

# Spinorial formulation of the GW-BSE equations and spin properties of excitons in two-dimensional transition metal dichalcogenides

Margherita Marsili <sup>1</sup>, Alejandro Molina-Sánchez <sup>2</sup>, Maurizia Palumbo <sup>3,4</sup>, Davide Sangalli <sup>5,4</sup> and Andrea Marini <sup>5,4</sup>

<sup>1</sup>*DISC, Dipartimento di Scienze Chimiche, University of Padova, Via Marzolo 1, I-35131 Padova, Italy*

<sup>2</sup>*Institute of Materials Science (ICMUV), University of Valencia, Catedrático Beltrán 2, E-46980 Valencia, Spain*

<sup>3</sup>*Dipartimento di Fisica, Università di Roma Tor Vergata, Via della Ricerca Scientifica 1, 00133 Rome, Italy*

<sup>4</sup>*European Theoretical Spectroscopy Facilities (ETSF)*

<sup>5</sup>*Istituto di Struttura della Materia and Division of Ultrafast Processes in Materials (FLASHit) of the National Research Council, via Salaria Km 29.3, I-00016 Monterotondo Stazione, Italy*



(Received 6 October 2020; revised 18 February 2021; accepted 1 April 2021; published 27 April 2021)

In many paradigmatic materials, such as transition metal dichalcogenides, the role played by the spin degrees of freedom is as important as the one played by the electron-electron interaction. Thus an accurate treatment of the two effects and of their interaction is necessary for an accurate and predictive study of the optical and electronic properties of these materials. Despite the fact that the GW-BSE approach correctly accounts for electronic correlations, the spin-orbit coupling effect is often neglected or treated perturbatively. Recently, spinorial formulations of GW-BSE have become available in different flavors in material-science codes. However, an accurate validation and comparison of different approaches is still missing. In this work, we go through the derivation of the noncollinear GW-BSE approach. The scheme is applied to transition metal dichalcogenides comparing the perturbative and full spinorial approaches. Our calculations reveal that dark-bright exciton splittings are generally improved when the spin-orbit coupling is included nonperturbatively. The exchange-driven intravalley mixing between the A and B excitons is found to play a role for Mo-based systems, being especially strong in the case of MoSe<sub>2</sub>. We finally compute the excitonic spin and use it to sharply analyze the spinorial properties of transition metal dichalcogenide excitonic states.

DOI: [10.1103/PhysRevB.103.155152](https://doi.org/10.1103/PhysRevB.103.155152)

## I. INTRODUCTION

The investigation of the excited-state properties of materials by means of modern *ab initio* theories is a rapidly developing field that has yielded notable progress in our understanding of bulk, surfaces, nanostructures, molecules, and disordered systems [1]. At the same time, an increasing number of experimental and technological applications have been highlighting the central role played by spin degrees of freedom in the explanation of novel and intriguing physical processes such as the chirality effect, observed in low-dimensional systems that lack structural inversion symmetry [2]. Moreover, spin-orbit coupling (SOC) is essential to correctly describing electronic and optical excitations not only for transition metal dichalcogenides (TMDs) and topological insulators (TIs) but also nanomaterials with light atoms, such as carbon nanotubes [3]. Finally, an accurate description of spin dynamics is essential in the fields of spintronics and magneto-optics [4]. Despite this fact, most of the many-body calculations based on the GW and Bethe-Salpeter-equation (BSE) methods [1] have been carried out within a spin-independent or spin-polarized [5] framework, generally neglecting the SOC. At present, several electronic structure codes [6–9], which include SOC within the GW-BSE approach either perturbatively [10,11] or fully considering the spinorial nature of the electronic wave functions, have been

applied in a number of cases [7,8,12–17]. However, the mathematical derivation and a complete discussion of the spinorial formulation of GW-BSE is not present in the literature, and an accurate comparison of the full spinorial GW-BSE with the perturbative approach, benchmarking the latter, is also missing.

In addition to their importance for fundamental and applicative reasons [18–20], TMD monolayers are ideal for testing different levels of SOC inclusion because both the presence of strong spin-orbit interaction and the enhancement of many-body effects are the basis of their intriguing electronic and optical properties. The strong spin-orbit interaction determines the macroscopic features of the optical properties of these materials, such as the presence of spin-split peaks and valley-selective optical transitions in their absorption spectra. Moreover, SOC is responsible of finer details that are nevertheless crucial for applications of TMDs in optoelectronic devices. These are the splitting and the energetic order of spin-allowed (bright) and spin-forbidden (dark) excitons, which are involved in the exciton dynamics. From this standpoint, TMDs are an optimal class of materials against which we can compare a full and a perturbative SOC formulation up to the GW-BSE level by using the same set of computational parameters. Despite the fact that most of the equations can be obtained as a direct extension of the standard formulation

[21], we provide a detailed derivation of noncollinear GW-BSE equations. In doing so, we aim at providing a complete reference with all spin-indexes carefully included and with an extensive discussion of the different approximations chosen.

We show how this formulation allows us to include SOC in a natural way at the level of the ground-state calculation, and it is nicely suited to the study of the optical properties of any many-body quantum-mechanical system in which the dependence from the spin can be described as a nonlocal term in the Hamiltonian. Finally, we apply this formalism to the calculation of the electronic and optical properties of group-VI TMD monolayers ( $MX_2$ , with  $M = \text{Mo, W}$  and  $X = \text{S, Se}$ ), and we accurately compare the no-, perturbative, and full SOC schemes. Furthermore, a full analysis of the spin character of its excitons, which is relevant for envisaged valleytronics applications but is intrinsically not achievable in a perturbative approach, is presented.

