Orbital Rashba effect and its detection by circular dichroism angle-resolved photoemission spectroscopy

Jin-Hong Park
Department of Physics and BK21 Physics Research Division, Sungkyunkwan University, Suwon 440-746, Korea

Choong H. Kim
Department of Physics and Astronomy, Seoul National University, Seoul 151-742, Korea

Jun-Won Rhim
School of Physics, Korea Institute for Advanced Study, Seoul 130-722, Korea

Jung Hoon Han*
Department of Physics and BK21 Physics Research Division, Sungkyunkwan University, Suwon 440-746, Korea and Asia Pacific Center for Theoretical Physics, Pohang University of Science and Technology, Pohang, 790-784, Korea
(Received 12 December 2011; revised manuscript received 28 December 2011; published 2 May 2012)

We show, by way of tight-binding and first-principles calculations, that a one-to-one correspondence between an electron's crystal momentum **k** and nonzero orbital angular momentum (OAM) is a generic feature of surface bands. The OAM forms a chiral structure in momentum space much as its spin counterpart in Rashba model does, as a consequence of the inherent inversion symmetry breaking at the surface but not of spin-orbit interaction. This is the orbital counterpart of conventional Rashba effect and may be called the "orbital Rashba effect." The circular dichroism (CD) angle-resolved photoemission (ARPES) method is an efficient way to detect this new order, and we derive formulas explicitly relating the CD-ARPES signal to the existence of OAM in the band structure. The cases of degenerate p- and d-orbital bands are considered.

DOI: 10.1103/PhysRevB.85.195401 PACS number(s): 79.60.−i, 72.25.−b, 73.20.At

I. INTRODUCTION

Rashba effect,[1] in the usual sense employed in surface science, refers to the breaking of spin degeneracy due to the inversion symmetry breaking (ISB) at the surface and the consequent formation of chiral spin angular momentum (SAM) structure in momentum space. How the loss of space symmetry leads to the breaking of spin degeneracy is customarily understood in the relativistic picture of a moving electron with momentum **k** in the xy plane. In its rest frame, the moving electron sees the perpendicular electric field $\mathbf{E} = \mathcal{E}\hat{z}$ as the magnetic field $\mathbf{B} = -(\mathbf{v}/c^2) \times \mathbf{E} = -(\hbar\mathbf{k}/mc^2) \times \mathbf{E}$. The Zeeman energy, $-\mu_B \boldsymbol{\sigma} \cdot \mathbf{B}$ ($\mu_B = e\hbar/2m$ = Bohr magneton), for the electron then reads

$$\left(\frac{\hbar}{mc}\right)^2 \frac{e}{2}\boldsymbol{\sigma} \cdot \mathbf{k} \times \mathbf{E} = (a_B\alpha_f)^2 \frac{e}{2}\boldsymbol{\sigma} \cdot \mathbf{k} \times \mathbf{E}. \qquad (1.1)$$

Bohr radius and fine structure constant are introduced as a_B and α_f, respectively. On a typical surface its strength may be estimated as $\sim\alpha_f^2(a_Bk)(e\mathcal{E}a_B)$, with $a_Bk \lesssim 1$ and $e\mathcal{E}a_B$ of the order of the surface work function, eV. The estimated Zeeman energy (Rashba-splitting energy) is therefore α_f^2 times the work function, admittedly a tiny splitting.

Experiments show a contrasting picture. Surface energy splitting due to the putative Rashba effect has been experimentally observed on the surface of several simple elements, such as Au(111) (Refs. 2–5), Bi (Ref. 6), Sb (Ref. 7), and some alloys as well,[8] all of which produced energy scales in excess of 100 meV. A lot of theoretical effort in the last decade in surface science has been to bridge the discrepancy in the energy scales between experiments and the naive Rashba theory.[9]

On the conceptual side, the prototype Rashba model assumes a free electron model even though typical surface systems where Rashba splitting has been observed consist of multiple p orbitals. Besides, crystal field splitting is minimal with simple elements like Au, Bi, and Sb, so all three atomic p orbitals will be nearly degenerate in their on-site energies. Such orbital degeneracy naturally opens up the possibility of the internal degrees of freedom of the orbital system, that is, orbital angular momentum (OAM), being an active participant in the determination of surface band structure. The main theme of this paper is to exploit this point in relation to the emergence of chiral OAM structure. It will be shown that electrostatic interaction leads to the momentum (**k**)-dependent lifting of the orbital degeneracy, in a way that gives rise to chiral OAM structure and what may be termed the "orbital Rashba effect." For strongly spin-orbit-coupled systems the chiral OAM gives way to chiral structure of the total angular momentum,[10] but with the basic energetics still dictated by the same electrostatics consideration. We emphasize that the phenomenon of chiral OAM can take place in the absence of spin-orbit interaction (SOI), as the only symmetry requirement for its existence is the loss of inversion symmetry. The mechanism by which the chiral SAM structure (conventional Rashba effect) arises on top of the pre-existing chiral OAM (our orbital Rashba effect) as the SOI is added is explained.

An important and essential ingredient of our proposal is the detection scheme by which one can unambiguously identify the surface bands that carry OAM. Spin-resolved angle-resolved photoemission (SARPES) is often employed to confirm the predicted chiral spin polarization of the Rashba-split bands.[11,12] Here we show that a similar criterion exists with the dichroism experiment, in which left- and right-circularly polarized (LCP and RCP) lights are employed in the photoemission experiments. The dichroism signal, defined as

the difference of the photoelectron intensity from the $\mathbf{k}$-vector position in momentum space with the incident light being LCP or RCP, is shown to be proportional to the average OAM $\langle \boldsymbol{L} \rangle$ of the quasiparticle state at that momentum. Hence, the nonzero signal in the circular dichroism ARPES (CD-ARPES) is a direct proof of nonzero OAM in the band structure being probed.

This paper is organized as follows. In Sec. II the electrostatics view of both spin and orbital Rashba splitting is summarized. Following it, we present a microscopic tight-binding calculation supporting the general electrostatics-driven picture in Sec. III. A model first-principles calculation is made for a single layer of Bi atoms where a complete agreement in the band-dependent OAM structure between first-principles LDA calculation and tight-binding model result is obtained. Analysis of dichroism for OAM-carrying bands is given in Sec. IV where some CD-ARPES formulas for p- and d-orbital systems are presented. A brief summary is given in Sec. V.

II. ELECTROSTATIC ENERGY PERSPECTIVE

The tiny energy splitting expected from the relativistic free-electron argument and the contrasting observation of a significant Rashba-splitting energy of a few hundred meV on actual surfaces argue in favor of a completely different mechanism at work. A feature of the degenerate orbital system is that it allows the introduction of an orientation vector $\hat{\boldsymbol{n}}$ characterizing the degeneracy manifold. One can show that Bloch states formed from Wannier states of different $\hat{\boldsymbol{n}}$ generally carry different dipolar moments. In the presence of the electrostatic field normal to the surface, such dipolar moment tends to be oriented along the normal to save electrostatic energy. This, in turn, leads to the momentum-dependent selection of preferred $\hat{\boldsymbol{n}}$ direction of the quasiparticle state. We show that the resultant energy splitting is exactly of the Rashba form, $\sim \hat{\boldsymbol{n}} \times \mathbf{k} \cdot \hat{z}$.

First we need to define the internal quantum number $\hat{\boldsymbol{n}}$ of the state in some fashion. In a single-band model, Bloch eigenstate is derived from the Wannier state as (N is the number of lattice sites)

$$|\mathbf{k}\rangle = \frac{1}{\sqrt{N}} \sum_i e^{i\mathbf{k}\cdot\mathbf{r}_i} |i\rangle, \tag{2.1}$$

where $\phi(\mathbf{r}-\mathbf{r}_i) = \langle \mathbf{r}|i\rangle$ would be a featureless, localized wave function centered at the atomic site $\mathbf{r}_i$. The situation changes dramatically in a degenerate, multiorbital system because the degenerate Wannier orbitals can be mixed in a coherent fashion. For illustration we choose $J = 1/2$ doublet where the two Wannier states are[10]

$$\begin{aligned} |u\rangle &= \frac{1}{\sqrt{3}}(|p_x\downarrow\rangle + i|p_y\downarrow\rangle + |p_z\uparrow\rangle) = \frac{1}{\sqrt{3}}(|0\uparrow\rangle - \sqrt{2}|1\downarrow\rangle), \\ |d\rangle &= \frac{1}{\sqrt{3}}(|p_x\uparrow\rangle - i|p_y\uparrow\rangle - |p_z\downarrow\rangle) = \frac{1}{\sqrt{3}}(\sqrt{2}|\bar{1}\uparrow\rangle - |0\downarrow\rangle), \end{aligned} \tag{2.2}$$

expressed in the p-orbital and angular-momentum bases, respectively. One can easily check that $\boldsymbol{J} = \boldsymbol{L} + (1/2)\boldsymbol{\sigma}$ satisfies all the standard spin-$\frac{1}{2}$ algebra in this restricted Hilbert space:

$$\begin{aligned} &J^z|u\rangle = \tfrac{1}{2}|u\rangle, \quad J^z|d\rangle = -\tfrac{1}{2}|d\rangle, \\ &J^+|u\rangle = 0, \quad J^+|d\rangle = |u\rangle, \\ &J^-|u\rangle = |d\rangle, \quad J^-|d\rangle = 0. \end{aligned} \tag{2.3}$$

