Quasiparticle energy bands and Fermi surfaces of monolayer NbSe$_2$

Sejoong Kim$^{1,2}$ and Young-Woo Son$^1$

$^1$Korea Institute for Advanced Study, Hoegiro 85, Seoul 02455, Korea
$^2$University of Science and Technology, Gajeong-ro 217, Daejeon 34113, Korea

(Dated: September 15, 2017)

A quasiparticle band structure of a single layer $2H$-NbSe$_2$ is reported by using first-principles $GW$ calculation. We show that a self-energy correction increases the width of a partially occupied band and alters its Fermi surface shape when comparing those using conventional mean-field calculation methods. Owing to a broken inversion symmetry in the trigonal prismatic single layer structure, the spin-orbit interaction is included and its impact on the Fermi surface and quasiparticle energy bands are discussed. We also calculate the doping dependent static susceptibilities from the band structures obtained by the mean-field calculation as well as $GW$ calculation with and without spin-orbit interactions. A complete tight-binding model is constructed within the three-band third nearest neighbour hoppings and is shown to reproduce our $GW$ quasiparticle energy bands and Fermi surface very well. Considering variations of the Fermi surface shapes depending on self-energy corrections and spin-orbit interactions, we discuss the formations of charge density wave (CDW) with different dielectric environments and their implications on recent controversial experimental results on CDW transition temperatures.

I. INTRODUCTION

After the first experimental success in isolating various single layers from layered materials such as graphite, hexagonal boron nitrides, transition metal dichalcogenides (TMD) and high-temperature cuprate superconductors, there have been tremendous efforts to understand physical properties of two-dimensional (2D) crystals. Notably, owing to the reduced spatial dimension compared to their bulk counterparts, the Coulomb interaction in the 2D crystals as well as screening of substrates on which they are placed have played very important roles in modifying their electronic structures. For example, it is now well established that many-body interactions alter low-energy bands in graphene significantly. Moreover, the substrate screening also changes the nature of Coulomb interactions in graphene on top of either dielectric materials or metals and hence modifies the bands further.

Besides graphene, several recent experiments have revealed new interesting physical properties in a mono- and a few-layer TMDs such as series of phase transitions and novel superconductivity, being different from those shown in their bulk forms. Formation of charge density waves (CDWs) in three-dimensional metallic TMDs has attracted interests for the last couple of decades and origins of CDWs in some materials are still not settled yet. Therefore, the current efforts in investigating physical properties of thin flakes of TMDs may shed light on the origin of CDW phase in three dimensional TMDs and open a way to find the characteristic new collective phenomena in 2D crystals.

Among metallic TMDs, the stacked trigonal prismatic structure of niobium diselenide ($2H$-NbSe$_2$) is one of the most studied materials and an ideal system to study phase transitions as functions of temperature and dopings. It has been known for a long time that three-dimensional stacking structure of $2H$-NbSe$_2$ is metallic at room temperature and undergoes a CDW transition at 33 K before becoming a superconductor at 7.2 K although there has been the controversy regarding on the origin of CDW and the competition between CDW and superconducting (SC) states. After a few earlier attempts to investigate physical properties of its thin flakes, a couple of recent works have reported successful isolations of its single layer form on top of various substrates and measure their CDW and SC phase transitions.

While all experiments hitherto have shown that the superconducting transition temperature decreases but does not diminish completely when the thickness of $2H$-NbSe$_2$ decreases to a single layer limit, the transition temperature ($T_{\text{CDW}}$) from metal to CDW phase differs from each other significantly. The work by Xi et al. measured $T_{\text{CDW}}$ of 145 K, more than four times larger than the bulk $T_{\text{CDW}}$ of 33 K whereas the work by Ugeda et al. reported that $T_{\text{CDW}}$ is similar to or less than that of the bulk. Moreover, the former attribute the strong coupling mechanism to the formation of CDW at the very high temperature while the latter measured a small CDW energy gap of 4 meV together with CDW modulation under high biases pointing to the puzzling dual nature (strong and weak) of CDW formation. It is noticeable that the former measured the transition in a sample on top of silicon substrate while the latter on top of epitaxial bilayer graphene (BLG) grown on $6H$-SiC(0001) surface.

Previous theoretical studies have shown that the bulk CDW phase of $2H$-NbSe$_2$ does not have a prominent sharp peak at the specific CDW wave vector, rather showing a broad peak in the real part of the bare susceptibility and its imaginary part does not peak at the CDW wave vector at all. These suggest a possible strong electron-phonon coupling mechanism to form the CDW phase, ruling out the Fermi surface nesting mechanism, or the saddle-point singularity driven...
CDW phase\textsuperscript{31,52}. A recent first-principles calculation using a semi-local exchange-correlation functional predicts the CDW instability in the single layer structure with an enhancement of the electron-phonon interaction at a specific CDW wave vector differing from that of the bulk\textsuperscript{53}. However, the predicted 4×1 CDW state is inconsistent with the recent measurement showing 3×3 CDW state by Ugeda and coworkers\textsuperscript{51}. Another recent calculation using the similar method also suggests a good metallic behavior for a monolayer undistorted structure\textsuperscript{54} while a transport measurement shows a semimetallic nature\textsuperscript{55}.

