Metal-semiconductor transition and Fermi velocity renormalization in metallic carbon nanotubes

Yan Li* and Umberto Ravaioli
Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA

Slava V. Rotkin
Department of Physics, Lehigh University, Bethlehem, Pennsylvania 18015, USA
(Received 21 September 2005; published 9 January 2006)

Angular perturbations modify the band structure of armchair (and other metallic) carbon nanotubes by breaking the tube symmetry and may induce a metal-semiconductor transition when certain selection rules are satisfied. The symmetry requirements apply for both the nanotube and the perturbation potential, as studied within a nonorthogonal π-orbital tight-binding method. Perturbations of two categories are considered: an on-site electrostatic potential and a lattice deformation which changes the off-site hopping integrals. Armchair nanotubes are proved to be robust against the metal-semiconductor transition in second-order perturbation theory due to their high symmetry, but can develop a nonzero gap by extending the perturbation series to higher orders or by combining potentials of different types. An assumption of orthogonality between π orbitals is shown to lead to an accidental electron-hole symmetry and extra selection rules that are weakly broken in the nonorthogonal theory. These results are further generalized to metallic nanotubes of arbitrary chirality.

DOI: 10.1103/PhysRevB.73.035415 PACS number(s): 73.63.Fg, 73.22.−f, 61.46.−w, 71.15.−m

I. INTRODUCTION

The subject of metal-insulator transitions has been studied for decades.[1] It is well understood that a metal-insulator transition is typically related to the breaking of a specific symmetry of the system, and carbon nanotubes are particularly interesting to study, due to their low dimensionality and special helical symmetry. In this paper we investigate the symmetry breaking in single-wall nanotubes (SWNTs) due to an external potential which depends on the angular coordinate along the SWNT circumference. This perturbation may induce a transition in a SWNT changing the type of its electronic structure by opening (closing) the band gap in the metallic (semiconducting) nanotube.[2–6,8] It is important for applications in which such gap engineering can potentially be controlled locally—for instance, by the field of a sharp tip, by a small molecule, or by a local gate.

SWNT lattice symmetry depends on two parameters, diameter and chirality, which determine the type of band structure.[9] About one-third of possible SWNTs are metallic, with one-dimensional energy subbands crossing at the Fermi level, as confirmed by experiments.[10] In this study, we consider metallic nanotubes and focus mainly on the special case of armchair SWNT (A-SWNT), which has a higher symmetry prohibiting an energy gap at the Fermi level[11] for typical nonchiral perturbations such as stretching, uniform electric field, impurity potentials, and many-body interactions. The same perturbations would open a small "secondary" band gap in metallic nanotubes of different symmetry. The high symmetry of A-SWNTs is also responsible for the absence of backscattering in the conduction channels.[7] Such ballistic transport is very attractive for future electronic applications,[9,10] and a method to control the conductance of A-SWNTs would be particularly desirable. Different perturbations have been attempted to modify the electronic structure of A-SWNTs.[2–6,8,9,12–15] Our goal is to demonstrate, using symmetry arguments, whether a particular perturbation can open a gap at all and how the gap depends on the magnitude of the perturbation potential. We will show below that, with minor exceptions, this cannot be a linear dependence. We call this transition a "metal-semiconductor transition" (MST) since the gap size is smaller than in typical insulators.[12]

A nonorthogonal tight-binding (TB) approach is used to model the SWNT electronic structure. Despite its simplicity, the TB approach may include as much important physics as more sophisticated models with the right choice of empirical parameters.[16] In addition, it possesses a great advantage for analytical derivations. Combining the TB approach and the summation of perturbation series with a group theory technique, it was shown in previous work[12] that mirror symmetry breaking is a necessary condition to mix the two crossing subbands (π and π^*) and open a band gap in A-SWNTs. Here, we find several additional significant results: (1) Due to the high lattice symmetry, the second-order contributions *always* cancel out and no second-order gap opening occurs. One notes that the first-order process is suppressed unless very specific selection rules are satisfied (see below). (2) Under potentials of a single angular Fourier component, $\mathbf{V}_q \cos q\theta$, the lowest contributing coupling order between π and π^* bands, μ_0, is determined by the angular momentum of the potential, q, and the index n of the (n,n) A-SWNT as $\mu_0 = 2n/\gcd(2n,q)$, in which gcd is the greatest common divisor. The band gap opening is proportional to $|\mathbf{V}_q|^{\mu_0}$ for small perturbation $|\mathbf{V}_q| \ll \upsilon_F/R$, where υ_F and R are the nanotube Fermi velocity and radius. In a typical experiment, when the perturbation has a small angular momentum ($q \sim 1$), the coupling order is high ($\mu_0 \sim n$) and the gap is small if any. To observe a linear effect ($\mu_0 = 1$), a high-q potential must be applied ($q = 2n$). The high coupling order can be reduced by choosing combinations of several angular Fourier components or different types of perturbation. (3) Additional symmetry of a particular model may lead to extra selection

rules for the band gap opening. For example, gaping is forbidden for A-SWNTs with even n if perfect electron-hole symmetry is assumed, as in an orthogonal basis. When the slight asymmetry between the conduction and valence bands is included, a gap proportional to the asymmetry parameter occurs. (4) Significant changes in the A-SWNT density of states (DOS) is observed even when the gap is absent. Modification of the low-energy band structure is mostly determined by the second-order perturbation. The DOS is enhanced near the Fermi level, and simultaneously v_F decreases. Peaks of the first pair of van Hove singularities are brought closer, resulting in a smaller excitation energy between these subbands.

Our study is restricted to the case of potentials which are uniform along the tube axis, but certain results are easily generalized for perturbations with even and odd axial dependence.

This paper is organized as follows. We first formulate the model and introduce the interaction matrix elements between TB wave functions in Sec. II. Using nearly degenerate perturbation theory (Appendix) and symmetry-based selection rules, we derive analytically the coupling between π and π^* subbands of A-SWNTs for both scalar potentials (Sec. III) and tensor potentials (Sec. IV). Comparisons are made with numerical results from TB band structure calculation. These results are further extended to metallic nanotubes of arbitrary chirality in Sec. VI. Finally, we summarize our results in Sec. VII.

II. MODEL FORMULATION

A. Perturbation series in the TB description

The electronic states Ψ_i are obtained within a nonorthogonal single π-orbital TB method by solving the stationary Schrödinger equation

$$H\Psi_i = E_i S\Psi_i, \tag{1}$$

where H and S are the Hamiltonian matrix and overlap matrix, respectively. For the sake of simplicity, only the nearest-neighbor hopping integral $\gamma_0=-3.033$ eV and overlap integral $s=0.129$ are considered.[17]

The wave function of an unperturbed A-SWNT can be expressed as a linear combination of the two periodic functions $u_\xi(\mathbf{k})=(1/\sqrt{N})\Sigma_i e^{i\mathbf{k}\cdot(\mathbf{r}_{i\xi}-\mathbf{r})}\varphi(\mathbf{r}-\mathbf{r}_{i\xi})$, where $\xi=A,B$ label the two sublattices and $\varphi(\mathbf{r}-\mathbf{r}_{i\xi})$ is the atomic orbital function localized at $\mathbf{r}_{i\xi}$:

$$\Psi_\sigma(\mathbf{k}) = \frac{e^{i\mathbf{k}\cdot\mathbf{r}}}{\sqrt{2[1-s\sigma|f(\mathbf{k})|]}}[e^{i\phi_\sigma(\mathbf{k})}u_A(\mathbf{k}) + e^{-i\phi_\sigma(\mathbf{k})}u_B(\mathbf{k})],$$

$$E_\sigma(\mathbf{k}) = \gamma_0\frac{-\sigma|f(\mathbf{k})|}{1-s\sigma|f(\mathbf{k})|}, \quad f(\mathbf{k}) = \sum_{\lambda=1}^{3} e^{i\mathbf{k}\cdot\mathbf{r}_\lambda}, \tag{2}$$

where $\sigma=\pm1$ denote the conduction and valence bands. $\mathbf{r}_\lambda$'s are the nearest-neighbor bond vectors, and we refer to $\lambda=1$ as the circumferential direction of A-SWNTs in the following. The wave vector $\mathbf{k}$ is composed of an axial component k_t and a quantized circumferential component $k_c=m/R$, with m the angular momentum. $2\phi_\sigma(\mathbf{k})\equiv\arg[-\sigma f(\mathbf{k})]$ indicates the phase difference of the coefficients before $u_A(\mathbf{k})$ and $u_B(\mathbf{k})$, which make a pseudospinor (see also Sec. VI). $\phi_\sigma(\mathbf{k})$ is constant through the whole range of k_t for the two crossing subbands: $\phi_\pi(k_t)=\pi/3$ and $\phi_{\pi^*}(k_t)=\phi_\pi-\pi/2$. Clearly, the π and π^* subbands are orthogonal due to this phase difference.

