Excitons in boron nitride single layer

Thomas Galvani,[1] Fulvio Paleari,[1] Henrique P. C. Miranda,[1] Alejandro Molina-Sánchez,[1] Ludger Wirtz,[1] Sylvain Latil,[2] Hakim Amara,[3] and François Ducastelle[3]

[1]*Physics and Materials Science Research Unit, University of Luxembourg, 162a avenue de la Faïencerie, L-1511 Luxembourg, Luxembourg*

[2]*SPEC, CEA, CNRS, Université Paris-Saclay, CEA Saclay 91191 Gif sur Yvette, France*

[3]*Laboratoire d'Etude des Microstructures, ONERA-CNRS, BP 72, 92322 Châtillon Cedex, France*

(Received 31 May 2016; revised manuscript received 4 August 2016; published 6 September 2016)

Boron nitride single layer belongs to the family of two-dimensional materials whose optical properties are currently receiving considerable attention. Strong excitonic effects have already been observed in the bulk and still stronger effects are predicted for single layers. We present here a detailed study of these properties by combining *ab initio* calculations and a tight-binding Wannier analysis in both real and reciprocal space. Due to the simplicity of the band structure with single valence (π) and conduction (π^*) bands the tight-binding analysis becomes quasiquantitative with only two adjustable parameters and provides tools for a detailed analysis of the exciton properties. Strong deviations from the usual hydrogenic model are evidenced. The ground-state exciton is not a genuine Frenkel exciton, but a very localized tightly bound one. The other ones are similar to those found in transition-metal dichalcogenides and, although more localized, can be described within a Wannier-Mott scheme.

DOI: 10.1103/PhysRevB.94.125303

I. INTRODUCTION

Two-dimensional (2D) materials are currently the object of many investigations concerning their electronic and optical properties. Graphene is the most known example [1] but hexagonal boron nitride [2–9] and transition-metal dichalcogenides (TMD) [10–14], as well as new materials such as phosphorene, [15,16] silicene, germanene, etc. [17,18] are receiving considerable attention. Unlike graphene, these materials are semiconductors with new optical properties that are absent in 3D semiconductors. Excitonic effects, in particular, are more pronounced in two dimensions than in three dimensions, with the exciton binding energies being of the order of 0.1–1 eV or more. The spatial extension of the excitons remains fairly large in general so that they are frequently considered as Wannier-Mott excitons. However, it has been quickly noticed that the usual hydrogenic model does not apply in two dimensions because of the different screening processes involved [19–27]. There is thus a need to understand more precisely these excitonic effects. The case of hexagonal boron nitride (hBN) is more specific still. Even in its bulk hexagonal form, very strong excitonic effects have been reported very early based on *ab initio* calculations [28]. The theoretical interpretation of very strongly bound excitons (0.7 eV for the ground-state exciton in bulk hBN) was confirmed by various experiments [2,3,5,6,8] and was refined in theoretical calculations making use of symmetry arguments [29–31]. The reason for the strong binding energy is the quasi-2D nature of the hBN structure [28] consisting of weakly interacting stackings of hexagonal layers. Furthermore, hBN has a very large gap, >6 eV, leading to a rather weak dielectric screening, so that all ingredients conspire to enhance these excitonic effects. They have been studied recently, but the experiments are difficult because of the necessity to work in the far UV range.

The current interest in 2D materials and the development of techniques to handle few-layer materials suggests to study the properties of hBN as a function of the number of layers, as has been done in the case of graphene and TMD. In the case of TMD it has been shown that the nature of the gap—indirect or direct—depends on the number of layers. What about hBN? Preliminary experimental studies are already available [32], but much remains to be done. First, a precise knowledge of the single-layer (SL) properties is required.

We present a detailed theoretical study of the first excitonic levels and characterize their energies and shapes by combining *ab initio* calculations and a simple tight-binding model. The *ab initio* approach is the usual one, based on a GW plus Bethe-Salpeter approach. The tight-binding approach is close to the approach put forward by Wannier long ago [33–36]. As it turns out we have to take into account just one π valence band and one π^* conduction band. Furthermore, close to the gap, the corresponding Bloch states are concentrated on the nitrogen (N) and boron (B) atoms respectively, so that the π orbitals can be considered as Wannier functions. It is then possible to work out the Wannier equations in real space in a simple but accurate way.

The paper is organized as follows: Sec. II is devoted to the electronic structure of hBN-SL, which is calculated using standard *ab initio* techniques and fitted to a simple tight-binding model. Section III contains the main discussion of the various excitons and of their symmetry, using in particular imaging tools in real and reciprocal space. Finally, Sec. IV is concerned with the calculation of optical matrix elements. This is followed by a discussion (Sec. V) and several appendixes.

II. ELECTRONIC STRUCTURE OF hBN SINGLE LAYER

A. Band structure

We first specify a few notations. The structure of the hBN single layer is shown in Fig. 1.

The band structure of the hBN single layer is shown in Fig. 2. The density functional theory (DFT) calculations have been made using the QUANTUM ESPRESSO code with the local density approximation (LDA) for the exchange-correlation functional [38]. The GW corrections were computed in the G_0W_0 approximation, using the YAMBO code [39] with the

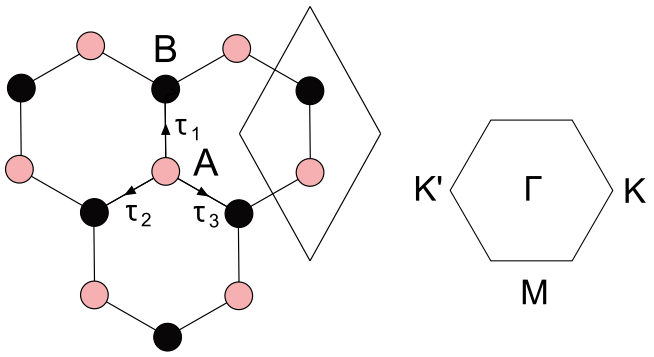

FIG. 1. Left: Honeycomb structure with its two triangular sublattices A and B occupied in hBN by nitrogen and boron atoms, respectively. $\boldsymbol{\tau}_1, \boldsymbol{\tau}_2, \boldsymbol{\tau}_3$ are the vectors joining first neighbors between the two sublattices. The vectors are opposite if the origin is taken on a B atom. The unit cell contains one nitrogen and one boron atom. Right: Brillouin zone.

plasmon-pole approximation for the frequency dependence of the dielectric function. These corrections have been applied to the last four valence bands and up to the first twenty conduction ones (only the first two are shown in Fig. 2). The lattice parameter has been fixed at the optimized lattice constant $a = 4.72$ a.u. (2.50 Å). The computational details are the same used for the subsequent Bethe-Salpeter calculation and are found in Sec. III. The gap, equal to 7.25 eV, is direct between the π and π^* band at point K in the Brillouin zone, while the bands are very flat along the KM lines. This is in agreement with several previous calculations [40–45]. Notice however that the G_0W_0 approximation is known to underestimate large band gaps. In addition, after the GW corrections are made, the bottom of the conduction band at the Γ point is lower than at K. These are nearly free electron states [40] that are forming around isolated layers of hBN—we checked that the energy position of these states does not depend on the vacuum space between repeated periodic images of the system. However, the transition matrix elements from localized valence band states into these states are very low, so, for all practical purposes, the

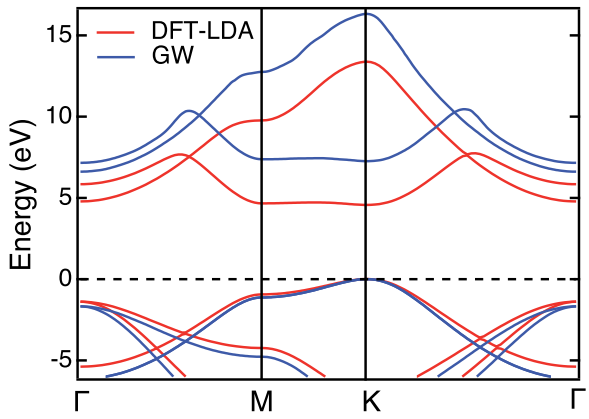

FIG. 2. DFT-LDA (red) and GW (blue) *ab initio* band structures of a single hBN layer: in the MK region, the gap is direct at point K between the flat π bands. The GW corrections were interpolated for this figure with the WANNIER90 code. [37].

isolated sheet of hBN can be considered to be a direct band gap material. This means that regardless of the nature of the quasiparticle gap (direct or indirect), the optical gap is direct at point K. The *ab initio* results confirm that the contributions to all exciton states of interest in this work come from transitions near K and away from Γ (see, for example, the reciprocal space plots in Figs. 5 and 10).

As in the case of graphene, the two π bands can be reproduced fairly well using a simple tight-binding model. Let us denote as $|\boldsymbol{n}\rangle$ the p_z atomic state at site $\boldsymbol{n}$. The corresponding atomic orbital is $\phi(\boldsymbol{r}-\boldsymbol{n}) = \langle \boldsymbol{r}|\boldsymbol{n}\rangle$. Then, we define the two Bloch functions on the A and B sublattices:

$$|\boldsymbol{k}\, A(B)\rangle = \frac{1}{\sqrt{N}} \sum_{\boldsymbol{n}\in A(B)} e^{i\boldsymbol{k}\cdot\boldsymbol{n}}\, |\boldsymbol{n}\rangle,$$

where N is the number of unit cells, i.e., half the number of atoms. As usual, in most cases we just keep first-neighbor hopping integrals $-t, t > 0$, and the nitrogen and boron atoms are distinguished by their on-site matrix elements, equal to $-\Delta$ on the A sites for the N atoms, and to Δ on the B sites for the B atoms. The matrix elements of the Hamiltonian in the Bloch basis are therefore written as:

$$\begin{aligned}
\langle \boldsymbol{k}A|H|\boldsymbol{k}A\rangle &= -\Delta \\
\langle \boldsymbol{k}B|H|\boldsymbol{k}B\rangle &= +\Delta \\
\langle \boldsymbol{k}A|H|\boldsymbol{k}B\rangle &= \langle \boldsymbol{k}B|H|\boldsymbol{k}A\rangle^* = -t\,\gamma(\boldsymbol{k}) \\
\gamma(\boldsymbol{k}) &= \sum_{\alpha=1,2,3} e^{i\boldsymbol{k}\cdot\boldsymbol{\tau}_\alpha}. \qquad (1)
\end{aligned}$$

The energy eigenvalues E are then given by:

$$E = sE_{\boldsymbol{k}}; \quad E_{\boldsymbol{k}} = \sqrt{\Delta^2 + t^2|\gamma(\boldsymbol{k})|^2}; \quad s = \mathrm{sgn}\,(E),$$

and the eigenstates are

$$|\boldsymbol{k}\, s\rangle = C_s^A\, |\boldsymbol{k}A\rangle + C_s^B\, |\boldsymbol{k}B\rangle.$$

Finally, up to a phase factor the coefficients C_s^A and C_s^B are given by:

$$\begin{aligned}
C_s^A &= \langle \boldsymbol{k}A|\boldsymbol{k}s\rangle = -s\frac{\gamma(\boldsymbol{k})}{|\gamma(\boldsymbol{k})|}\sqrt{\frac{E_{\boldsymbol{k}} - s\Delta}{2E_{\boldsymbol{k}}}} \\
C_s^B &= \langle \boldsymbol{k}B|\boldsymbol{k}s\rangle = \sqrt{\frac{E_{\boldsymbol{k}} + s\Delta}{2E_{\boldsymbol{k}}}}. \qquad (2)
\end{aligned}$$

Thus, the π electronic structure of hBN-SL can be characterized by only two parameters, t and Δ. Their order of magnitude is $t \simeq \Delta \simeq 3$ eV but more precise values can be obtained by fitting the valence and conduction bands $E_{\boldsymbol{k}} = \pm\sqrt{\Delta^2 + t^2|\gamma(\boldsymbol{k})^2|}$ to those provided by *ab initio* calculations. Δ is fixed so that the gap 2Δ is equal to the *ab initio* one, $\Delta = 3.625$ eV, and t is then obtained using standard fitting procedures. Different values are obtained depending on the energy range where the fit is optimized. A global fit, disregarding the nearly free electron states, leads to $t = 3.0$ eV, but here we are more interested to have a good fit along the MK line, in which case we obtain $t = 2.30$ eV (similar to the values for recent fitting of band structures in Ref. [42]).

The fit is better for the valence band than for the conduction band (see Fig. 3). The fit can easily be improved by adding

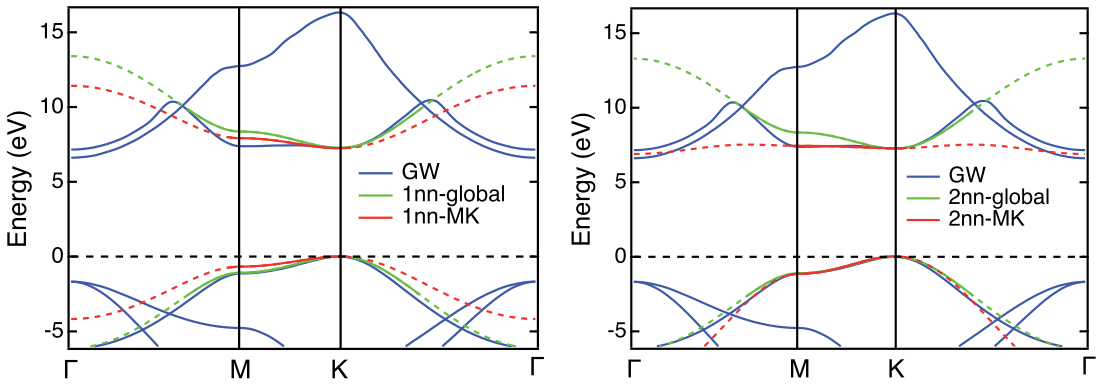

FIG. 3. Tight-binding fit to the *ab initio* bands including first-neighbor interactions (left) and first- and second-neighbor interactions (right). Solid lines denote the region of the fit: the global fit is made for all π bands except the nearly free electron states, whereas the MK fits are optimized for the local band structure along the MK lines.

further neighbor interactions. Neighbors on the same sublattice contribute to diagonal matrix elements whereas neighbors on different sublattices contribute to off-diagonal elements.