A coherent state can be formed in the standard fashion:

$$|\hat{\boldsymbol{n}}\rangle = \cos\frac{\theta}{2}|u\rangle + e^{i\phi}\sin\frac{\theta}{2}|d\rangle, \quad \boldsymbol{J}\cdot\hat{\boldsymbol{n}}|\hat{\boldsymbol{n}}\rangle = \frac{1}{2}|\hat{\boldsymbol{n}}\rangle, \tag{2.4}$$

where $\hat{\boldsymbol{n}} = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$. With such a coherent Wannier state $|\hat{\boldsymbol{n}}, i\rangle$ localized at each site i, a Bloch state can be constructed as

$$|\hat{\boldsymbol{n}}, \mathbf{k}\rangle = \frac{1}{\sqrt{N}} \sum_i e^{i\mathbf{k}\cdot\mathbf{r}_i} |\hat{\boldsymbol{n}}, i\rangle. \tag{2.5}$$

As long as the underlying Hamiltonian preserves the p-orbital degeneracy, the energy of the Bloch state will be invariant with respect to the orientation of $\hat{\boldsymbol{n}}$. As the following exercise shows, however, the surface-normal electric field breaks this degeneracy to give rise to $\hat{\boldsymbol{n}}$ dependence in the energy.

In first-quantized form the Bloch wave function labeled by two quantum numbers $\mathbf{k}$ and $\hat{\boldsymbol{n}}$ becomes

$$\boldsymbol{\psi}_{\hat{\boldsymbol{n}},\mathbf{k}}(\mathbf{r}) = \sum_i e^{i\mathbf{k}\cdot\mathbf{r}_i} \boldsymbol{\phi}_{\hat{\boldsymbol{n}}}(\mathbf{r}-\mathbf{r}_i). \tag{2.6}$$

Boldface is used to emphasize the spinor nature of the wave function, for example,

$$\boldsymbol{\phi}_{\lambda,\uparrow}(\mathbf{r}) = \begin{pmatrix} \phi_\lambda(\mathbf{r}) \\ 0 \end{pmatrix}, \quad \boldsymbol{\phi}_{\lambda,\downarrow}(\mathbf{r}) = \begin{pmatrix} 0 \\ \phi_\lambda(\mathbf{r}) \end{pmatrix}. \tag{2.7}$$

Each $\phi_\lambda(\mathbf{r}-\mathbf{r}_i) \sim (\mathbf{r}-\mathbf{r}_i)_\lambda f(|\mathbf{r}-\mathbf{r}_i|)$ is the λth orbital Wannier function, with $\lambda = x, y, z$. The density of the Bloch wave function is

$$\begin{aligned} \boldsymbol{\psi}^\dagger_{\hat{\boldsymbol{n}},\mathbf{k}}(\mathbf{r})\boldsymbol{\psi}_{\hat{\boldsymbol{n}},\mathbf{k}}(\mathbf{r}) = &\sum_i \boldsymbol{\phi}^\dagger_{\hat{\boldsymbol{n}}}(\mathbf{r}-\mathbf{r}_i)\boldsymbol{\phi}_{\hat{\boldsymbol{n}}}(\mathbf{r}-\mathbf{r}_i) \\ &+ \sum_{i\neq j} e^{i\mathbf{k}\cdot(\mathbf{r}_i-\mathbf{r}_j)} \boldsymbol{\phi}^\dagger_{\hat{\boldsymbol{n}}}(\mathbf{r}-\mathbf{r}_j)\boldsymbol{\phi}_{\hat{\boldsymbol{n}}}(\mathbf{r}-\mathbf{r}_i). \end{aligned} \tag{2.8}$$

The first term is the sum of the local densities of individual Wannier orbitals, which turns out to be isotropic and independent of $\hat{\boldsymbol{n}}$. The second part, on the other hand, depends explicitly on $\hat{\boldsymbol{n}}$ and carries nonzero dipole moment. The electrostatic energy due to surface electric field $\mathbf{E} = \mathcal{E}\hat{z}$ can be written as $E_{\rm es}(\hat{\boldsymbol{n}}, \mathbf{k}) = -e\mathcal{E}\int d^3\mathbf{r} z \boldsymbol{\psi}^\dagger_{\hat{\boldsymbol{n}},\mathbf{k}}(\mathbf{r})\boldsymbol{\psi}_{\hat{\boldsymbol{n}},\mathbf{k}}(\mathbf{r})$ for each Bloch state. Upon detailed calculation one finds the energy

$$E_{\rm es}(\hat{\boldsymbol{n}}, \mathbf{k}) = \eta(e\mathcal{E}a) \sum_{i\neq j} \sin(\mathbf{k}\cdot\mathbf{r}_{ij})(\hat{\boldsymbol{n}}\times\mathbf{r}_{ij}\cdot\hat{z}) e^{-\frac{1}{4a^2}(\mathbf{r}_i-\mathbf{r}_j)^2}. \tag{2.9}$$

In order to arrive at this formula we assumed a Gaussian envelope $\phi_\lambda(\mathbf{r}) \sim x_\lambda e^{-r^2/2a^2}$ and the atomic layer in the xy plane with $z_i = 0$. With a more general prescription $\phi_\lambda(\mathbf{r}) \sim x_\lambda f(|\mathbf{r}|)$ we obtain some other fast-decaying envelope function replacing the Gaussian in the above formula. The numerical constant η of order unity can be fixed as well with an explicit envelope function $f(|\mathbf{r}|)$.

The above sum can be brought to a transparent form if we restrict it to the nearest-neighbor sites. For two-dimensional square lattice they are $\mathbf{r}_{ij} = a\hat{x}$ and $a\hat{y}$, and

$$\begin{aligned} E_{\rm es}(\hat{\boldsymbol{n}}, \mathbf{k}) &\sim (e\mathcal{E}a)[n_x \sin(k_y a) - n_y \sin(k_x a)] \\ &\sim (e\mathcal{E}a^2)\hat{\boldsymbol{n}}\times\mathbf{k}\cdot\hat{z}, \end{aligned} \tag{2.10}$$

the second line following from the first in the small-$\mathbf{k}$ limit. To save electrostatic energy, $\hat{\boldsymbol{n}} = \langle \mathbf{k}|\boldsymbol{J}|\mathbf{k}\rangle$ needs to be oriented

along $\mathbf{k}\times\hat{z}$, which will result in the chiral angular momentum pattern $\hat{\boldsymbol{n}} \parallel \mathbf{k}\times\hat{z}$ in the Brillouin zone close to the Γ point. In an effective spin-$\frac{1}{2}$ system considered now, only two orientations of $\hat{\boldsymbol{n}}$, parallel and antiparallel to $\mathbf{k}\times\hat{z}$, will be allowed. They form the upper and lower branches of the Rashba-split bands. In the strong-SOI band, it is therefore the total spin $\hat{\boldsymbol{n}}$ which should behave as anticipated in the original Rashba picture. Being electrostatic in origin, however, the real driving force for Rashba splitting is not the electron's spin, but rather its orbital dipole moment.[10]

It is not surprising then that Rashba-like splitting can occur even in the complete absence of SOI. Indeed, the above simple energetic argument rests only on the ISB effect arising from the surface-normal electric field. To check this idea we assume spinless p-orbital system where the coherent state is formed in terms of $|1\rangle,|0\rangle,|\bar{1}\rangle \equiv |-1\rangle$ basis states as

$$|\hat{\boldsymbol{n}}\rangle = e^{-i\phi}\cos^2\frac{\theta}{2}|1\rangle + \sqrt{2}\cos\frac{\theta}{2}\sin\frac{\theta}{2}|0\rangle + e^{i\phi}\sin^2\frac{\theta}{2}|\bar{1}\rangle. \tag{2.11}$$

This state satisfies the coherent-state condition $\boldsymbol{L}\cdot\hat{\boldsymbol{n}}|\hat{\boldsymbol{n}}\rangle = |\hat{\boldsymbol{n}}\rangle$. Repeating the same steps as before, we obtain the dipolar energy of the Bloch state that depends on $\hat{\boldsymbol{n}}$ in exactly the same way as Eq. (2.10). In this extreme, it is the OAM, not SAM, that forms a chiral pattern in the momentum space.