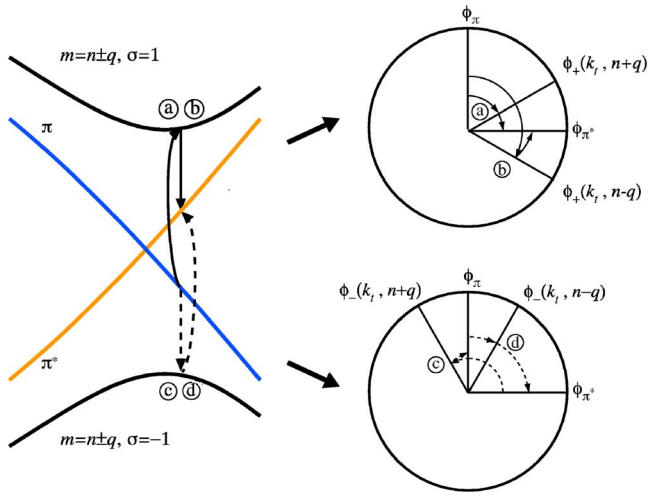

FIG. 1. (Color online). Schematics of the second-order coupling between π and π^* subbands and the corresponding phase angles of intermediate states.

Consider a perturbation H_1 which is uniform in the axial direction; then, k_t is conserved. To obtain the low-energy behavior of π and π^* subbands, we write the effective 2×2 perturbation Hamiltonian matrix using nearly degenerate perturbation theory (see the Appendix):

$$H_{\text{eff}}(k_t) = \begin{bmatrix} E_\pi(k_t)+H_{\pi\pi}(k_t) & H_{\pi\pi^*}(k_t) \\ H^*_{\pi\pi^*}(k_t) & E_{\pi^*}(k_t)+H_{\pi^*\pi^*}(k_t) \end{bmatrix}, \tag{3}$$

where the matrix element $H_{\alpha\beta}(k_t)$, with $\alpha(\beta)=\pi$ or π^*, can be represented by the sum of the perturbation series over all possible coupling orders μ as

$$H_{\alpha\beta}(k_t) = \sum_\mu \sum_{\{\Psi_i\}} H^{(\mu)}_{\alpha\beta}(\{\Psi_i\}), \quad \Psi_i \equiv \Psi_{\sigma_i}(k_t, m_i),$$

$$H^{(\mu)}_{\alpha\beta}(\{\Psi_i\}) = \langle\Psi_\alpha|H_1|\Psi_1\rangle\frac{\prod_{i=2}^{\mu-1}\langle\Psi_{i-1}|H_1|\Psi_i\rangle}{\prod_{i=1}^{\mu-1}(-E_i)}\langle\Psi_{\mu-1}|H_1|\Psi_\beta\rangle. \tag{4}$$

One notes that all intermediate states Ψ_i $(i=1,\ldots,\mu-1)$ are different from Ψ_π and Ψ_{π^*} by definition. Figure 1 illustrates an example of second-order coupling between π and π^* subbands via four different paths. Also shown is the phase angle of the intermediate states, $\phi_\pm(k_t,n\pm q)$, relative to ϕ_π and ϕ_{π^*}. The relation between these phase angles will be discussed in detail in the following sections.

B. Interaction matrix element between Bloch states

Below, we derive the interaction matrix element $\langle\Psi|H_1|\Psi'\rangle$ of the angular perturbation of a single Fourier

component, $H_1=\mathbf{V}_q\cos q(\theta-\theta_0)$, where $\mathbf{V}_q$ could be either scalar or tensor. θ_0 is defined as the minimum angular displacement between the vertical mirror planes (or glide planes) of the A-SWNT and the mirror planes of the potential.[12] Assuming that $\langle\varphi_1|H_1|\varphi_2\rangle$ is nonzero only when φ_1 and φ_2 are centered on the same atom or nearest-neighbor atoms, the interaction matrix element can be decomposed into an on-site term and an overlap term:

$$\langle\Psi|H_1|\Psi'\rangle=\frac{\delta_{m-m',q}e^{-iq\theta_0}+\delta_{m-m',-q}e^{iq\theta_0}}{2\sqrt{(1-s\sigma|f|)(1-s\sigma'|f'|)}}(\langle H_1\rangle_{\text{on-site}}+\langle H_1\rangle_{\text{overlap}}),$$

$$\langle H_1\rangle_{\text{on-site}}=\langle\varphi(\mathbf{r})|\mathbf{V}_q e^{iq\theta}|\varphi(\mathbf{r})\rangle\cos(\phi-\phi'),$$

$$\langle H_1\rangle_{\text{overlap}}=\sum_{\lambda=1}^{3}\langle\varphi(\mathbf{r}-\mathbf{r}_\lambda/2)|\mathbf{V}_q e^{iq\theta}|\varphi(\mathbf{r}+\mathbf{r}_\lambda/2)\rangle\cos\left(\phi+\phi'-\frac{m+m'}{2}\theta_\lambda-k_t z_\lambda\right),\tag{5}$$

where it is used that $\varphi(\mathbf{r})$ is real and invariant under $\theta\leftrightarrow-\theta$ inversion. $\delta_{m-m',\pm q}$ arises from the conservation of angular momentum. This conservation law also allows q to differ by multiples of $2n$, an angular analog of the reciprocal lattice vector, which is not considered here (see the discussion in Sec. VI). The interaction matrix can be further simplified depending on the type of $\mathbf{V}_q$:

Scalar potential: $\mathbf{V}_q$=const. Define $\langle\varphi(\mathbf{r})|\mathbf{V}_q e^{iq\theta}|\varphi(\mathbf{r})\rangle=U_q$; then,

$$\langle\varphi(\mathbf{r}-\mathbf{r}_\lambda/2)|V_q e^{iq\theta}|\varphi(\mathbf{r}+\mathbf{r}_\lambda/2)\rangle\approx\frac{1}{2}s[\langle\varphi(\mathbf{r}+\mathbf{r}_\lambda/2)|V_q e^{iq\theta}|\varphi(\mathbf{r}+\mathbf{r}_\lambda/2)\rangle+\langle\varphi(\mathbf{r}-\mathbf{r}_\lambda/2)|V_q e^{iq\theta}|\varphi(\mathbf{r}-\mathbf{r}_\lambda/2)\rangle]=sU_q\cos\frac{q\theta_\lambda}{2},\tag{6}$$

$$\langle H_1\rangle_{\text{overlap}}=\frac{1}{4}sU_q\left[e^{i(\phi+\phi')}\sum_{\lambda=1}^{3}(e^{-im\theta_\lambda-ik_t z_\lambda}+e^{-im'\theta_\lambda+ik_t z_\lambda})+\text{c.c.}\right]=-\frac{1}{2}sU_q\cos(\phi-\phi')(\sigma|f|+\sigma'|f'|),\tag{7}$$

and Eq. (5) is reduced to

$$\langle H_1\rangle_{\text{on-site}}+\langle H_1\rangle_{\text{overlap}}\approx U_q\cos(\phi-\phi')\left[1-\frac{1}{2}s(\sigma|f|+\sigma'|f'|)\right].\tag{8}$$

Tensor potential: The Fourier component of a tensor potential can be expressed in the second quantization formalism as

$$\sum_{\lambda=1}^{3}\delta\gamma_{\lambda,q}\sum_{\langle i,j\rangle^\lambda}e^{iq(\theta_i+\theta_j)/2}c_i^\dagger c_j,\tag{9}$$

where pairs $\langle i,j\rangle^\lambda$ are confined to first nearest neighbors with bonds along the $\mathbf{r}_\lambda$ direction and $\delta\gamma_{\lambda,q}$ is the corresponding change of the hopping integral. The on-site term is absent while the overlap term is reduced to

$$\langle H_1\rangle_{\text{overlap}}\approx\sum_{\lambda=1}^{3}g_\lambda(k_t;m,\sigma;m',\sigma')=\sum_{\lambda=1}^{3}\delta\gamma_{\lambda,q}\cos\left(\phi+\phi'-\frac{m+m'}{2}\theta_\lambda-k_t z_\lambda\right).\tag{10}$$

Comparing Eqs. (8) and (10), one can conclude that the interaction matrix elements from a scalar potential and a tensor potential have quite a different dependence on the phase angle ϕ and quantum numbers k_t,m. We will show in next section that certain selection rules can be derived for scalar potentials of general form and also for simple tensor potentials allowing summation over λ.