Considering significant changes in the low energy bands of the semimetallic and semiconducting 2D crystals with a proper inclusion of electronic self-energy correction as well as its modification by the substrate screening\textsuperscript{31,41,42,43,45}, a plain mean field calculation is not sufficient and it is necessary to investigate effects of suitable corrections from the many-body Coulomb interactions on the low energy bands of monolayer NbSe\textsubscript{2}. Although the CDW formation is quite sensitive to a shape of Fermi surface, its variation with the self energy corrections and alternations by the substrate screening have not been investigated yet fully. In this work, motivated by recent rapid progress in this field and experiments reporting different CDW formations\textsuperscript{54,55}, we report first-principles density functional calculation and $GW$ approximation results of single layer NbSe\textsubscript{2} in its normal metallic state which provides comprehensive pictures of the low energy electronic structures, a prerequisite to understand the controversial CDW features.

This paper is organized as follows. In Sec. III, we introduce a model system and calculation details. In Sec. III, our calculation results for the low energy band structures based on mean-field calculation methods and $GW$ approximation with and without spin-orbit interactions are presented. The doping dependent variations of Fermi surfaces are also discussed within various levels of approximations. The bare susceptibilities with calculated energy band structures are also presented in this section. Conclusions and discussion on CDW formation mechanism in the single layer are in Sec. IV. In addition, the effects of including semi core orbitals, detailed derivation of atomic SOC s in this system, constructions of a tight binding Hamiltonian within the three-band third nearest neighbor hopping between $d$-orbital of niobiums fit for the DFT-GGA and $GW$ energy bands, respectively, and discussions on $GW$ energy bands of the bulk 2H-NbSe\textsubscript{2} are presented in Appendix.

II. SYSTEMS AND CALCULATION DETAILS

Figure 1 displays the crystal structure of a single layer of 2H-NbSe\textsubscript{2}. The single layer is composed of three sub-layers: a triangular lattice of Nb at the center, which is sandwiched by two outermost layers of Se triangular lattice, forming a trigonal prismatic structure. The two Se sub-layers of the 2H structure are mirror symmetric to each other with respect to the plane of the Nb layer as shown in Fig. (a). From a top view, the single layer of 2H-NbSe\textsubscript{2} looks like a hexagonal lattice.

We simulate the single layer of 2H-NbSe\textsubscript{2} by taking a vacuum layer of 12 Å. We increase the vacuum layer up to 20 Å and find no changes in calculation results. Our relaxation calculation gives that the lattice constant of the layer is $a = 3.45$ Å and the distance between Nb and Se sublayers is 1.68 Å, agreeing well with the previous studies\textsuperscript{54,55}. We performed density functional theory (DFT) calculation by adopting the PBE generalized gradient approximation (GGA)\textsuperscript{56} for the exchange-correlation functional and the norm-conserving pseudopotentials\textsuperscript{57} with a non-linear core correction\textsuperscript{57}. Considering a crucial role of core levels in estimating self energy\textsuperscript{58,59}, we treated the semicore $4s^24p^6$ electrons of Nb atoms as valence electrons and discuss its impact on the quasiparticle energy bands in Fig. 1 in Appendix. We used the plane-wave DFT code QUANTUM-ESPRESSO\textsuperscript{60} with a cutoff of 55 Ry, a $40 \times 40 \times 1$ k-point grid, and a smearing temperature $k_B\tau = 0.005$ Ry. Quasiparticle energies were calculated within the level of $G_0W_0$ approximation\textsuperscript{51,52} implemented in the BERKELEYGW code\textsuperscript{61} (hereafter, we will call the quasiparticle bands from $G_0W_0$ approximation as $GW$ energy bands for convenience). The slab truncation scheme was used to treat the Coulomb interaction for the single layer geometry\textsuperscript{62}. We used unoccupied bands up to 5 Ry above the Fermi energy for the dielectric function calculation to achieve the convergence of calculations.

III. RESULTS

A. Band Structure

Figure 2 shows electronic band structures calculated by DFT-GGA (red dashed line) and $GW$ approximation (blue solid line) respectively along symmetric lines of the irreducible Brillouin zone (IBZ). Our DFT-GGA band structure shows a quite good agreement with previous

FIG. 1. (Color online) Crystal structure and Brillouin zone of a single layer of 2H-NbSe\textsubscript{2}. (a) 2H structure of NbSe\textsubscript{2}. A transition metal atom (Nb, blue sphere) are surrounded by six chalcogens (Se, yellow sphere). Three of them lie on the upper plane, and the other three on the lower one. (b) Crystal structure from a top view. Primitive lattice vectors $a_1$ and $a_2$ are denoted. (c) The first Brillouin zone and reciprocal lattice vectors $b_1$ and $b_2$. 