Although we do not extend the calculation any further, it is clear that the dipolar energy $E_{\rm es}(\hat{\boldsymbol{n}},\mathbf{k}) \sim W\hat{\boldsymbol{n}}\times\mathbf{k}\cdot\hat{z}$ is a generic result, with $W \sim e\mathcal{E}a$ of the order of the work function, for surfaces subject to ISB. For one thing it is the only scalar quantity involving three vectors $\hat{\boldsymbol{n}}$, $\mathbf{k}$, and $\hat{z}$ and preserves time-reversal symmetry but breaks inversion. When SOI is very weak, the role of $\hat{\boldsymbol{n}}$ is taken up by the OAM $\boldsymbol{L}$. For strong SOI, it is replaced by the total angular momentum $\boldsymbol{J}$. In both instances the orbital degeneracy is an important ingredient in forming the degenerate manifold of states spanned by orientations of $\hat{\boldsymbol{n}}$. The breaking of such degeneracy due to the surface potential gradient then results in the Rashba-type energy splitting. In the following sections we largely focus on cases of weak SOI where chiral OAM effect dominates and provide further microscopic justifications for its existence, as well as a theory for its detection.

III. *p*-ORBITAL SURFACE

A generic tight-binding Hamiltonian consisting of p_x, p_y, and p_z orbitals can be constructed with the aid of Slater-Koster parameters,

$$\begin{aligned}
(p_x|p_x) &= -V_1\cos^2\theta_0 + V_2\sin^2\theta_0,\\
(p_y|p_y) &= -V_1\sin^2\theta_0 + V_2\cos^2\theta_0,\\
(p_x|p_y) &= (V_2-V_1)\cos\theta_0\sin\theta_0 = (p_y|p_x),\\
(p_z|p_z) &= V_2.
\end{aligned} \tag{3.1}$$

We choose the convention

$$(\cos\theta_0,\sin\theta_0) = (x_0,y_0)/\sqrt{x_0^2+y_0^2} \tag{3.2}$$

for the relative vector $\mathbf{r}_0 = (x_0,y_0)$ of the two Wannier orbitals. The symbol $(p_i|p_j)$ refers to the matrix element $\langle p_i|H|p_j\rangle$ between adjacent p orbitals, with the p_i orbital at the origin and p_j at $\mathbf{r}_0$. Surface character is implicitly introduced by assuming a Hamiltonian that breaks inversion symmetry, which therefore allows the following matrix elements:

$$\begin{aligned}
(p_z|p_x) &= \tfrac{3}{2}\gamma\cos\theta_0 = -(p_x|p_z),\\
(p_z|p_y) &= \tfrac{3}{2}\gamma\sin\theta_0 = -(p_y|p_z).
\end{aligned} \tag{3.3}$$

Nonzero γ is an indicator of the ISB. To the tight-binding model we include the local spin-orbit Hamiltonian, $H_{\rm SO} = (\alpha/2)\sum_i \boldsymbol{L}_i\cdot\boldsymbol{\sigma}_i$ to complete the construction of the model. A similar Hamiltonian was studied previously.[13] We analyze two extreme situations in the two sections below, one where SOI is dominant ($\alpha \gg \gamma$) and the other where ISB prevails ($\gamma \gg \alpha$). Most of the results in Sec. III A are already reported in Ref. 10. They are included here for completeness and to define the notations to be used through the rest of the paper.

A. SOI over ISB

In the limit where SOI interaction dominates over the energy scales arising from ISB, $J = 3/2$ atomic states can be regarded as completely decoupled from those spanned by $J = 1/2$ manifold. When restricted to the $J = 1/2$ subspace, the tight-binding Hamiltonian, on a square lattice of unit spacing, becomes

$$H = -2\sum_{\mathbf{k}} C_{\mathbf{k}}^\dagger \begin{pmatrix} t(c_x+c_y) & \gamma(s_y+is_x)\\ \gamma(s_y-is_x) & t(c_x+c_y)\end{pmatrix} C_{\mathbf{k}}, \tag{3.4}$$

with $t = (V_1 - 3V_2)/3$, and shorthand notations are $c_{x(y)} = \cos k_{x(y)}$, $s_{x(y)} = \sin k_{x(y)}$. The effect of SOI is already incorporated through the $J = 1/2$ eigenstates forming the basis states. $C_{\mathbf{k}}$ refers to the Fourier transform

$$C_{\mathbf{k}} = \begin{pmatrix} u_{\mathbf{k}}\\ d_{\mathbf{k}}\end{pmatrix} = \frac{1}{\sqrt{N}}\sum_i e^{-i\mathbf{k}\cdot\mathbf{r}_i}\begin{pmatrix} u_i\\ d_i\end{pmatrix}. \tag{3.5}$$

Here u_i and d_i are the annihilation operators for the $|u\rangle$ and $|d\rangle$ Wannier states introduced in Eq. (2.2). Note that the off-diagonal elements are entirely due to ISB.

Defining $e^{i\phi_{\mathbf{k}}} = (-s_y + is_x)/\sqrt{s_x^2+s_y^2}$, the eigenstates are

$$|\mathbf{k},\pm\rangle = \frac{1}{\sqrt{2}}(|u_{\mathbf{k}}\rangle \pm e^{i\phi_{\mathbf{k}}}|d_{\mathbf{k}}\rangle) \tag{3.6}$$

for the upper and lower bands with energy $-2t(c_x + c_y)\pm 2\gamma(s_x^2+s_y^2)^{1/2}$, respectively. Near the Γ point the two energy bands are split as $t\mathbf{k}^2 \pm 2\gamma|\mathbf{k}|$ in accordance with the phenomenological Rashba theory. Average SAM and OAM for $|\mathbf{k},\pm\rangle$ are given by the expectation values of

$$\boldsymbol{\sigma} = \frac{1}{N}\sum_i \boldsymbol{\sigma}_i, \quad \boldsymbol{L} = \frac{1}{N}\sum_i \boldsymbol{L}_i. \tag{3.7}$$

Here $\boldsymbol{\sigma}_i$ and $\boldsymbol{L}_i$ are the spin and OAM operators acting on the particular atomic site $\mathbf{r}_i$. It is easily shown that $\langle\mathbf{k},\pm|\sigma^z|\mathbf{k},\pm\rangle = \langle\mathbf{k},\pm|L^z|\mathbf{k},\pm\rangle = 0$, while

$$\begin{aligned}
\langle\mathbf{k},\pm|\sigma^+|\mathbf{k},\pm\rangle &= \pm\tfrac{1}{6}e^{-i\phi_{\mathbf{k}}},\\
\langle\mathbf{k},\pm|L^+|\mathbf{k},\pm\rangle &= \mp\tfrac{2}{3}e^{-i\phi_{\mathbf{k}}},
\end{aligned} \tag{3.8}$$

$L^+ = L^x + iL^y$, $\sigma^+ = \sigma^x + i\sigma^y$. The upper Rashba-split band carries SAM/OAM which are opposite to those carried

by the lower band. Within each band, SAM and OAM are oppositely oriented as a result of the strong SOI that favors antiparallel alignment of $\boldsymbol{L}$ and $\boldsymbol{\sigma}$. The electrostatics argument of the previous section helps clarify that the magnitude of γ is of the order of the work function.

B. ISB over SOI

It is useful to consider $\alpha = 0$ band structure as a zeroth-order solution and include effects of α perturbatively when SOI strength α is much smaller than the ISB parameter γ. We find that the $\alpha = 0$, $\gamma \neq 0$ band already supports the chiral OAM structure. The perturbative influence of SOI is to spin split each OAM-carrying band according to the conventional Rashba scheme. The result is a pair of bands having the common OAM orientation, with opposite SAM orientations in each band.

Spin states are degenerate for $\alpha = 0$, and we obtain spinless 3×3 Hamiltonian on a square lattice,

$$H = \sum_{\mathbf{k}} \mathbf{C}_{\mathbf{k}}^{\dagger} \begin{pmatrix} 2(V_2 c_y - V_1 c_x) & 0 & -3\gamma i s_x \\ 0 & 2(V_2 c_x - V_1 c_y) & -3\gamma i s_y \\ 3\gamma i s_x & 3\gamma i s_y & 2V_2(c_x + c_y) \end{pmatrix} \mathbf{C}_{\mathbf{k}}, \tag{3.9}$$

where, in obvious notation,

$$\mathbf{C}_{\mathbf{k}} = \begin{pmatrix} p_{x,\mathbf{k}} \\ p_{y,\mathbf{k}} \\ p_{z,\mathbf{k}} \end{pmatrix} = \frac{1}{\sqrt{N}} \sum_i e^{-i\mathbf{k}\cdot\mathbf{r}_i} \begin{pmatrix} p_{x,i} \\ p_{y,i} \\ p_{z,i} \end{pmatrix}. \tag{3.10}$$