## II. THE NONCOLLINEAR MANY-BODY PROBLEM

We start from the many-body total Hamiltonian of the system, including first-order relativistic corrections

$$\hat{H} = \hat{H}^0 + \hat{H}^{(e-e)} + \hat{H}^{(\text{RK})} + \hat{H}^{(\text{SOI})}. \quad (1)$$

Here  $\hat{H}^0$  is the nonrelativistic one-body Hamiltonian, composed of a kinetic term and the atomic scalar external potential,  $\hat{H}^{(e-e)}$  is the electron-electron Coulomb interaction, while  $\hat{H}^{(\text{RK})} + \hat{H}^{(\text{SOI})}$  are the first-order relativistic corrections.  $\hat{H}^{(\text{RK})}$  is the mass-velocity term plus the Darwin term, while  $\hat{H}^{(\text{SOI})}$  is the spin-orbit-interaction (SOI) term [22].

In the position, momentum, and spin of each electron  $\hat{\mathbf{x}}, \hat{\mathbf{p}}, \hat{\mathbf{s}}$ , the different terms can be expressed as

$$\hat{H}^0 = H^0[\{\hat{\mathbf{x}}_n, \hat{\mathbf{p}}_n\}] = \sum_i h^0(\hat{\mathbf{x}}_i, \hat{\mathbf{p}}_i), \quad (2a)$$

$$\hat{H}^{(e-e)} = H^{(e-e)}[\{\hat{\mathbf{x}}_n\}] = \frac{1}{2} \sum_{i \neq j} \frac{1}{|\hat{\mathbf{x}}_i - \hat{\mathbf{x}}_j|}, \quad (2b)$$

$$\hat{H}^{(\text{RK})} = H^{(\text{RK})}[\{\hat{\mathbf{p}}_n\}] = \sum_i h^{(\text{RK})}(\hat{\mathbf{p}}_i), \quad (2c)$$

$$\hat{H}^{(\text{SOI})} = \sum_i H^{(\text{SOI})}[\{\hat{\mathbf{x}}_n, \hat{\mathbf{p}}_n\}](\hat{\mathbf{s}}_i). \quad (2d)$$

Here bold symbols indicate spatial vectors,  $\hat{h}^0 = \hat{t} + \hat{v}^{\text{ext}}$  is the sum of the kinetic and one-body external potential, while  $\hat{H}^{(\text{RK})}$  is the sum of one-body terms; for its explicit expression, we refer the reader to Ref. [23], Eqs. 9.2.2. By following Ref. [24], we know that the SOI term is composed of three different contributions:

$$\hat{H}^{(\text{SOI})} = \hat{H}^{[\text{SO}(N)]} + \hat{H}^{[\text{SO}(2e)]} + \hat{H}^{(\text{SOO})}. \quad (3)$$

Here  $\hat{H}^{[\text{SO}(2e)]}$  and  $\hat{H}^{(\text{SOO})}$  are many-body terms, called ‘‘two-electrons spin-orbit’’ and ‘‘spin-other-orbit,’’ respectively.  $\hat{H}^{[\text{SO}(N)]}$ , or ‘‘one-electron spin-orbit,’’ is a purely one-body term:

$$\hat{H}^{[\text{SO}(N)]} = \sum_i \underline{v}^{(\text{SOC})}(\hat{\mathbf{x}}_i, \hat{\mathbf{p}}_i), \quad (4)$$

with  $\underline{v}$  a  $2 \times 2$  matrix in the  $\uparrow, \downarrow$  spin space. What is relevant here is that  $H^{(\text{SOI})}$  can be expressed as a sum of terms that depend on a single spin operator and can thus be conveniently written as a sum of a  $2 \times 2$  matrix in  $\uparrow, \downarrow$  spin space. Since all other terms are spin-independent, this implies that the whole Hamiltonian can be expressed as a sum of  $2 \times 2$  matrices. We denote as  $\underline{v}$  such matrices in  $\uparrow, \downarrow$  spin space.

### A. The single-particle part of the Hamiltonian within density functional theory

Within DFT, the whole Eq. (1) is replaced with a mean-field representation:

$$\hat{H} \implies \sum_i \underline{h}^{\text{KS}}[\underline{\rho}](\hat{\mathbf{x}}_i, \hat{\mathbf{p}}_i). \quad (5)$$

The one-body KS Hamiltonian depends on the electronic density matrix,  $\underline{\rho}$ , defined in (11), and it reads

$$h_{ss'}^{\text{KS}}[\underline{\rho}](\hat{\mathbf{x}}, \hat{\mathbf{p}}) = h(\hat{\mathbf{x}}, \hat{\mathbf{p}})\delta_{ss'} + v_{ss'}^{(\text{SOC})}(\hat{\mathbf{x}}) + v_{ss'}^{(H_{\text{xc}})}[\underline{\rho}](\hat{\mathbf{x}}), \quad (6)$$

with  $s$  the spin index. In Eq. (6),  $\hat{h} = \hat{h}^0 + \hat{h}^{(\text{RK})}$  while  $\hat{v}^{(\text{SOC})}$  is the one-electron spin-orbit-coupling nucleus term, which comes from the  $\text{SO}(N)$  term defined in Eq. (4), and  $v_{ss'}^{(H_{\text{xc}})}$  is the sum of the exchange-correlation and Hartree potential. The more general form of  $\hat{v}^{(\text{SOC})}$  (if the vector potential is zero) is

$$\underline{v}^{(\text{SOC})}(\hat{\mathbf{x}}, \hat{\mathbf{p}}) = \frac{1}{2c^2} \underline{\sigma} \cdot [\nabla v^{(\text{ext})}(\hat{\mathbf{x}}) \times \hat{\mathbf{p}}]. \quad (7)$$

Here  $\underline{\sigma}$  is the three-dimensional vector whose components are the Pauli matrices. In a plane-wave representation, the SOC is, in practice, accounted for by the use of pseudopotentials [25], and also  $\hat{h}^{(\text{RK})}$  corrections are taken into account for the kinetic energy of core electrons in the generation of the pseudopotential [26]. The SOC contribution from the pseudopotential captures what is called ‘‘local SOC,’’ which is due to the electrons orbiting around the nuclei. It neglects instead the itinerant SOC, which cannot be easily captured in periodic boundary conditions. Relativistic corrections to the kinetic energy of valence (and conduction) electrons are instead usually neglected.