FIG. 2. Electronic band structures calculated by DFT-GGA (red dashed line) and $GW$ approximation (blue solid line) respectively along symmetric lines of the irreducible Brillouin zone (IBZ). Our DFT-GGA band structure shows a quite good agreement with previous...
The shape of unoccupied DFT-GGA band structure and their GW corrections shown in Fig. 2 looks quite similar to those of MoS$_2$, although the semiconducting MoS$_2$ has a significant band gap enhancement from the self-energy corrections$^{10}$. Here, the spectral gap between the partially occupied band and unoccupied bands does not increases with the GW correction but the band width of partially occupied band crossing the $E_F$ is enlarged by 17% with GW approximations.

B. Fermi Surface

The Fermi surface also changes when the GW correction is added to the DFT-GGA result. As shown in Fig. 3, calculated Fermi surfaces using the two methods basically contain two distinct hole pockets at the $K$ and $\Gamma$-points respectively. In the DFT-GGA Fermi surface shown in Fig. 3 (a), a hexagonal hole pocket is at the $\Gamma$-point and a rounded triangular hole pocket at the $K$-point agreeing well with previous calculations$^{18,50,53,54}$. In contrast, the triangular hole pockets in the DFT-GGA band become to be a rounded hexagonal pocket in the GW band as shown in Fig. 3 (b). So, flat sides of triangular pockets facing corners of hexagonal ones in the DFT-GGA Fermi surfaces protrude toward the $\Gamma$ point with GW corrections. Then, the distance between the two band crossing points along the $\Gamma K$ line decreases with GW corrections. This corresponds to the fact that the saddle point of the GW band in the $MK$ line shifts up, as seen in Fig. 2, thus enabling one approach to the saddle point easier with slight hole doping.

C. Noninteracting Susceptibility

In order to investigate the implication of the change in the electronic structure on a formation of the charge density wave, we have calculated the real part of the noninteracting static susceptibility $\chi'_0(\textbf{q})$.$^{56}$$^{58}$

$$\chi'_0(\textbf{q}) = \frac{1}{2} \sum_{n,n'} \sum_{\textbf{k}} \sum_{\textbf{k}'} \frac{f(\epsilon_{nk}) - f(\epsilon_{n'k+q})}{\epsilon_{nk} - \epsilon_{n'k+q}} \left| \langle n\textbf{k}|e^{i\textbf{q}\cdot\textbf{r}}|n'\textbf{k}+\textbf{q} \rangle \right|^2,$$

where $\epsilon_{nk}$ and $|n\textbf{k}\rangle$ are the energy of the $n$-th band at the crystal momentum $\textbf{k}$ and its corresponding Bloch state, and $f(\epsilon_{nk})$ is the Fermi-Dirac distribution function. Other quantum numbers such as spin are implicitly included in $n$. In principle, the matrix element is needed to calculate but it is known that the constant matrix approximation$^{50,53,56,58}$, in which the matrix element is set to be unity, is good for the transition metal dichalcogenide. The $\textbf{k}$-point grid used for the calculation is $300 \times 300 \times 1$.

Figure 4 shows the real part of the bare static susceptibility $\chi'_0(\textbf{q})$ of (a) DFT-GGA and (b) GW calcu-
lations. We have checked that $\chi'_0(q)$ of DFT-GGA has a broad maxima extending approximately from $2/5\Gamma M$ to $4/5\Gamma M$. Our calculation agrees well with other previous studies on the bulk structure\cite{18,19} as well as the single layer\cite{23}. With GW approximation, a general shape of $\chi'_0(q)$ is quite similar to that with DFT-GGA [Fig. 4 (b)] in spite of apparent differences between Fermi surfaces from the two calculation methods. We note that a peak is at $\left(\frac{1}{2}\vec{a}_1 + \frac{1}{2}\vec{a}_2\right)$ point of IBZ not along the high symmetric line unlike the case with DFT-GGA.

In order to see the temperature dependence of $\chi'_0(q)$, we have repeated $\chi'_0(q)$ calculations by changing temperature $k_B T$. Figure 5 shows the bare susceptibility $\chi'_0(q)$ along the IBZ boundary at several temperatures. At $k_B T \leq 10$ meV, the maximum plateau of the DFT-GGA susceptibility has a subpeak at $q \approx 2/5\Gamma M$. As temperature increases, this subpeak disappears and the broad plateau between $2/5\Gamma M$ and $4/5\Gamma M$ remains. For the GW case, the trend is quite similar to those with DFT-GGA along the high symmetric lines of BZ.

D. Doping Effects

Next we consider the doping effect on the bare static susceptibility. The doping effect might be taken into account by rigidly shifting the Fermi energy. $\delta\epsilon_F$ denotes the Fermi energy shift with respect to the charge neutral level due to the doping effect. Figure 6 shows Fermi surfaces for two $p$-doped cases [(a) $\delta\epsilon_F = -150$ meV, (b) $\delta\epsilon_F = -77$ meV], the undoped case [(c) $\delta\epsilon_F = 0$ meV], and two $n$-doped cases [(d) $\delta\epsilon_F = 80$ meV, (e) $\delta\epsilon_F = 150$ meV]. First we consider $p$-doped cases [Fig. 6 (a) and (b)]. As seen in Fig. 2 the saddle point becomes closer to the Fermi level as the system is more $p$-doped. Thus, one can expect that the distance between the two hole pockets gets shorter. In particular, Fig. 6 (b) displays the special case where the Fermi level touches the saddle point.
on the ΓK line. In this case, two hole pockets (triangle and hexagon) are connected to each other. The Fermi surface in Fig. 6(b) is exactly the same one discussed in Ref. 51. So, a slight hole doping enables the system to undergo a CDW phase transition through the logarithmic divergence of the susceptibility at the CDW wave vector connecting the saddle points.51 Being more p-doped, the topology of the Fermi surface is totally different from that of the undoped one. As shown in Fig. 6(a), triangular and hexagonal hole pockets are no longer observed, but there is only one rounded rectangular electron pocket, thus changing its carrier type.