Before considering second-neighbor interactions explicitly, let us see a simple way to deduce the band structure of hBN-SL from that of graphene. Let H^0 be the Hamiltonian of graphene with only t hopping integrals. The full Hamiltonian H of hBN-SL is given by $H = H^0 + \hat{\Delta}$ where $\hat{\Delta}$ is the atomic diagonal Hamiltonian with matrix elements equal to $-\Delta$ on the A sublattice and to $+\Delta$ on the B sublattice. It is clear that $H^0\hat{\Delta} + \hat{\Delta}\,H^0 = 0$ so that $H^2 = (H^0)^2 + \Delta^2$ where Δ^2 is a simple constant (multiplied by the unit matrix). $(H^0)^2$ on the other hand is a Hamiltonian connecting sites entirely on sublattice A or on sublattice B. On these triangular lattices $(H^0)^2$ connects the first neighbors with a hopping integral equal to t^2 but it has also diagonal on-site matrix elements equal to $3t^2$. In the Bloch basis $(H^0)^2$ is therefore diagonal:

$$(H^0)^2 = \sum_{\boldsymbol{k}} |\boldsymbol{k}A\rangle|t\,\gamma(\boldsymbol{k})|^2\langle\boldsymbol{k}A| + |\boldsymbol{k}B\rangle|t\,\gamma(\boldsymbol{k})|^2\langle\boldsymbol{k}B|,$$

and the eigenvalues are indeed equal to $\Delta^2 + t^2|\gamma(\boldsymbol{k})|^2$.

Adding second-neighbor interactions $-t_2$ in the hBN-SL structure is then equivalent to introducing first-neighbor interactions on the triangular sublattices, and the eigenvalues are therefore given by:

$$E_{\boldsymbol{k}} = -t_2(|\gamma(\boldsymbol{k})|^2 - 3) \pm \sqrt{\Delta^2 + t^2|\gamma(\boldsymbol{k})|^2}.$$

Since both t and t_2 are positive, second-neighbor interactions induce an asymmetry between the valence and the conduction band: The conduction band becomes flatter than the valence band, in agreement with *ab initio* calculations (Fig. 3). The best local fit along MK is provided by $t = 2.30\,\text{eV}$, $t_2 = 0.096\,\text{eV}$. Actually, under the approximation that the valence and conduction bands are pure N and B states, as shown below, only the energy difference between these two bands enters the tight-binding excitonic Hamiltonian derived in Sec. III, and the second-nearest-neighbors hopping term does not contribute to this difference. For this reason we limit our tight-binding model for the excitons to first-nearest-neighbor hopping and keep the simplest previous fit with $\Delta = 3.625\,\text{eV}, t = 2.30\,\text{eV}$.

B. Wave functions, densities of states

Many electronic properties of hBN-SL only depend on the electronic states close to the gap, i.e., in energy ranges where $t|\gamma(\boldsymbol{k})|$ is small compared to the gap 2Δ. This means that in a first approximation, the coefficients $|C_s^i|, i = A, B$ are equal to one or zero. In other words close to the gap, the valence states are concentrated on the N sites whereas the conduction states are concentrated on the B sites, and the eigenvalues can be approximated by:

$$E_{\boldsymbol{k}} \simeq \pm\left(\Delta + \frac{t^2}{2\Delta}|\gamma(\boldsymbol{k})|^2\right).$$

To examine the validity of this approximation, we have calculated the local densities of states $n_{N(B)}(E)$ on both N and B sites. They are shown in Fig. 4. In our simple tight-binding model, one can easily see that $n_B(-E) = n_N(E)$. Furthermore, $n_B(E)$ shows a step-function-like onset at $E = +\Delta$, whereas $n_N(E)$ has its onset at $E = -\Delta$. As a result the states are

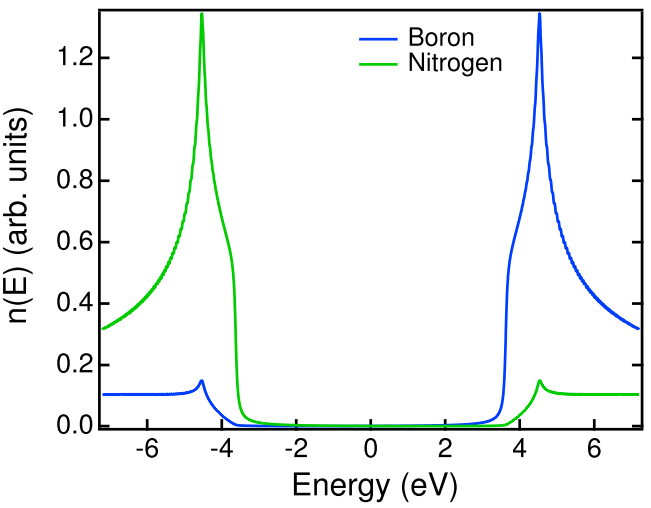

FIG. 4. Tight-binding local densities of states on N and B sites calculated using the recursion method ($t = 2.30\,\text{eV}$, $\Delta = 3.625\,\text{eV}$). The boron (nitrogen) density of states is discontinuous at the upper (lower) band edge $E = \Delta\,(-\Delta)$.

indeed quasipure B states in a fairly broad energy range above Δ, and of course they are quasipure N states below $-\Delta$.

To summarize we can assume that the Wannier functions associated with the valence band and the conduction band can be identified, to lowest order, with the atomic π functions centered on the corresponding sites of their triangular lattices. The effective Hamiltonians H_v for the π band and H_c for the π^* band are then given by:

$$H_v = -\sum_{\boldsymbol{n}} |\boldsymbol{n}A\rangle\,(\Delta + 3t_v)\,\langle \boldsymbol{n}A| - \sum_{\boldsymbol{n},\boldsymbol{m}}{}' |\boldsymbol{n}A\rangle t_v \langle \boldsymbol{m}A| \quad (3)$$

$$H_c = \sum_{\boldsymbol{n}} |\boldsymbol{n}B\rangle\,(\Delta + 3t_c)\,\langle \boldsymbol{n}B| + \sum_{\boldsymbol{n},\boldsymbol{m}}{}' |\boldsymbol{n}B\rangle t_c \langle \boldsymbol{m}B|, \quad (4)$$

where the primes indicate sums over nearest neighbors on the triangular lattices, and $t_v = t_c = t^2/2\Delta$.

III. EXCITONS IN hBN SINGLE LAYER

Ab initio excitonic calculations are based on the Bethe-Salpeter formalism, [36,46,47] which in practice leads to an effective Schrödinger or Wannier equation for electron-hole pairs[1]:

$$(E_{\boldsymbol{k}c} - E_{\boldsymbol{k}v})\Phi_{\boldsymbol{k}vc} + \sum_{\boldsymbol{k}'v'c'} \langle \boldsymbol{k}vc|K_{\rm eh}|\boldsymbol{k}'v'c'\rangle \Phi_{\boldsymbol{k}'v'c'} = E\ \Phi_{\boldsymbol{k}vc},$$

where $E_{\boldsymbol{k}c}$ and $E_{\boldsymbol{k}v}$ are the conduction and valence band energy, respectively, $K_{\rm eh}$ is the electron-hole interaction kernel and $\Phi_{\boldsymbol{k}vc}$ is the electron-hole wave function in $\boldsymbol{k}$ space. In this paper, we only consider vertical excitations where the electron and the hole have the same wave vector $\boldsymbol{k}$, i.e., we consider excitons with vanishing wave vector $\boldsymbol{Q}$ of their center of mass. The $\Phi_{\boldsymbol{k}vc}$ are the coefficients in the expansion of the excitonic state $|\Phi\rangle$ in terms of electron-hole excitations:

$$|\Phi\rangle = \sum_{\boldsymbol{k}} \Phi_{\boldsymbol{k}vc}\ a^\dagger_{c\boldsymbol{k}} a_{v\boldsymbol{k}} |\emptyset\rangle,$$

where the vacuum state $|\emptyset\rangle$ is the state where, at zero temperature, all valence states are full and all conduction states are empty. Only singlet states are considered here so that spin indices are omitted.

The Bethe-Salpeter equation has been solved using the YAMBO code [39]. A Coulomb cutoff of the screened potential in the vertical direction has been used in order to avoid long-range interaction between repeated copies of the monolayer [48]. In this way, we find that both the GW corrections and the first excitonic peaks are already converged (with about 0.01 eV accuracy) with an interlayer separation of 40 atomic units. The same level of convergence was achieved by sampling the two-dimensional Brillouin zone with a $24 \times 24 \times 1$ k-point grid. The additional parameters for many-body perturbation theory (MBPT) calculations as implemented in YAMBO were converged with the same level of accuracy. We also carried out calculations with a $36 \times 36 \times 1$ k-point grid in order to show the higher-energy excitonic wave functions in real space without any overlap between repeated copies on the same monolayer. We verified furthermore that choosing an interlayer separation of 80 a.u. does not modify the results.

A. Ground-state exciton

As an introduction, we present *ab initio* results concerning the ground-state exciton level. Its binding energy measured with respect to the bottom of the conduction band is huge: 1.9 eV. In Fig. 5 we show an image of the excitonic wave function $\Phi(\boldsymbol{r}_h,\boldsymbol{r}_e)$, where the hole (at $\boldsymbol{r}_h$) is localized just above a nitrogen atom. The plot represents the total probability $|\Phi(\boldsymbol{r}_h,\boldsymbol{r}_e)|^2$, i.e., the probability to find the electron at position $\boldsymbol{r}_e$ if the hole is located at $\boldsymbol{r}_h$. Since this exciton is doubly degenerate, we sum the total probability over the two

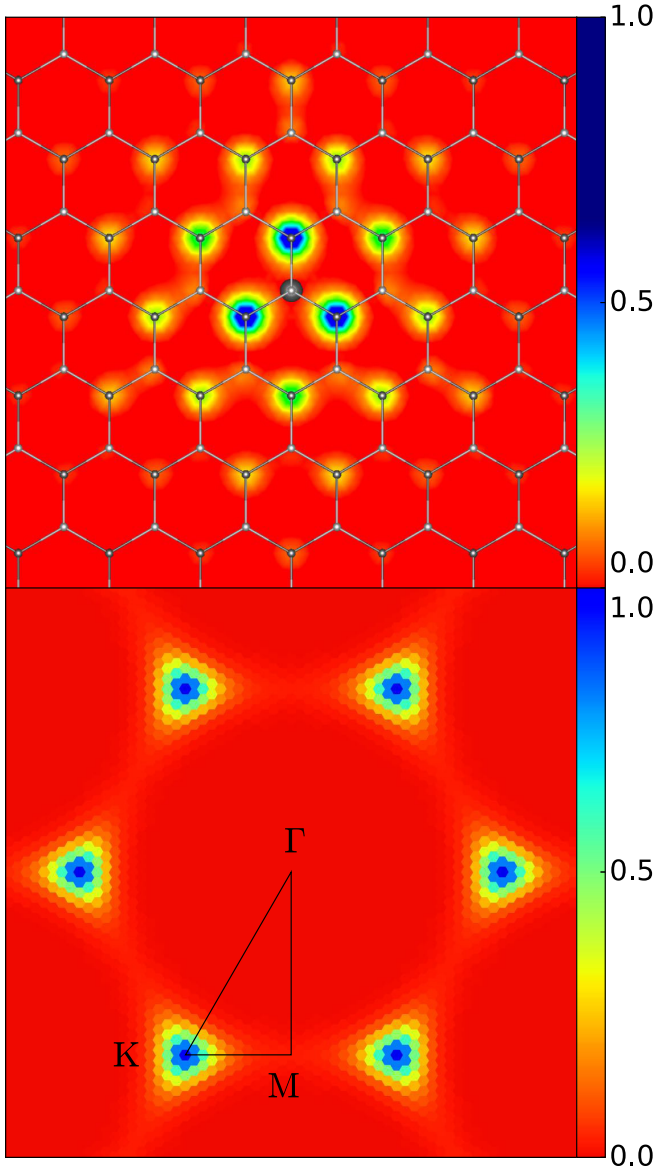

FIG. 5. Top: Total probability density $|\Phi(\boldsymbol{r}_h,\boldsymbol{r}_e)|^2$ of the ground-state degenerate exciton states. The hole is located 0.37 Å above the nitrogen atom in the center (black circle) and $\boldsymbol{r}_e$ is at the same altitude. Bottom: Corresponding Fourier intensity $\sum_{vc} |\Phi_{\boldsymbol{k}cv}|^2$. All intensities are summed over the two degenerate components.

[1]Standard treatments of excitons can be found, for example, in Refs. [34,35], and [36].

degenerate states in order to preserve the trigonal symmetry of the crystal lattice [30]. As expected, the electron density is centered on the boron atoms, with a high probability—about 30%—on the first-nearest neighbors. Although not in the genuine Frenkel limit, the exciton is well localized in real space. The shape of this exciton is actually quite similar to that found for the 3D hBN crystal, which is not surprising since in the latter case the exciton was already found to be well confined in a single layer although with a lower binding energy [28–30,41]. We also show the wave function in reciprocal space. Here we plot the (summed) weight $\sum_{vc} |\Phi_{\boldsymbol{k}vc}|^2$ of the electron-hole pairs of wave vector $\boldsymbol{k}$ that constitute the bound exciton. The distribution is peaked around the high-symmetry point K but extends along the KM boundaries of the Brillouin zone.