III. SCALAR POTENTIAL

Assume that a scalar perturbation in the form of $H_1=V_q\cos q(\theta-\theta_0)$ is applied to the (n,n) A-SWNT. Using the interaction matrix element from Sec. II, we can now derive the μth-order perturbation matrix elements within nearly degenerate perturbation theory:

$$H_{\alpha\beta}^{(\mu)}(\{\Psi_i\})=e^{-i\Delta\mu q\theta_0}\left(\frac{U_q}{2}\right)^\mu\frac{P_{\alpha\beta}(\{\Psi_i\})Q(\{\Psi_i\})}{\prod_{i=1}^{\mu-1}(-E_i^0)},$$

$$P_{\alpha\beta}(\{\Psi_i\})=\left[\prod_{i=1}^{\mu-1}\cos(\phi_{i-1}-\phi_i)\right]\cos(\phi_{\mu-1}+\Delta\mu q\pi/3n-\phi_\mu),$$

$$Q(\{\Psi_i\})=\prod_{i=1}^{\mu}\left[1-\frac{1}{2}s(\sigma_{i-1}|f_{i-1}|+\sigma_i|f_i|)\right],\tag{11}$$

where the subscripts "0" and "μ" correspond to the initial state Ψ_α and final state Ψ_β, respectively, with $m_0=m_\mu=n$. We stress that $E_i^0\equiv E_{\sigma_i}^0(k_t,m_i)=-\sigma_i\gamma_0|f(k_t,m_i)|$, since the factor $(1-s\sigma_i|f_i|)^{-1}$ is canceled by those from the wave functions. $P_{\alpha\beta}$ is the total phase of the perturbation term and corresponds to inner products of the pseudospinors (see Sec. VI). Q corrects for contributions from a nonzero nearest-neighbor overlap. The intermediate states $\{\Psi_i\equiv\Psi_{\sigma_i}(k_t,m_i)\}$ satisfy the conservation law of angular momentum and have the following constraints:

$$m_{i-1}-m_i=\pm q,\quad i=1,\ldots,\mu-1,$$

$$m_{\mu-1}-m_\mu=\pm q+\text{multiple of }2n,\tag{12}$$

which leads to $\Sigma_{i=1}^{\mu}(m_{i-1}-m_i)=\Delta\mu q+\text{multiple of }2n=0$. The role of a nonzero $\Delta\mu$ can be seen from the extra phase factor

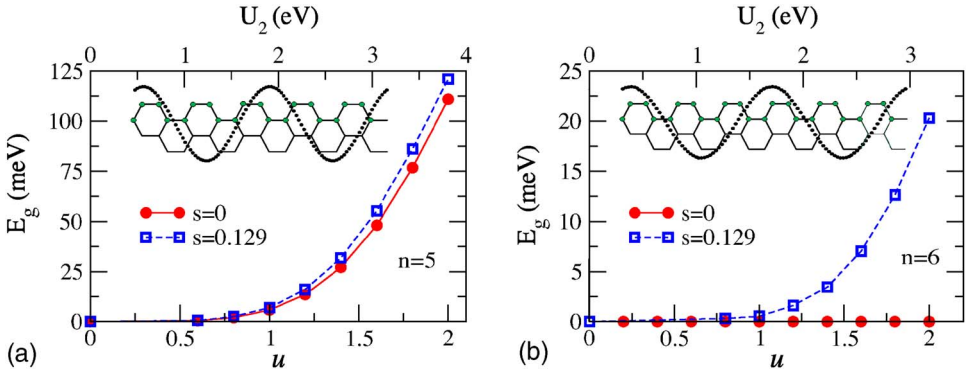

FIG. 2. (Color online). Band-gap variation of (a) (5, 5) and (b) (6, 6) A-SWNTs as a function of the applied angular potential with $q=2$. Insets: the unwrapped unit cell and schematics of the potential. Mirror planes of the potential pass through atomic sites so that all vertical mirror reflection and glide reflection symmetries are simultaneously broken.

$e^{-i\Delta\mu q\theta_0}$ in $H^{(\mu)}_{\alpha\beta}$ and the corresponding term $\Delta\mu q\pi/3n$ in $P_{\alpha\beta}$. Direct evaluation of Eq. (11) with all possible sets of $\{\Psi_i\}$ is formidable; nevertheless, one can get useful information by applying symmetry arguments.

A. Off-diagonal coupling: Gapping of A-SWNTs

We first study the off-diagonal term $H^{(\mu)}_{\pi\pi^*}$ and replace m_i with $\tilde{m}_i=2n-m$ in Eq. (11), which is allowed by conservation of angular momentum. The energy denominators and the function Q remain unchanged because $|f(k_t,\tilde{m})|=|f(k_t,m)|$, while the sign before $\Delta\mu q$ changes. Notice that for A-SWNTs, $f(k_t,\tilde{m})=e^{i4\pi/3}f^*(k_t,m)$, so that by the definition of ϕ, $\cos(\tilde{\phi}-\tilde{\phi}')=\sigma\sigma'\cos(\phi-\phi')$. Define $\tilde{\Psi}_i\equiv\Psi_{\sigma_i}(k_t,2n-m_i)$; then,

$$P_{\pi\pi^*}(\{\tilde{\Psi}_i\})=\left(\prod_{i=1}^{\mu}\sigma_{i-1}\sigma_i\right)P_{\pi\pi^*}(\{\Psi_i\})$$
$$=\sigma_\pi\sigma_{\pi^*}P_{\pi\pi^*}(\{\Psi_i\})=-P_{\pi\pi^*}(\{\Psi_i\}), \quad (13)$$

which leads to

$$H^{(\mu)}_{\pi\pi^*}(\{\Psi_i\})+H^{(\mu)}_{\pi\pi^*}(\{\tilde{\Psi}_i\})\propto\left(\frac{U_q}{2}\right)^{\mu}\sin(\Delta\mu q\theta_0)P_{\pi\pi^*}(\{\Psi_i\}). \quad (14)$$

Since the above relation is true for all possible sets of intermediate states at all coupling orders, one can conclude the following: (a) The coupling between π and π^* subbands is *always* zero if $\theta_0=0$—i.e., when a mirror plane of the potential overlaps with one of the vertical mirror planes (or glide planes) of the A-SWNT. This is an explicit result of the mirror symmetry requirement.[12] (b) The coupling is zero if $\Delta\mu=0$, which results from reflection symmetry of the energy bands: $E_\sigma(k_t,m)=E_\sigma(k_t,2n-m)$. This excludes the possibility of any nonzero second-order contribution—i.e., $\mu=2,\Delta\mu=0$. In other words, a *second-order* band gap is forbidden in A-SWNTs. (c) The next *lowest possible* $\Delta\mu$ satisfying the angular momentum conservation in Eq. (12) is $\mu_0\equiv 2n/\gcd(2n,q)$, which is also the lowest contributing order of the perturbation series.[12] For small U_q, the band-gap opening will be at least the μ_0th order in U_q,

$$E_g\approx 2|H_{\pi\pi^*}(k_F)|$$
$$\sim\frac{v_F}{R}u^{\mu_0}\sin(\mu_0 q\theta_0)h(q,n)+(\text{terms of }\mu>\mu_0), \quad (15)$$

where $k_F=2\pi/3a$ is the Fermi point with $a\simeq 2.5$ Å. $u=U_qR/v_F$ is the dimensionless potential, and $h(q,n)$ is a complicated function depending on the angular momentum of the potential q, as well as the A-SWNT index n.