In contrast, as shown in Figs. 6(d) and (e), the topology of the n-doped system does not change compared with the undoped case, except for the fact that the size of the two pockets is reduced.

Such a difference between Fermi surfaces of p-doped and n-doped cases leads to a qualitative change in the bare static susceptibility $\chi_0(q)$. For the n-doped cases where the topology of the Fermi surface does not change, the landscape of the noninteracting susceptibility $\chi_0(q)$ resembles that of the undoped system. See Fig. 7(c) and (d) for $\chi_0(q)$ of $\delta\epsilon_F = 80$ meV and 150 meV, respectively. Compared with the undoped case, the minor difference is that the peak moves slightly toward the ΓM line, and the broad maximum from 3/5ΓM to 4/5ΓM is enhanced.

On the contrary, for the p-doped cases where the relative position of the saddle point is critical, the topological change in the Fermi surface gives rise to a qualitatively different susceptibility pattern. For low doping, i.e., $\delta\epsilon_F > -77$ meV, the susceptibility pattern does not greatly deviate from that of the undoped system [See Fig. 7(a)]. However, when the system is more p-doped such that the Fermi energy is below the saddle point, quite a different pattern emerges. For example, see Fig. 6(b) where $\delta\epsilon_F = -150$ meV. As shown in the Fermi surface for the case of $\delta\epsilon_F = -77$ meV, the peak of $\chi_0(q)$ for this case is at the $M$-point indicating its logarithmic divergence shown in Fig. 7(a). One noticeable change for the case of further p-doping is that the largest peak lies neither at the $M$ point nor on the ΓM line [Fig. 7(b)]. Rather, the main peak is found on the ΓK line for the case of $\delta\epsilon_F = -150$ meV.

\begin{figure}[h]
\centering
\includegraphics[width=0.8\textwidth]{fig7.png}
\caption{(Color online) $\chi_0(q)$ of p-doped cases for GW bands: (a) $\delta\epsilon_F = -77$ meV and (b) $\delta\epsilon_F = -150$ meV. $\chi_0(q)$ of n-doped cases for GW bands: (c) $\delta\epsilon_F = 80$ meV and (d) $\delta\epsilon_F = 150$ meV.}
\end{figure}

E. Effect of Spin-Orbit Coupling

So far the effect of spin-orbit coupling (SOC) has not been taken into account. Unlike the bulk 2H structure where the combination of time-reversal and inversion symmetries prohibits band splitting due to SOC, a single layer of the 2H structure can be significantly affected by SOC, since there is no center of inversion symmetry. One might expect that a new feature can emerge on the Fermi surface, depending on whether the band splitting due to SOC is comparable to the energy of the saddle point with respect to the Fermi level.

For DFT calculations fully relativistic pseudopotentials are used. SOC are included on the final stage of the GW approximation by adding DFT-SOC results since the effect of SOC to GW results is marginal as shown in Ref. 70. We also note that a similar calculation method for the SOC effects has been used for explaining optical spectrum of a single layer MoS$_2$ successfully. Figure 8 shows the electronic band structure corrected by SOC around the Fermi energy from (a) DFT-GGA and (b) GW schemes. As expected, the single band around the Fermi level is split into two bands for both cases. Note that the electronic band on the ΓM line is not affected by SOC for the both cases. This can be understood by the fact that there is a mirror symmetry with respect to a plane perpendicular to the lattice vector $a_1$ in Fig. 1(b). This mirror symmetry together with time-reversal symmetry prohibits band splitting due to SOC along the ΓM line.

With SOC corrections, we found that the general shape of energy bands are quite similar for the both calculation methods except the local Fermi surfaces shape near the saddle point. As shown in Fig. 8(a), SOC introduces two rounded triangular pockets around $K$ and two hexagons around Γ to the Fermi surface of DFT-GGA, which are just two copies of the Fermi surface without SOC. In contrast, with GW approximation, the two pockets along
We next consider the bare static susceptibility \( \chi_0(\mathbf{q}) \) for bands corrected by SOC. Special cares are needed when the constant matrix approximation is applied to Eq. (1) for the SOC-corrected band structure. For spin-degenerate bands without SOC, the oscillator strength matrix element between different spin components is obviously zero. Thus, the spin degree of freedom, which is implicit in Eq. (1), just gives the factor two to \( \chi_0^0(\mathbf{q}) \). In contrast, the electronic bands corrected by SOC cannot be labelled by a definite spin quantum number over \( \mathbf{k} \) space, but generally have a spin texture varying on \( \mathbf{k} \) space. Thus it is likely that the interband matrix element of the oscillator strength might vanish for some points or regions of \( \mathbf{k} \) space. It might not be correct to replace the interband matrix element of the oscillator strength just as unity for the whole \( \mathbf{k} \) space without justification.