Each $p_{\lambda,i}$ ($\lambda = x,y,z$) expresses an annihilation operator for the λ orbital at site i. It proves useful to study the linearized Hamiltonian, valid near the Γ point, and return to the full problem numerically. For the linearized Hamiltonian,

$$H = \sum_{\mathbf{k}} \mathbf{C}_{\mathbf{k}}^{\dagger} \begin{pmatrix} 2(V_2 - V_1) & 0 & -3\gamma i k_x \\ 0 & 2(V_2 - V_1) & -3\gamma i k_y \\ 3\gamma i k_x & 3\gamma i k_y & 4V_2 \end{pmatrix} \mathbf{C}_{\mathbf{k}}, \tag{3.11}$$

the eigenstates and energies can be obtained in closed form. Using the degeneracy of p_x and p_y orbitals, the three $\gamma = 0$ eigenstates can be organized as

$$\begin{aligned} |\mathrm{I}\rangle &= \frac{k_y}{k}|p_x\rangle - \frac{k_x}{k}|p_y\rangle, \\ |\mathrm{II}\rangle &= \frac{k_x}{k}|p_x\rangle + \frac{k_y}{k}|p_y\rangle, \\ |\mathrm{III}\rangle &= |p_z\rangle, \end{aligned} \tag{3.12}$$

where $k = |\mathbf{k}|$. The γ terms in Eq. (3.11) mix states $|\mathrm{II}\rangle$ with $|\mathrm{III}\rangle$, leaving the state $|\mathrm{I}\rangle$ decoupled from the rest. The influence of γ is summarized as a 2×2 Hamiltonian,

$$H = (c_{\mathrm{II}}^{\dagger} \quad c_{\mathrm{III}}^{\dagger}) \begin{pmatrix} 2(V_2 - V_1) & -3i\gamma k \\ 3i\gamma k & 4V_2 \end{pmatrix} \begin{pmatrix} c_{\mathrm{II}} \\ c_{\mathrm{III}} \end{pmatrix}, \tag{3.13}$$

where $c_{\mathrm{II}}^{\dagger}, (c_{\mathrm{II}}), c_{\mathrm{III}}^{\dagger}, (c_{\mathrm{III}})$ denote the creation (annihilation) operator of the states $|\mathrm{II}\rangle, |\mathrm{III}\rangle$. It assumes the massive Dirac Hamiltonian where the mass term is a consequence of different hopping amplitudes among the in-plane (p_x and p_y) and out-of-plane (p_z) orbitals. To arrive at the Dirac form, however, the basis states themselves are chosen to corotate with the $\mathbf{k}$ orientation as prescribed in Eq. (3.12). It differs significantly from the conventional way of writing down the Rashba Hamiltonian, which is done in the fixed spin-up and spin-down basis states. There is a potentially intriguing issue of the Berry phase related to the chiral orbital structure and its physical manifestation, which we hope to pursue in a future presentation.

Finally, the eigenstates and their energies of the linearized Hamiltonian are obtained, $\Delta = V_1 + V_2$,

$$\begin{aligned} |\mathbf{k},1\rangle &= |\mathrm{I}\rangle, \\ |\mathbf{k},2\rangle &= |\mathrm{II}\rangle - \frac{3i\gamma k}{2\Delta}|\mathrm{III}\rangle, \\ |\mathbf{k},3\rangle &= |\mathrm{III}\rangle - \frac{3i\gamma k}{2\Delta}|\mathrm{II}\rangle, \\ E_1(\mathbf{k}) &= -2V_1 + 2V_2, \\ E_2(\mathbf{k}) &= -2V_1 + 2V_2 - \frac{9\gamma^2 k^2}{2\Delta}, \\ E_3(\mathbf{k}) &= 4V_2 + \frac{9\gamma^2 k^2}{2\Delta}. \end{aligned} \tag{3.14}$$

Two of the bands carry nonzero OAM,

$$\langle \mathbf{k},2|L^+|\mathbf{k},2\rangle = \frac{3i\gamma}{\Delta}(k_x + ik_y) = -\langle \mathbf{k},3|L^+|\mathbf{k},3\rangle. \tag{3.15}$$

The remaining band has $\langle \mathbf{k},1|L^+|\mathbf{k},1\rangle = 0$ and the net OAM is zero: $\sum_{n=1}^{3} \langle \mathbf{k},n|\mathbf{L}|\mathbf{k},n\rangle = 0$. OAM obtained from numerical diagonalization of the full Hamiltonian (3.9) is shown in Fig. 1. The net OAM is exactly zero everywhere in the Brillouin zone.

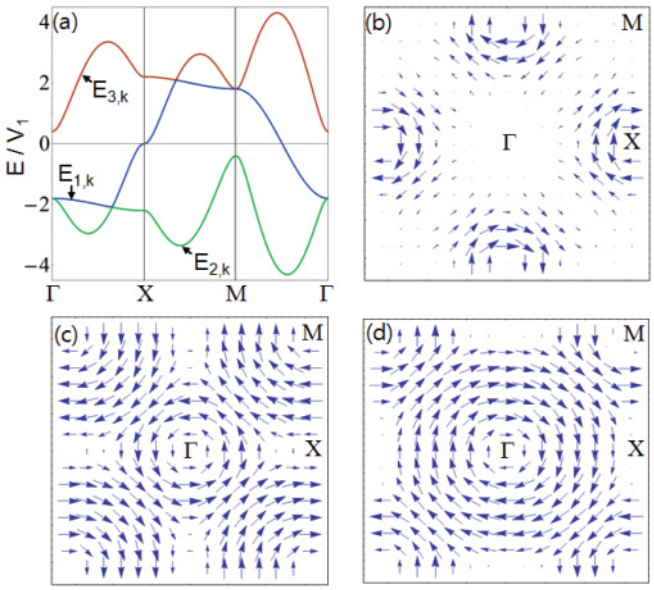

FIG. 1. (Color online) (a) Band structure obtained from the Hamiltonian (3.11) with the parameter $(V_1, V_2, \gamma, \alpha) = (1, 0.1, 1, 0)$. (b)–(d) Calculated OAM in the Brillouin zone $[-\pi,\pi] \times [-\pi,\pi]$ for each band. $E_{1,k}$ through $E_{3,k}$ denoting the different bands are indicated with arrows in (a) and their corresponding OAM are shown in (b) through (d), respectively. The total OAM summed over all bands is zero for each $\mathbf{k}$ in the Brillouin zone.

As the above perturbative and numerical results indicate, the presence of OAM is tied to nonzero γ, or ISB. The amount of OAM grows with γ, as shown in Eq. (3.15). This is in contrast to the spin Rashba model which displays perfect spin polarization irrespective of the strength of the Rashba parameter. The difference between the continuous growth of chiral OAM with the ISB parameter γ and the sudden emergence of chiral SAM can be understood as follows. Even the non-Rashba-split bands already have perfectly spin-polarized quasiparticle state. It is the energy degeneracy of the oppositely oriented spin states which mask their spin-polarized nature. Now, the ISB removes the energy degeneracy in such a way that spin orientations of opposite chiralities occur at different energies. In the spin-Rashba case it is the energy difference of spin-split bands that scales with γ. On the contrary, chiral OAM is endowed upon nondegenerate bands, such as $E_2(\mathbf{k})$ and $E_3(\mathbf{k})$ in Eq. (3.14). The initial OAM value for each band is strictly zero for inversion symmetric case. Upon imposing $\gamma \neq 0$, OAM grows in proportion to it.

Let us insert SOI $H_{\rm so} = (\alpha/2)\sum_i \boldsymbol{L}_i \cdot \boldsymbol{\sigma}_i$ as a perturbation now. For each band obtained above, $H_{\rm so}$ couples states of the same momentum $\mathbf{k}$ but different spins (we ignore interband coupling induced by $H_{\rm so}$). The matrix elements arising from $H_{\rm so}$ are

$$\begin{aligned}
&\langle \mathbf{k},1;\sigma'|H_{\rm so}|\mathbf{k},1;\sigma\rangle = 0,\\
&\langle \mathbf{k},2;\sigma'|H_{\rm so}|\mathbf{k},2;\sigma\rangle = -\frac{3\alpha\gamma}{2\Delta}\begin{pmatrix} 0 & k_y+ik_x \\ k_y-ik_x & 0\end{pmatrix},\\
&\langle \mathbf{k},3;\sigma'|H_{\rm so}|\mathbf{k},3;\sigma\rangle = +\frac{3\alpha\gamma}{2\Delta}\begin{pmatrix} 0 & k_y+ik_x \\ k_y-ik_x & 0\end{pmatrix}.
\end{aligned} \tag{3.16}$$

One recovers exactly the Rashba-type spin splitting as induced by SOI. Combined with the previous conclusions on OAM, we arrive at the following picture. The OAM-carrying band has, say, a CCW (counterclockwise) sense of OAM near the Γ point. The two spin-split bands in turn have one CCW and one CW (clockwise) rotation of spins. Each band is therefore characterized by a pair of rotations for OAM and SAM, which reads (CCW, CCW) and (CCW, CW), respectively. For the other OAM-carrying band the assignments would be (CW, CCW) and (CW, CW). For the non-OAM-carrying band there is no splitting of levels by $H_{\rm so}$ and hence no spin-Rashba splitting. This calculation thus points out that the presence of OAM is a prerequisite in the further splitting of spin degeneracy induced by SOI. To obtain Rashba spin splitting in the orbital-quenched band (such as graphene) is still possible, but its magnitude will have to be derived from a truly relativistic argument which, as we saw in the Introduction, gives a prohibitively small number.