B. Wannier tight-binding model

We use now the result derived in Sec. II that the valence and conduction states are pure A (N atoms) and B (B atoms) states, respectively. This approximation is fully justified when looking at the *ab initio* results of the previous section. Then we can write:

$$a^\dagger_{\boldsymbol{k}c} a_{\boldsymbol{k}v}|\emptyset\rangle \simeq a^\dagger_{\boldsymbol{k}B} a_{\boldsymbol{k}A}|\emptyset\rangle \simeq \frac{1}{N}\sum_{\boldsymbol{n},\boldsymbol{m}} a^\dagger_{\boldsymbol{m}B} a_{\boldsymbol{n}A}\, e^{i\boldsymbol{k}\cdot(\boldsymbol{m}-\boldsymbol{n})}|\emptyset\rangle.$$

The sum over the electron and hole positions can be decomposed into a sum over the hole position $\boldsymbol{n}$ and over the electron-hole distance $\boldsymbol{R}$, which is then a vector joining a site on the A (hole) sublattice to a B (electron) sublattice site. For simplicity we use "bra" and "ket" notations:

$$|\boldsymbol{k}vc\rangle = a^\dagger_{\boldsymbol{k}c} a_{\boldsymbol{k}v}|\emptyset\rangle = \frac{1}{\sqrt{N}}\sum_{\boldsymbol{R}} e^{i\boldsymbol{k}\cdot\boldsymbol{R}}|\boldsymbol{R}vc\rangle \quad (5)$$

$$|\boldsymbol{R}vc\rangle = \frac{1}{\sqrt{N}}\sum_{\boldsymbol{n}} a^\dagger_{\boldsymbol{n}A+\boldsymbol{R}}\, a_{\boldsymbol{n}A}|\emptyset\rangle. \quad (6)$$

We see that $|\boldsymbol{R}vc\rangle$ is the linear superposition of exciton amplitudes for pairs separated by $\boldsymbol{R}$. This is nothing but the Bloch wave function for excitons with a wave vector $\boldsymbol{Q} = 0$. We can then rewrite the Bethe-Salpeter-Wannier equation in real space using the $|\boldsymbol{R}vc\rangle$ basis and the wave function coefficients $\Phi_{\boldsymbol{R}} = \langle \boldsymbol{R}vc|\Phi\rangle = \frac{1}{\sqrt{N}}\sum_{\boldsymbol{k}} e^{i\boldsymbol{k}\cdot\boldsymbol{R}}\,\Phi_{\boldsymbol{k}}$. The kinetic energy term $(E_{\boldsymbol{k}c} - E_{\boldsymbol{k}v})\Phi_{\boldsymbol{k}vc}$, diagonal in $\boldsymbol{k}$ space, becomes a tight-binding-like term in $\boldsymbol{R}$ space equal to $\sum_{\boldsymbol{R}'} h_{vc}(\boldsymbol{R}-\boldsymbol{R}')\Phi_{\boldsymbol{R}'}$ with $h_{vc}(\boldsymbol{R}) = \frac{1}{N}\sum_{\boldsymbol{k}} e^{i\boldsymbol{k}\cdot\boldsymbol{R}}\,(E_{c\boldsymbol{k}} - E_{v\boldsymbol{k}})$. To be more precise let H_{vc} be the Hamiltonian acting in the excitonic space. The kinetic energy (or free single-particle) term is the difference of the two Hamiltonians (3) and (4). We rewrite it here, $H^0_{vc} = H_c \otimes 1 - 1 \otimes H_v$, to indicate that each term in the right-hand side acts either on the electronic or on the hole component of the electron-hole states. Dropping now the v,c indices within the bras and kets, the matrix elements $\langle \boldsymbol{R}|H_{vc}|\boldsymbol{R}'\rangle$ of H^0_{vc} are equal to $h_{vc}(\boldsymbol{R}-\boldsymbol{R}')$ and are therefore equal to $2\Delta + 3t^2/\Delta$ if $\boldsymbol{R} = \boldsymbol{R}'$, and to t^2/Δ if $\boldsymbol{R}$ and $\boldsymbol{R}'$ are first neighbors on the triangular lattice.

The electron-hole interaction kernel, on the other hand, contains a direct term and an exchange contribution. Let us first consider the direct term, which is the most important one, as will be checked later (see Table I). The corresponding integral can be expanded in real space and involves integrals of type:

$$-\int d\boldsymbol{r}d\boldsymbol{r}'\varphi_v(\boldsymbol{r}-\boldsymbol{R}_n)\varphi_v(\boldsymbol{r}-\boldsymbol{R}_p) \times \frac{e^2}{|\boldsymbol{r}-\boldsymbol{r}'|}\varphi_c(\boldsymbol{r}'-\boldsymbol{R}_m)\varphi_c(\boldsymbol{r}'-\boldsymbol{R}_q),$$

where the $\varphi_{v(c)}(\boldsymbol{r})$ are (real) valence or conduction orbitals.

TABLE I. The five first excitons in the order fixed by the *ab initio* calculations. Energies are in eV.

	Exciton				
	1 (×2)	2 (×2)	3	4	5 (×2)
Ab initio	−1.932	−1.076	−1.045	−0.980	−0.892
Ab initio without exchange	−2.018	−1.095	−1.045	−1.358	−0.898
Tight binding	−1.932	−1.053	−0.999	−1.0944	−0.830
Symmetry	E	E	A_2	A_1	E

Usually the largest integral is the one where all indices are identical, but here this on-site integral is forbidden since the conduction and valence orbitals belong to different sublattices (actually, this is not completely true as will be discussed below in Sec. III C). The next most important integrals are those where $\boldsymbol{R}_p = \boldsymbol{R}_n$ and $\boldsymbol{R}_q = \boldsymbol{R}_m$ in the previous integral, and finally the (direct) Coulomb term $\sum_{\boldsymbol{k}'}\langle \boldsymbol{k}vc|K^d_{\rm eh}|\boldsymbol{k}'vc\rangle\Phi_{\boldsymbol{k}'vc}$ becomes $\sum_{\boldsymbol{R}\neq 0} U_{\boldsymbol{R}}\Phi_{\boldsymbol{R}}$ where:

$$U_{\boldsymbol{R}} = \langle \boldsymbol{R}|K^d_{\rm eh}|\boldsymbol{R}\rangle = -\int d\boldsymbol{r}d\boldsymbol{r}'\varphi_c^2(\boldsymbol{r})\frac{e^2}{|\boldsymbol{r}-\boldsymbol{r}'|}\varphi_v^2(\boldsymbol{r}'-\boldsymbol{R}),$$

which means that $U_{\boldsymbol{R}}$ acts as a local potential on site $\boldsymbol{R}$. The Coulomb potential should also be screened but here the $U_{\boldsymbol{R}}$ will just be considered as parameters to be fitted to *ab initio* data.

We have therefore reduced our problem to a very simple tight-binding problem for the relative motion of the electron and of the hole. Since the motion is relative we can fix the hole at the origin of the A sublattice. The $\boldsymbol{R}$ vectors lie on the B sublattice: our problem becomes the problem of an electron moving on the B sublattice in the presence of a hole at the origin, which plays the part of an impurity, source of the attractive potential $U_{\boldsymbol{R}}$. To summarize, when exchange effects are neglected, we have to handle the standard tight-binding equations:

$$E\Phi_{\boldsymbol{R}} = \sum_{\boldsymbol{R}'} h_{\rm eh}(\boldsymbol{R}-\boldsymbol{R}')\Phi_{\boldsymbol{R}'} + \sum_{\boldsymbol{R}} U_{\boldsymbol{R}}\Phi_{\boldsymbol{R}},$$

which therefore depends only on $t_{\rm ex} = t^2/\Delta$ and on $U_{\boldsymbol{R}}$.

C. Discussion of the Wannier model

Although standard, the Wannier equations are difficult to solve in many cases because several valence and conduction bands are involved. As a consequence the Wannier functions have no longer direct relationships with the atomic orbitals. On the other hand, in the case of strong screening and small

gaps, the potential does not perturb the single-particle Bloch states too much and $\boldsymbol{k}\cdot\boldsymbol{p}$ expansions can be used, leading to the familiar hydrogenic model where the underlying lattice can finally be forgotten. This model is *a posteriori* justified when the extension of the excitonic states is large compared to the lattice parameter. As shown in Fig. 5 this is not the case here, but the simplicity of the electronic structure of the boron nitride single sheet will allow us to take lattice effects fully in account.

1. A very simple model

As shown above, we have to solve an impurity problem in a simple tight-binding basis. In the case of localized potentials, Green's function or direct diagonalization techniques are known to be very efficient. Actually, we just have to adapt the methods used to study deep impurity centers in semiconductors [49,50]. Although the Coulomb potential is a long-range $1/R$ potential we expect that the energy of the lowest bound state can reasonably be obtained using truncated potentials in real space. Let us recall that the electron is moving on a triangular lattice with first-neighbor hopping integrals in the presence of an impurity located at the origin taken at a lattice point of the hole sublattice, i.e., at the center of a triangle of the B sublattice (see Fig. 6). We have therefore

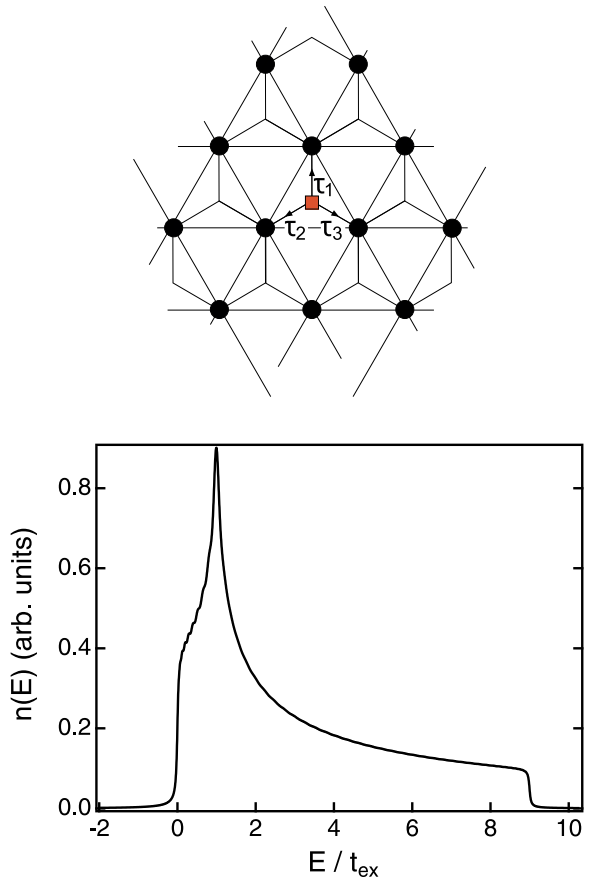

FIG. 6. Top: The hole is at the origin (square) and the electron is moving on a triangular sublattice. Bottom: Density of states corresponding to the spectrum of the excitonic Hamiltonian without Coulomb interactions, $H^0_{\rm eh}-2\Delta$, so that the spectrum starts at $E=0$. The Van Hove singularity is at $E=t_{\rm ex}=1.46\,$eV.

to diagonalize the following Hamiltonian:

$$\begin{aligned} H_{\rm eh} &= H^0_{\rm eh}+U \\ \langle \boldsymbol{R}|H^0_{\rm eh}|\boldsymbol{R'}\rangle &= 2\Delta+3t^2/\Delta \quad \text{if} \quad \boldsymbol{R}=\boldsymbol{R}' \\ &= t^2/\Delta \quad \text{if } \boldsymbol{R} \text{ and } \boldsymbol{R}' \text{ are first neighbors} \\ &= 0 \quad \text{otherwise} \\ U &= \sum_{\boldsymbol{R}} |\boldsymbol{R}\rangle U_{\boldsymbol{R}} \langle \boldsymbol{R}|, \end{aligned}$$

where the potential is to be fitted to *ab initio* data, and where exchange terms are neglected. The spectrum of $H^0_{\rm eh}$ is known since it is the spectrum of the triangular lattice with (positive) first-neighbor hopping integrals (Fig. 6). The lowest eigenvalue (taken as the origin in Fig. 6) is at 2Δ, which corresponds to the energy gap in this model. In the presence of a localized attractive potential, we expect that bound excitonic states appear when the potential is strong enough. This is indeed what happens. A brief analytical discussion is presented in Appendix A. Here, we present results obtained with a potential fitted up to the 28th neighboring shell to *ab initio* data, and which is discussed below, in Sec. III D. We have diagonalized our tight-binding Hamiltonian using a box containing about 10^3 sites on the triangular lattice. Many exciton states have been studied, but since our model becomes inaccurate at higher energies, and when the extension of the exciton increases, we only considered five states in detail. The advantage of our procedure in real space is that we handle real wave functions and so we can easily display the wave function themselves. Furthermore in the case of degenerate states it is easy to show components of definite symmetry.

2. Ground-state exciton and exciton symmetries

We show first in Fig. 7 the results concerning the ground-state exciton, which is doubly degenerate. In the tight-binding case, we show the two components, which are clearly antisymmetric or symmetric with respect to the y axis. The agreement with the *ab initio* result is very good. At this point it is useful to comment on the symmetry of this state. Since we have fixed the position of the hole, we can use the point symmetry of the triangular lattice with respect to the origin located at a center of a triangle. In principle the problem is not purely a 2D one since the π orbitals extend in the z direction and are odd with respect to a $z\to -z$ reflection. Apart from this trivial symmetry, we have only to consider the C_{3v} symmetry with its threefold rotation axis and its mirror planes σ_v. This group is known to have three different representations. Beyond the identity one, A_1, we have the familiar two-dimensional representation $E(x,y)$, and a second representation of dimension 1, A_2, characterized by an odd character for the σ_v reflections. Our exciton clearly has the E symmetry with two (chiral) components, which can be taken to vary as $x+iy$ or $x-iy$.