The summation over all possible intermediate states can be further simplified by combining the original process with $\{\Psi_i\}$ and the reversal process with $\{\Psi^R_i\equiv\Psi_{-\sigma_{\mu-i}}(k_t,2n-m_{\mu-i})\}$. The sign change of the energy results in an extra factor of $(-1)^{\mu-1}$ in the denominator and the function Q changes accordingly. It can be proved that $P_{\pi\pi^*}(\{\Psi^R_i\})=P_{\pi\pi^*}(\{\Psi_i\})$, and therefore at small s the relation holds:

$$H^{(\mu)}_{\pi\pi^*}(\{\Psi_i\})+H^{(\mu)}_{\pi\pi^*}(\{\Psi^R_i\})\propto[1-(-1)^{\mu}]-s[1+(-1)^{\mu}]\sum_{i=1}^{\mu-1}\sigma_i|f_i|+O(s^2). \quad (16)$$

Within the orthogonal model ($s=0$), only the first term in Eq. (16) exists and is nonzero when μ_0, and therefore μ, is odd. This constraint on μ_0 results from the invariance of the inner product of pseudospinors upon reversal operation, in combination with the electron-hole symmetry $E^0_{-\sigma}(k_t,m)=-E^0_\sigma(k_t,m)$. The latter, however, is not an intrinsic property of SWNTs, but rather due to the nearest-neighbor approximation. For instance, the energy band symmetry is broken when the second-nearest-neighbor hopping integral is included or, equivalently, if $s\neq 0$. In the latter case, $Q(\{\Psi_i\})$ and $-Q(\{\Psi^R_i\})$ do not cancel out and a nonzero band gap proportional to s opens for even μ_0. Figure 2 plots the variation of band gap at $q=2$ calculated within the orthogonal and nonorthogonal TB models. At $s=0$, the (6,6) A-SWNT remains metallic, because the coupling order $\mu_0=6$ is forbidden. At nonzero s, a small band gap occurs and increases as

a power law of u. It is also found that the band gap grows linearly with the magnitude of s, consistent with the prediction of Eq. (16). In contrast, the band-gap curve of a (5,5) A-SWNT only shows a slight increase at nonzero s, because the corrections is of the order of s^2.

B. Diagonal coupling: Renormalization of the Fermi velocity

Except for a few special cases—for instance, with q =multiples of $2n$—the coupling order between π and π^* subbands is about the same order of n and the resulting band gap remains small. For example, $\mu_0=2n$ at $q=1$ and $\mu_0=n$ at $q=2$. However, the diagonal coupling matrix elements $H_{\pi\pi}$ and $H_{\pi^*\pi^*}$ are not necessarily small. The same symmetry arguments can be applied here. Upon reflection or reversal operation, $P_{\pi\pi}$ (or $P_{\pi^*\pi^*}$) remains the same and one obtains

$$H^{(\nu)}_{\alpha\alpha}(\{\Psi_i\}) + H^{(\nu)}_{\alpha\alpha}(\{\tilde{i}_i\}) \propto \left(\frac{U_q}{2}\right)^{\nu} \cos(\Delta\nu q\theta_0) P_{\alpha\alpha}(\{\Psi_i\}), \tag{17}$$

$$H^{(\nu)}_{\alpha\alpha}(\{\Psi_i\}) + H^{(\nu)}_{\alpha\alpha}(\{\Psi^R_i\}) \propto [P_{\alpha\alpha} - (-1)^{\nu} P_{\beta\beta}] - s[P_{\alpha\alpha} + (-1)^{\nu} P_{\beta\beta}] \sum_{i=1}^{\nu-1} \sigma_i |f_i| + O(s^2), \tag{18}$$

where $\alpha(\beta)=\pi,\pi^*$ and $\alpha\neq\beta$. Since $\cos(\Delta\nu q\theta_0)$ is always unity whenever $\Delta\nu=0$, the lowest contributing coupling order is therefore $\nu=2$. So unlike the off-diagonal coupling, nonzero diagonal terms $H_{\pi\pi}$ and $H_{\pi^*\pi^*}$ can always be obtained from low-order angular perturbation. In an orthogonal basis ($s=0$), only the first term in Eq. (18) remains, which corresponds to an energy shift for π and π^* subbands in the same direction when ν=odd and the opposite direction when ν=even. If $s\neq 0$, a relative shift between the two subbands always occurs. The values of $H_{\pi\pi}$ and $H_{\pi^*\pi^*}$ do not contribute to the band-gap opening of the A-SWNT, but may influence the Fermi-point position as well as the DOS near the Fermi level. Assume that $s=0$; then, the second-order perturbation summation is reduced to

$$\sum_{\{\Psi_i\}} H^{(2)}_{\pi\pi} = 2\left(\frac{U_q}{2}\right)^2 \frac{\cos^2[\phi_\pi - \phi_+(k_t,n+q)] - \cos^2[\phi_{\pi^*} - \phi_+(k_t,n+q)]}{-E_+(k_t,n+q)} = \frac{U_q^2}{2|\gamma_0|} F(k_t,q) = -\sum_{\{\Psi_i\}} H^{(2)}_{\pi^*\pi^*},$$

$$F(k_t,q) \approx -\frac{1}{2[1+2\cos(q\pi/n)]} + \frac{\sqrt{3}\cos(2q\pi/3n)}{8\sin^2(q\pi/2n)}(|k_t|-k_F)a + O(\Delta k^2). \tag{19}$$

The π subband is a *decreasing* function of $|k_t|$, and since $F(k_t,q)$ is an *increasing* function of $|k_t|$ at small q values, it becomes flattened near $\pm k_F$ as a result of the second-order perturbation. A similar trend can be found for the π^* subband. The new Fermi points move toward $k_t=0$ as $F(\pm k_F,q)<0$, and the renormalized Fermi velocity is given by

$$\bar{v}_F \approx \left(1+\frac{u^2}{2q^2}\right)^{-1}\left(v_F - \frac{\sqrt{3}n^2a}{4q^2\pi^2}\frac{U_q^2}{|\gamma_0|}\right) \approx \left(1-\frac{u^2}{q^2}\right)v_F, \tag{20}$$

where the prefactor $(1+u^2/2q^2)^{-1}$ is due to the normalization of the perturbed wave function. $\bar{v}_F$ can also be estimated from a chiral gauge transformation[18] as $J_0(2u/q)v_F \approx (1-u^2/q^2)v_F$, with J_0 the Bessel function of the first kind. The renormalized Fermi velocity of a (10, 10) A-SWNT is plotted in Fig. 3, and excellent agreement is found between the TB results and the analytical predictions. Under potentials of small q's, the shape of π and π^* subbands is strongly perturbed and the low-energy DOS is greatly enhanced (see the insets of Fig. 3), which becomes more evident for large radius A-SWNTs even at a relatively weak perturbation.