For the single layer of metallic TMDs, the bare susceptibility is mainly determined by two bands splitted by SOC around the Fermi level. It is known that these bands are composed primarily of \( d_{2z^2} \), \( d_{xy} \), and \( d_{x^2−y^2} \) orbitals of transition metal. In this effective subspace \{\( d_{2z^2}, d_{xy}, d_{x^2−y^2} \}\), the SOC term does not mix spin components. In other words, the two bands splitted by SOC have the opposite spin over \( \mathbf{k} \) space. Therefore, interband components of the oscillator strength between the two bands might vanish, and the constant matrix approximation can be applied to intraband components. Note that this is the zeroth approximation, and this argument might be modified when \( p \) orbitals of chalcogens and their SOC are involved. For detailed discussions on the effective tight-binding model of \( d \) orbitals and the atomic SOC, see Appendix B and C.

Figure 10 shows the bare susceptibility \( \chi_0^0(\mathbf{q}) \) calculated from DFT-GGA and \( GW \) methods based on the above argument. The bare susceptibility of SOC-corrected bands looks similar to that of electronic bands without SOC in Fig. 4 except for that there are more substructures for the SOC-corrected bands. For example, the maximum plateau of DFT-GGA bands on the \( \Gamma M \) line has more steps as shown in Fig. 10(a). For \( GW \) bands, the bare susceptibility follows a similar variation with the case of DFT-GGA [Fig. 10(b)].

IV. DISCUSSION AND CONCLUSION

Having established the quasiparticle energy bands and the Fermi surface of a monolayer \( \text{NbSe}_2 \) using \( GW \) approximation, now we discuss their implications on a possible CDW phase in the single layer on top of various substrates. Following previous discussions about ef-
The Fermi surface variation within the $GW$ approximation alone cannot explain the large discrepancy of $T_{CDW}$ between the two experiments. Since a giant phonon softening should occur in the vicinity of $M$-point when the saddle points touch the Fermi energy, this will enhance the transition temperature beyond the simple mean-field estimation but the more comprehensive studies to understand electron-phonon interaction with strong Coulomb interaction are needed to understand this phenomena, that is beyond the scope of current work. A recent detailed calculation\cite{30,31} for the single layer on top of graphene indicates a negligible interlayer interaction between them. Contrary to the experiment interpretation\cite{31}, the detailed density of state analysis using DFT-GGA calculation\cite{30} suggests a strong coupling mechanism for the CDW in the single layer limit but the origin of small energy gap at the Fermi energy is still not clear yet.

In conclusion, we have calculated quasiparticle band structures of a single layer $2H$-NbSe$_2$ by using first-principles $GW$ calculation. We found that the width of a partially occupied band increases and its Fermi surface shape changes significantly compared with those obtained using DFT-GGA calculation method. The SOC changes the Fermi surfaces further and the resulting energy bands have the singular saddle points very close to the Fermi surface, enabling the system undergo CDW phase transition. Considering a relatively easy control of charge doping in 2D crystals, we think that the present system is particularly interesting in realizing doping-dependent phase transition. We also provide a simple tight-binding model with a basis of three $d$-orbitals of niobium that can be useful for many-body calculation of large hybrid systems involving NbSe$_2$ and other layered systems.
ACKNOWLEDGMENTS

We thank Antonio Castro Neto, Cheol-Hwan Park, Ting Cao and Steven Louie for discussion. Y.-W.S. was supported by the National Research Foundation of the Ministry of Science, ICT and Future Planning Grants of Korean government (QMMRC, No. R11-2008-053-01002-0). Computations were supported by the Center for Advanced Computation of Korea Institute for Advanced Study.

Appendix A: Effects of semicore states on GW approximations

Here we compare band calculation results of the GW approximation with and without including Nb semicore states in the pseudopotential. It has been investigated that GW energy bands can be significantly changed by the inclusion of core states of the transition metal, for example, in II-VI semiconductors like CdS or bulk copper. In these materials, usual DFT band structures rarely change by including core states whose binding energies are well separated from those of valence states. Despite a distinct separation between core and valence states, when core states have a large spatial overlap with valence states, core states can strongly interact with valence states when the exchange part of the GW self-energy is calculated. This strong interaction between core and valence states indeed leads to a significant modification on the GW band structure.

\[ \text{FIG. 11. (Color online) (a) GW bands for a single layer NbSe}_2 \text{ with and without semicore levels of 4s}^24p^6 \text{ of Nb atom, which are denoted by blue solid lines and red dashed ones. (b) Fermi surfaces of GW bands with semicore states (blue solid lines), GW without semicore states (red dashed lines), and DFT-GGA bands (black dotted lines).} \]

Regarding this effect, we have generated pseudopotentials of Nb atoms with the semicore 4s^24p^6 electrons as valence ones and without them, and we have applied the GW approximation to the two cases. Almost flat bands originating from 4s and 4p states are located at about 54.03 eV and 30.39 eV, respectively, below the Fermi level, which are separated from 4d bands around the Fermi level. All-electron calculation of the Nb atom reveals that radial wave functions of 4s, 4p, and 4d states mostly spread over the similar radial range, thereby implying that there is a strong interaction between core (4s and 4p) and valence (4d) states in the exchange interaction of the GW approximation.