C. Triangular lattice

Several elements with partially filled p orbitals showing the Rashba phenomena have the fcc crystal structure and have the triangular-lattice surface along the [111] direction.[2,6,14] We show that the previous chiral OAM structure found in the square lattice also exists in the triangular lattice. In the same basis used to define the square lattice model $C_{\mathbf{k}}$ [Eq. (3.10)], we find for the triangular lattice the small-$\mathbf{k}$ Hamiltonian

$$H_{\mathbf{k}} = \begin{pmatrix} \alpha k_x^2+\beta k_y^2 & (\alpha-\beta)k_xk_y & -i\frac{3}{2}\gamma k_x \\ (\alpha-\beta)k_xk_y & \alpha k_y^2+\beta k_x^2 & -i\frac{3}{2}\gamma k_y \\ i\frac{3}{2}\gamma k_x & i\frac{3}{2}\gamma k_y & 4(\alpha-\beta)-\frac{3}{2}V_2k^2 \end{pmatrix}. \tag{3.17}$$

Here, $\alpha = 3(3V_1 - V_2)/8$, $\beta = 3(V_1 - 3V_2)/8$, and $k^2 = k_x^2 + k_y^2$. The lattice constant is taken to be unity. To diagonalize $H_{\mathbf{k}}$, as before, it is convenient to choose a new set of basis vectors,

$$\begin{aligned}
|\mathrm{I}\rangle &= (k_y/k)|p_x\rangle - (k_x/k)|p_y\rangle,\\
|\mathrm{II}\rangle &= (k_x/k)|p_x\rangle + (k_y/k)|p_y\rangle,\\
|\mathrm{III}\rangle &= e^{-i\phi_{\mathbf{k}}}|p_z\rangle,
\end{aligned} \tag{3.18}$$

$k = |\mathbf{k}|$, $e^{i\phi_{\mathbf{k}}} = (k_x + ik_y)/k$. The state $|\mathrm{I}\rangle$ remains decoupled at energy $E_{1,\mathbf{k}} = 3V_2 - 3V_1 + 3(V_1 - 3V_2)\mathbf{k}^2/8$, while $|\mathrm{II}\rangle$ and $|\mathrm{III}\rangle$ combine to form eigenstates which, to leading order of γ/Δ, $\Delta = V_1 + V_2$, are

$$\begin{aligned}
|\mathbf{k},2\rangle &\simeq |\mathrm{II}\rangle - \frac{i\gamma(k_x - ik_y)}{2\Delta}|\mathrm{III}\rangle,\\
|\mathbf{k},3\rangle &\simeq |\mathrm{III}\rangle - \frac{i\gamma(k_x + ik_y)}{2\Delta}|\mathrm{II}\rangle,
\end{aligned} \tag{3.19}$$

with energies

$$\begin{aligned}
E_{2,\mathbf{k}} &\simeq 3(V_2 - V_1) + 3[(3V_1 - V_2)/8 - \gamma^2/4\Delta]k^2,\\
E_{3,\mathbf{k}} &\simeq 6V_2 + (3\gamma^2/4\Delta - 3V_2/2)k^2.
\end{aligned} \tag{3.20}$$

As in the square lattice case, two of the bands obtained above carry nonzero, chiral OAM:

$$\langle \mathbf{k},2|L^+|\mathbf{k},2\rangle = \frac{i\gamma}{\Delta}(k_x + ik_y) = -\langle \mathbf{k},3|L^+|\mathbf{k},3\rangle. \tag{3.21}$$

The two OAM-carrying states $|\mathrm{II},\mathbf{k}\rangle$ and $|\mathrm{III},\mathbf{k}\rangle$ obey a reduced 2×2 Hamiltonian,

$$H_{\rm OAM} = -4\beta I_{2\times 2} + M^{-1}\mathbf{k}^2 + \tfrac{3}{2}\hat{z}\cdot(\gamma\mathbf{k}\times\boldsymbol{\tau} - \Delta\tau), \tag{3.22}$$

with

$$M^{-1} = \begin{pmatrix} \alpha & 0 \\ 0 & -3V_2/2 \end{pmatrix}$$

the effective mass tensor and $\boldsymbol{\tau}$ the pseudospin matrix.

D. Comparison to LDA

A check on the existence of chiral OAM is performed by employing the first-principles local-density approximation (LDA) calculation for a Bi single layer forming a triangular lattice. The choice is inspired by Bi being a prototypical p-orbital band material. An external electric field of 3 V/Å perpendicular to the layer was imposed by hand to mimic the surface potential gradient without having the complication of dealing with the bulk states. We also performed calculations for the physically more realistic case of a Bi bilayer[14] with a perpendicular electric field, with results that are entirely in accord with the statements made below for the single-layer case regarding the emergence of chiral OAM. For the density-functional theory (DFT) calculations within the LDA, we used the DFT code, OPENMX,[15] based on the linear-combination-of-pseudo-atomic-orbitals (LCPAO) method.[16]

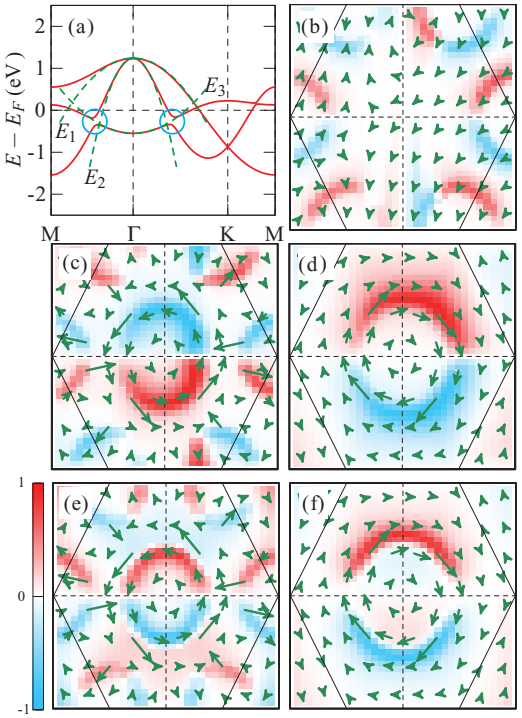

FIG. 2. (Color online) OAM and CD from first-principles and tight-binding calculations of Bi monolayer without SOI. (a) LDA band structure for Bi monolayer with SOI turned off. A perpendicular electric field of 3 V/Å was imposed externally. Three dashed curves represent the tight-binding energy dispersions around the Γ point. (b)–(d) OAM vectors (green arrows) and NCD signals (color backgrounds) for the three bands, E_1 (b), E_2 (c), and E_3 (d), over the whole Brillouin zone are marked by solid hexagons. The largest OAM has a magnitude $\approx 1\hbar$ for bands E_2 and E_3. NCD is calculated with $k_{F,z} = 2.27$ Å^{-1}. (e),(f) NCD calculated with $k_{F,z} = 0$ in Eq. (4.8) for E_2 (e) and E_3 (f) bands. The opposite color assignments between (c) and (e) is a consequence of photon energy dependence of the scattering intensity.

LCPAO coefficients at the specific k points were used to calculate the OAM. To emphasize the relevance of ISB we again chose to investigate the spin-degenerate case by turning off SOI in the LDA calculation. The resulting electronic structure for a spinless case consisting of three p-orbital-derived bands is shown in Fig. 2(a). As the external electric field is turned on, a level repulsion between the middle [E_2 in Fig. 2(a)] and the bottom [E_3 in Fig. 1(a)] bands occurs as indicated by circles in Fig. 2(a). These two bands exhibit the chiral OAM patterns with the maximum OAM vector $|\langle \boldsymbol{L} \rangle| \approx 0.96\hbar$, as shown in Figs. 2(c) and 2(d), while the third one, shown in Fig. 2(b), carries much less OAM around the Γ point. The OAM chiralities of the two bands are opposite, in accordance with the previous TB analysis. An excellent fit of the LDA band structure near the Γ point was possible with the TB parameters $V_1 = -0.725$ eV, $V_2 = -0.11$ eV, and $\gamma = 0.2623$ eV [Fig. 2(a)]. The OAM magnitude is seen to decrease continuously upon approaching the Γ point in the LDA calculation [Fig. 2(c) and 2(d)], as predicted by the TB calculation [Eq. (3.15)]. It reaches a maximum value around the $\mathbf{k}$ points, where the level repulsion is the greatest [blue circles in Fig. 2(a)].

IV. CIRCULAR DICHROISM

In the previous sections, emergence of OAM was predicted to be a general occurrence in surface bands due to the lack of inversion symmetry. In this section we argue that such OAM structure can be probed readily with the current ARPES technique, provided one uses two opposite circular polarizations of incoming lights to measure the photoelectron intensity.