Neglecting the relativistic corrections  $\hat{v}^{(H_{\text{xc}})}$  is the mean-field replacement of  $\hat{H}^{(e-e)}$ . When relativistic corrections are taken into account instead, one should in principle account for the effects of the many-body terms  $\hat{H}^{[\text{SO}(2e)]}$  and  $\hat{H}^{(\text{SOO})}$ . A short cut [24] is to add a term constructed replacing  $v^{\text{ext}}$  with  $v^{(H_{\text{xc}})}$  into Eq. (7). Doing so partially accounts for the physics of  $\hat{H}^{[\text{SO}(2e)]}$  while the effects described by  $\hat{H}^{(\text{SOO})}$  are neglected, as is commonly done in standard DFT calculations.

Thus there are two terms entering the  $\hat{h}^{\text{KS}}$  Hamiltonian that are nondiagonal in spin-space and are responsible for inducing spinorial eigenstates:  $\hat{v}^{(\text{SOC})}$  and  $\hat{v}^{(H_{\text{xc}})}$  [27]. The spin-dependent formulation of the density functional theory [28] in its local spin density approximation (LSDA) is currently implemented in several *ab initio* codes and is at the basis of the present excited-state calculations. It gives rise to a renormalization of the SOC splitting. In Appendix A, we also show how the noncollinear form of the Hamiltonian can be rewritten in terms of density and magnetization coupling with

the xc scalar potential  $\phi_{xc}$  plus an xc magnetic field  $\mathbf{B}_{xc}$ . If the local magnetization is zero everywhere, then  $\mathbf{B}_{xc} = \mathbf{0}$ .

The eigenstates of  $\hat{h}^{KS}$  are vectors in  $\uparrow, \downarrow$  spin space:

$$\langle \mathbf{x} | \hat{c}_{\mathbf{I}}^\dagger | 0 \rangle = \vec{\psi}_{\mathbf{I}}(\mathbf{x}) = \langle \mathbf{x} | \vec{n}\mathbf{k} \rangle \equiv \begin{pmatrix} \phi_{n\mathbf{k}\uparrow}(\mathbf{x}) \\ \phi_{n\mathbf{k}\downarrow}(\mathbf{x}) \end{pmatrix}, \quad (8)$$

with  $n$  the spinor band index, and  $\mathbf{k}$  a Brillouin zone generic point, which from now on we group in the index  $\mathbf{I} \equiv (n, \mathbf{k})$  to simplify the notation.  $\vec{\psi}_{\mathbf{I}}$  satisfies the matrix equation

$$\underline{h}^{KS}(\hat{\mathbf{x}}) \vec{\psi}_{\mathbf{I}}(\mathbf{x}) = \epsilon_{\mathbf{I}} \vec{\psi}_{\mathbf{I}}(\mathbf{x}). \quad (9)$$

Thanks to Eq. (8), we can define the fermionic field operators

$$\hat{\Psi}(\mathbf{x}, s) = \sum_{\mathbf{I}} \phi_{\mathbf{I}s}(\mathbf{x}) \hat{c}_{\mathbf{I}}, \quad (10)$$

with the operators written in the Heisenberg representation. In Eq. (10), we have embodied in the spinorial wave function,  $\phi_{\mathbf{I}s}$ , the  $\frac{1}{\sqrt{N_{\mathbf{k}}}}$  prefactor, with  $N_{\mathbf{k}}$  the number of  $\mathbf{k}$ -points. In this way we can write, for example, the spin components of the density matrix as

$$\rho_{ss'}(\mathbf{x}) = \sum_{\mathbf{I}} \sum_{\alpha=0,3} \phi_{\mathbf{I}s}^*(\mathbf{x}) [\sigma_{\alpha}]_{ss'} \phi_{\mathbf{I}s'}(\mathbf{x}), \quad (11)$$

where  $\phi_{\mathbf{I}s}^*$  is the complex conjugate of  $\phi_{\mathbf{I}s}$ , and  $\sigma_0$  is the identity. The properties of the spinorial field operators can be easily obtained from some anticommutation rules of the fermionic creation and annihilation operators. We have that

$$\{\hat{c}_{\mathbf{I}_1}, \hat{c}_{\mathbf{I}_2}^\dagger\} = \delta_{\mathbf{I}_1, \mathbf{I}_2}, \{\hat{c}_{\mathbf{I}_1}, \hat{c}_{\mathbf{I}_2}\} = 0, \quad (12a)$$

$$\sum_s \int d\mathbf{x} \phi_{\mathbf{I}_1 s}^*(\mathbf{x}) \phi_{\mathbf{I}_2 s}(\mathbf{x}) = \delta_{\mathbf{I}_1, \mathbf{I}_2}, \quad (12b)$$

$$\sum_{\mathbf{I}} \phi_{\mathbf{I}s_1}^*(\mathbf{x}_1) \phi_{\mathbf{I}s_2}(\mathbf{x}_2) = \delta_{s_1 s_2} \delta(\mathbf{x}_1 - \mathbf{x}_2), \quad (12c)$$

where  $\delta_{\mathbf{I}, \mathbf{J}} \equiv \delta_{\mathbf{k}, \mathbf{k}'} \delta_{n, n'}$ .

### B. The interaction term and the perturbative expansion

A crucial point of the present formulation is that by replacing the bare single-particle Hamiltonian,  $\hat{h}$ , with the Kohn-Sham (KS) Hamiltonian,  $\hat{h}^{KS}$ , the whole perturbative expansion is done on top of the KS energies and eigenvectors.

A subtle but essential point, however, is that in order to prevent double-counting problems, the pure electron-electron interaction needs to be amended. Without relativistic corrections, this means that

$$\sum_i \hat{h}_i^0 + \hat{H}^{(e-e)} \rightarrow \sum_i \hat{h}_i^{KS} + \Delta \hat{H}^{(e-e)}, \quad (13)$$

with

$$\Delta \hat{H}^{(e-e)} = \hat{H}^{(e-e)} - \sum_i [\hat{v}_i^{Hxc}], \quad (14)$$

which is done in practice by subtracting  $v^{(xc)}$  from the many-body self-energy  $\Sigma$ .

