K}, \quad v'_f = 5 \text{ meV}, \quad v_f = 20 \text{ meV}, \quad \gamma = 0.5, \quad \text{and} \quad E_0 + V_0 = E_p, \) where \( V_0 \) is a reference bias for \( V_a \). For simplicity, \( V_0 = 0. \) The tight-binding parameters are assumed to scale according to the \( 1/R^2 \) rule,\(^{29} \) where \( R \) is the separation between two QDs. Thus, we have \( v'_f = v_f(a/b)^2 \) and \( v_i = v'_i(a/b)^2 \). When \( a = b \), we have the square lattice case. In Fig. 3 we plot the auxiliary functions \( f_1 = (\Gamma_L + \Gamma_R)/\Gamma_L[N_d/(1 - N_d)] \) and \( f_2 \), which are the left- and right-hand-side of eq. (14) versus \( N_d \) for various values of applied bias, \( V_a \). The results obtained by using the constant DOS and realistic DOS (from a tight-binding model) are shown in Figs. 3(a) and 3(b), respectively. The two sets of results are similar, except that the curves obtained with realistic DOS contain some sharp features caused by the van Hove singularities in the DOS. The occupation number per spin component in a given conduction band, \( N_s \), as a function of \( N_d \) is also shown in this figure (see dotted lines). The dotted lines from bottom to top correspond to the applied bias 40, 60, 80, 100, and 110 mV, respectively. The intersection of \( f_1 \)

\[ f_1(N_d) \]

\[ f_2(N_d) \]

\[ v_a = 40 \text{ mV} \]

\[ 110 \text{ mV} \]

\[ 80 \text{ mV} \]

\[ 60 \text{ mV} \]

\[ 100 \text{ mV} \]

\[ 80 \text{ mV} \]

Fig. 2. (Color online) Schematic diagram for \( p_{x,y} \) orbitals on a rectangular lattice with lattice constants \( a \) and \( b \), which form the conduction bands of interest. \( v_f(\nu_f) \) and \( v_i(\nu_i) \) denote, respectively, the electron hopping integrals of \( p_x(p_y) \) orbital along the \( a \)- and \( b \)-axes.

Fig. 3. (Color online) Auxiliary functions \( f_1(N_d) \) and \( f_2(N_d) \) vs \( N_d \) for various values of applied bias: (a) using eq. (14) with constant DOS, (b) using the DOS obtained from a tight-binding model. \( N_s \) denotes the occupation number per unit cell for the \( i \)-th conduction band. The dotted lines from bottom to top correspond to the applied bias 40, 60, 80, 100, and 110 mV, respectively.
20 meV. The dash-dotted line includes the current through and constant DOS. Solid line and dashed line denote, respectively, tight binding DOS and constant DOS.

and $f_2$ gives the solution to $N_d$. $f_2$ arising from the tight binding DOS is slightly shifted from that of constant DOS. Such a shifting results from the DOS-dependent occupation, $N_c = \sum N_d$. We find that at low or high bias regime, only one root to the equation $f_1 = f_2$ exists, while for moderate biases ($80 \leq V_d < 100$ mV), we find three roots to the equation $f_1 = f_2$. One of these three roots (the middle one) is an unstable solution. The stable solutions are more interesting; therefore, we plot the occupation number, $N_d$ as a function of the applied bias at zero temperature (shown as solid line) in Fig. 4. The result remains very similar if we use the constant DOS approximation as described in eqs. (16) and (17) with the same bandwidth, $W = 4(v_f + v_i)$. Although there are van Hove singularities in tight binding DOS, the structure of bias-dependent occupation number does not exhibit an abnormal feature. This is because $p_z$ orbital is correlated with $p_{x,y}$ orbital via $N_c$, which is related to the integral over the DOS.

We see from Fig. 4 that the occupation number, $N_d$ has bistable roots. Although the QDA crystal structure reported in refs. 10, 13 has a triangle lattice, the results of Fig. 4 indicate that the hysteresis behavior will not be sensitive to the detailed shape of the band structure or density of states. Once the occupation numbers are solved, the tunneling current through a single $p_z$ orbital can be obtained by the relation $J_f = \frac{e}{h} \Gamma_c N_d / J_0 N_d$, which is valid at zero temperature and when the carrier injection from the right lead can be ignored. Consequently, the curve for $N_d$ versus applied bias also illustrates the tunneling current characteristics. The roots for the turn-on and turn-off processes in Fig. 4 are determined by selecting a root closest to the root corresponding to the previous value of $V_d$ when multiple roots are allowed.