One may note that, for A-SWNTs, the finite curvature shifts the Fermi point further toward $k=0$ while inclusion of

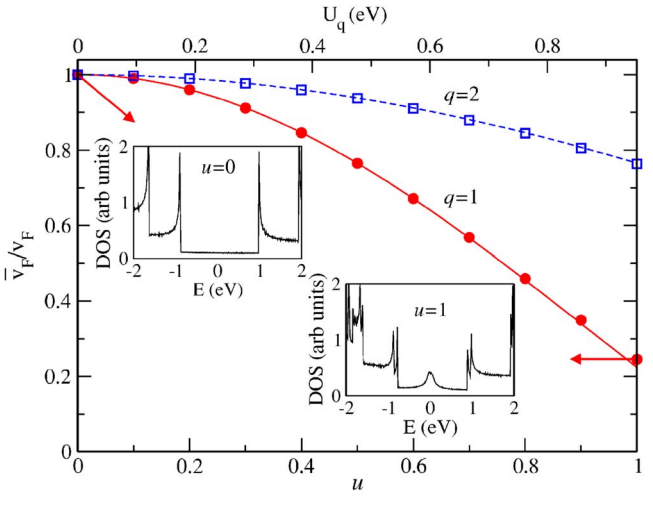

FIG. 3. (Color online). Renormalized Fermi velocity $\bar{v}_F$ of a (10, 10) A-SWNT as a function of u with $q=1$ (circles) and $q=2$ (squares). Solid and dashed lines are corresponding predictions from $J_0(2u/q)$. Insets show the DOS structure near the Fermi level at $u=0$ and $u=1$, with $q=1$.

σ orbitals may also modify the magnitude of the band-gap opening.[19,20] For the radius range considered in this paper, the correction remains small, and since our symmetry arguments are based on the lattice geometry of A-SWNTs (mirror or angular symmetry), the selections rules are not affected.

IV. TENSOR POTENTIAL: $q=2$

Another type of perturbation is realized by changing the hopping integrals between neighboring atomic orbitals—for example, by elastic deformations. Due to the high flexibility, carbon nanotubes can sustain remarkable deformations, which cause drastic changes in the electronic properties. Stretching, twisting, and squashing the nanotubes have been attempted both theoretically[3–5,13–15,21,22] and experimentally.[23–26] Here, we apply a radial deformation across the A-SWNT and derive the condition to open a secondary band gap.

Assume that a uniform stress is applied along the y direction, which is rotated from the A-SWNTs vertical mirror planes (or glide planes) by θ_0. The cross section of the A-SWNT is distorted into an ellipse with the two axes given by

$$R_y = R_0(1-\epsilon), \quad R_x = R_0(1+\nu\epsilon), \tag{21}$$

where ϵ is the strain along the y direction and ν is the cross-section Poisson ratio. At small strain, the hopping integral between nearest neighbors can be linearly expanded as

$$\gamma_{\lambda,i} - \gamma_0 \approx \left(\frac{\partial\gamma}{\partial r}\right)_{r_0} \Delta r_{\lambda,i} = \delta\gamma_\lambda \left(\frac{1-\nu}{1+\nu} + \cos 2\theta_{\lambda,i}\right), \tag{22}$$

$$\delta\gamma_\lambda = -2\epsilon r_0(1+\nu)\left(\frac{R_0}{r_0}\sin\frac{\theta_\lambda}{2}\right)^2\left(\frac{\partial\gamma}{\partial r}\right)_{r_0}, \tag{23}$$

where $\theta_{\lambda,i}$ corresponds to the midpoint of the ith bond along the $\mathbf{r}_\lambda$ direction. The first term in brackets in Eq. (22) has no angular dependence and does not break mirror symmetry because $\theta_2=\theta_3$ and thus $\delta\gamma_2=\delta\gamma_3$. Since the Poisson ratio was found to be close to unity,[13] the second term dominates, which corresponds to an angular momentum $q=2$.

The perturbation matrix element $H_{\alpha\beta}$ can be expanded into a perturbation series in a similar fashion as in Eq. (11), but now with $Q_{\alpha\beta}(\{\Psi_i\})=1$ and

$$P_{\alpha\beta}(\{\Psi_i\}) = \prod_{i=1}^{\mu}\sum_{\lambda=1}^{3} g_\lambda(k_t; m_{i-1}, \sigma_{i-1}; m_i, \sigma_i), \tag{24}$$

where g_λ is defined in Eq. (10). Using the fact that $\delta\gamma_2=\delta\gamma_3$, one can prove that

$$\sum_{\lambda=1}^{3} g_\lambda(k_t; m, \sigma, m', \sigma') = \sigma\sigma'\sum_{\lambda=1}^{3} g_\lambda(k_t; 2n-m, \sigma, 2n-m', \sigma') \tag{25}$$

$$= -\sum_{\lambda=1}^{3} g_\lambda(k_t; 2n-m', -\sigma'; 2n-m, -\sigma). \tag{26}$$

By applying the reflection operation $\Psi\to\tilde{\Psi}$ on the intermediate states, Eq. (13) and (14) are recovered, as required by mirror symmetry conservation. The lowest contributing order is therefore $\mu_0=2n/\gcd(2n,2)=n$. Now apply a reversal operation $\Psi\to\Psi^R$ as in Sec. III. Since $E^0_{-\sigma}(k_t,2n-m)=-E^0_\sigma(k_t,m)$, from Eq. (26) one has

$$H^{(\mu)}_{\pi\pi^*}(\{\Psi_i\}) + H^{(\mu)}_{\pi\pi^*}(\{\Psi^{\mathrm{R}}_i\}) = H^{(\mu)}_{\pi\pi^*}(\{\Psi_i\})\left[1+\frac{(-1)^\mu}{(-1)^{\mu-1}}\right] = 0. \tag{27}$$

This means that a band gap does not occur in any perturbation order and such a strain cannot induce MST in A-SWNTs. Certainly, a hidden symmetry—namely, the electron-hole symmetry $E^0_{-\sigma}(k_t,-m)=-E^0_\sigma(k_t,m)$—forbids the π and π^* subband mixing. By including the second-nearest-neighbor interactions, this symmetry can be weakly broken and a finite band gap occurs.[27] The magnitude of this band gap depends strongly on the parity of the A-SWNT index n. It was reported earlier that squashing a (6, 6) or (8, 8) A-SWNT does not induce a MST in the range of elastic deformation,[13,14] and the nanotube remains metallic until the two opposite walls are brought close enough to form new bonds. We prove that the vanishing band gap is caused by the high coupling order between π and π^* subbands, $\mu_0=n$, and the additional smallness of the overlap integral s and higher neighbor interactions, which make it impossible to observe the MST effect until the A-SWNT collapses. The situation is quite different, for example, for a (5, 5) A-SWNT which has an odd number index and a smaller coupling order, and a finite band gap was observed at moderate deformation.[5] On the other hand, when the radial deformation is large enough to induce a strong π-σ interaction, the single π-orbital description may no longer be sufficient.

V. COMBINATION OF DIFFERENT TYPES OF POTENTIALS

One way to reduce the coupling order μ_0 is by combining potentials of different angular momenta. For example, we have shown that by applying a scalar potential of the form of $V_0(\sin\theta+\sin 2\theta)$, μ_0 can be reduced to 3 for all values of n.[12] The combination of elastic radial deformation and uniaxial electrostatic potential will have a similar effect. By choosing appropriate angular momentum and relative position of the two components, the coupling order can be even lowered to $\mu_0=2$, as shown below.

Assume that a scalar potential (denoted as U) is applied on an A-SWNT together with a tensor perturbation (denoted as $\mathcal{E}\propto\delta\gamma_q$) of the same angular momentum q, but with an angular difference θ_d between the mirror planes of these two components. As shown in Secs. III and IV, the second-order contribution from either component is zero, but the cross

terms do not necessarily vanish. For the second-order coupling between π and π^* subbands, $\Psi_\pi \xrightarrow{\mathcal{E},U} \Psi_\sigma(k_t,m) \xrightarrow{U,\mathcal{E}} \Psi_{\pi^*}$, there are eight different cross terms with $m=n\pm q$ and $\sigma=\pm 1$. At $k_t=k_F$, the cross terms add up to

$$\sum H^{(2)}_{\pi\pi^*,\mathrm{cross}}(k_F) = \sin(q\theta_d)\sum_{\lambda=1}^{3}\frac{U_q\delta\gamma_{\lambda,q}\cos(k_F z_\lambda)}{i|E^0_+(k_F,n+q)|}\sin[2\phi_+(k_F,n+q)-(n+q/2)\theta_\lambda] \propto u\epsilon\sin(q\theta_d)\frac{|\gamma_0|}{q}, \quad q\ll n, \tag{28}$$

where $u=U_qR/v_F$ and $\epsilon\sim\delta\gamma/\gamma_0$ are the dimensionless potential and strain, respectively. According to Eq. (28), the coupling between π and π^* subbands is largest when $\theta_d=\pi/2q$ and always vanishes whenever the mirror planes of U and $\mathcal{E}$ overlap. The magnitude of the band gap, $E_g\approx 2|\Sigma H^{(2)}_{\pi\pi^*,\mathrm{cross}}(k_F)|$, is linear in both u and ϵ, and the dependence on the A-SWNT radius R (or index n) is very weak. This differs from the situation of mixed scalar potentials of different angular momenta, in which case E_g decreases with R as an inverse power law.[12]