In the presence of relativistic corrections, one needs to replace in Eq. (13),

$$\hat{h}^0 \rightarrow \hat{h}^0 + \hat{h}^{(RK)} + \hat{v}^{SOC}, \quad (15)$$

$$\hat{H}^{(e-e)} \rightarrow \hat{H}^{(e-e)} + \hat{H}^{[SO(2e)]} + \hat{H}^{(SOO)}. \quad (16)$$

The resulting effective electron-electron interaction is non-diagonal in spin-space, and one should in principle follow the derivation of Ref. [21] to define the proper many-body self-energy at the diagrammatic level. Here we neglect such a complication, and we focus on the effect of using a fully noncollinear noninteracting Hamiltonian while keeping a standard spin-independent interaction at the diagrammatic level. This means that we will use the standard definition of the many-body self-energy and still rely on Eq. (14) for applying corrections of MBPT on top of DFT.

### III. SPINORIAL FORMULATION OF HEDIN'S EQUATIONS

When a noncollinear potential is present in  $\underline{h}$ , this implies that the whole formulation of the many-body problem must be rewritten in the spinorial basis. We start from the standard Hedin equation in the space and spin,  $(\mathbf{x}, s)$ , basis. Then, using the definition of the spinorial field operators of Eq. (9), we expand all terms in the theory in the spinorial basis. In Appendix B, we give a short review of Hedin's equations, which solve exactly the problem. To rewrite the MBPT in the spinorial representation, we note that, in principle by using Eq. (10), the different components of Hedin's equations can be conveniently rotated.

In practice, we define two *maps*,  $M_2$  and  $M_4$ :

$$M_2 : F(1, 2) \equiv \phi_{\mathbf{I}_1 s_1}^*(\mathbf{x}_1) F(1, 2) \phi_{\mathbf{I}_2 s_2}(\mathbf{x}_2) = F_{\mathbf{I}_1 \mathbf{I}_2}(t_1, t_2) \quad (17)$$

and

$$\begin{aligned} M_4 : F(1, 2, 3, 4) \\ \equiv \phi_{\mathbf{I}_1 s_1}^*(\mathbf{x}_1) \phi_{\mathbf{I}_2 s_2}(\mathbf{x}_2) F(1, 2, 3, 4) \phi_{\mathbf{I}_3 s_3}(\mathbf{x}_3) \phi_{\mathbf{I}_4 s_4}^*(\mathbf{x}_4) \\ = F_{\mathbf{I}_1 \mathbf{I}_2 \mathbf{I}_3 \mathbf{I}_4}(t_1, t_2, t_3, t_4). \end{aligned} \quad (18)$$

In Eqs. (17) and (18),  $n \equiv (\mathbf{x}_n, s_n, t_n)$  and repeated variables are either integrated or summed up. Thanks to these two maps, we can easily define the representations in the spinorial basis of the different components of Hedin's equations. More specifically, following Ref. [29], we define

$$G_{\mathbf{I}_1 \mathbf{I}_2}(t_1, t_2) = M_2 : G(1, 2), \quad (19a)$$

$$\Sigma_{\mathbf{I}_1 \mathbf{I}_2}(t_1, t_2) = M_2 : \Sigma(1, 2), \quad (19b)$$

$$V_{\mathbf{I}_1 \mathbf{I}_2 \mathbf{I}_3 \mathbf{I}_4} = M_4 : v(1, 3) \delta(1, 2) \delta(3, 4), \quad (19c)$$

$$W_{\mathbf{I}_1 \mathbf{I}_2 \mathbf{I}_3 \mathbf{I}_4}(t_1, t_2) = M_4 : W(1, 2) \delta(1, 3) \delta(2, 4), \quad (19d)$$

$$\tilde{\Gamma}_{\mathbf{I}_1 \mathbf{I}_2 \mathbf{I}_3 \mathbf{I}_4}(t_1, t_2; t_3) = M_4 : \tilde{\Gamma}(1, 2; 3) \delta(3, 4), \quad (19e)$$

$$\tilde{\chi}_{\mathbf{I}_1 \mathbf{I}_2 \mathbf{I}_3 \mathbf{I}_4}(t_1, t_2) = M_4 : \tilde{\chi}(1, 2) \delta(1, 3) \delta(2, 4). \quad (19f)$$

The kind of map to be applied depends on the number of field operators involved in the definition of the corresponding quantity and not on the number of space-time or spin arguments. This implies the need for  $\delta$  functions that extend “contracted quantities.” For example, the response function  $\chi(1, 2)$  is a contraction of the more general two-particle Green function  $L(1, 3; 2, 4)$  with  $\tilde{\chi}(1, 2) = \tilde{L}(1, 1; 2, 2)$ . The proof of each rotation is given in Appendix C and referenced here when necessary. As a simple example, we see that the most elemental ingredient of MBPT is the GF. This can be rotated by simply using Eq. (10),

$$G(1, 2) = \sum_{\mathbf{I}_1, \mathbf{I}_2} \phi_{\mathbf{I}_1 s_1}(\mathbf{x}_1) G_{\mathbf{I}_1 \mathbf{I}_2}(t_1, t_2) \phi_{\mathbf{I}_2 s_2}^*(\mathbf{x}_2). \quad (20)$$