This exciton is very similar to the so-called A or B excitons met in TMDs. In this case, the Wannier-Mott approach in the $\boldsymbol{k}\cdot\boldsymbol{p}$ approximation is generally used, and the symmetry of the excitons is frequently defined as follows: The exciton wave function is written in the form $\Phi(\boldsymbol{r}_h,\boldsymbol{r}_e)=\phi_{\boldsymbol{k}_0 c}(\boldsymbol{r}_e)\phi_{\boldsymbol{k}_0 h}(\boldsymbol{r}_v)g(\boldsymbol{r}_e-\boldsymbol{r}_h)$, where the ϕ_{k_0} are the

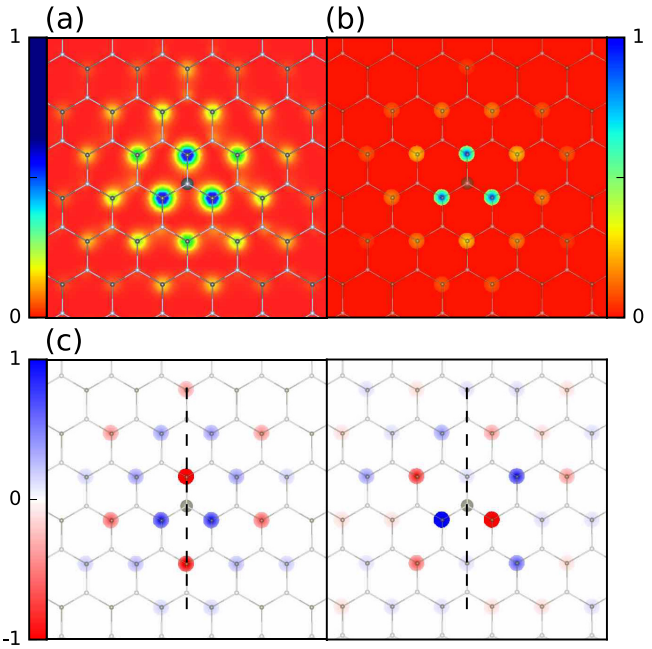

FIG. 7. Results for the ground-state exciton. (a) *Ab initio* intensity. (b) Tight-binding intensity, and (c) tight-binding amplitudes for the two degenerate states, symmetric and antisymmetric, with respect to the y axis. Blue and red colors in (c) correspond to opposite signs.

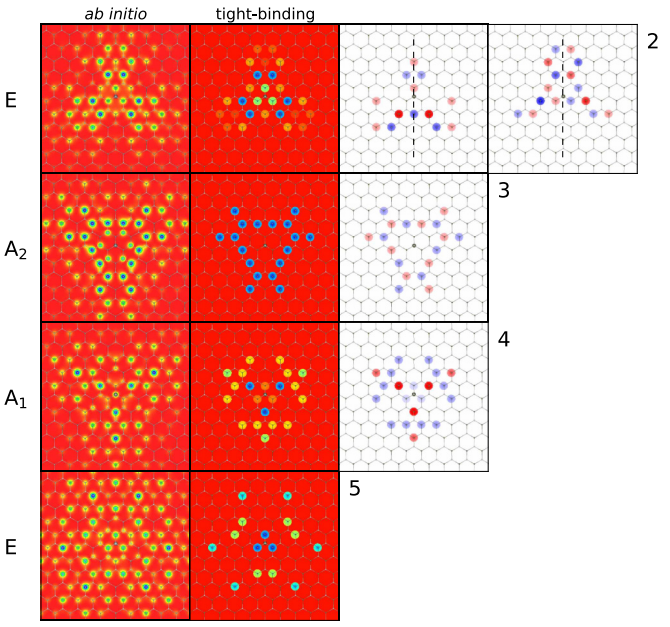

FIG. 8. Results for the excitons Nos. 2–5. Left: *ab initio* intensity. Right: tight-binding intensities and amplitudes. The hole is at the center of the central triangle. Blue and red colors in these plots correspond to opposite signs. Exciton No. 2 has two components as the ground-state exciton. The tight-binding analysis shows clearly that exciton No. 3 has the A_2 symmetry with an antisymmetric behavior with respect to the three σ_v mirrors. Notice in the *ab initio* image the low intensity in the interior hexagon, i.e., on N sites. This is a signal that the simple tight-binding model which forbids this possibility begins to fail. But otherwise, the agreement is very good. Notice also that, due to the A_2 symmetry, the intensity strictly vanishes on the first neighbors and more generally on the symmetry axis. Exciton No. 4 on the other hand as the full C_{3v} symmetry typical of the identity representation A_1. The *ab initio* results show a significant intensity on the central N site. Finally, exciton No. 5 has an E symmetry. Since the amplitude images are fairly complex only the tight-binding intensity is shown.

single-particle Bloch functions at point $\boldsymbol{k_0}$ corresponding to the considered direct gap, and $g(\boldsymbol{r})$, the envelope function, is the solution of the hydrogeniclike excitonic equation for the relative coordinate $\boldsymbol{r} = \boldsymbol{r}_e - \boldsymbol{r}_h$ [51]. The full excitonic symmetry is the symmetry of this product, but notations generally use the symmetry of $g(\boldsymbol{r})$. This decomposition makes sense only if the $\boldsymbol{k} \cdot \boldsymbol{p}$ expansion around $\boldsymbol{k}_0$ is valid, which is not necessarily the case here. In our case, the direct gap occurs at points K and K'. Neglecting intervalley coupling (which may not be valid either), we see that $\boldsymbol{k}_0 = \boldsymbol{K}$ and the product of Bloch functions $\phi_{\boldsymbol{k_0}c}(\boldsymbol{r}_e)\phi_{\boldsymbol{k_0}h}(\boldsymbol{r}_v)$ varies as $e^{i\boldsymbol{K}\cdot\boldsymbol{r}}$, i.e., as one E component of the representation of the C_3 symmetry at point K. Now, since the conduction and valence bands are nondegenerate at point K, the ground-state envelope function $g(\boldsymbol{r})$ is isotropic and of symmetry s. This is why the A exciton is denoted a $1s$ exciton. With similar arguments we obtain that the exciton at K' has also s symmetry modulated by the Bloch function proportional to $e^{-i\boldsymbol{K}\cdot\boldsymbol{r}}$. So, in this description we obtain two degenerate $1s$ excitons, but if they are considered together they form a doubly degenerate exciton of symmetry E. Both descriptions are equivalent in the case of large excitons, which can be associated separately to points K and K' [14]. In our case, where the exciton is more localized in real space (and consequently more delocalized in reciprocal space), using directly the full point symmetry of the exciton is more accurate.

3. *Other excitons*

a. Analysis in real space. At higher energy, a group of six states appears. All of them as well as the previous state have similar energies within 0.1 eV. Actually, they do not appear in the same order in *ab initio* and tight-binding calculations. Their wave functions are however very similar. We follow here the order provided by the *ab initio* calculations. Both *ab initio* and tight-binding approaches find first a similar exciton with again a twofold degeneracy (exciton No. 2 in Fig. 8). It has therefore also an E symmetry. The agreement between both calculations is still fairly good. The two following ones are nondegenerate. The TB method shows unambiguously that the first one transforms according to the A_2 representation (exciton No. 3, Fig. 8) and the second one according to the A_1 one (exciton No. 4, Fig. 8). Finally, the two upper states are degenerate and belong to the E symmetry (fifth exciton in Table I). We summarize in Table I the energies and symmetries of these excitons. The next excitons are found more than 0.2 eV above this group in the *ab initio* calculations.

Although the overall agreement between *ab initio* and TB calculations is fairly good, the behavior of the A_1 exciton seems particular. This is still more obvious if we compare *ab initio* calculations performed with and without the exchange contribution (Table I). Whereas the energy variation between the two calculations for the other excitons is of a few percent, the A_1 exciton is strongly perturbed, its energy moving from -0.98 eV to -1.358 eV when the (repulsive) exchange

contribution is suppressed. This is quite unusual but can be related to the fact that the A_1 exciton is, by symmetry, the only one where the electronic intensity at the origin is nonvanishing. Actually, although we have neglected this possibility in our simplified TB model, the *ab initio* calculations do show such a nonvanishing intensity (Fig. 8). The point is that even if this intensity is low, it introduces a perturbation proportional to the on-site exchange term $2J_0$, which is very large. Actually, the local Coulomb and exchange integrals U_0 and J_0 are of the same order of magnitude since they both involve similar π orbitals. They are of opposite signs however, which explains why our TB scheme, which neglects on-site Coulomb interactions, is not too bad even in this case.

b. 2s and 2p states: Analysis in reciprocal space. At this point it is instructive to compare our results with those obtained for dichalcogenides. In TMD the Wannier-Mott model is valid provided appropriate anisotropic potentials are used. It is convenient to analyze the excitons in each valley in terms of $s, p, \ldots$ symmetries. The usual sequence is a $1s$ level, and then a $2s$ level nearly degenerate with a $2p$ level. This $2p$ level gives rise to four states because of the valley degeneracy. A careful examination of the symmetry of the π and π^* states close to the K points based on the so-called massive Dirac model has shown that actually the degeneracy within each valley is lifted, but time-reversal symmetry between K and K' insures that the $2p$ states are split into two doubly degenerate states. Furthermore, the $2s$ level is found to be above the $2p$ levels [14,52–54]. In our case where lattice effects and therefore intervalley effects are included a further splitting occurs. As pointed out above, the full symmetry of the exciton states is obtained from the product of the envelope function (E symmetry for p states) and of the Bloch functions at points K of symmetry E also. Now, the decomposition of the $E \times E$ representation gives rise precisely to the observed one: $E \times E = E + A_1 + A_2$. In the $s, p, \ldots$ language, the p states are first split into $p_x \pm ip_y$ states whose chiralities are equal or opposite to those of the Bloch functions at points K and K'. Hence, two states (one in each valley) have a vanishing global chirality. Forming bonding and antibonding states between these states leads to the A_1 and A_2 states. The two other ones remain degenerate and form an E state. We expect the bonding state A_1 to be below A_2, but, as argued before, the repulsive exchange contribution neglected in this discussion pushes the A_1 upwards. This is described in Fig. 9.

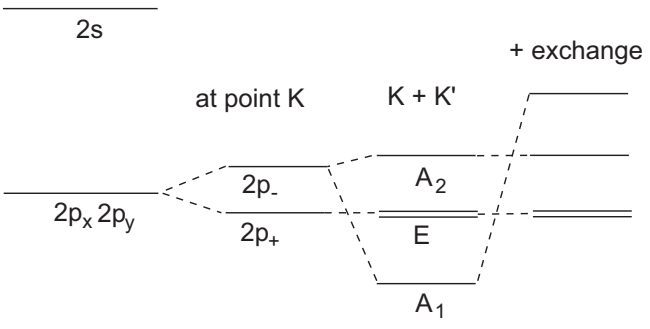

FIG. 9. Schematic splitting scheme of the $2p$ levels. Depending on the calculations, *ab initio* or TB, the level separations are of the order of 0.1 eV or less. The $1s$ state is about 1 eV below.

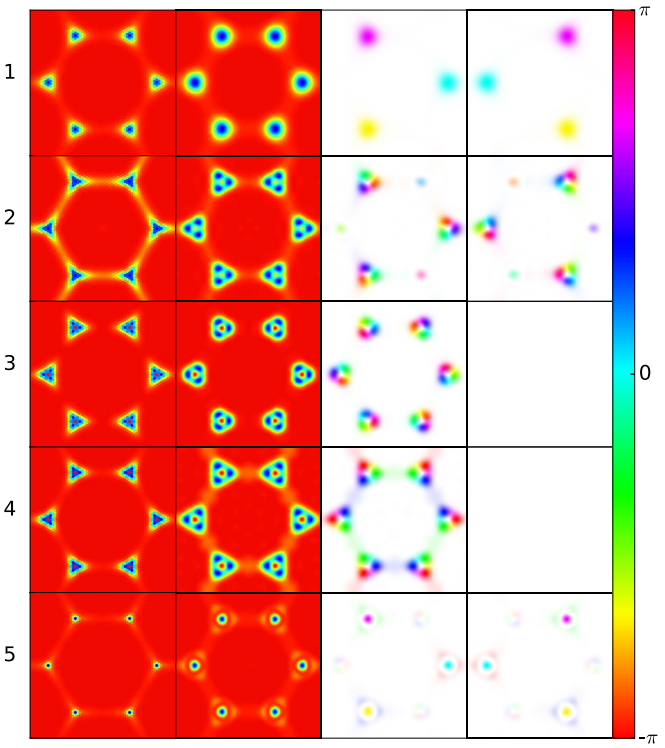

FIG. 10. *Ab initio* and TB results for the intensities and wave functions in reciprocal space for the excitons No. 1 to No. 5. First column: *Ab initio* results. The first and last excitons show intensities peaked at K and K', more concentrated for the last $2s$ state than for the first $1s$ one, as expected. This is actually the way state No. 5 is identified as $2s$. The three other states (four, including degeneracy) have similar shapes with significant trigonal warping effects. The TB results are shown in the second column and are very similar. The last two columns show the TB wave function with a modulus-phase representation: a circle with an opacity proportional to the intensity at this k point is placed on each point of a grid. The color of the circle is related to the phase between $-\pi$ and $+\pi$ as indicated in the color bar. Two plots are shown in case of degenerate states, which have been filtered according to the chiral symmetries. It is clear that the phases are equal to 0 and $\pm 2\pi/3$ at these points for the $1s$ and $2s$ states. The four other states have more rich structures with phases rotating within each triangular spot located at the K points and, as explained in the main text, are a signature of $2p$ states.