Circular dichroism (CD) refers to phenomena in which the physical response of a system to probing light depends systematically on the light polarization being LCP or RCP.[17–19] In ARPES, CD manifests itself as different scattering intensities of the photoelectrons depending on the helicity of incident light being RCP or LCP. The CD-ARPES signal can be quantified through the normalized CD (NCD) defined as

$$D(\mathbf{k}) = \frac{\sum_\sigma \left[I_\sigma^{\mathrm{RCP}}(\mathbf{k}) - I_\sigma^{\mathrm{LCP}}(\mathbf{k}) \right]}{\sum_\sigma \left[I_\sigma^{\mathrm{RCP}}(\mathbf{k}) + I_\sigma^{\mathrm{LCP}}(\mathbf{k}) \right]}. \tag{4.1}$$

The sum over the final state spin σ reflects the spin-integrated nature of the detection scheme. The spin index is restored in the following derivation of the NCD formula.

The initial state $|\mathrm{I}\rangle$ is the Bloch state of momentum $\mathbf{k}$ constructed as $|\mathbf{k},\boldsymbol{m}\rangle = N^{-1/2} \sum_i e^{i\mathbf{k}\cdot\mathbf{r}_i} |i,\boldsymbol{m}\rangle$. In turn, the Wannier state $|i,\boldsymbol{m}\rangle$ at site $\mathbf{r}_i$ is given by

$$|i,\boldsymbol{m}\rangle = \sum_{\lambda,\sigma} m_{\lambda,\sigma}(\mathbf{k}) |i,\lambda,\sigma\rangle, \tag{4.2}$$

as a linear combination of constituent atomic orbitals labeled by λ, and spin σ, with $\mathbf{k}$-dependent coefficients $m_{\lambda,\sigma}(\mathbf{k})$. Plane-wave forms are assumed for the final state[20]: $\psi_F \sim e^{i\mathbf{k}_F\cdot\mathbf{r}}$.

The transition amplitude into the final state of spin σ is evaluated as

$$\langle \mathrm{F},\sigma | \mathbf{p}\cdot\mathbf{A} | \mathrm{I} \rangle \sim \langle \mathrm{F},\sigma | \mathbf{r}\cdot\mathbf{A} | \mathrm{I} \rangle \sim \sum_{i,\lambda} m_{\lambda,\sigma}(\mathbf{k}) \langle \mathrm{F} | \mathbf{r}\cdot\mathbf{A} | i,\lambda \rangle e^{i\mathbf{k}\cdot\mathbf{r}_i}. \tag{4.3}$$

Given the localized nature of the Wannier state, it is useful to rewrite $\mathbf{r}\cdot\mathbf{A} = (\mathbf{r}-\mathbf{r}_i)\cdot\mathbf{A} + \mathbf{r}_i\cdot\mathbf{A}$. One immediately finds that $\langle \mathrm{F}|\mathbf{r}_i\cdot\mathbf{A}|\mathrm{I}\rangle = (\mathbf{r}_i\cdot\mathbf{A})\langle \mathrm{F}|\mathrm{I}\rangle$ is zero from the presumed orthogonality of the initial and the final states. To proceed further, we treat the cases of p-orbital and d-orbital bands separately as they require somewhat different strategies for evaluation of the NCD formula. For p orbitals the Wannier states are assigned the hydrogenic wave function $\langle \mathbf{r}|i,\lambda\rangle \sim (\mathbf{r}-\mathbf{r}_i)_\lambda f(|\mathbf{r}-\mathbf{r}_i|)$ ($\lambda = x,y,z$), and with these we find that the transition amplitude becomes

$$\begin{aligned} &\langle \mathrm{F},\sigma | \mathbf{r}\cdot\mathbf{A} | \mathrm{I} \rangle \\ &\quad \propto \sum_i e^{i(\mathbf{k}-\mathbf{k}_F)\cdot\mathbf{r}_i} \left(\sum_\lambda m_{\lambda,\sigma}(\mathbf{k}) \int e^{-i\mathbf{k}_F\cdot\mathbf{r}} [\mathbf{r}\cdot\mathbf{A}] x_\lambda f(|\mathbf{r}|) \right) \\ &\quad \propto \sum_\lambda m_{\lambda,\sigma}(\mathbf{k}) \int e^{-i\mathbf{k}_F\cdot\mathbf{r}} [\mathbf{r}\cdot\mathbf{A}]\, x_\lambda f(|\mathbf{r}|). \end{aligned} \tag{4.4}$$

The sum $\sum_i e^{i(\mathbf{k}-\mathbf{k}_F)\cdot\mathbf{r}_i} = \delta(\mathbf{k}_F^{\parallel} - \mathbf{k} + \mathbf{G})$ simply yields in-plane momentum conservation up to the reciprocal lattice

vector $\mathbf{G}$ and is dropped in the last line. Making use of the fact that $f(|\mathbf{r}|)$ depends only on the radial distance, one can rewrite the second line of Eq. (4.4) as

$$-A_\alpha m_{\beta,\sigma}\partial_\alpha\partial_\beta\int e^{-i\mathbf{k}_F\cdot\mathbf{r}}f(|\mathbf{r}|)d^3\mathbf{r}=-\mathbf{A}\cdot\nabla_{\mathbf{k}_F}g_\sigma(\mathbf{k}_F), \quad (4.5)$$

where we have introduced

$$\begin{aligned} f(|\mathbf{k}_F|)&=\int e^{-i\mathbf{k}_F\cdot\mathbf{r}}f(|\mathbf{r}|)d^3\mathbf{r},\\ g_\sigma(\mathbf{k}_F)&=\boldsymbol{m}_\sigma\cdot\nabla_{\mathbf{k}_F}f(|\mathbf{k}_F|),\end{aligned} \quad (4.6)$$

and $\boldsymbol{m}_\sigma=(m_{x,\sigma},m_{y,\sigma},m_{z,\sigma})$. The set of coefficients $\boldsymbol{m}_\sigma$ introduced in the p-orbital Bloch state can be used to calculate its OAM $\langle\boldsymbol{L}\rangle$. A simple result is obtained:

$$\langle\mathbf{k},\boldsymbol{m}|\boldsymbol{L}|\mathbf{k},\boldsymbol{m}\rangle\equiv\langle\boldsymbol{L}\rangle=i\sum_\sigma\boldsymbol{m}_\sigma(\mathbf{k})\times\boldsymbol{m}_\sigma^*(\mathbf{k}). \quad (4.7)$$

Of great importance is the fact that nonzero OAM is possible even with spin degeneracy, $\boldsymbol{m}_\sigma=\boldsymbol{m}$, where one would obtain $\langle\boldsymbol{L}\rangle=2i\boldsymbol{m}\times\boldsymbol{m}^*$. It is commonly perceived that complex coefficients $\boldsymbol{m}_\sigma\neq\boldsymbol{m}_\sigma^*$ are associated with the SOI. Our theory of the previous sections shows, on the other hand, that only ISB is required to generate complex coefficients in forming the Bloch state, which then gives rise to nonzero OAM.

We can now proceed to show that NCD is fundamentally related to $\langle\boldsymbol{L}\rangle$. Equations (4.4) and (4.6) allow the NCD formula (4.1) to be recast in the compact, suggestive form:

$$D(\mathbf{k})=\frac{(\mathbf{A}\times\mathbf{A}^*)\cdot\sum_\sigma\nabla g_\sigma\times\nabla g_\sigma^*}{\sum_\sigma[(\mathbf{A}\cdot\nabla g_\sigma)(\mathbf{A}^*\cdot\nabla g_\sigma^*)+(\mathbf{A}^*\cdot\nabla g_\sigma)(\mathbf{A}\cdot\nabla g_\sigma^*)]}. \quad (4.8)$$

Gradients in the above are with respect to $\mathbf{k}_F$. The vector potentials are $\mathbf{A}=(\boldsymbol{\varepsilon}_1+i\boldsymbol{\varepsilon}_2)/\sqrt{2}$ for RCP and $\mathbf{A}^*$ for LCP, with $\mathbf{A}\times\mathbf{A}^*=-i\boldsymbol{\varepsilon}_1\times\boldsymbol{\varepsilon}_2=-i\hat{k}_{\rm ph}$ giving the incident photon direction. The remaining task is to evaluate the quantity $\sum_\sigma\nabla g_\sigma\times\nabla g_\sigma^*$ which governs the CD response of the given initial state. The denominator, by definition, is always positive and plays a minor role in characterizing the OAM structure.

For p orbitals, inserting $g_\sigma=\boldsymbol{m}_\sigma\cdot\nabla f$ gives

$$\begin{aligned}&\sum_\sigma\nabla g_\sigma\times\nabla g_\sigma^*\\ &\quad=\frac{1}{2}\varepsilon^{\alpha\beta\gamma}\left(\sum_\sigma\boldsymbol{m}_\sigma\times\boldsymbol{m}_\sigma^*\right)_\alpha\nabla\partial_\beta f\times\nabla\partial_\gamma f\\ &\quad=-\frac{i}{2}\varepsilon^{\alpha\beta\gamma}\langle L_\alpha\rangle\nabla\partial_\beta f\times\nabla\partial_\gamma f,\end{aligned} \quad (4.9)$$

where Eq. (4.7) has been adopted in the final line. Clearly, this is proportional to the components of OAM in the initial state. The remaining "form factor" $\nabla\partial_\beta f\times\nabla\partial_\gamma f$ depends on $\mathbf{k}_F$, which in turn depends on the incoming photon energy. Both $\langle\boldsymbol{L}\rangle$ and the form factors can be obtained by faithful LDA calculations of the wave functions. Analytically, a reasonable choice of $f(r)\sim e^{-r/a}$ would yield $f(k_F)\sim 1/(1+k_F^2a^2)^2$ and all the form factors can be worked out. The central feature of our NCD formula, of course, is its proportionality to OAM carried by the initial state.