Then the Dyson equation reads

$$G_{\mathbf{I}_1 \mathbf{I}_2}(t_1, t_2) = G_{\mathbf{I}_1 \mathbf{I}_2}^{(0)}(t_1, t_2) + G_{\mathbf{I}_1 \mathbf{I}_3}^{(0)}(t_1, t_3) \Sigma_{\mathbf{I}_3 \mathbf{I}_4}^{Hxc}(t_3, t_4) G_{\mathbf{I}_4 \mathbf{I}_2}(t_4, t_2) \quad (21)$$

with  $\Sigma^{Hxc} = v^H + \Sigma$ ; the Hartree potential and the self-energy are defined as

$$v_{\mathbf{I}_3 \mathbf{I}_4}^H(t_3) = -iV_{\mathbf{I}_3 \mathbf{I}_4} G_{\mathbf{I}_3 \mathbf{I}_4}(t_3, t_3^+), \quad (22a)$$

$$\Sigma_{\mathbf{I}_1 \mathbf{I}_2}(t_1, t_2) = -iG_{\mathbf{I}_1' \mathbf{I}_3}(t_1, t_3) \tilde{\chi}_{\mathbf{I}_3 \mathbf{I}_2}(t_3, t_2; t_4) W_{\mathbf{I}_4 \mathbf{I}_2}(t_4, t_1). \quad (22b)$$

The equation of motion for the vertex can be derived by using a generalized chain rule written in the spinorial basis; this is derived in Appendix C 3:

$$\begin{aligned} \tilde{\Gamma}_{\mathbf{I}_3 \mathbf{I}_2}^{\mathbf{I}_4 \mathbf{I}_5}(t_3, t_2; t_4) &= \delta_{\mathbf{I}_3, \mathbf{I}_4} \delta_{\mathbf{I}_2, \mathbf{I}_5} \delta(t_3 - t_4) \delta(t_4 - t_2) \\ &+ \frac{\delta \Sigma_{\mathbf{I}_3 \mathbf{I}_2}(t_3, t_2)}{\delta G_{\mathbf{I}_6 \mathbf{I}_7}(t_6, t_7)} G_{\mathbf{I}_6 \mathbf{I}_8}(t_6, t_8) \\ &\times \tilde{\Gamma}_{\mathbf{I}_8 \mathbf{I}_9}^{\mathbf{I}_4 \mathbf{I}_5}(t_8, t_9; t_4) G_{\mathbf{I}_9 \mathbf{I}_7}(t_9, t_7). \end{aligned} \quad (23)$$

From Eqs. (23) and (B10) we obtain the equation for the response function in the spinorial basis. Indeed,

$$\tilde{L}_{\mathbf{I}_1 \mathbf{I}_1'}^{\mathbf{I}_2 \mathbf{I}_2'}(t_1, t_2) = G_{\mathbf{I}_1 \mathbf{I}_3}(t_1, t_3) \tilde{\Gamma}_{\mathbf{I}_3 \mathbf{I}_4}^{\mathbf{I}_2 \mathbf{I}_2'}(t_3, t_4; t_2) G_{\mathbf{I}_4 \mathbf{I}_1'}(t_4, t_1). \quad (24)$$

It is crucial to observe that  $\tilde{L}$  is a two times and four space points function that can be contracted to define  $\tilde{\chi}$ , as explained in Appendix C 4. Equations (21)–(24) represent the spinorial form of Hedin’s equations.

### A. The GW approximation

Starting from Eqs. (22b) and (23), the GW approximation follows from choosing

$$\tilde{\Gamma}_{\mathbf{I}_3 \mathbf{I}_2}^{\mathbf{I}_4 \mathbf{I}_5}(t_3, t_2; t_4) \approx \delta_{\mathbf{I}_3, \mathbf{I}_4} \delta_{\mathbf{I}_2, \mathbf{I}_5} \delta(t_3 - t_4) \delta(t_4 - t_2), \quad (25)$$

from which

$$\Sigma_{\mathbf{I}\mathbf{J}}^{\text{GW}}(t_1, t_2) = -iG_{\text{LM}}(t_1, t_2) W_{\mathbf{J}\mathbf{M}}^{\text{RPA}}(t_1, t_2). \quad (26)$$

Starting from the GW self-energy, different flavors of the scheme can be considered. The  $G_0W_0$  flavor assumes  $G_{\mathbf{I}\mathbf{J}}(t_1, t_2) \sim G_{\mathbf{I}\mathbf{J}}^{\text{KS}}(t_1, t_2)$  and also  $W^{\text{RPA}}$  functional of the KS states only.

From this point up to the end of the present section, we will use the extended form of the spinor indexes,  $\mathbf{I} \rightarrow (n\mathbf{k})$ , together with the translational invariance. Thus we assume to be in a perfectly periodic system, where the Coulomb interaction and the response function are represented as a Fourier expansion in terms of plane waves,  $\mathbf{G}$ , and transferred momenta,  $\mathbf{q}$ ,

$$v(\mathbf{x} - \mathbf{x}') = \sum_{\mathbf{G}} \int \frac{d\mathbf{q}}{(2\pi)^3} \frac{4\pi}{|\mathbf{q} + \mathbf{G}|^2} e^{i(\mathbf{q} + \mathbf{G}) \cdot (\mathbf{x} - \mathbf{x}')} \quad (27a)$$

and

$$W^{\text{RPA}}(\mathbf{x}, \mathbf{x}'; \omega) = \sum_{\mathbf{G}_1, \mathbf{G}_2} \int \frac{d\mathbf{q}}{(2\pi)^3} W_{\mathbf{G}_1 \mathbf{G}_2}^{\text{RPA}}(\mathbf{q}; \omega) e^{i(\mathbf{q} + \mathbf{G}_1) \cdot \mathbf{x}} e^{-i(\mathbf{q} + \mathbf{G}_2) \cdot \mathbf{x}'}, \quad (27b)$$

with  $W^{\text{RPA}}(\mathbf{x}_1, \mathbf{x}_2; \omega)$  the Fourier transform of  $W^{\text{RPA}}(1, 2)$ . The extended forms of Eq. (27) can be found in several references; see, for example, Ref. [1]. By using Eq. (27b), we finally get that