In the case of radial deformation (Sec. IV), the tensor perturbation on the hopping integrals and the effective on-site potential[27] have overlapping mirror planes so that no second-order coupling occurs. Assume that one applies on the A-SWNT a scalar potential with the $q=2$ component shifted $\theta_d=\pi/4$ relative to the stress—for example, by changing the electrostatic environment around the A-SWNT. The resulting band gap is plotted in Fig. 4 for A-SWNTs of different radius as a function of u. The hopping integral under deformation is assumed to change as $\gamma\propto r^{-2}$, where the new bond length r is calculated from Eq. (21). The Poisson ratio is taken to be unity. The numerical values of E_g clearly follow a linear dependence on u and ϵ when the perturbation is weak. The radius dependence is barely seen even at large u, which means that one can always generate a substantial band gap in a large-radius A-SWNT using only a moderate external potential. Recently, it was found that the conductance of a carbon nanotube can be controlled by tuning the voltage of the local gate placed near mechanical defects on the tube (kinks or bends).[28] This scenario is reminiscent of the combination of scalar and tensor potentials of $q=1$, and might be related to the resulting second-order band gap.

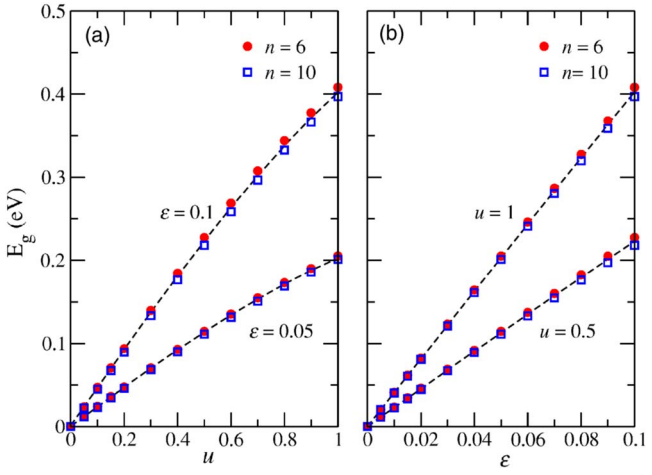

FIG. 4. (Color online). Band-gap variations of A-SWNTs (a) as a function of u at $\epsilon=0.05$, 0.1 and (b) as a function of ϵ at $u=0.5$, 1. Dashed lines are to guide the eye.

VI. METALLIC SWNTs WITH ARBITRARY CHIRALITY

In this section, we generalize our derivation of subband coupling to arbitrary metallic SWNTs by expanding the TB wave functions near the Fermi point. Only angular perturbations will be considered here; i.e., the axial wave vector is always conserved. An analogy is made to the cancellation rules of the backscattering process in SWNTs.[7] For the sake of clarity of the derivation, the overlap integral s is assumed zero. We recall that some cancellation rules will be weakly broken if the electron-hole symmetry is lifted, for example, by $s\neq 0$.

First, the electronic states are approximated as the product of a plane-wave part and a pseudospinor part by expanding the wave vector near the Fermi point $\mathbf{K}$ of two-dimensional graphite:

$$\Psi_\sigma(\hat{\mathbf{k}}) = \frac{e^{i\hat{\mathbf{k}}\cdot\mathbf{r}}}{\sqrt{2}}\begin{pmatrix} e^{i\phi_{\hat{\mathbf{k}},\sigma}} \\ e^{-i\phi_{\hat{\mathbf{k}},\sigma}}\end{pmatrix}, \quad 2\phi_\sigma(\hat{\mathbf{k}}) = -\sigma\frac{\pi}{2}+\arg(\hat{k}_t+i\hat{k}_c)+\eta, \tag{29}$$

where $\hat{\mathbf{k}}=(\hat{k}_t,\hat{k}_c=\hat{m}/R)$ is measured from $\mathbf{K}$ and η is the chiral angle. It can be proved that the definition of $\phi_\sigma(\hat{\mathbf{k}})$ here is consistent with that of $\phi_\sigma(\mathbf{k})$ in Sec. II. In contrast to Ref. 7, the phase difference due to the sign of σ is absorbed in the definition of $\phi_\sigma(\hat{\mathbf{k}})$ so that the product of two pseudospinors is always real. In addition, the wave functions of the two crossing subbands of metallic SWNTs remain continuous as a function of $\hat{k}_t$ when the wave vector passes through the Fermi point.

Assume applying an arbitrary angular scalar potential $V(\theta)=\Sigma_q V_q e^{iq\theta}$, where V_q is the angular Fourier component of the potential: $V_q\equiv(2\pi)^{-1}\oint d\theta V(\theta)e^{-iq\theta}$. The *direct* coupling matrix element between two states $\Psi_{\sigma_1}(\hat{\mathbf{k}})$ and $\Psi_{\sigma_2}(\hat{\mathbf{k}}')$, with $\hat{k}_t=\hat{k}'_t$, can be reduced to

$$M(\hat{\mathbf{k}},\sigma;\hat{\mathbf{k}}',\sigma') \equiv \langle \hat{\mathbf{k}},\sigma|V(\theta)|\hat{\mathbf{k}}',\sigma'\rangle \approx \sum_{Q} U_{\hat{m}-\hat{m}'+Q} e^{iQ(\theta_{0A}+\theta_{0B})/2} \cos[\phi-\phi' + Q(\theta_{0B}-\theta_{0A})/2], \tag{30}$$

where θ_{0A} and θ_{0B} are angular coordinates of any A and B atoms. Angular quantum number Q corresponds to the angular analog of the reciprocal lattice vector of two-dimensional graphite and accounts for contributions from the short-wavelength component of the potential. When all nonzero Q's are neglected, the matrix element $M(\hat{\mathbf{k}},\sigma;\hat{\mathbf{k}}',\sigma')$ is reduced to the product of the Fourier transform of the potential and the inner product of two pseudospinors, comparable with the on-site term in Eq. (5) for A-SWNTs.

If the perturbation is a tensor potential—i.e., affecting the off-site hopping integrals instead of the on-site energies—one can similarly write the coupling matrix as

$$M(\hat{\mathbf{k}},\sigma;\hat{\mathbf{k}}',\sigma') \equiv \langle \hat{\mathbf{k}},\sigma|\delta\gamma^{\mathrm{op}}|\hat{\mathbf{k}}',\sigma'\rangle \approx \sum_{\lambda=1}^{3} \delta\gamma_{\lambda,\hat{m}-\hat{m}'} \cos\left[\phi+\phi'-\left(\frac{\hat{\mathbf{k}}+\hat{\mathbf{k}}'}{2}+\mathbf{K}\right)\cdot\mathbf{r}_{\lambda}\right], \tag{31}$$

where $\delta\gamma_{\lambda,q}$ is the discrete angular Fourier transform of the change in hopping integrals: $\delta\gamma_{\lambda,q}=N^{-1}\Sigma_i \delta\gamma_{\lambda,i} e^{-iq\theta_{\lambda,i}}$. Comparing with Eq. (30), one finds that matrix elements for the two types of perturbations have a different dependence on the phase angle $\phi_{\sigma}(\hat{\mathbf{k}})$.