It remains to determine which E exciton belongs to this $2p$ family. This might be exciton No. 2, as shown in the figure, or exciton No. 5. It is not obvious to decide from the plots of the wave functions in real space, but we show now that an analysis in reciprocal space provides the answer. In the first column of Fig. 10 are shown the intensities $|\Phi_k|^2$ of the five excitons considered previously. It is clear at once that excitons Nos. 2, 3, and 4 belong to the same family and are therefore the expected $2p$ excitons, which means that the relevant E exciton is exciton No. 2 as depicted in Fig. 9. A consequence is that exciton No. 5 is the $2s$ exciton. Furthermore, we check that the $2s$ state, which is more spread out around the origin in real space than the $1s$ state, is more concentrated on the K points in reciprocal space.

Tight-binding calculations lead to quite similar results as can be seen in the second column of Fig. 10. The advantage of

the TB method is that we can easily obtain the wave functions themselves. It is possible also to have a modulus-phase representation by placing on each point of a grid a circle with color related to the phase between $-\pi$ and $+\pi$ with an opacity proportional to the intensity at this k point, which is shown in the last columns of Fig. 10. The four $2p$ states have more rich structures with phases rotating within each triangular spot located at the K points. They can be explained if we use the model recalled above where the symmetry of the exciton states is governed by the product of the Bloch functions and an envelope function so that the exciton wave function is proportional to $e^{i\boldsymbol{K_1}\cdot\boldsymbol{r}}g(\boldsymbol{r})$, where $\boldsymbol{K_1}$ is one particular $\boldsymbol{K}$ vector among the three equivalent ones. In the discrete tight-binding model $\boldsymbol{r}$ is the sum of a vector of the triangular lattice and of any first-neighbor vector $\boldsymbol{\tau}$. Assume now that $g(\boldsymbol{r})$ is a s envelope function and only depends on the modulus r of $\boldsymbol{r}$. Under a rotation of angle $2\pi/3$, $\boldsymbol{K}_1$ is transformed into another equivalent vector, modulo a vector $\boldsymbol{G}$ of the reciprocal lattice. The exciton state is therefore multiplied by a phase factor equal to $e^{i\,\boldsymbol{G}\cdot\boldsymbol{\tau}} = \omega$ or ω^2, where $\omega = e^{2i\pi/3}$ is the cubic root of unity, depending on the initial orientation of the lattice with respect to the origin. As mentioned previously, the wave function transforms as the component E_+ of positive chirality of the representation E, and the wave function $\Phi_{\boldsymbol{k}}$ is given, up to a constant by:

$$\Phi_{\boldsymbol{k}} \propto \int d\boldsymbol{r}\; e^{-i\,(\boldsymbol{K}_1-\boldsymbol{k})\cdot\boldsymbol{r}}\, g(r)n(\boldsymbol{r}),$$

where $n(\boldsymbol{r})$ is the site density, i.e., the sum of Dirac functions on the triangular lattice sites, proportional to $\sum_{\boldsymbol{G}} e^{i\boldsymbol{G}\cdot(\boldsymbol{r}-\boldsymbol{\tau})}$.

For $\boldsymbol{k}$ close to $\boldsymbol{K}_1$, $\boldsymbol{k} = \boldsymbol{K}_1 + \boldsymbol{q}$, we can neglect the variation of $n(\boldsymbol{r})$, i.e., keep only the $\boldsymbol{G} = 0$ term, and the integral over the angle yields $\Phi_{\boldsymbol{k}} \propto \int r dr J_0(qr)g(r)$, where $J_0(x)$ is the Bessel function of zero order. If $g(r)$ is peaked at some average value $\bar{r}$ (remember that $\bar{r}$ is at least equal to the minimum hole-electron distance $a/\sqrt{3}$), $\Phi_{\boldsymbol{k}} \propto J_0(q\bar{r})$ and is therefore peaked at $q = 0$ and the intensity is maximum in a circle of radius $\sim 1/\bar{r}$. If $\boldsymbol{k}$ is close to another vector, say $\boldsymbol{K_2}$, then the integral is multiplied by a factor $e^{-i\boldsymbol{G}\cdot\boldsymbol{\tau}}$, where $\boldsymbol{G} = \boldsymbol{K_2} - \boldsymbol{K_1}$, i.e., by a factor ω^2 or ω. Actually the functions in reciprocal space have the same symmetry properties as in real space.

Assume now that $g(\boldsymbol{r})$ has a p symmetry, so that $g(\boldsymbol{r}) = g(r)e^{\pm i\varphi_{\boldsymbol{r}}}$, where $\varphi_{\boldsymbol{r}}$ is the angle between $\boldsymbol{r}$ and the x axis. Here also, the integral over this angle can be performed explicitly, so that $\Phi_{\boldsymbol{k}} \propto \pm J_1(q\bar{r})e^{\pm i\varphi_{\boldsymbol{q}}}$, where $J_1(x)$ is the Bessel function of order one, and $\varphi_{\boldsymbol{q}}$ is now the angle of $\boldsymbol{q}$ with the x axis. We conclude that the phase rotates within each circle centered on the K points. This is clearly as shown in Fig. 10. This proves definitely that the states Nos. 2–4 are of symmetry p. More precisely, consider first the nondegenerate exciton No. 3 and No. 4. Exciton No. 4 does show the A_1 symmetry already evidenced in real space. Furthermore, one can notice that the phases rotates in opposite directions around the K and K' points, and that these rotations are counterbalanced by the rotations between different points of the same family (K or K'), in full agreement with the arguments put forward above. The same is true for exciton No. 3 except that the amplitudes are odd with respect to the σ_v mirrors (phase shift of π). One can also notice the signature of the J_1 Bessel functions: The intensities vanish at the origin of the spots, disappear at larger distances than in the case of s states, and are maximum in between. Actually, warping effects along the Γ-M lines transform the circles into triangles. The case of the degenerate E state is more complex, since each component seems to mix the behaviors of s and p states. Mixing between different E states is allowed indeed (this is also the case for the $2s$ exciton). It is clear however that the main features correspond to rotating phases within circles, and one can check that here the rotations *within* the circles and those *between* the circles are in the same direction.

To summarize, although the first excitons considered here are fairly localized, with important lattice effects, they can be classified to some extent within a scheme borrowed from the 2D atomic terminology ($1s$,$2s$,$2p$, ... states) and which has already been successfully applied to TMD. The genuine symmetry of the exciton states is, however, more precisely related to the representations of the triangular point group. We have also seen that exchange effects are unusually strong for fully invariant states.

D. Fit of the potential

In a first approach we have tested a screened Coulomb potential, but it was quickly apparent that it was not possible in this way to reproduce accurately more than the first exciton. Meanwhile, several developments in the literature were convincingly arguing that it is not possible in two dimensions to use such a potential and that a genuine 2D electrostatic potential [55] should be used instead [14,21,22,27,56–58]. We have therefore used the Keldysh potential[2]:

$$V_{2D}(r) = \frac{\pi e^2}{2r_0}\left[H_0\left(\frac{r}{r_0}\right) - Y_0\left(\frac{r}{r_0}\right)\right],$$

where the only parameter is the screening length r_0, directly related to the 2D polarizability. Finally, since we are dealing with relative binding energies, our model only depends on two parameters, the excitonic hopping integral $t_{\rm ex} = t^2/\Delta$ and r_0. Of course, the Keldysh potential is still defined within a continuous approach, which has no reason to apply exactly here where lattice effects are important. In the best fit, the hopping integral $t_{\rm exc} = t^2/\Delta$ is found equal to 1.50 eV, so that $t = 2.33$ eV, which is completely consistent with our value $t = 2.30$ eV deduced from *ab initio* band structures. Finally, we find $r_0 = 10.0$ Å. When r is much larger than r_0 the potential tends to an unscreened Coulomb potential. Below r_0 the potential is screened and becomes logarithmic. Since the first electronic shell around the hole is at a first-neighbor (B-N) distance, about 1.45 Å, we see in Fig. 11 that we are here in the screened regime where the potential is slowly varying. In Fig. 11 we show the effective distance-dependent dielectric constant defined from $V(r) = e^2/\epsilon(r)r$. Notice also that the distance between the two first excitons is completely different from that predicted by the 2D hydrogenic model. This is due in part to these screening effects, but also and more importantly to lattice effects coupled with the specific electronic structure of hBN, since such deviations have already been observed

[2]More precisely, we have used the simplified form proposed in Ref. [21].

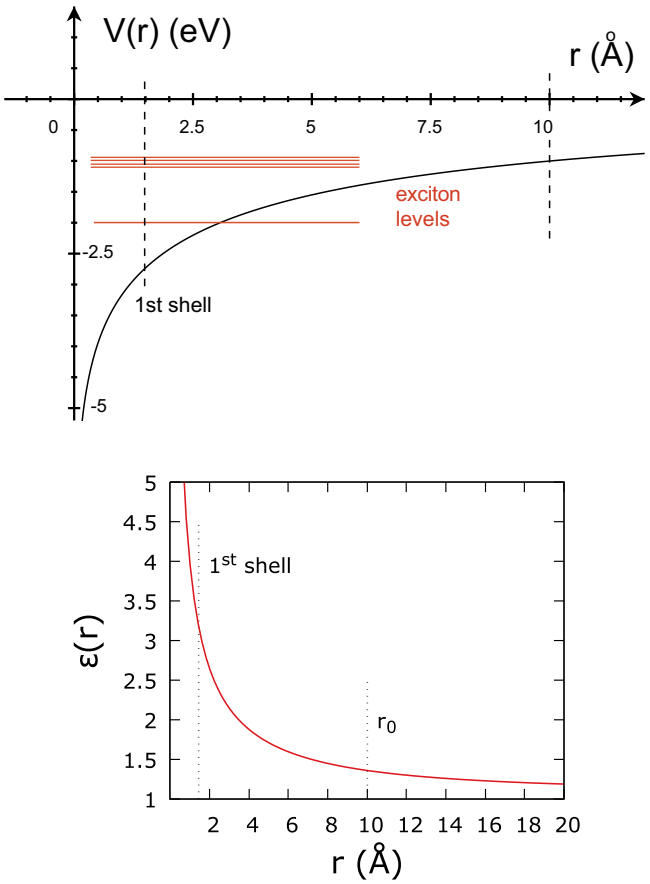

FIG. 11. Keldysh potential corresponding to the case of hBN-SL; $r_0 = 10.0$ Å and effective dielectric constant $\epsilon(r)$ defined from $V(r) = e^2/\epsilon(r)r$.

with a fixed dielectric constant (see also Ref. [14]). The main reason is probably that it is nearly forbidden for the hole and the electron to be at the same position, which penalizes principally the binding energy of the ground-state exciton.

We have neglected the exchange term here. Actually, its short-range part contributes to the splitting between the spin singlet and triplet states, the latter being dark in the absence of spin-orbit coupling. This is a repulsive (positive effect) absent in the triplet term whose level should therefore be below the singlet one. *Ab initio* calculations predict a splitting about 90 meV for the $1s$ exciton [59]. In a first-order perturbation calculation this splitting is equal to the average of $K^x_{\rm eh}$ in the considered excitonic state. In the case of the ground-state exciton, the excitonic wave function is concentrated on the first-neighbor shell, so that this splitting is equal to a fraction of $J_{\boldsymbol{\tau}}$. As expected then $J_{\boldsymbol{\tau}} \simeq 0.1 - 0.3$ eV $\ll U_{\boldsymbol{\tau}}$. This is also completely consistent with the variation shown in Table I of the exciton energy when suppressing the exchange term. The particular case of the A_1 exciton, which is very dependent on the intra-atomic values U_0 and J_0 of the Coulomb and exchange potentials has been discussed above. A simple perturbation method improving the TB model used here can be derived to discuss this effect in more detail and is described in Appendix C.

To summarize, the simple tight-binding model for the excitons in hBN-SL is remarkably successful, even at a quantitative level and the comparison with *ab initio* calculations shows that the effective screened Coulomb potential to be used in a continuous model is really a potential of the Keldysh type in its strongly screened regime.

IV. OPTICAL MATRIX ELEMENTS

The optical absorption is related to transitions from the ground state (energy $E_\emptyset$) to final states of energy $E_i = E_\emptyset + \hbar\omega$ and to the corresponding matrix elements of the perturbation induced by the electromagnetic field. Each transition i is characterized by an oscillator strength $f_i = 2m|\langle\emptyset|\boldsymbol{v}\cdot\hat{\boldsymbol{e}}|i\rangle|^2/\hbar\omega_i$. Here $\hat{\boldsymbol{e}}$ is the (unit) vector of the light polarization, and $\boldsymbol{v}$ is the velocity operator.