$$\begin{aligned} \Sigma_{nm\mathbf{k}}^{G_0W_0}(\omega) &= -i \int \frac{d\omega'}{2\pi} \int \frac{d^3\mathbf{q}}{(2\pi)^3} \\ &\times \sum_{ij\mathbf{G}_1\mathbf{G}_2} G_{ij\mathbf{k}-\mathbf{q}}^{\text{KS}}(\omega - \omega') W_{\mathbf{G}_1\mathbf{G}_2}^{\text{RPA}}(\mathbf{q}, \omega') \\ &\times \rho_{n\mathbf{k}}^{\mathbf{q}}(\mathbf{G}_1) \rho_{j\mathbf{m}\mathbf{k}}^{\mathbf{q}*}(\mathbf{G}_2) \end{aligned} \quad (28)$$

with

$$\rho_{nm\mathbf{k}}^{\mathbf{q}}(\mathbf{G}) = \sum_s \int d\mathbf{x} \phi_{n\mathbf{k}s}(\mathbf{x}) \phi_{m\mathbf{k}-\mathbf{q}s}^*(\mathbf{x}) e^{i(\mathbf{q} + \mathbf{G}) \cdot \mathbf{x}}. \quad (29)$$

The Fourier transform of  $G_{nm\mathbf{k}}^{\text{KS}}(t_1, t_2)$  can be conveniently written as

$$G_{nm\mathbf{k}}^{\text{KS}}(\omega) = \delta_{n,m} \left[ \frac{(1 - f_{n\mathbf{k}})}{\omega - \epsilon_{n\mathbf{k}}^{\text{KS}} + iO^+} + \frac{f_{n\mathbf{k}}}{\omega - \epsilon_{n\mathbf{k}}^{\text{KS}} - iO^+} \right]. \quad (30)$$

We have now all the ingredients to calculate the self-energy. Indeed, thanks to the definition Eq. (29), all can be recast in the product of simple oscillators that can be efficiently calculated via fast Fourier transformation techniques.

The use of the KS Hamiltonian as a zeroth-order term of the total Hamiltonian implies also that  $\Sigma_{nm\mathbf{k}}(\omega)$  needs to be replaced by  $\Sigma_{nm\mathbf{k}}(\omega) - v_{nm\mathbf{k}}^{\text{xc}}$  in the Dyson equation for  $G$ . It follows then that the Dyson equation reads

$$G_{nm\mathbf{k}}(\omega) = G_{nm\mathbf{k}}^{\text{KS}}(\omega) + G_{n\mathbf{k}}^{\text{KS}}(\omega) [\Sigma_{j\mathbf{k}}(\omega) - v_{j\mathbf{k}}^{\text{xc}}] G_{j\mathbf{m}\mathbf{k}}(\omega). \quad (31)$$

The last approximation in the  $G_0W_0$  flavor is to assume that only the energies need to be corrected and not the wave functions. This implies

$$\Sigma_{nmk}(\omega) \approx \delta_{nm}\Sigma_{nk}(\omega), \quad (32a)$$

$$G_{nmk}(\omega) \approx \delta_{nm}G_{nk}(\omega). \quad (32b)$$

From Eq. (32b) we obtain the final form of the Dyson equation, Eq. (B4) used in this work,

$$\epsilon_{nk}^{\text{GW}} \approx \epsilon_{nk}^{\text{KS}} + [\Sigma_{nmk}(\epsilon_{nk}^{\text{GW}}) - v_{nmk}^{\text{xc}}]. \quad (33)$$

#### IV. THE BETHE-SALPETER EQUATION

The spinorial Bethe-Salpeter equation can be derived from the general spinorial Hedin's equations by some manipulations that we outline in the following. Let us start by introducing the static limit of the GW self-energy, the so called screened exchange (SEX) approximation

$$\Sigma_{\mathbf{I}_1\mathbf{I}_2}^{\text{SEX}}(t) = -iG_{\mathbf{I}_4\mathbf{I}_3}(t)W_{\mathbf{I}_1\mathbf{I}_1, \mathbf{I}_3\mathbf{I}_2}^{\text{st}}, \quad (34)$$

with  $W^{\text{st}} = \delta(t_1 - t_2)W^{\text{RPA}}(t_1, t_2)$ . The approximation introduced by Eq. (34) is crucial in turning the BSE, an equation for a four-point Green's function  $\tilde{L}$ , into a simpler equation for a two-time-point function. Still the general solution is a four-index function  $\tilde{L}_{\mathbf{I}_1\mathbf{I}_1', \mathbf{I}_2\mathbf{I}_2'}(\omega)$ .

We can now easily calculate the functional derivative  $\frac{\delta\Sigma_{\mathbf{I}_3\mathbf{I}_2}}{\delta G_{\mathbf{I}_6\mathbf{I}_7}}$ , which defines the kernel of the BSE for the vertex function:

$$\frac{\delta\Sigma_{\mathbf{I}_3\mathbf{I}_2}(t_3, t_2)}{\delta G_{\mathbf{I}_6\mathbf{I}_7}(t_6, t_7)} \approx \delta(t_3 - t_6)\delta(t_2 - t_7)W_{\mathbf{I}_3\mathbf{I}_6, \mathbf{I}_7\mathbf{I}_2}^{\text{st}}, \quad (35)$$

where we have assumed the derivative of  $W^{\text{st}}$  to be negligible. Thanks to the approximation Eq. (34), the equation of motion for the vertex acquires a simple form that can be solved in subspace of single-frequency vertex functions,  $\Gamma(\omega)$ . From Eq. (35) and by using Eq. (32b) we can work out the BSE for the spinorial vertex in the SEX approximation:

$$\Gamma_{\mathbf{I}_3\mathbf{I}_2, \mathbf{I}_4\mathbf{I}_5}(\omega) = \delta_{\mathbf{I}_3, \mathbf{I}_4}\delta_{\mathbf{I}_2, \mathbf{I}_5} + i \left[ W_{\mathbf{I}_3\mathbf{I}_6, \mathbf{I}_7\mathbf{I}_2}^{\text{st}} - V_{\mathbf{I}_3\mathbf{I}_6, \mathbf{I}_7\mathbf{I}_2} \right] \times G_{\mathbf{I}_6\mathbf{I}_6}(\omega)\tilde{\Gamma}_{\mathbf{I}_6\mathbf{I}_7, \mathbf{I}_4\mathbf{I}_5}(\omega)G_{\mathbf{I}_7\mathbf{I}_7}(\omega). \quad (36)$$