It is crucial to clarify how the system selects the high conductivity state (larger $N_d$) or the low conductivity state (smaller $N_d$) as the applied bias is turned on or off. In Fig. 5(a), we plot the bistable current for various strengths of $U_{dL}$. Curves 1, 2, and 3 denote, respectively, $U = 50, 30, $ and $20$ meV. The dash-dotted line includes the current through 2D conduction band with tunneling rates $\Gamma_{cL} = \Gamma_{cR} = \Gamma_{dL}/20 = \Gamma_{dR}/20$. It is seen that due to the small coupling strength between the electrodes and the 2D conduction band, the current through the 2D conduction band only slightly modifies the total tunneling current, and its effect on the bistability behavior is negligibly small. For simplicity, such a current is ignored in the analysis that follows. The bistable current vanishes for weaker Coulomb interactions. The critical Coulomb interaction needed to maintain a bistable current depend on the physical parameters, such as bandwidth of the 2D conduction band, tunneling rates between dot array and electrodes, and temperature. For the application as a memory device, larger $U$ (corresponding to smaller dot size) is favored because the bistability behavior will be more robust against the increase of temperature and broadening due to stray fields. Although Fig. 5(a) exhibits the bistable current, it does not show any negative differential conductivity (NDC), unlike the bistable current in quantum well systems, which is typically associated with an NDC. For the same set of Coulomb interaction strengths...
as in Fig. 5(a). Understanding the behavior of \( N_c \) is useful for clarifying the bistable mechanism of \( N_d \) as explained below.

The physical mechanism for this bistability behavior can be explained as follows. As we gradually raise the Fermi level in the left electrode from below the resonance level \( E_R \), the allowed solution to the occupation number, \( N_d \), remains small during the “turn-on” process, as a result of inter-level Coulomb blockade (which pushes up the \( p_i \) level by the amount \( N_c U_{dc} \), where the 2D conduction band occupation number \( N_c \) is appreciable). Once the applied bias reaches a critical value (the inter-level Coulomb blockade is overcome), \( N_d \) increases quickly to a value around 1/3. Now, charges accumulate in the localized \( p_i \) orbitals, which lead to an increase of the self-energy of the conduction band states by \( 2U_{dc}N_d \), and hence reduces \( N_c \) to a much smaller value (from \( N_{c2} \) to \( N_{c3} \) as illustrated in Fig. 5). This in turn causes the \( p_i \) energy level to decrease during the “turn-off” process while \( N_d \) maintains around 1/3 to keep the transcendental equation self-consistent. When the applied bias continues to decrease to the lower critical value, \( N_c \) is switched from \( N_{c4} \) to \( N_{c1} \) [see Fig. 5(b)], and \( N_d \) quickly goes back to the lower value which becomes the only allowed self-consistent solution to the transcendental equation. The results of Fig. 5 demonstrate that the bistable tunneling current arises from the on-site repulsive Coulomb interactions \( U_{dc} \). In the polaron model adopted in refs. 6 and 7, the bistable mechanism arises from an attractive potential of the form \(-2\gamma p_i d_{\sigma i}d_{\sigma i-d}\) where \( \gamma \) and \( W_0 \) are the electron–phonon interaction strength and the phonon frequency. This two particle interaction term should be corrected as \( U = -2\gamma p_i d_{\sigma i}d_{\sigma i-d} \) when the intra-level Coulomb interaction is included in a single QD or molecule. Due to the large repulsive intralevel Coulomb interaction \( U \), a net attractive electron–electron interaction mediated by phonon is difficult to achieve, which may explain why a bistable tunneling current through a single QD junction has not been observed.

To realize high density nanoscale memory structures, we need to examine whether this bistable current also exists in 1D arrays. Figure 6 shows the tunneling current through the \( p_i \) orbital for various ratios of \( b/a \). Other parameters are the same as those for Fig. 4. We find that from two dimensional array to 1D array, the bistability behavior is sustained (although somewhat weaker in the 1D limit). When \( b/a \) is greater than 3, (results not shown here) the bistable current behaves essentially the same as in the \( b/a = 3 \) case. This indicates that we already reached the 1D limit for \( b/a \approx 3 \), and the bistable current still exists. If we assume \( a = 3 \) nm, \( b = 9 \) nm, and 50 coupled QDs along the chain in the \( x \) direction are needed to establish a band-like behavior, then the density of the memory device is around \( 1/(1350 \text{nm}^2) \approx 0.5 \text{Tbyte/in}^2 \), which is much higher than the existing nonvolatile memory device density. It is conceivable if 50 QDs are arranged in a close-packed 2D array, the band-like behavior for the \( p_{os} \) states should remain, and the memory density can be even higher (by a factor more than 2). The operating voltage needed is less than 100 mV, which indicates very low power consumption. Because this is a quantum device, the switching time is expected to be comparable to the tunneling rate, which is on the order of 1 THz. We have also examined the dependence of the bistability behavior on the band width, \( W \) (not shown here). We find that if the band width becomes too small (approaching the limit of isolated QDs), the bistability disappears.

Finally, we show in Fig. 7 the tunneling current for various tunneling rate ratios \( \Gamma_L/\Gamma_R \) for the 1D limit (\( b/a = 3 \)). It is seen that the bistability disappears for \( \Gamma_L = 0.1 \text{ meV} \) and \( \Gamma_R = 1 \text{ meV} \) (shell-tunneling condition). In this case, charges are unable to accumulate in the \( p_i \) orbitals. Consequently, the effect of \( U_{dc} \) is suppressed. On the other hand, for \( \Gamma_L = 1 \text{ meV} \) and \( \Gamma_R = 0.1 \text{ meV} \) (shell-filling condition), the effect of \( U_{dc} \) is enhanced. This leads to a wider voltage range for bistable current. The results of Fig. 7 indicates that we can control the bistable current by adjusting the tunneling rate ratio.