A. Matrix elements between single-particle states

In the absence of excitonic effets we have just to calculate the matrix elements between valence and conduction Bloch states with identical $\boldsymbol{k}$ vectors. It is not difficult to calculate them in the general case [54,60,61], but here we just detail the calculation for states close to the gap where we know that the Bloch functions live on separate triangular sublattices. Then

$$\langle \boldsymbol{k}v|\boldsymbol{v}|\boldsymbol{k}c\rangle = \frac{1}{N}\sum_{\boldsymbol{n},\boldsymbol{m}} e^{i\boldsymbol{k}\cdot(\boldsymbol{m}-\boldsymbol{n})}\langle \boldsymbol{n}A|\boldsymbol{v}|\boldsymbol{m}B\rangle.$$

There is no unique way of calculating the matrix elements of $\boldsymbol{v}$, depending on whether we express it using the momentum operator [51] or the relation $\boldsymbol{v} = [\boldsymbol{r}, H]/i\hbar$. Both methods are equivalent in an exact treatment but not when using an incomplete basis as in our tight-binding basis. The second method has the disadvantage to use the $\boldsymbol{r}$ operator, which is not always well defined in periodic systems. This is not the case in our model and the advantage is precisely to work in real space, assuming that $\boldsymbol{r}|\boldsymbol{n}\rangle \simeq \boldsymbol{n}|\boldsymbol{n}\rangle$, and therefore:

$$\langle \boldsymbol{n}A|\boldsymbol{v}|\boldsymbol{m}B\rangle = -\frac{1}{i\hbar}(\boldsymbol{m}-\boldsymbol{n})t, \tag{7}$$

if $\boldsymbol{m}$ and $\boldsymbol{n}$ are first neighbors (on the honeycomb lattice), and zero otherwise. Then

$$\langle \boldsymbol{k}v|\boldsymbol{v}|\boldsymbol{k}c\rangle = \frac{1}{N}\sum_{\boldsymbol{n},\boldsymbol{m}} e^{i\boldsymbol{k}\cdot(\boldsymbol{m}-\boldsymbol{n})}\frac{it}{\hbar}(\boldsymbol{m}-\boldsymbol{n})$$
$$= \frac{it}{\hbar}\sum_{\alpha} e^{i\boldsymbol{k}\cdot\boldsymbol{\tau}_\alpha}\boldsymbol{\tau}_\alpha = \frac{it}{\hbar}\nabla_k\gamma(\boldsymbol{k}).$$

In the limit $\boldsymbol{k}\to K$, one finds $\nabla_k\gamma(\boldsymbol{k}) \simeq -\frac{3}{2}ia(\hat{\boldsymbol{x}} + i\hat{\boldsymbol{y}})$, where a is here the nearest-neighbor distance, i.e., the lattice parameter divided by $\sqrt{3}$ and $\hat{\boldsymbol{x}}$ and $\hat{\boldsymbol{y}}$ are the unit vectors along the x axis and the y axis, respectively, so that finally:

$$\langle \boldsymbol{K}v|\boldsymbol{v}\cdot\hat{\boldsymbol{e}}|\boldsymbol{K}c\rangle| \simeq v|e_x + ie_y|; \quad \hbar v = 3at/2.$$

Notice that Δ/v^2 is the effective mass m^* of the conduction and valence bands at point K. Using $\Delta \simeq 3$ eV and $v \simeq 1$ km/s (as the Fermi velocity of graphene is precisely given by $3at/2\hbar$), we obtain $m^*/m \simeq 0.54$.

At point K', $e_x + ie_y$ is replaced by $e_x - ie_y$. The matrix elements are maximum for circularly polarized light and, for linearly polarized light, the matrix element is constant and equal to v. The oscillator strength is equal to $(m/m^*)(2\Delta/\hbar\omega)$

when $\hbar\omega$ is larger than the gap, i.e., about 2 close to the gap (equal to 2Δ). The absorption, proportional to $\sum_i f_i/\hbar\omega_i$, is therefore proportional to the density of states of the triangular lattice divided by $(\hbar\omega)^2$. Its shape is characterized by a discontinuity at the edge and a Van Hove singularity at a distance equal to $t_{\rm exc}$ above the edge, as shown in Fig. 6.

B. Matrix elements between excitonic states

In the presence of excitons we have now to calculate the matrix element $\langle\emptyset|\boldsymbol{v}\cdot\hat{\boldsymbol{e}}|\Phi\rangle$, where $|\Phi\rangle$ is the exciton state. From (6) and (7), we see that:

$$\boldsymbol{v}|\emptyset\rangle = \frac{-t}{i\hbar}\sum_{\boldsymbol{n},\boldsymbol{m}}{}'(\boldsymbol{m}-\boldsymbol{n})a^\dagger_{\boldsymbol{m}B}a_{\boldsymbol{n}A}|\emptyset\rangle = \frac{it}{\hbar}\sqrt{N}\sum_\alpha \boldsymbol{\tau}_\alpha|\boldsymbol{\tau}_\alpha\rangle,$$

so that:

$$\langle\emptyset|\boldsymbol{v}\cdot\hat{\boldsymbol{e}}|\Phi\rangle = \frac{-it}{\hbar}\sqrt{N}\sum_\alpha \hat{\boldsymbol{e}}\cdot\boldsymbol{\tau}_\alpha\langle\boldsymbol{\tau}_\alpha|\Phi\rangle.$$

Defining the dipole $\boldsymbol{d}_\Phi$ associated with the exciton Φ through:

$$\boldsymbol{d}_\Phi = \sum_\alpha \boldsymbol{\tau}_\alpha\langle\boldsymbol{\tau}_\alpha|\Phi\rangle,$$

we see that $|\langle\emptyset|\boldsymbol{v}\cdot\hat{\boldsymbol{e}}|\Phi\rangle| = (t\sqrt{N}/\hbar)|\hat{\boldsymbol{e}}\cdot\boldsymbol{d}_\Phi|$. Thus, only the local components of the exciton wave function contribute to the optical matrix element. This is completely equivalent to the statement that, within the usual hydrogenic model, only s states contribute (Elliott theory, see Refs. [35,50]). Here we have an equivalent selection rule: The dipole $\boldsymbol{d}_\Phi$ should not vanish; in particular the wave function $\langle\boldsymbol{R}|\Phi\rangle$ should have finite components on the first neighbors $\boldsymbol{R}=\boldsymbol{\tau}_\alpha$.

C. Application to the five first excitons of hBN-SL

Consider first the ground-state E exciton. It has two components Φ^+ and Φ^-, which can be chosen as those corresponding to circular polarizations, so that the components are the cubic roots of unity, $\langle\boldsymbol{\tau}_\alpha|\Phi^\pm\rangle = C_\Phi e^{\pm\frac{2i\pi}{3}(\alpha-1)}$, and $\hat{\boldsymbol{e}}\cdot\boldsymbol{d}_\Phi = -\frac{3}{2}a\,C_\Phi(e_x \pm ie_y)$, and finally $|\langle\emptyset|\boldsymbol{v}\cdot\hat{\boldsymbol{e}}|\Phi\rangle|/\sqrt{N} = C_\Phi v|e_x+ie_y|$. Here, C_Φ is the amplitude of the exciton state on the first neighbors, at most equal to $1/\sqrt{3}$. In the case of single-particle transitions the oscillator strength was of the order of mv^2/Δ for a transition close to the gap; hence, a total oscillator strength of the order of N times this value. We see here that the oscillator strength of the exciton is of the same order of magnitude if C_Φ is large, i.e., if the exciton is strongly localized, which is the case here. Actually, from *ab initio* calculations, the weight C_Φ^2 is found about one third its maximum value 1/3. In other words 30% of the weight of the ground state is concentrated on the first triangular shell. TB calculations on the other hand find a weight of about 50%. The oscillator strength of the other excitons are smaller. The second exciton as well as the last one (No. 5) has the same symmetry as the first one but their amplitude on the first neighbors is weak. The two other ones studied above (Nos. 3–4) are dark because their symmetry are characterized by representations A_1 and A_2 different from the vectorial representation E, so that $\boldsymbol{d}_\Phi = 0$ and this is confirmed by the *ab initio* calculations. Finally, the ground-state exciton takes almost all the oscillator strength.

V. DISCUSSION

The excitons of hBN-SL have been characterized in detail. The first one, of lowest energy, is particularly localized. Is it a Frenkel or Wannier-Mott exciton? This discussion is somewhat semantic. It is in some sense similar to the longstanding debate between the Heitler-London (atomic) approach and the Hund-Mulliken (molecular) approach to single-particle properties. In practice, it turns out that in solids the Hund-Mulliken band approach is more fruitful since it can deal with many situations except when correlations effects are very strong. Even then, specific approaches in the Hubbard style can be used and compete with the methods of quantum chemistry (interaction configuration approach). In between, the tight-binding method has proven very efficient to deal with electrons sharing itinerant properties (conductivity) and localized ones (magnetism, chemical bonding). We are certainly here in a similar situation. The localized excitons of hBN-SL can be described within a TB-Wannier framework, but cannot be described accurately within a $\boldsymbol{k}\cdot\boldsymbol{p}$ approach similar to the nearly free electron approach of electronic properties. They could simply be described as tightly bound excitons. We have shown indeed that in the case of hBN-SL, which is a genuine case study, the tight-binding approach can be very accurate by fitting to *ab initio* data a few parameters.

On the experimental side optical properties of hBN-SL are not available yet, but there are already some indications that the expected main exciton is observed. In the case of bulk hBN, stacking effects induce splittings of this exciton level which are observed. Progress in the analysis of these stacking effects is being made. Finally, dispersion effects as well as exciton-phonon coupling remain to be studied.

Note added in proof. A recent paper presents a model for excitons in dichalcogenides whose spirit is quite similar to our tight-binding model [62].

ACKNOWLEDGMENTS

A.M.-S. and L.W. acknowledge support from the National Research Fund, Luxembourg (Projects C14/MS/773152/FAST-2DMAT and INTER/ANR/13/20/NANOTMD). F.D. and H.A. are indebted to L. Schué, J. Barjon, and A. Loiseau for numerous fruitful discussions. The research leading to these results has received funding from the European Union Seventh Framework Programme under Grant Agreement No. 604391 Graphene Flagship. We acknowledge funding by the French National Research Agency through Project No. ANR-14-CE08-0018.

APPENDIX A: A VERY SIMPLE MODEL FOR THE GROUND-STATE EXCITON

The ground-state exciton is so localized that its properties do not depend too much on the long-range part of the potential. It is useful, therefore, to examine the properties of a model where the range of the potential is limited to the three first neighbors of the central hole. We have then to diagonalize the

following Hamiltonian:

$$\begin{aligned} H_{\rm eh} &= H^0_{\rm eh} + U \\ \langle \boldsymbol{R}|H^0_{\rm eh}|\boldsymbol{R}'\rangle &= 3t_{\rm exc} \quad \text{if} \quad \boldsymbol{R} = \boldsymbol{R}' \\ &= t_{\rm exc} \quad \text{if} \quad \boldsymbol{R} \quad \text{et} \quad \boldsymbol{R}' \quad \text{are first neighbors} \\ &= 0 \quad \text{otherwise,} \\ U &= \sum_{\boldsymbol{R}=1,2,3} |\boldsymbol{R}\rangle u \langle \boldsymbol{R}|, \end{aligned}$$

where the three sites surrounding the hole at the origin are labeled 1,2,3.

1. Green's functions

The resolvent or Green's function corresponding to this Hamiltonian is $G(z) = (z - H_{\rm eh})^{-1}$. $G^0(z) = (z - H^0_{\rm eh})^{-1}$ is then the Green's function of the triangular lattice. With the chosen origin of energies, the spectrum of $H^0_{\rm eh}$ starts at $E = 0$ with a jump equal to $\pi\sqrt{3}\,t_{\rm exc}$. In the presence of the attractive potential u we can have a bound state if the determinant of the operator $(1 - G^0U)$ within the space of dimension 3 generated by the three states $|1\rangle$,$|2\rangle$, and $|3\rangle$ vanishes. Let now F_0 and F_1 be the diagonal and off-diagonal matrix elements of G^0, respectively $F_0 = \langle 1|G^0|1\rangle = \langle 2|G^0|2\rangle = \langle 3|G^0|3\rangle$; $F_1 = \langle 1|G^0|2\rangle = \langle 2|G^0|3\rangle = \langle 3|G^0|1\rangle$. We find a double solution $(F_0 - F_1) = 1/u$ and a simple solution $(F_0 + 2F_1) = 1/u$. Using standard methods one can determine the behavior of F_0 and F_1 close to the origin. First, assuming a constant density of states $n(E) \simeq 1/W$, we find that:

$$F_0 = \frac{1}{W} \log \frac{z}{z - W},$$

so that $F_0(E < 0) \simeq \frac{1}{W} \log(|E|/W)$ when E is close to 0. Similarly, F_1 is found to behave as $-F_0/2$. Then $F_0 - F_1$ is negative and diverges logarithmically below $E = 0$: $F_0 - F_1 \simeq (3/2W)\log(|E|/W)$ whereas $F_0 + 2F_1$ tends to a constant. As a result the first equation has always a negative solution $E_{\rm exc}$ for E, such that $F_0(E) - F_1(E) = -1/|u|$. With the previous model for F_0 and F_1, we obtain, for small values of $|u|$, $|E_{\rm exc}|/W = \simeq \exp(-\frac{2W}{3|u|})$. If $|u|$ is large, $E_{\rm exc} \simeq u$. The corresponding eigenstates are, as expected, the chiral states $|\phi^+\rangle \propto |1\rangle + \omega|2\rangle + \omega^2|3\rangle$ and $|\phi^\pm\rangle \propto |1\rangle + \omega^2|2\rangle + \omega|3\rangle$, where ω is the cubic root of unity, $\omega = e^{2i\pi/3}$. We recover our exciton of symmetry E. The components of $|\Phi^\pm\rangle$ beyond the (1, 2, 3) cluster can be obtained from the equation $|\phi^\pm\rangle = G^0U|\phi^\pm\rangle$, i.e., $\langle \boldsymbol{R}|\phi^\pm\rangle = \sum_{\boldsymbol{R}'=1,2,3} \langle \boldsymbol{R}|G^0|\boldsymbol{R}'\rangle u \langle \boldsymbol{R}'|\phi^\pm\rangle$.