Figure 11(a) shows the resulting GW band structures of the NbSe_2 monolayer by using pseudopotentials with and without 4s and 4p semicore states. One distinguished feature induced by the inclusion of semicore states is the shape of the partially occupied band in the neighborhood of the symmetric point M. Without semicore states the minimum of the partially occupied band is located at M. It is in a sharp contrast to the fact that the partially occupied band has the energy minimum at the middle of the ΓM line in the GW approximation with semicore states. The energy difference of the two cases at M is approximately 233 meV. Fermi surfaces of GW bands with and without semicore states are shown in Fig. 11(b) together with the Fermi surface of DFT-GGA bands. Due to the energy shift around M caused by semicore states, the crossing point of this band with the Fermi level on the MK line moves toward the M point, resulting in a rounded hexagonal Fermi surface around M. Without semicore states, the Fermi surface around M looks like a triangular shape whose flat sides face the M point, so the triangular pocket is distinguished from that of the DFT-GGA calculation in terms of the orientation. Except for the energy correction around M, energies and shapes of partially occupied bands from the two calculations are almost the same.

When it comes to energy bands far from the Fermi level, it is first found that four lowest unoccupied bands originating from 4d orbitals of Nb do not significantly change. In contrast higher unoccupied bands and fully occupied ones below the Fermi energy are shifted upwards overall.

Appendix B: Atomic Spin-Orbit Coupling in monolayer NbSe_2

We here discuss the effect of SOC and the spin texture of the partially occupied band. Here we consider SOC as the atomic one that gives only on-site terms,

\[ H_{so} = \lambda_{TM} \mathbf{S} \cdot \sum_{m \in Nb} \mathbf{L}_m + \lambda_{CH} \mathbf{S} \cdot \sum_{m \in Se} \mathbf{L}_m, \]  

(B1)

where \( \lambda_{TM} \) and \( \lambda_{CH} \) are SOC constants for transition metal and chalcogen respectively, \( \mathbf{S} \) is the spin 1/2 operator, and \( \mathbf{L}_m \) is the angular momentum operator of an atom \( m \). First of all, for the d-orbital space, the SOC term at transition metal atom \( m \) reads
modified, since there is the SOC term of eigenstates for the partially occupied band in the density of states calculation, the wavefunction of the partially occupied band is well separated from the other four bands at about 20%.

Without SOC, it is shown that two subspaces of d orbitals \{d_{xz}, d_{xy}, d_{x^2-y^2}\} and \{d_{yz}, d_{xz}\} are not coupled in the group-theoretical construction of the d-orbital TB model\(^{(21,22)}\). Equation (B2) tells us that SOC leads to mixing of the two subspaces \{d_{xz}, d_{xy}, d_{x^2-y^2}\} and \{d_{yz}, d_{xz}\}. Furthermore, \(|\uparrow\rangle\) and \(|\downarrow\rangle\) are no longer spin eigenstates, since \(\langle \psi_1, \uparrow | \hat{S}_z \cdot \mathbf{L}_m | \psi_2, \downarrow \rangle \neq 0\) in which \(\psi_1 \in \{d_{xz}, d_{xy}, d_{x^2-y^2}\}\) and \(\psi_2 \in \{d_{yz}, d_{xz}\}\).

When we consider the partially occupied band around the Fermi energy coming from \{d_{xz}, d_{xy}, d_{x^2-y^2}\}, this band is well separated from the other four bands at least by 2 eV. Therefore, from the viewpoint of the perturbation theory, the mixing term of Eq. (B2) between \{d_{xz}, d_{xy}, d_{x^2-y^2}\} and \{d_{yz}, d_{xz}\} might give a negligible contribution to the partially occupied band. In the subspace \{d_{xz}, d_{xy}, d_{x^2-y^2}\}, the SOC term [Eq. (B2)] approximately reads

\[
\lambda_{TM} \hat{S} \cdot \mathbf{L}_m \approx \lambda_{TM} \begin{bmatrix}
0 & 0 & 0 & 0 & 0 & 0 & i\sqrt{3}/2 & -\sqrt{3}/2 \\
0 & i & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & i\sqrt{3}/2 & -\sqrt{3}/2 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
-\sqrt{3}/2 & i\sqrt{3}/2 & 0 & 0 & 0 & 0 & 0 & 0 \\
-\sqrt{3}/2 & -i\sqrt{3}/2 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix},
\]

with the basis \(|d_{xz}, \uparrow\rangle, |d_{xy}, \uparrow\rangle, |d_{x^2-y^2}, \uparrow\rangle, |d_{yz}, \uparrow\rangle, |d_{xz}, \downarrow\rangle, |d_{xy}, \downarrow\rangle, |d_{x^2-y^2}, \downarrow\rangle, |d_{yz}, \downarrow\rangle, |d_{xz}, \downarrow\rangle\). Here \(|\uparrow\rangle\) and \(|\downarrow\rangle\) are spin eigenstates of the \(\hat{S}_z\) operator.