To connect Eq. (36) to an equation of motion for the response function, the BSE, we now move from the  $\mathbf{I}$  basis to the explicit  $(n\mathbf{k})$  presentation. We start by introducing, for the general representation of  $\tilde{L}_{\mathbf{I}_1\mathbf{I}_1', \mathbf{I}_2\mathbf{I}_2'}(t)$  evaluated for a given transferred momentum  $\mathbf{q}$ ,

$$\mathbf{I}_1 \equiv (n\mathbf{k}), \quad (37a)$$

$$\mathbf{I}_1' \equiv (n'\mathbf{k} - \mathbf{q}), \quad (37b)$$

$$\mathbf{I}_2 \equiv (m\mathbf{p}), \quad (37c)$$

$$\mathbf{I}_2' \equiv (m'\mathbf{p} - \mathbf{q}). \quad (37d)$$

We denote as  $\tilde{L}_{\mathbf{I}_1\mathbf{I}_1', \mathbf{I}_2\mathbf{I}_2'}^{nn'\mathbf{k}}(\mathbf{q}, t)$  the response function whose scattering geometry is defined by Eq. (37). Equation (36) now defines an equation for  $\tilde{L}_{\mathbf{I}_1\mathbf{I}_1', \mathbf{I}_2\mathbf{I}_2'}^{nn'\mathbf{k}}(q, \omega)$  as, following the notation

Eq. (37), we can write

$$\tilde{L}_{\mathbf{I}_1\mathbf{I}_1', \mathbf{I}_2\mathbf{I}_2'}^{nn'\mathbf{k}}(\mathbf{q}, \omega) = G_{\mathbf{I}_1\mathbf{I}_1'}(\omega - \bar{\omega})G_{\mathbf{I}_2'\mathbf{I}_2}(\bar{\omega})\Gamma_{\mathbf{I}_1\mathbf{I}_1', \mathbf{I}_2\mathbf{I}_2'}^{nn'\mathbf{k}}(\omega). \quad (38)$$

By putting together Eqs. (38) and (36), we get the final equation for  $\tilde{L}$ :

$$L_{\mathbf{I}_1\mathbf{I}_1', \mathbf{I}_2\mathbf{I}_2'}^{nn'\mathbf{k}}(\mathbf{q}, \omega) = L_{\mathbf{I}_1\mathbf{I}_1', \mathbf{I}_2\mathbf{I}_2'}^{0, nn'\mathbf{k}}(\mathbf{q}, \omega)L_{\mathbf{I}_1\mathbf{I}_1', \mathbf{I}_2\mathbf{I}_2'}^{0, nn'\mathbf{k}}(\mathbf{q}, \omega)K_{\mathbf{I}_1\mathbf{I}_1', \mathbf{I}_2\mathbf{I}_2'}^{nn'\mathbf{k}}(\mathbf{q})\tilde{L}_{\mathbf{I}_1\mathbf{I}_1', \mathbf{I}_2\mathbf{I}_2'}^{mm'\mathbf{p}}(\mathbf{q}, \omega) \quad (39)$$

with

$$L_{\mathbf{I}_1\mathbf{I}_1', \mathbf{I}_2\mathbf{I}_2'}^{0, nn'\mathbf{k}}(\mathbf{q}, \omega) \equiv \delta(\mathbf{k} - \mathbf{p})\delta_{nm}\delta_{n'm'}G_{nmk}(\omega - \bar{\omega})G_{n'n'k-q}(\bar{\omega}), \quad (40)$$

$$-iK_{\mathbf{I}_1\mathbf{I}_1', \mathbf{I}_2\mathbf{I}_2'}^{nn'\mathbf{k}}(\mathbf{q}) \equiv \left[ W_{\mathbf{I}_3\mathbf{I}_6, \mathbf{I}_7\mathbf{I}_2}^{\text{st}}(\mathbf{k} - \mathbf{p}) - V_{\mathbf{I}_3\mathbf{I}_6, \mathbf{I}_7\mathbf{I}_2}^{nn'\mathbf{k}}(\mathbf{q}) \right]. \quad (41)$$

$\int \frac{d^3\bar{\mathbf{k}}}{(2\pi)^3}$ ,  $\int \frac{d^3\bar{\mathbf{p}}}{(2\pi)^3}$ , and  $\int \frac{d\bar{\omega}}{(2\pi)}$  are implicit. Equation (39) is the spinorial Bethe-Salpeter equation written for a generic transferred momentum  $\mathbf{q}$ . In the  $\mathbf{q} \rightarrow \mathbf{0}$  (optical) limit it reduces to the optical BSE, which we will use to study the optical properties of group VI TMDs in the next section.

As in the  $G_0W_0$  case, Eq. (39) looks the same as the scalar BSE. Indeed, the only difference is the definition of the oscillators defined in Eq. (29). It also follows that, as in the spin-independent case, the solution of Eq. (39) can be recast in an eigenvalue problem. To show this, it is enough to solve Eq. (39) by noting that  $L^0(\omega)$  is a sum of simple single-pole functions. Carefully separating the resonant and the antiresonant term, the eigenvalue problem can be defined,

$$H^{\text{exc}}A^\lambda = E_\lambda MA^\lambda, \quad (42)$$

in terms of an excitonic matrix  $H^{\text{exc}}$  [1] and a metric tensor,

$$M = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (43)$$

The final form of  $L_{\mathbf{I}_1\mathbf{I}_1', \mathbf{I}_2\mathbf{I}_2'}^{nn'\mathbf{k}}(\omega)$  can be expressed in terms of the eigenstate of the BSE matrix. It has a particularly simple expression if the resonant-only contribution is considered:

$$L_{\mathbf{I}_1\mathbf{I}_1', \mathbf{I}_2\mathbf{I}_2'}^{nn'\mathbf{k}}(\omega) = \left( \sum_{\lambda} \frac{A_{\mathbf{I}_1\mathbf{I}_1', \mathbf{I}_2\mathbf{I}_2'}^{\lambda, *} A_{\mathbf{I}_1\mathbf{I}_1', \mathbf{I}_2\mathbf{I}_2'}^{\lambda}}{\omega - E_\lambda + i0^+} \right). \quad (44)$$