A. First-order coupling

The first-order subband coupling corresponds to direct mixing between $\Psi_+(\hat{\mathbf{k}})$ and $\Psi_-(\hat{\mathbf{k}})$ and has a straightforward description within nearly degenerate perturbation theory, where $\hat{\mathbf{k}}=(\hat{k}_t,0)$ for metallic SWNTs. Since $\phi_{\sigma}(\hat{\mathbf{k}})$ is continuous in the vicinity of $\hat{k}_t=0$, the diagonal matrix element $M(\hat{k}_t,\sigma;\hat{k}_t,\sigma)$ is continuous as well and merely shifts the location of the Fermi point and renormalizes the Fermi velocity. The change of the band gap is therefore determined by the off-diagonal term $M(\hat{k}_t,\sigma;\hat{k}_t,-\sigma)$.

According to Eq. (30), only Fourier components of the scalar potential with $q=Q$'s contribute to the direct coupling. For example, for (n,n) armchair or $(n,0)$ zigzag nanotubes, it can be an angular perturbation of the form $\cos(2n\theta)$. Such a Fourier component can be obtained by applying torsion, using chemical and biological decoration of the tube surface or the high multipoles of inhomogeneous potentials. On the other hand, since the pseudospinors of $\Psi_+(\mathbf{k})$ and $\Psi_-(\mathbf{k})$ are always orthogonal, the matrix element in Eq. (30) is reduced to

$$M(\hat{k}_t,+;\hat{k}_t,-) = \sum_{Q} U_Q e^{iQ(\theta_{0A}+\theta_{0B})/2} \sin\frac{Q(\theta_{0B}-\theta_{0A})}{2} = i(\tilde{U}_A-\tilde{U}_B)/2, \tag{32}$$

with $\tilde{U}_{A,B}=\Sigma_Q U_Q e^{iQ\theta_{0A,0B}}$. Equation (32) indicates that such potential components must be distinguishable at the two sublattices so as to mix the two orthogonal pseudospinors directly.

For the tensor potential, the symmetry is lower and even a uniform deformation ($q=0$) can result in the first-order coupling. The matrix element in Eq. (31) is reduced to

$$\begin{aligned} M(\hat{k}_t,+;\hat{k}_t,-) &= \sum_{\lambda=1}^{3} \delta\gamma_{\lambda} \cos[\arg(\hat{k}_t)+\eta-(\hat{\mathbf{k}}+\mathbf{K})\cdot\mathbf{r}_{\lambda}] \\ &\overset{\hat{k}_t\to 0}{=} \mathrm{sgn}(\hat{k}_t)[\delta\gamma_1\cos(\eta-2\pi/3)+\delta\gamma_2\cos\eta+\delta\gamma_3\cos(\eta+2\pi/3)] \\ &= \mathrm{sgn}(\hat{k}_t)\left[\cos\eta\left(\delta\gamma_2-\frac{\delta\gamma_1+\delta\gamma_3}{2}\right)+\frac{\sqrt{3}}{2}\sin\eta(\delta\gamma_1-\delta\gamma_3)\right], \end{aligned} \tag{33}$$

and more specifically for metallic achiral SWNTs,

$$M(0,+;0,-) \sim \begin{cases} \delta\gamma_2-\delta\gamma_3, & \text{armchair}, \\ 2\delta\gamma_2-(\delta\gamma_1+\delta\gamma_3), & \text{zigzag}, \end{cases} \tag{34}$$

which is consistent with previous findings about band-gap changes in metallic SWNTs under uniaxial and torsional strain.[4]

B. High-order coupling

When the first-order coupling between states $\Psi_+(\hat{\mathbf{k}})$ and $\Psi_-(\hat{\mathbf{k}})$ is forbidden, one has to turn to higher orders of the perturbation. Deriving the coupling could be tedious but some general rules can be built using the symmetry of pseudospinors. Similar to the case of A-SWNTs, coupling between the two crossing subbands of arbitrary metallic SWNT can be represented by $H_{+-}(\hat{k}_t)=\Sigma_{\mu}\Sigma_{\{\Psi_i\}}H_{+-}^{(\mu)}(\{\Psi_i\})$, and the lowest contributing coupling order μ_0 can be determined by the dominating Fourier components of the potential.

We first discuss the case of a scalar potential $V(\theta)$ and restrict $Q=0$ in Eq. (30). The cancellation rule is similar to those which led to Eq. (16). Assume there is a μth-order coupling process between $\Psi_+(\hat{\mathbf{k}})$ and $\Psi_-(\hat{\mathbf{k}})$ through intermediate states $\{\Psi_i \equiv \Psi_{\sigma_i}(\hat{k}_t,\hat{m}_i)\}$ with $i=1,\ldots,\mu-1$:

$$H_{+-}(\{\Psi_i\}) = \frac{\prod_{i=1}^{\mu} U_{\hat{m}_{i-1}-\hat{m}_i} \cos[\phi_{\sigma_{i-1}}(\hat{k}_t,\hat{m}_{i-1})-\phi_{\sigma_i}(\hat{k}_t,\hat{m}_i)]}{\prod_{i=1}^{\mu-1}[-E^0_{\sigma_i}(\hat{k}_t,\hat{m}_i)]}, \tag{35}$$

where the subscripts "0" and "μ" correspond to $\Psi_+(\hat{\mathbf{k}})$ and $\Psi_-(\hat{\mathbf{k}})$, respectively, with $\hat{m}_0=\hat{m}_\mu=0$. Define a reversal process with intermediate states $\{\Psi_i^R \equiv \Psi_{-\sigma_{\mu-i}}(\hat{k}_t,-\hat{m})\}$. After rearrangement of the summation orders and using the relation $\phi_{-\sigma}(\hat{k}_t,-\hat{m})=\eta-\phi_\sigma(\hat{k}_t,\hat{m})$, one arrives at

$$H_{+-}(\{\hat{k}_t,-\hat{m}_{\mu-i},-\sigma_{\mu-i}\})=\frac{\prod_{i=1}^{\mu} U_{\hat{m}_{i-1}-\hat{m}_i}\cos[\phi_{-\sigma_{i-1}}(k_t,-m_{i-1})-\phi_{-\sigma_i}(\hat{k}_t,-\hat{m}_i)]}{\prod_{i=1}^{\mu-1}[-E^0_{-\sigma_i}(\hat{k}_t,-\hat{m}_i)]}\approx(-1)^{\mu-1}H_{+-}(\{\hat{k}_t,\hat{m}_i,\sigma_i\}), \tag{36}$$

which cancels out with Eq. (35) for even μ. The approximation sign for the subband coupling process in Eq. (36) arises from assumptions on the reflection symmetry $E^0_\sigma(\hat{k}_t,\hat{m})=E^0_\sigma(\hat{k}_t,-\hat{m})$ and reversal symmetry $E^0_\sigma(\hat{k}_t,\hat{m})=-E^0_{-\sigma}(\hat{k}_t,\hat{m})$, which can be derived from the linear dispersion approximation—i.e., $E^0_\sigma(\hat{\mathbf{k}})=\sigma v_F|\hat{\mathbf{k}}|$. More generally, the reflection symmetry does not hold except for A-SWNTs due to the trigonal warping effect. For example, $E_\sigma(\hat{k}_t,1)\neq E_\sigma(\hat{k}_t,-1)$ for metallic zigzag SWNTs, and a secondary band gap $E_g\propto R^{-2}$ always opens under a uniform electric field perpendicular the nanotube radius.[6] Here $\hat{m}=\pm 1$ is measured relative to the Fermi point **K**. When a nonzero overlap s or high-order nearest-neighbor interaction is included, the reversal symmetry can also be weakly broken. We conclude that the selection rules for arbitrary metallic SWNTs are similar to those for A-SWNTs, but may acquire a chirality dependence beyond the linear dispersion approximation.

For general tensor potentials, $M(\hat{k}_t;\hat{m}_i,\sigma_i;\hat{m}_j,\sigma_j)$ and $M(\hat{k}_t,-\hat{m}_j,-\sigma_j;-\hat{m}_i,-\sigma_i)$ usually have different magnitude and no simple cancellation rule can be built. An exception is when $\gamma_2=\gamma_3$ for A-SWNT, e.g., under a radial deformation. In that case, the coupling between π and π^* subbands is reduced to zero as shown in Sec. IV due to cancellation from the reversal process within the nearest-neighbor approximation.