The SOC term on a chalcogen m is

\[
\lambda_{CH} \hat{S} \cdot \mathbf{L}_m \approx \lambda_{CH} \begin{bmatrix}
0 & -i & 0 & 0 & 0 & 0 & \frac{1}{2} & 0 \\
\frac{1}{2} & 0 & 0 & 0 & 0 & -\frac{1}{2} & 0 & 0 \\
0 & 0 & 0 & -\frac{1}{2} & 0 & \frac{1}{2} & -\frac{1}{2} & 0 \\
0 & 0 & 0 & \frac{1}{2} & 0 & -\frac{1}{2} & \frac{1}{2} & 0 \\
0 & 0 & 0 & -\frac{1}{2} & 0 & -\frac{1}{2} & \frac{1}{2} & 0 \\
0 & 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} & -\frac{1}{2} & 0
\end{bmatrix},
\]

which is ordered in the basis of \(|p_x, \uparrow\rangle, |p_y, \uparrow\rangle, |p_z, \uparrow\rangle, |p_x, \downarrow\rangle, |p_y, \downarrow\rangle, |p_z, \downarrow\rangle\). Two subspaces of p orbitals \{p_x, p_y\} and \{p_z\} are mixed under SOC. Further, \(|\uparrow\rangle\) and \(|\downarrow\rangle\) are not exact spin eigenstates. Eigenstates of the total angular momentum operator \(J = L + S\) could be true ones since orbital and spin angular momenta are correlated via SOC. Considering contributions of p orbitals, it might be expected that there is relatively small mixing between \(|\uparrow\rangle\) and \(|\downarrow\rangle\) components for bands splitted by SOC.

### Appendix C: Three-Band Tight-Binding Model

In this section, we construct a tight-binding (TB) model which closely reproduces GW bands, especially the partially occupied band around the Fermi energy. The partially occupied band comes mostly from \(d_{xz}, d_{xy},\) and \(d_{x^2-y^2}\) orbitals of transition metal p orbitals from chalcogen atoms gives the next contribution to the band. Although it is more precise to construct the tight-binding model by including the p orbitals, we would like to have a minimal model capturing the main physics so that it can be used for other theoretical investigations. For this purpose, we start with the three-orbital TB model including only \(d_{xz}, d_{xy},\) and \(d_{x^2-y^2}\) orbitals.

The form of the TB model and the number of independent parameters can be determined by the lattice symmetry. In particular, Ref. \(^{(21)}\) provides a complete table of matrix components of the TB model of d orbitals for honeycomb structure, which is relevant for TMDs. For detailed discussion on the group-theoretical construction of the TB model, see Refs. \(^{(72)}\) and \(^{(73)}\).
The effective TB model of \( d \) orbitals has been applied to TMDs in the literature\(^{24,25}\). Recently, this model has been extensively used in order to fit DFT band structures for monolayers of group-VIB TMDs\(^{22}\). In Ref.\(^{75}\), the authors have extended the three \( d \)-orbital TB model up to the third nearest neighbor (TNN) hoppings. Using the TNN TB model, they have successfully reproduced DFT electronic bands.

As mentioned in Refs.\(^{74,75}\), the \( d-d \) interactions in the effective TB model contain the direct hoppings between \( d \) orbitals of transition metal and the indirect \( d-d \) hoppings via \( p \) orbitals of chalcogens. However, within this effective model, one cannot know how much contribution \( p \) orbitals give to the energy band of our interest, which might be important for some aspects, for example, the effect of SOC.

Using the three-band TNN model of Ref.\(^{75}\) we fit DFT-GGA and GW bands on the IBZ boundary. For this, we first adopt the least-square fitting procedure, and finely tune parameters in order to fit the saddle point of the \( MK \) line. Figure 12(a) shows energy bands of the three-band TNN TB model together with DFT-GGA bands. Following notations of Ref.\(^{75}\) fitting parameters of the TB model for DFT-GGA bands are \( \epsilon_1 = 1.408, \epsilon_2 = 2.048, t_0 = -0.128, t_1 = 0.115, t_2 = -0.466, t_{11} = 0.115, t_{12} = 0.122, t_{22} = 0.036, r_0 = 0.025, r_1 = 0.194, r_2 = -0.079, r_{11} = 0.021, r_{12} = 0.096, u_0 = -0.031, u_1 = -0.037, u_2 = -0.002, u_{11} = 0.258, u_{12} = -0.179, \) and \( u_{22} = -0.167 \) in units of eV. The energy bands of the TB model are in a very good agreement with DFT-GGA bands except for a small deviation at the energy minimum of the partially occupied band. As shown in Fig. 12(b), the fitted TB model well reproduces the Fermi surface of the DFT-GGA calculation.