Next we turn to the case of degenerate d-orbital bands. Now $g_\sigma=(\boldsymbol{m}_\sigma\cdot\mathbf{D})f$ involves the inner product between the five-dimensional coefficients $\boldsymbol{m}_\sigma$ and the corresponding differential operators D_α matching the given orbital basis. Explicitly, they are $D_{xy}=\partial_x\partial_y$ and its two permutations, $D_{3z^2-r^2}=(2\partial_z^2-\partial_x^2-\partial_y^2)/2\sqrt{3}$ and $D_{x^2-y^2}=(\partial_{x^2}-\partial_{y^2})/2$. Although the dichroism formula (4.8) in its general form still applies to the d-orbital case, we are no longer able to transform its numerator to a simple shape like Eq. (4.9).

At this point, however, recall that in realistic ARPES experiment the incident photons carry energies of several tens of eV, delivering, however, at most an eV of energy to the occupied electrons.[20] After subtracting what amounts to the surface potential energy barrier, there is still a lot of energy imparted to the final photoelectron, whose energy is typically in excess of 10 eV. Due to in-plane momentum conservation (ignoring higher-order scattering), the in-plane component of photoelectron momentum can carry only a small fraction of this energy, which means most of the kinetic energy is contained in the z component, $(k_{F,z})^2/2m$. As a result, the typical situation in ARPES experiment is the one in which $k_{F,z}$ dominates over the planar components, and $k_{F,z}a$ is rather larger than unity, a being the typical radius of the orbital wave function. In computing $g_\sigma=(\boldsymbol{m}_\sigma\cdot\mathbf{D})f$, therefore, the orbitals containing at least one power of z will be dominant over those that contain none, due to extra powers of $k_{F,z}a$ produced by the differentiation. Such reasoning reduces the number of relevant orbitals from five to three, that is, d_{zx}, d_{yz}, and $d_{3z^2-r^2}$. This is irrespective of particular crystal-field symmetries of the d orbitals and is rather dictated by the experimental conditions of ARPES. We also introduce the notion of "partial OAM," which is obtained by assuming only the three, labeled $1\equiv zx$, $2\equiv yz$, and $3\equiv 3z^2-r^2$, out of the five d orbitals contribute to the wave function:

$$\begin{aligned}\langle L_x\rangle'&=\sqrt{3}i\sum_\sigma(m_{yz,\sigma}m_{3z^2-r^2,\sigma}^*-m_{yz,\sigma}^*m_{3z^2-r^2,\sigma}),\\ \langle L_y\rangle'&=\sqrt{3}i\sum_\sigma(m_{3z^2-r^2,\sigma}m_{zx,\sigma}^*-m_{3z^2-r^2,\sigma}^*m_{zx,\sigma}),\\ \langle L_z\rangle'&=i\sum_\sigma(m_{zx,\sigma}m_{yz,\sigma}^*-m_{zx,\sigma}^*m_{yz,\sigma}).\end{aligned} \quad (4.10)$$

The prime is a reminder that contributions to OAM from d_{xy} and $d_{x^2-y^2}$ are being ignored. For completeness we give the full expression for OAM in the d-orbital band which reads

$$\begin{aligned}\langle L_x\rangle&=\langle L_x\rangle'+i\sum_\sigma[m_{xy,\sigma}m_{zx,\sigma}^*-m_{xy,\sigma}^*m_{zx,\sigma}\\ &\quad+m_{yz,\sigma}m_{x^2-y^2,\sigma}^*-m_{yz,\sigma}^*m_{x^2-y^2,\sigma}],\\ \langle L_y\rangle&=\langle L_y\rangle'+i\sum_\sigma[m_{yz,\sigma}m_{xy,\sigma}^*-m_{yz,\sigma}^*m_{xy,\sigma}\\ &\quad+m_{zx,\sigma}m_{x^2-y^2,\sigma}^*-m_{zx,\sigma}^*m_{x^2-y^2,\sigma}],\\ \langle L_z\rangle&=\langle L_z\rangle'+2i\sum_\sigma[m_{x^2-y^2,\sigma}m_{xy,\sigma}^*-m_{x^2-y^2,\sigma}^*m_{xy,\sigma}].\end{aligned} \quad (4.11)$$

Following the same calculation procedure as in the p-orbital case, we find the CD formula for d orbitals valid for $k_{F,z}a\gg 1$,

assuming $f(r) \sim e^{-r/a}$,

$$D_d(\mathbf{k}) = \frac{\hat{k}_{\rm ph} \cdot (2\langle L_x\rangle' \hat{x} + 2\langle L_y\rangle' \hat{y} - \langle L_z\rangle' \hat{z})}{\sum_\sigma [A_x m_{1,\sigma} + A_y m_{2,\sigma} - 2\sqrt{3} A_z m_{3,\sigma}][(A_x)^* m^*_{1,\sigma} + (A_y)^* m^*_{2,\sigma} - 2\sqrt{3}(A_z)^* m^*_{3,\sigma}] + (\mathbf{A} \leftrightarrow \mathbf{A}^*)}. \quad (4.12)$$

On a technical note, the form factors are expanded in powers of $(k_F a)^{-2}$, so that even a moderate factor like $k_F a \sim 2$ is adequate to justify the large-$k_F a$ formula above. The OAM shown in the numerator refers to those in partial OAM. Again, nontrivial CD-ARPES signal is predicated on the existence of OAM. In the same limit, $k_{F,z} a \gg 1$, the p-orbital formula simplifies to

$$D_p(\mathbf{k}) = \frac{5\hat{k}_{\rm ph} \cdot \langle \boldsymbol{L}\rangle - 6(\hat{k}_{\rm ph} \cdot \hat{k}_F)(\hat{k}_F \cdot \langle \boldsymbol{L}\rangle)}{\sum_\sigma [6(\hat{k}_F \cdot \mathbf{A})(\hat{k}_F \cdot \boldsymbol{m}_\sigma) - (\boldsymbol{m}_\sigma \cdot \mathbf{A})][6(\hat{k}_F \cdot \mathbf{A}^*)(\hat{k}_F \cdot \boldsymbol{m}^*_\sigma) - (\boldsymbol{m}^*_\sigma \cdot \mathbf{A}^*)] + (\mathbf{A} \leftrightarrow \mathbf{A}^*)}. \quad (4.13)$$

The case of t_{2g} bands involving the three orbitals, $\boldsymbol{m}_\sigma = (m_{yz,\sigma}, m_{zx,\sigma}, m_{xy,\sigma})$, can be worked out as well. Referring to the general OAM expression in Eq. (4.11), OAM in the t_{2g} band reads $\langle \mathbf{L}\rangle = -i \sum_\sigma \boldsymbol{m}_\sigma \times \boldsymbol{m}^*_\sigma$. [Note the opposite sign compared to the p orbital result in Eq. (4.7).] The factor $\sum_\sigma \nabla g_\sigma \times \nabla g^*_\sigma$ in Eq. (4.8) becomes

$$\frac{1}{2}\varepsilon^{\alpha\beta\gamma}\left(\sum_\sigma \boldsymbol{m}_\sigma \times \boldsymbol{m}^*_\sigma\right)_\alpha \nabla D_\beta f \times \nabla D_\gamma f = \frac{i}{2}\varepsilon^{\alpha\beta\gamma}\langle L_\alpha\rangle \nabla D_\beta f \times \nabla D_\gamma f, \quad (4.14)$$

where $\mathbf{D} = (D_{yz}, D_{zx}, D_{xy})$. Again, nontrivial OAM is responsible for NCD for t_{2g} bands. The actual task of evaluating the form factors leads to cumbersome expressions which we omit. For practical applications, it is better to evaluate form factors such as $\nabla D_\beta f \times \nabla D_\gamma f$ in the above numerically using LDA-obtained wave functions and their Fourier transforms.