4. Conclusions

In conclusion, we have illustrated a novel mechanism for generating the bistable tunneling by using a junction

![Graph](image-url)
involving a 2D periodic array of QDs with six-fold degenerate p-like states, which consists of localized p_z orbitals interacting with non-localized p_x,y orbitals via the on-site Coulomb interaction. Due to the interplay of Coulomb blockade effect for the localized state and the self-energy correction to the 2D conduction bands formed by the p_x,y orbitals, a bistable tunneling current with well defined hysteresis behavior can be achieved. This bistable current is sensitive to the charging energy, the band width, and the ratio of incoming to outgoing tunneling rate. Such a hysteresis behavior arises from a collective effect, not observable in a single QD. It is shown that this hysteresis behavior can also exist in one dimensional QDA and very high density high-speed integrated memory circuits can be realized.

Acknowledgment

This work was supported by the Academia Sinica, Taiwan.

1) V. J. Goldman, D. C. Tsui, and J. E. Cunningham: Phys. Rev. Lett. 58 (1987) 1256.
2) R. Egger, H. Grabert, A. Koutouza, H. Saleur, and F. Siano: Phys. Rev. Lett. 84 (2000) 3682.
3) J. O. Ouyang, C. W. Chu, C. R. Szmanda, L. P. Ma, and Y. Yang: Nat. Mater. 3 (2004) 918.
4) R. J. Tseng, J. Huang, J. Ouyang, R. B. Kaner, and Y. Yang: Nano Lett. 5 (2005) 1077.
5) A. S. Alexandrov and A. M. Bratkovsky: Phys. Rev. B 67 (2003) 235312.
6) M. Galperin, M. A. Ratner, and A. Nitzan: Nano Lett. 5 (2005) 125.
7) A. L. Magna and I. Deretzis: Phys. Rev. Lett. 99 (2007) 136404.
8) S. Moon, W. Song, J. S. Lee, N. Kim, J. Kim, S. G. Lee, and M. S. Choi: Phys. Rev. Lett. 99 (2007) 176804.
9) P. Liljeroth, L. Jdira, K. Overgaag, B. Granddidier, S. Speller, and D. Vannaaelbergh: Phys. Chem. Chem. Phys. 8 (2006) 3845.
10) C. B. Murray, C. R. Kagan, and M. G. Bawendi: Science 270 (1995) 1335.
11) C. P. Collier, R. J. Saykally, J. J. Shiang, S. E. Henrichs, and J. R. Heath: Science 277 (1997) 1978.
12) F. Remacle, K. C. Beverly, J. R. Heath, and R. D. Levine: J. Phys. Chem. B 107 (2003) 13892.
13) D. V. Talapin and C. B. Murray: Science 310 (2005) 86.
14) H. E. Romero and M. Dndic: Phys. Rev. Lett. 95 (2005) 156801.
15) T. S. Mentzel, V. J. Porter, S. Geyer, K. MacLean, M. G. Bawendi, and M. A. Kastner: Phys. Rev. B 77 (2008) 075316.
16) D. M. T. Kuo and Y. C. Chang: Phys. Rev. Lett. 99 (2007) 086803.
17) Y. C. Chang and D. M. T. Kuo: Phys. Rev. B 77 (2008) 245412.
18) D. M. T. Kuo, G. Y. Guo, and Y. C. Chang: Appl. Phys. Lett. 79 (2001) 3851.
19) L. M. Falicov and J. C. Kimball: Phys. Rev. Lett. 22 (1969) 997.
20) R. Ramirez, L. M. Falicov, and J. C. Kimball: Phys. Rev. B 2 (1970) 3383.
21) J. K. Freericks and V. Zlatic: Rev. Mod. Phys. 75 (2003) 1333, and references therein.
22) L. V. Keldysh: Zh. Eksp. Teor. Fiz. 47 (1964) 1515 [Translation: Sov. Phys. JETP 20 (1965) 1018].
23) H. Haug and A. P. Jauho: Quantum Kinetics in Transport and Optics of Semiconductors (Springer, Heidelberg, 1996).
24) J. Zang and J. L. Birman: Quantum Kinetics in Transport and Optics of Semiconductors (Springer, Heidelberg, 1996).
25) S. Datta: Electronic Transport in Mesoscopic System (Cambridge University Press, Cambridge, U.K., 1995).
26) D. K. Ferry and S. M. Goodnick: Transport in Nanostructures (Cambridge University Press, Cambridge, U.K., 1995).
27) Q. F. Sun and H. Guo: Phys. Rev. B 66 (2002) 155308.
28) C. Niu, L.-J. Liu, and T.-H. Lin: Phys. Rev. B 51 (1995) 5130.
29) W. A. Harrison: Applied Quantum Mechanics (World Scientific, Singapore, 2000).