2. Reciprocal space

We can also express $|\phi^\pm\rangle$ in reciprocal space:

$$\begin{aligned} \Phi^\pm_{\boldsymbol{k}} = \langle \boldsymbol{k}|\phi^\pm\rangle &= \frac{1}{\sqrt{N}} \sum_{\boldsymbol{R}} e^{-i\boldsymbol{k}\cdot\boldsymbol{R}} \langle \boldsymbol{R}|\phi^\pm\rangle \\ &\simeq \frac{1}{\sqrt{3N}} \gamma(\pm\boldsymbol{K} - \boldsymbol{k}), \end{aligned}$$

where we have limited the sum to the first neighbors and taken into account that $e^{i\boldsymbol{K}.\boldsymbol{R}} = 1, \omega, \omega^2$ when $\boldsymbol{R} = 1,2,3$ provided K is chosen along the x axis, as in Fig. 1. Thus,

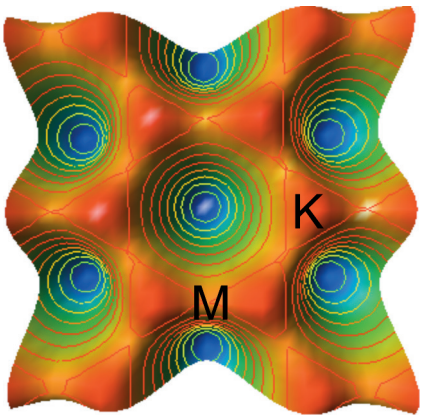

FIG. 12. Tight-binding weight of the ground-state exciton in reciprocal space.

up to a normalization constant the weight $|\Phi^\pm_{\boldsymbol{k}}|^2$ is equal to $|\gamma(\boldsymbol{k} \mp \boldsymbol{K})|^2$. Since $|\gamma(\boldsymbol{k})|^2$ is maximum when $\boldsymbol{k} = 0$, we see that $|\Phi^+_{\boldsymbol{k}}|^2$ and $|\Phi^-_{\boldsymbol{k}}|^2$ are peaked at points K and $K' = -K$, respectively. As expected the sum is maximum on the boundary of the Brillouin zone, as shown in Fig. 12.

APPENDIX B: EXCHANGE CONTRIBUTION

The exchange contribution involves integrals of type:

$$+ \int d\boldsymbol{r} d\boldsymbol{r}' \varphi_v(\boldsymbol{r} - \boldsymbol{R}_n)\varphi_v(\boldsymbol{r}' - \boldsymbol{R}_p) \times \frac{2}{|\boldsymbol{r} - \boldsymbol{r}'|} \varphi_c(\boldsymbol{r} - \boldsymbol{R}_m)\varphi_c(\boldsymbol{r}' - \boldsymbol{R}_q).$$

The largest integrals correspond to cases where the overlap is minimum for the $\boldsymbol{r}$ and $\boldsymbol{r}'$ integrations. However, because we have forbidden site coincidence for valence and conduction orbitals, the best we can do is to consider first-neighbor overlap between $\varphi_v(\boldsymbol{r} - \boldsymbol{R}_n)$ and $\varphi_c(\boldsymbol{r} - \boldsymbol{R}_m)$ and between $\varphi_v(\boldsymbol{r}' - \boldsymbol{R}_p)$ and $\varphi_c(\boldsymbol{r}' - \boldsymbol{R}_q)$. Therefore, we only keep the terms $\boldsymbol{R}_m = \boldsymbol{R}_n + \boldsymbol{\tau}$ and $\boldsymbol{R}_q = \boldsymbol{R}_p + \boldsymbol{\tau}'$, where $\boldsymbol{\tau}$ and $\boldsymbol{\tau}'$ are first neighbors on the honeycomb lattice. The largest terms occur when $n = p$, and the integral becomes a function $J(\boldsymbol{\tau}, \boldsymbol{\tau}'\boldsymbol{\rho})$ of $\boldsymbol{\tau}, \boldsymbol{\tau}'$, and $\boldsymbol{\rho}$ where $\boldsymbol{\rho}$ measures the separation between the pairs $\boldsymbol{\tau}$ and $\boldsymbol{\tau}'$: (see Fig. 13)

$$J(\boldsymbol{\tau}, \boldsymbol{\tau}'\boldsymbol{\rho}) = \int d\boldsymbol{r} d\boldsymbol{r}' \varphi_v(\boldsymbol{r})\varphi_c(\boldsymbol{r} - \boldsymbol{\tau}) \times \frac{2}{|\boldsymbol{r} - \boldsymbol{r}'|} \varphi_v(\boldsymbol{r}' - \boldsymbol{\rho})\varphi_c(\boldsymbol{r}' - \boldsymbol{\rho} - \boldsymbol{\tau}'),$$

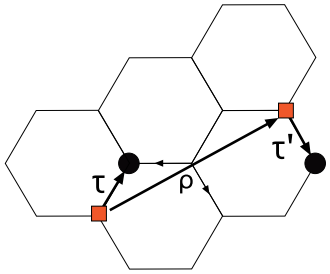

FIG. 13. Schematic representation of the geometry of the integral $J(\boldsymbol{\tau}, \boldsymbol{\tau}'\boldsymbol{\rho})$. Hole positions are shown as red squares and electron ones as black circles.

which induces in the tight-binding Hamiltonian an effective overlap integral $\langle \boldsymbol{\tau}|K^x|\boldsymbol{\tau}'\rangle$, which depends on the first neighbors of the origin $\boldsymbol{\tau}$ and $\boldsymbol{\tau}'$:

$$\langle \boldsymbol{\tau}|K^x_{\rm eh}|\boldsymbol{\tau}'\rangle = \sum_{\boldsymbol{\rho}} J(\boldsymbol{\tau},\boldsymbol{\tau}'\boldsymbol{\rho}).$$

To lowest order, when $\boldsymbol{\tau} = \boldsymbol{\tau}'$, this adds a local term $J_{\boldsymbol{\tau}}$ to the direct term on the first neighbors:

$$\begin{aligned} J_{\boldsymbol{\tau}} &= J(\boldsymbol{\tau},\boldsymbol{\tau},\boldsymbol{\rho}=\mathbf{0}) \\ &= \int d\boldsymbol{r}d\boldsymbol{r}'\varphi_v(\boldsymbol{r})\varphi_c(\boldsymbol{r}-\boldsymbol{\tau})\frac{2}{|\boldsymbol{r}-\boldsymbol{r}'|}\varphi_v(\boldsymbol{r}')\varphi_c(\boldsymbol{r}'-\boldsymbol{\tau}). \end{aligned}$$

On the other hand, when we sum all contributions corresponding to all separations of the distant $\boldsymbol{\tau}$ and $\boldsymbol{\tau}'$ pairs we obtain a sum of dipolar contributions which are known to be singular here ($\boldsymbol{Q} \to 0$ limit). In three dimensions this produces the so-called longitudinal-transversal splitting [35,63]. In two dimensions the singularity is weaker, with terms varying as $|\boldsymbol{Q}|$ [14,64]. This will be discussed elsewhere.

APPENDIX C: AN IMPROVED WANNIER MODEL

The tight-binding model developed in the main text is based on several approximations. To lowest order in t/Δ the Wannier functions corresponding to the valence and conduction bands are taken as the π orbitals on the nitrogen and boron sites, respectively. Then the kinetic energy part of the exciton Hamiltonian is approximated by its second-order term in t/Δ and finally the Coulomb matrix elements $U_{\boldsymbol{R}}$ are calculated using atomic orbitals, i.e., Wannier function of zero order. We will see that higher-order terms induce corrections of order $(t/2\Delta)^2 \simeq 0.1$. This is not negligible but has been implicitly taken into account via our fitting procedure for the interactions where $\boldsymbol{R} \neq 0$ on the triangular lattice. Problems arise because, to higher order, other interactions become allowed, in particular when $\boldsymbol{R} = 0$. Let us then define more accurate Wannier states $|\boldsymbol{n}\pm\rangle_w$ from the exact eigenstates $|\boldsymbol{k}\pm\rangle$ defined in Eq. (2). We choose the phases such that these Wannier states reduce to the atomic states when $t \to 0$. Then, to linear order in t/Δ:

$$\begin{aligned} |\boldsymbol{m}+\rangle_w &= \frac{1}{\sqrt{N}}\sum_{\boldsymbol{k}} e^{-i.\boldsymbol{k}.\boldsymbol{m}}|\boldsymbol{k}+\rangle \\ &\simeq |\boldsymbol{m}B\rangle - \frac{t}{2\Delta}\sum_{\boldsymbol{\tau}}|\boldsymbol{m}B-\boldsymbol{\tau}\rangle \\ |\boldsymbol{n}-\rangle_w &= \frac{1}{\sqrt{N}}\sum_{\boldsymbol{k}} e^{-i.\boldsymbol{k}.\boldsymbol{n}}|\boldsymbol{k}-\rangle \\ &\simeq |\boldsymbol{n}A\rangle + \frac{t}{2\Delta}\sum_{\boldsymbol{\tau}}|\boldsymbol{n}A+\boldsymbol{\tau}\rangle. \end{aligned} \tag{C1}$$

The sites $\boldsymbol{n}A(B)$ are on the $A(B)$ sublattices and the sites $\boldsymbol{m}B-\boldsymbol{\tau}$ ($\boldsymbol{n}A+\boldsymbol{\tau}$) are also on the $A(B)$ sublattices. These Wannier states remain centered on boron (B) and nitrogen (A) sites, respectively, so that we can continue to use sublattice labels A,B instead of band labels $\pm$, but they spread on the neighboring sites on the other sublattices. The excitonic kinetic energy term calculated to second order in t/Δ has exactly the form derived previously, but we are now able to calculate the corrections to the Coulomb (direct) term. The Wannier functions on neighbor sites overlap so that the Coulomb matrix elements between electron and hole Wannier states labeled by $\boldsymbol{n}A$ and $\boldsymbol{m}B$ involve not only the usual Coulomb integrals $U_{\boldsymbol{nm}}$ but also integrals involving sites $\boldsymbol{n}A+\boldsymbol{\tau}$ and $\boldsymbol{m}B-\boldsymbol{\tau}$. More precisely, let us define the Wannier electron-hole states $|\boldsymbol{R}\rangle_w$:

$$|\boldsymbol{R}\rangle_w = \frac{1}{\sqrt{N}}\sum_{\boldsymbol{n}} w^\dagger_{\boldsymbol{n}A+\boldsymbol{R}}\, w_{\boldsymbol{n}A}|\emptyset\rangle,$$

where $w^\dagger_{\boldsymbol{m}}$ is the creation operator in the Wannier state $|\boldsymbol{m}B\rangle_w \equiv |\boldsymbol{m}+\rangle_w$, and $w_{\boldsymbol{n}}$ is the destruction operator in the Wannier state $|\boldsymbol{n}A\rangle_w \equiv |\boldsymbol{n}-\rangle_w$. $\boldsymbol{R}$ is as previously a vector between the two sublattices. The matrix element of the direct Coulomb kernel now becomes:

$${}_w\langle \boldsymbol{R}|K^d_{\rm eh}|\boldsymbol{R}'\rangle_w \simeq U_{\boldsymbol{R}}\,\delta_{\boldsymbol{R},\boldsymbol{R}'} + \frac{t^2}{2\Delta^2}\sum_{\boldsymbol{\tau},\boldsymbol{\tau}'}\delta_{\boldsymbol{R}-\boldsymbol{\tau},\boldsymbol{R}'-\boldsymbol{\tau}'}U_{\boldsymbol{R}-\boldsymbol{\tau}}\,.$$

We will keep only the corrective terms involving U_0. Thus these terms only correct the matrix elements between neighbors of the origin:

$${}_w\langle \boldsymbol{\tau}|K^d_{\rm eh}|\boldsymbol{\tau}'\rangle_w \simeq U_{\boldsymbol{\tau}}\,\delta_{\boldsymbol{\tau},\boldsymbol{\tau}'} + \frac{t^2}{2\Delta^2}\,U_0.$$

Actually, second-order terms in the Wannier function expansions also contribute, but they do not involve U_0. This perturbation expansion is the counterpart in real space of the developments in reciprocal space (and within the $\boldsymbol{k}\cdot\boldsymbol{p}$ approximation), performed in Refs. [14], [52], and [53].

Although U_0 cannot be derived from the continuous Keldysh potential, it should be significantly larger than its value at the first-neighbor positions, about 3 eV (see Fig. 11), and the perturbation is not negligible *a priori*. Let us estimate to lowest order the correction δE to the energy of the excitonic state $|\Phi\rangle$:

$$\delta E \simeq \langle\Phi|\delta K^d_{\rm eh}|\Phi\rangle = \frac{t^2}{2\Delta^2}\,U_0\left|\sum_{\boldsymbol{\tau}}\langle\boldsymbol{\tau}|\Phi\rangle\right|^2 < 0.$$

The sum of the amplitude on the first shell, $\sum_{\boldsymbol{\tau}}\langle\boldsymbol{\tau}|\Phi\rangle$ is nonvanishing only for states of full symmetry A_1. This is the case of the exciton No. 4 discussed in the main text. As discussed there, however, the exchange contribution has also to be taken into account and the corrective term proportional to $2J_0$ (the calculation is similar to that derived above for the direct term) is here more important than the usual $J_{\boldsymbol{\tau}}$ (see Appendix B). Since $2J_0$ is a positive contribution at least equal to $|U_0|$, there is a compensation effect and the global correction to our TB model should be slightly positive, which is consistent with the results shown in Table I. On the other hand this discussion shows that neglecting exchange effects, which is common practice, is not valid here for hBN when dealing with fully symmetric excitons. In this case the singlet-triplet splitting is huge as can be seen in the same table.

[1] A. H. Castro Neto, F. Guinea, N. M. R. Peres, K. S. Novoselov, and A. K. Geim, The electronic properties of graphene, Rev. Mod. Phys. **81**, 109 (2009).

[2] K. Watanabe, T. Taniguchi, T. Kuroda, and H. Kanda, Band-edge luminescence of deformed hexagonal boron nitride single crystals, Diam. Relat. Mater. **15**, 1891 (2006).

[3] P. Jaffrennou, J. Barjon, J.-S. Lauret, A. Loiseau, F. Ducastelle, and B. Attal-Tretout, Origin of the excitonic recombinations in hexagonal boron nitride by spatially resolved cathodoluminescence spectroscopy, J. Appl. Phys. **102**, 116102 (2007).