We also construct the three-band TNN model of GW bands, whose fitting parameters are \( \epsilon_1 = 1.148, \epsilon_2 = 2.379, t_0 = -0.118, t_1 = -0.386, t_2 = -0.366, t_{11} = 0.167, t_{12} = 0.243, t_{22} = -0.075, r_0 = 0.094, r_1 = 0.043, r_2 = -0.152, r_{11} = 0.055, r_{12} = -0.012, u_0 = -0.061, u_1 = -0.010, u_2 = 0.002, u_{11} = 0.140, u_{12} = -0.077, \) and \( u_{22} = -0.014 \) in units of eV. Energy bands reproduced by the TNN tight-binding model and GW bands are shown in Fig. 12(c). The energy bands of the TB model agree well with those of the GW calculation, but there is some deviation of the partially occupied band on the \( MK \) line. While hexagonal pockets around \( \Gamma \) nicely match, the triangular pocket of the TB model is less rounded than that of GW bands. This is due to the small mismatch of the two bands on the \( MK \) line in Fig. 12(c).

Figure 13 displays orbital-projected bands calculated from the three-orbital TNN TB model of the GW bands. \( d_{z^2} \) orbital is dominant for the lowest band around the Fermi level, while \( d_{xy} \) and \( d_{z^2-y^2} \) orbitals give main con-
tributions to the other two unoccupied bands.

Appendix D: GW bands of bulk $2H$-NbSe$_2$

We have extended DFT-GGA and GW calculations to energy bands of the bulk $2H$-NbSe$_2$ by using the same pseudo-potential including semi-core states. For detailed calculations we have used the plane-wave basis set with a cutoff of 55 Ry, a $20 \times 20 \times 5$ k-point grid, and a smearing temperature $k_B \tau = 0.005$ Ry$^{-1}$. The GW bands of the bulk $2H$-NbSe$_2$ is calculated within the level of the $G_0W_0$ approximation by including about 400 unoccupied bands, which are up to 10 Ry above the Fermi energy.

Figure 14(a) shows DFT-GGA and GW electronic bands of the bulk $2H$-NbSe$_2$ along symmetric lines of the IBZ, which are denoted by red dashed lines and blue solid ones respectively. Their corresponding Fermi surfaces are depicted in Fig. 14(b)-(d). Figure 14(b) and (c) show Fermi surfaces of DFT-GGA bands on the ΓMK and ALH planes, respectively, while Fig. 14(d) and (e) are Fermi surfaces of GW bands on ΓMK and ALH planes respectively. As discussed in Sec. IV the Fermi surface of the bulk $2H$-NbSe$_2$ GW bands on the ALH plane is similar to that of the single-layer DFT-GGA bands on the ΓMK plane. In particular the Fermi surface around $H$ is a rounded triangular pocket whose flat lines face the symmetric point $A$, but not a rounded hexagonal pocket as seen in the Fermi surface of the monolayer GW calculation. The relative large screening effect of neighboring layers could lead to this similarity between the Fermi surface of bulk GW bands and that of monolayer DFT-GGA bands.
13

9, 2485 (1974).
54 J. Rath and A. J. Freeman, Phys. Rev. B 11, 2109 (1975).
55 R. P. Gupta and A. J. Freeman, Phys. Rev. B 13, 4376 (1976).
56 H. W. Myron and A. J. Freeman, Phys. Rev. B 11, 2735 (1975).
57 H. W. Myron, J. Rath, and A. J. Freeman, Phys. Rev. B 15, 885 (1977).
58 Y. Ge and A. Y. Liu, Phys. Rev. B 86, 104101 (2012).
59 O. V. Yazyev, E. Kioupakis, J. E. Moore, and S. G. Louie, Phys. Rev. B 85, 161101 (2012).
60 M. Miasek, Phys. Rev. 107, 92 (1957).
61 R. F. Egorov, B. I. Reser, and V. P. Shirokovskii, Phys. Status Solidi B 26, 391 (1968).
62 V. N. Kuznetsov and A. N. Men, Phys. Status Solidi B 85, 95 (1978).
63 L. F. Mattheiss, Phys. Rev. Lett. 30, 784 (1973).
64 G.-B. Liu, W.-Y. Shan, Y. Yao, W. Yao, and D. Xiao, Phys. Rev. B 88, 085433 (2013).
65 P. R. Gray, P. J. Hurst, S. H. Lewis, and R. G. Meyer, Analysis and Design of Analog Integrated Circuits (Wiley, New York, 1984).
66 T. Ohta, A. Bostwick, J. L. McChesney, T. Seyller, K. Horn, and E. Rotenberg, Phys. Rev. Lett. 98, 206802 (2007).
67 S. Y. Zhou, G.-H. Gweon, A. V. Fedorov, P. N. First, W. A. de Heer, D.-H. Lee, F. Guinea, A. H. C. Neto, and A. Lanzara, Nat. Mat. 6, 770 (2007).
68 S. Kim, J. Ihm, H. J. Choi, and Y.-W. Son, Solid State Comm. 175-176, 83 (2013).
69 J. A. Silva-Guillén, P. Ordejón, F. Guinea, and E. Canadell, 2D Mater. 3, 035028 (2016).