VII. CONCLUSION

In this paper we employ the group theory approach to clarify the issue of MST in armchair and other metallic SWNTs under angular perturbations. We study the symmetry requirements of MST on the nanotube and the perturbation potential, and demonstrate that the smallness of the MST effect is related to the symmetry of the pseudospinor components of the electron wave functions. Namely, the spinors of the crossing subbands are orthogonal; thus, any interaction between them is strongly weakened. For A-SWNTs, the gap is diminishing for almost any pure angular perturbation with a *single* angular Fourier component, due to the high coupling order. The coupling order is proportional to the number of atoms along the tube circumference for both types of perturbation studied here: the on-site (scalar) potential and off-site (tensor) deformation. The MST effect can be greatly enhanced by combining perturbations of different types and/or different angular momenta.

We formulate selection rules for the band-gap opening and its dependence on the perturbation strength. The combination of the diagrammatic derivation of interaction matrix elements and group theory technique allows one to predict the scaling of the band gap on the potential: $E_g\propto|\mathbf{V}|^{\mu_0}$, where the scaling exponent μ_0 can be easily calculated for an arbitrary metallic SWNT for given symmetry of the potential. Corrections may arise due to refinement of the model TB Hamiltonian—e.g., the electron-hole asymmetry, the inclusion of σ orbitals, and others not considered here. As an example of such refined model, we calculate the gap dependence on the overlap integral s added to the classic orthogonal TB model.

We present the analytical expression for the renormalization of the Fermi velocity, which occurs even if no MST is observed. The decrease of the Fermi velocity due to the perturbation is also seen as the enhancement of the DOS close to the Fermi level.

The MST effect by SWNT symmetry breaking could have potential applications for nanoscale electronic and optoelectronic devices. Additionally, we emphasize the possibility of engineering the nanotube DOS even when MST is forbidden under given perturbations, which can be potentially employed for SWNT opticals as well as switching devices.

ACKNOWLEDGMENTS

This work was supported by ARMY DURINT Contract No. SIT 527826-08 and NSF Grant CCR 01-21616. S.V.R. acknowledges support by a start-up fund and Feigl Scholarship of Lehigh University, and the NSF Grant No. ECS 04-03489.

APPENDIX: NEARLY DEGENERATE PERTURBATION THEORY

When calculating the coupling between two nearly degenerate states Ψ_α and Ψ_β, it is more convenient to treat them as degenerate states. Since $E_\alpha=E_\beta=0$ only at the crossing point, one can include the energy dispersion by redefining the un-

perturbed Hamiltonian and the external perturbation, for $|E_\alpha - E_\beta|$ being small:

$$\tilde{H}_0 = H_0 - E_\alpha |\Psi_\alpha\rangle\langle\Psi_\alpha| - E_\beta |\Psi_\beta\rangle\langle\Psi_\beta|,$$

$$\tilde{H}_1 = H_1 + H_0 - \tilde{H}_0. \tag{A1}$$

Ψ_α and Ψ_β now become degenerate states of $\tilde{H}_0$ with $\tilde{E}_{\alpha,\beta} = 0$. Their original energy difference is absorbed in the perturbation while other states Ψ' are not affected:

$$\langle\Psi_{\alpha,\beta}|\tilde{H}_1|\Psi_{\alpha,\beta}\rangle = E_{\alpha,\beta}, \quad \langle\Psi_\alpha|\tilde{H}_1|\Psi_\beta\rangle = \langle\Psi_\alpha|H_1|\Psi_\beta\rangle,$$

$$\langle\Psi'|\tilde{H}_1|\Psi_{\alpha,\beta}\rangle = \langle\Psi'|H_1|\Psi_{\alpha,\beta}\rangle, \quad \langle\Psi'|\tilde{H}_1|\Psi''\rangle = \langle\Psi'|H_1|\Psi''\rangle. \tag{A2}$$

As long as $\tilde{H}_1 - H_1$ is small, the rearrangement of H_0 and H_1 will not affect the results of the perturbation theory.

*Electronic address: yanli@uiuc.edu

[1] N. F. Mott, *Metal-Insulator Transitions* (Taylor & Francis, London, 1974).

[2] P. Delaney, H. J. Choi, J. Ihm, S. G. Louie, and M. L. Cohen, Nature (London) **391**, 466 (1998).

[3] C. L. Kane and E. J. Mele, Phys. Rev. Lett. **78**, 1932 (1997).

[4] L. Yang, M. P. Anantram, J. Han, and J. P. Lu, Phys. Rev. B **60**, 13874 (1999).

[5] C. J. Park, Y. H. Kim, and K. J. Chang, Phys. Rev. B **60**, 10656 (1999).

[6] Y. Li, S. V. Rotkin, and U. Ravaioli, Nano Lett. **3**, 183 (2003).

[7] T. Ando, T. Nakanishi, and R. Saito, J. Phys. Soc. Jpn. **67**, 2857 (1998); T. Ando and H. Suzuura, *ibid.* **71**, 2753 (2002).

[8] S. V. Rotkin and K. Hess, Appl. Phys. Lett. **84**, 3139 (2004).

[9] *Applied Physics of Nanotubes: Fundamentals of Theory, Optics and Transport Devices*, edited by S. V. Rotkin and S. Subramoney (Springer-Verlag, Berlin, 2005).

[10] *Carbon Nanotubes: Synthesis, Structure, Properties, and Applications*, edited by M. S. Dresselhaus, G. Dresselhaus, and P. Avouris (Springer-Verlag, Berlin, 2001).

[11] M. Damnjanović, T. Vuković, and I. Milošević, Solid State Commun. **116**, 265 (2000).

[12] Y. Li, S. V. Rotkin, and U. Ravaioli, Appl. Phys. Lett. **85**, 4178 (2004).

[13] O. Gülseren, T. Yildirim, S. Ciraci, and Ç. Kılıç, Phys. Rev. B **65**, 155410 (2002).

[14] J.-Q. Lu, J. Wu, W. Duan, F. Liu, B.-F. Zhu, and B.-L. Gu, Phys. Rev. Lett. **90**, 156601 (2003).

[15] H. Mehrez, A. Svizhenko, M. P. Anantram, M. Elstner, and T. Frauenheim, Phys. Rev. B **71**, 155421 (2005).

[16] S. Reich, J. Maultzsch, C. Thomsen, and P. Ordejón, Phys. Rev. B **66**, 035412 (2002).

[17] R. Saito, G. Dresselhaus, and M. S. Dresselhaus, *Physical Properties of Carbon Nanotubes* (Imperial College Press, London, 1998).

[18] D. S. Novikov and L. S. Levitov, cond-mat/0204499 (unpublished).

[19] N. Hamada, S.-I. Sawada, and A. Oshiyama, Phys. Rev. Lett. **68**, 1579 (1992).

[20] X. Blase, L. X. Benedict, E. L. Shirley, and S. G. Louie, Phys. Rev. Lett. **72**, 1878 (1994).

[21] A. Rochefort, P. Avouris, F. Lesage, and D. R. Salahub, Phys. Rev. B **60**, 13824 (1999).

[22] A. Charlier, E. McRae, R. Heyd, and M.-F. Charlier, J. Phys. Chem. Solids **62**, 439 (2001).

[23] J. Cao, Q. Wang, and H. Dai, Phys. Rev. Lett. **90**, 157601 (2003).

[24] E. D. Minot, Y. Yaish, V. Sazonova, J.-Y. Park, M. Brink, and P. L. McEuen, Phys. Rev. Lett. **90**, 156401 (2003).

[25] L.-J. Li, R. J. Nicholas, R. S. Deacon, and P. A. Shields, Phys. Rev. Lett. **93**, 156104 (2004).

[26] J. Wu *et al.*, Phys. Rev. Lett. **93**, 017404 (2004).

[27] The deformation also results in an effective on-site potential $V \propto \cos(2\theta)$, which contributes to the band-gap opening as given by Eqs. (14) and (16) with $\mu_0 = n$.

[28] M. J. Biercuk, N. Mason, J. M. Chow, and C. M. Marcus, Nano Lett. **4**, 2499 (2004).