[4] Y. Kubota, K. Watanabe, O. Tsuda, and T. Taniguchi, Deep ultraviolet light-emitting hexagonal boron nitride synthesized at atmospheric pressure, Science **317**, 932 (2007).

[5] K. Watanabe and T. Taniguchi, Jahn-Teller effect on exciton states in hexagonal boron nitride single crystal, Phys. Rev. B **79**, 193104 (2009).

[6] L. Museur, G. Brasse, A. Pierret, S. Maine, B. Attal-Tretout, F. Ducastelle, A. Loiseau, J. Barjon, K. Watanabe, T. Taniguchi, and A. Kanaev, Exciton optical transitions in a hexagonal boron nitride single crystal, Phys. Status Solidi - Rapid Res. Lett. **5**, 214 (2011).

[7] S. Galambosi, L. Wirtz, J. A. Soininen, J. Serrano, A. Marini, K. Watanabe, T. Taniguchi, S. Huotari, A. Rubio, and K. Hämäläinen, Anisotropic excitonic effects in the energy loss function of hexagonal boron nitride, Phys. Rev. B **83**, 081413 (2011).

[8] A. Pierret, J. Loayza, B. Berini, A. Betz, B. Plaçais, F. Ducastelle, J. Barjon, and A. Loiseau, Excitonic recombinations in hBN: From bulk to exfoliated layers, Phys. Rev. B **89**, 035414 (2014).

[9] G. Fugallo, M. Aramini, J. Koskelo, K. Watanabe, T. Taniguchi, M. Hakala, S. Huotari, M. Gatti, and F. Sottile, Exciton energy-momentum map of hexagonal boron nitride, Phys. Rev. B **92**, 165122 (2015).

[10] W. Zhao, Z. Ghorannevis, L. Chu, M. Toh, C. Kloc, P.-H. Tan, and G. Eda, Evolution of electronic structure in atomically thin sheets of WS_2 and WSe_2, ACS Nano **7**, 791 (2012).

[11] M. Xu, T. Liang, M. Shi, and H. Chen, Graphene-like two-dimensional materials, Chem. Rev. **113**, 3766 (2013).

[12] H. Zhang, Ultrathin two-dimensional nanomaterials, ACS Nano **9**, 9451 (2015).

[13] A. Molina-Sánchez, K. Hummer, and L. Wirtz, Vibrational and optical properties of mos2: From monolayer to bulk, Surf. Sci. Rep. **70**, 554 (2015).

[14] F. Wu, F. Qu, and A. H. MacDonald, Exciton band structure of monolayer MoS_2, Phys. Rev. B **91**, 075310 (2015).

[15] X. Ling, H. Wang, S. Huang, F. Xia, and M. S. Dresselhaus, The renaissance of black phosphorus, Proc. Natl. Acad. Sci. USA **112**, 4523 (2015).

[16] A. Favron, E. Gaufrès, F. Fossard, Anne-Laurence Phaneuf-L'Heureux, Nathalie Y.-W. Tang, Pierre L. Lévesque, A. Loiseau, R. Leonelli, S. Francoeur, and R. Martel, Photooxidation and quantum confinement effects in exfoliated black phosphorus, Nature Mater. **14**, 826 (2015).

[17] P. Vogt, P. De Padova, C. Quaresima, J. Avila, E. Frantzeskakis, M. C. Asensio, A. Resta, B. Ealet, and G. Le Lay, Silicene: Compelling Experimental Evidence for Graphenelike Two-Dimensional Silicon, Phys. Rev. Lett. **108**, 155501 (2012).

[18] L. Li, S.-z. Lu, J. Pan, Z. Qin, Y.-q. Wang, Y. Wang, G.-yu Cao, S. Du, and H.-J. Gao, Buckled germanene formation on pt(111), Adv. Mater. **26**, 4820 (2014).

[19] K. F. Mak, C. Lee, J. Hone, J. Shan, and T. F. Heinz, Atomically Thin MoS_2: A New Direct-Gap Semiconductor, Phys. Rev. Lett. **105**, 136805 (2010).

[20] A. Splendiani, L. Sun, Y. Zhang, T. Li, J. Kim, C.-Y. Chim, G. Galli, and F. Wang, Emerging photoluminescence in monolayer MoS_2, Nano Lett. **10**, 1271 (2010).

[21] P. Cudazzo, I. V. Tokatly, and A. Rubio, Dielectric screening in two-dimensional insulators: Implications for excitonic and impurity states in graphane, Phys. Rev. B **84**, 085406 (2011).

[22] A. Chernikov, T. C. Berkelbach, H. M. Hill, A. Rigosi, Y. Li, O. Burak Aslan, D. R. Reichman, M. S. Hybertsen, and T. F. Heinz, Exciton Binding Energy and Nonhydrogenic Rydberg Series in Monolayer WS_2, Phys. Rev. Lett. **113**, 076802 (2014).

[23] S. Huang, Y. Liang, and L. Yang, Exciton spectra in two-dimensional graphene derivatives, Phys. Rev. B **88**, 075441 (2013).

[24] A. Molina-Sánchez, D. Sangalli, K. Hummer, A. Marini, and L. Wirtz, Effect of spin-orbit interaction on the optical spectra of single-layer, double-layer, and bulk MoS_2, Phys. Rev. B **88**, 045412 (2013).

[25] D. Y. Qiu, F. H. da Jornada, and S. G. Louie, Optical Spectrum of MoS_2: Many-Body Effects and Diversity of Exciton States, Phys. Rev. Lett. **111**, 216805 (2013).

[26] K. He, N. Kumar, L. Zhao, Z. Wang, K. F. Mak, H. Zhao, and J. Shan, Tightly Bound Excitons in Monolayer WSe_2, Phys. Rev. Lett. **113**, 026803 (2014).

[27] O. Pulci, M. Marsili, V. Garbuio, P. Gori, I. Kupchak, and F. Bechstedt, Excitons in two-dimensional sheets with honeycomb symmetry, Phys. Status Solidi (b) **252**, 72 (2015).

[28] B. Arnaud, S. Lebègue, P. Rabiller, and M. Alouani, Huge Excitonic Effects in Layered Hexagonal Boron Nitride, Phys. Rev. Lett. **96**, 026402 (2006).

[29] L. Wirtz, A. Marini, and A. Rubio, Excitons in Boron Nitride Nanotubes: Dimensionality Effects, Phys. Rev. Lett. **96**, 126104 (2006).

[30] L. Wirtz, A. Marini, M. Grüning, C. Attaccalite, G. Kresse, and A. Rubio, Comment on "Huge Excitonic Effects in Layered Hexagonal Boron Nitride", Phys. Rev. Lett. **100**, 189701 (2008).

[31] B. Arnaud, S. Lebègue, P. Rabiller, and M. Alouani, Arnaud, Lebègue, Rabiller, and Alouani Reply:, Phys. Rev. Lett. **100**, 189702 (2008).

[32] L. Schue, B. Berini, A. C. Betz, B. Placais, F. Ducastelle, J. Barjon, and A. Loiseau, Dimensionality effects on the luminescence properties of hBN, Nanoscale **8**, 6986 (2016).

[33] G. H. Wannier, The structure of electronic excitation levels in insulating crystals, Phys. Rev. **52**, 191 (1937).

[34] R. S. Knox, *Theory of Excitons* (Academic Press, New York, 1963).

[35] Y. Toyozawa, *Optical Processes in Solids* (Cambridge University Press, Cambridge, 2003).

[36] F. Bechstedt, *Many-Body Approach to Electronic Excitations* (Springer, Berlin, 2015).

[37] A. A. Mostofi, J. R. Yates, Y.-S. Lee, I. Souza, D. Vanderbilt, and N. Marzari, wannier90: A tool for obtaining maximally-localised wannier functions, Comput. Phys. Commun. **178**, 685 (2008).

[38] P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G. L. Chiarotti, M. Cococcioni, I. Dabo, A. Dal Corso, S. de Gironcoli, S. Fabris, G. Fratesi, R. Gebauer, U. Gerstmann, C. Gougoussis, A. Kokalj, M. Lazzeri, L. Martin-Samos, N. Marzari, F. Mauri, R. Mazzarello, S. Paolini, A. Pasquarello, L. Paulatto, C. Sbraccia, S. Scandolo, G. Sclauzero, A. P. Seitsonen, A. Smogunov, P. Umari, and R. M. Wentzcovitch, Quantum espresso: A modular and open-source software project for quantum simulations of materials, J. Phys.: Condens. Matter **21**, 395502 (2009).

[39] A. Marini, C. Hogan, M. Grüning, and D. Varsano, YAMBO: An ab initio tool for excited state calculations, Comput. Phys. Commun. **180**, 1392 (2009).

[40] X. Blase, A. Rubio, S. G. Louie, and M. L. Cohen, Quasiparticle band structure of bulk hexagonal boron nitride and related systems, Phys. Rev. B **51**, 6868 (1995).

[41] L. Wirtz and A. Rubio, Optical and vibrational properties of boron nitride nanotubes, *B-C-N Nanotubes and Related Nanostructures* (Springer, New York, 2009), p. 105.

[42] R. M. Ribeiro and N. M. R. Peres, Stability of boron nitride bilayers: Ground-state energies, interlayer distances, and tight-binding description, Phys. Rev. B **83**, 235312 (2011).

[43] N. Berseneva, A. Gulans, A. V. Krasheninnikov, and R. M. Nieminen, Electronic structure of boron nitride sheets doped with carbon from first-principles calculations, Phys. Rev. B **87**, 035404 (2013).

[44] F. Hüser, T. Olsen, and K. S. Thygesen, Quasiparticle GW calculations for solids, molecules, and two-dimensional materials, Phys. Rev. B **87**, 235132 (2013).

[45] P. Cudazzo, L. Sponza, C. Giorgetti, L. Reining, F. Sottile, and M. Gatti, Exciton Band Structure in Two-Dimensional Materials, Phys. Rev. Lett. **116**, 066803 (2016).

[46] M. Rohlfing and S. G. Louie, Electron-hole excitations and optical spectra from first principles, Phys. Rev. B **62**, 4927 (2000).

[47] G. Onida, L. Reining, and A. Rubio, Electronic excitations: Density-functional versus many-body green's-function approaches, Rev. Mod. Phys. **74**, 601 (2002).

[48] C. A. Rozzi, D. Varsano, A. Marini, E. K. U. Gross, and A. Rubio, Exact coulomb cutoff technique for supercell calculations, Phys. Rev. B **73**, 205119 (2006).

[49] M. Lannoo and J. Bourgoin, *Point Defects in Semiconductors I* (Springer, Berlin, 1981).

[50] P. Y. Yu and M. Cardona, *Fundamentals of Semiconductors* (Springer, Berlin, 2010).

[51] A. Jorio, R. Saito, G. Dresselhaus, and M. S. Dresselhaus, Theory of excitons in carbon nanotubes, *Raman Spectroscopy in Graphene Related Systems* (Wiley-VCH, Weinheim, 2011), p. 223.

[52] A. Srivastava and A. Imamoğlu, Signatures of Bloch-Band Geometry on Excitons: Nonhydrogenic Spectra in Transition-Metal Dichalcogenides, Phys. Rev. Lett. **115**, 166802 (2015).

[53] J. Zhou, W.-Y. Shan, W. Yao, and D. Xiao, Berry Phase Modification to the Energy Spectrum of Excitons, Phys. Rev. Lett. **115**, 166803 (2015).

[54] T. C. Berkelbach, M. S. Hybertsen, and D. R. Reichman, Bright and dark singlet excitons via linear and two-photon spectroscopy in monolayer transition-metal dichalcogenides, Phys. Rev. B **92**, 085413 (2015).

[55] L. V. Keldysh, Coulomb interaction in thin semiconductor and semimetal films, JETP Lett. **29**, 658 (1979).

[56] T. C. Berkelbach, M. S. Hybertsen, and D. R. Reichman, Theory of neutral and charged excitons in monolayer transition metal dichalcogenides, Phys. Rev. B **88**, 045318 (2013).

[57] S. Latini, T. Olsen, and K. S. Thygesen, Excitons in van der waals heterostructures: The important role of dielectric screening, Phys. Rev. B **92**, 245123 (2015).

[58] A. S. Rodin, A. Carvalho, and A. H. Castro Neto, Excitons in anisotropic two-dimensional semiconducting crystals, Phys. Rev. B **90**, 075429 (2014).

[59] L. Wirtz, A. Marini, M. Gruning, and A. Rubio, Excitonic effects in optical absorption and electron-energy loss spectra of hexagonal boron nitride, arXiv:cond-mat/0508421 [cond-mat.mtrl-sci].

[60] Vl. A. Margulis, E. E. Muryumin, and E. A. Gaiduk, Optical conductivity of atomic hexagonal boron nitride layers, Physica B **407**, 4244 (2012).

[61] Vl. A. Margulis, E. E. Muryumin, and E. A. Gaiduk, Theoretical calculations of low-field electroreflectance of ultra-thin hexagonal bn films at the fundamental absorption edge, J. Phys.: Condens. Matter **26**, 045301 (2014).

[62] D. Gunlycke and F. Tseng, Triangular lattice exciton model, Phys. Chem. Chem. Phys. **18**, 8579 (2016).

[63] M. M. Denisov and V. P. Makarov, Longitudinal and transverse excitons in semiconductors, Phys. Status Solidi (b) **56**, 9 (1973).

[64] D. Y. Qiu, T. Cao, and S. G. Louie, Nonanalyticity, Valley Quantum Phases, and Lightlike Exciton Dispersion in Monolayer Transition Metal Dichalcogenides: Theory and First-Principles Calculations, Phys. Rev. Lett. **115**, 176801 (2015).