We illustrate how the various formulas derived in this section can be adopted in the LDA calculation to yield NCD for realistic materials. The OAM of the Bi bands shown in Fig. 2(a), for instance, can be used to obtain $D(\mathbf{k})$ according to Eqs. (4.8) and (4.9). Employing the exponential function $f(r) \sim e^{-r/a}$ with the radius $a = 1.6$ Å to evaluate the form factors, Figs. 2(b)–2(d) show $D(\mathbf{k})$ overlaid with local OAM when $k_F = 2.27$ Å^{-1} is used ($k_F a \sim 3.6$). In-plane component of $\mathbf{k}_F$ is chosen to match the quasiparticle momentum $\mathbf{k}$. The z component of $\mathbf{k}_F$ is then uniquely fixed. The next set of figures in (e) and (f) is obtained with $\mathbf{k}_F = 0$ Å^{-1} ($k_F a = 0$) for E_2 and E_3 bands. Incident photon direction is chosen $\hat{k}_{\rm ph} = (\cos 60^\circ, 0, -\sin 60^\circ)$ in both sets. The photon energy dependence of the NCD is obvious by comparing the two sets in Fig. 2. LDA calculation for light-element d-orbital bands of Cu are shown in Fig. 3. In this calculation the ISB electric field is generated spontaneously through self-consistent electronic structure calculation and not imposed externally as in the Bi monolayer case. Despite these differences we found a clear, chiral OAM pattern around the Γ point as shown in Fig. 3(b). Even though the makeup of the surface bands receives substantial contributions from both p and d orbitals, OAM arises predominantly from the d-orbital components; hence, we used the d-orbital formula in Eq. (4.12) to obtain 3(b). A recent CD-ARPES experiment on a Cu surface is consistent with our prediction,[21] where a similar CD result on a Au surface is also reported. We comment that a potential modification of the NCD formula may be necessary for heavy elements due to significant SOI in the Hamiltonian and the consequent modification of the current (momentum) operator coupling to $\mathbf{A}$ (Ref. 22). Our present results are, however, free from such complications as we exclusively focus on spin-degenerate band structures. Nontrivial OAM and NCD features are only due to the degenerate multiple-orbital nature of the eigenstates.

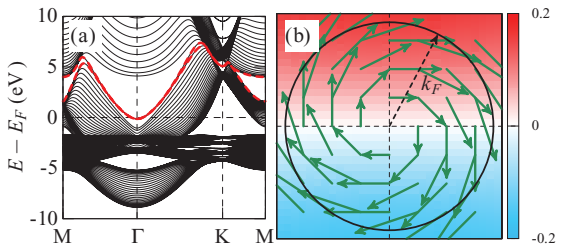

FIG. 3. (Color online) (a) First-principles electronic band structure of Cu(111) 30-layer slab. Red dashed lines indicate the two surface bands. (b) Calculated OAM and CD corresponding to outer energy surface bands. Inner surface band shows basically the same OAM and CD patterns. Maximum OAM vector is ~0.07$\hbar$.

V. SUMMARY

The general notion of "orbital Rashba effect" is presented in this paper. As a direct counterpart of the well-known spin Rashba effect, the orbital version arises in the presence of ISB as commonly occurs in typical surface bands and, as in the recent case of BiTeI,[23] in the bulk band of noncentrosymmetric crystals as well. Careful arguments are given as to why, in reality, the orbital Rashba effect is the more generic phenomenon associated with ISB and how the spin Rashba splitting follows on top of the pre-existing chiral orbital structure as a consequence of added SOI. We have further presented some formulas for CD-ARPES experiment which clearly indicate that CD-ARPES is but a direct probe of the existence of OAM, much as the spin-ARPES is the direct probe of spin polarization in the band structure. We believe both the general ideas of chiral orbital Rashba effect and its detection scheme through CD-ARPES is a widespread feature of metallic surfaces, interfaces, and non-centrosymmetric bulk bands. A related emergence of chiral OAM in a time-reversal symmetry-broken magnetic metallic surface is being investigated.

Before closing, we note the past extensive research in magnetic CD, as well as magnetic linear dichroism phenomena on magnetic surfaces.[19,24] The CD phenomena and its observation through CD-ARPES as discussed in this paper is a consequence of inversion symmetry alone, while for magnetic surfaces time-reversal symmetry is also violated. A related paper[25] pointed out the spin polarization of the photoelectrons coming off the nonmagnetic surface by the polarized incident photon. We, on the other hand, discuss the spin-integrated scattering intensity coming off the nonmagnetic surface. It is to be emphasized that our findings are unrelated to spin polarization effects.

ACKNOWLEDGMENT

H. J. H. is supported by NRF Grant No. 2011-0015631. Useful discussions with Changyoung Kim and Jaejun Yu are acknowledged.

*hanjh@skku.edu

[1] Y. A. Bychkov and E. I. Rashba, JETP Lett. **39**, 78 (1984).

[2] S. LaShell, B. A. McDougall, and E. Jensen, Phys. Rev. Lett. **77**, 3419 (1996).

[3] F. Reinert, G. Nicolay, S. Schmidt, D. Ehm, and S. Hüfner, Phys. Rev. B **63**, 115415 (2001).

[4] G. Nicolay, F. Reinert, S. Hüfner, and P. Blaha, Phys. Rev. B **65**, 033407 (2001).

[5] M. Hoesch, M. Muntwiler, V. N. Petrov, M. Hengsberger, L. Patthey, M. Shi, M. Falub, T. Greber, and J. Osterwalder, Phys. Rev. B **69**, 241401(R) (2004).

[6] Christian R. Ast and Hartmut Höchst, Phys. Rev. Lett. **87**, 177602 (2001); Yu. M. Koroteev, G. Bihlmayer, J. E. Gayone, E. V. Chulkov, S. Blügel, P. M. Echenique, and Ph. Hofmann, *ibid.* **93**, 046403 (2004); T. Hirahara, K. Miyamoto, I. Matsuda, T. Kadono, A. Kimura, T. Nagao, G. Bihlmayer, E. V. Chulkov, S. Qiao, K. Shimada, H. Namatame, M. Taniguchi, and S. Hasegawa, Phys. Rev. B **76**, 153305 (2007).

[7] K. Sugawara, T. Sato, S. Souma, T. Takahashi, M. Arai, and T. Sasaki, Phys. Rev. Lett. **96**, 046411 (2006).

[8] D. Pacile, C. R. Ast, M. Papagno, C. DaSilva, L. Moreschini, M. Falub, A. P. Seitsonen, and M. Grioni, Phys. Rev. B **73**, 245429 (2006); Christian R. Ast, J. Henk, A. Ernst, L. Moreschini, M. C. Falub, D. Pacile, P. Bruno, K. Kern, and M. Grioni, Phys. Rev. Lett. **98**, 186807 (2007); Christian R. Ast *et al.*, Phys. Rev. B **77**, 081407(R) (2008); Fabian Meier, V. Petrov, S. Guerrero, C. Mudry, L. Patthey, J. Osterwalder, and J. H. Dil, *ibid.* **79**, 241408(R) (2009).

[9] G. Bihlmayer, S. Blügel, and E. V. Chulkov, Phys. Rev. B **75**, 195414 (2007); Miki Nagano *et al.*, J. Phys.: Condens. Matter **21**, 064239 (2009); Tamio Oguchi and Tatsuya Shishidou, *ibid.* **21**, 092001 (2009); Koichiro Yaji *et al.*, Nat. Commun. **1**, 17 (2010); E. Frantzeskakis, S. Pons, and M. Grioni, Phys. Rev. B **82**, 085440 (2010).

[10] S. R. Park, C. H. Kim, J. Yu, J. H. Han, and C. Kim, Phys. Rev. Lett. **107**, 156803 (2011).

[11] J. H. Dil, J. Phys.: Condens. Matter **21**, 403001 (2009).

[12] A. Kimura *et al.*, Phys. Rev. Lett. **105**, 076804 (2010).

[13] L. Petersen and P. Hedegård, Surf. Sci. **459**, 49 (2000).

[14] Y. Liu and R. E. Allen, Phys. Rev. B **52**, 1566 (1995).

[15] The DFT code, OPENMX, is available at [http://www.openmx-square.org] in the constitution of the GNU GeneralPublic License.

[16] T. Ozaki, Phys. Rev. B **67**, 155108 (2003).

[17] B. Ritchie, Phys. Rev. A **13**, 1411 (1976).

[18] R. L. Dubs, S. N. Dixit, and V. McKoy, Phys. Rev. Lett. **54**, 1249 (1985).

[19] B. T. Thole and G. van der Laan, Phys. Rev. B **44**, 12424 (1991).

[20] A. Damascelli, Z. Hussain, and Z.-X. Shen, Rev. Mod. Phys. **75**, 473 (2003).

[21] B. Kim *et al.*, Phys. Rev. B **85**, 195402 (2012).

[22] Y. H. Wang, D. Hsieh, D. Pilon, L. Fu, D. R. Gardner, Y. S. Lee, and N. Gedik, Phys. Rev. Lett. **107**, 207602 (2011).

[23] K. Ishizaka *et al.*, Nat. Mater. **10**, 521 (2011).

[24] J. Henk, S. V. Halilov, T. Scheunemann, and R. Feder, Phys. Rev. B **50**, 8130 (1994); S. V. Halilov, J. Henk, T. Scheunemann, and R. Feder, *ibid.* **52**, 14235 (1995).

[25] E. Tamura and R. Feder, Europhys. Lett. **16**, 695 (1991).