Starting from the eigenvectors of the excitonic Hamiltonian, we can define the excitonic-state wave function as a linear

combination of electron-hole pairs:

$$|\vec{\lambda}\rangle = \sum_{nm\mathbf{k}} A_{nm\mathbf{k}}^{\lambda} |\vec{n\mathbf{k}}\rangle \otimes |\vec{m\mathbf{k}}\rangle. \quad (45)$$

The electron-hole pairs contributing to the excitonic wave function are vectors in spin space. In practice, this means that the exciton is, in the noncollinear case, a linear combination of the four possible spin orientations of the electron and hole, i.e., a tensor.

### A. Blocking of the BSE matrix and BSE spin structure

#### 1. The magnetic case

If  $\hat{v}^{(\text{SOC})}$  is weak, the noncollinearity of the KS eigenstates can be neglected as a first step and then the SOC correction treated perturbatively. In the collinear limit case,  $s_z$ , i.e., the spin projection of electrons (and holes), becomes a good quantum number and accordingly also  $S_z$ , i.e., the spin projection of the exciton. The extra quantum number can be added in Eq. (8), which becomes

$$\langle \mathbf{x} | \vec{n\mathbf{k}s} \rangle \equiv \vec{\psi}_{\mathbf{I}s}(\mathbf{x}) = \phi_{n\mathbf{k}s}(\mathbf{x}) |\vec{s}\rangle, \quad (46)$$

where  $|\vec{s}\rangle$  is either  $|\vec{\uparrow}\rangle = (1, 0)^t$  (for  $s = +1/2$ ) or  $|\vec{\downarrow}\rangle = (0, 1)^t$  (for  $s = -1/2$ ); the superscript “ $t$ ” indicates the transposition operation. If a collinear calculation is performed, in practice one can compute just  $\phi_{n\mathbf{k}s}(\mathbf{x})$  and reconstruct the spinorial wave functions via Eq. (46). This would be, in principle, also the result of a spinorial calculation with a collinear Hamiltonian. However, in this second case, whenever  $\epsilon_{n\mathbf{k}\uparrow} = \epsilon_{n\mathbf{k}\downarrow}$ , the resulting wave functions will be a random (but orthogonal) linear combination of  $\vec{\psi}_{\mathbf{I}\uparrow}(\mathbf{x})$  and  $\vec{\psi}_{\mathbf{I}\downarrow}(\mathbf{x})$  wave functions.

In the collinear limit,  $H^{\text{exc}}$  can be blocked in two matrices with half the size of the spinorial BSE: the spin-conserving ( $\Delta S_z = 0$ ) transitions or “excitons” ( $\lambda_e$ ) and the spin-flip ( $\Delta S_z = \pm 1$ ) transitions or “magnons” ( $\lambda_m$ ), with

$$|\lambda_e\rangle = \sum_{c\mathbf{v}\mathbf{k}s} A_{c\mathbf{v}\mathbf{k}s}^{\lambda_e} |c\mathbf{k}, s\rangle \otimes |v\mathbf{k}, s\rangle, \quad (47a)$$

$$|\lambda_m\rangle = \sum_{c\mathbf{v}\mathbf{k}s} A_{c\mathbf{v}\mathbf{k}s}^{\lambda_m} |c\mathbf{k}, s\rangle \otimes |v\mathbf{k}, -s\rangle. \quad (47b)$$

Excitons and magnons, in Eq. (47), distinguish the two possible spin combinations of the electron-hole pair. In the collinear case, only excitons, where the  $c$  and  $v$  states have the same spin, can be excited by the laser pulse. Magnons cannot be generated, since optical transitions between states with opposite spin are forbidden. Notice that the magnon block is also composed by two independent subblocks, which are in general different, i.e.,  $\Delta S_z = +1$  and  $-1$ . These are two independent sets of excitations when  $S_z$  is a good quantum number. For an analysis of BSE applied to the magnon channel, see, for example, Ref. [30]. Instead, the total spin is not yet a good quantum number, and the origin of this can be traced back to the fact that  $|\phi_{n\mathbf{k}\uparrow}(\mathbf{x})|^2 \neq |\phi_{n\mathbf{k}\downarrow}(\mathbf{x})|^2$ . This is well known in the literature of quantum chemistry, where the term “spin contamination” is used, and restricted calculations, which indeed impose  $\phi_{n\mathbf{k}\uparrow}(\mathbf{x}) = \phi_{n\mathbf{k}\downarrow}(\mathbf{x})$  and  $\epsilon_{n\mathbf{k}\uparrow} = \epsilon_{n\mathbf{k}\downarrow}$ , are sometimes performed. In extended systems, the breaking of spin

symmetry is instead regarded as less important; the exchange splitting  $\Delta_{n\mathbf{k}} = \epsilon_{n\mathbf{k}\uparrow} - \epsilon_{n\mathbf{k}\downarrow}$  is seen as a physical quantity.

#### 2. The nonmagnetic case

Instead if the system is nonmagnetic,  $\epsilon_{n\mathbf{k}\uparrow} = \epsilon_{n\mathbf{k}\downarrow}$  for any  $n\mathbf{k}$ , the ground state has total spin  $S = 0$ , and

$$\phi_{n\mathbf{k}s}(\mathbf{x}) = e^{i\alpha_{n\mathbf{k}s}} \phi_{n\mathbf{k}}(\mathbf{x}), \quad (48)$$

with  $\alpha_{n\mathbf{k}s}$  an arbitrary phase factor. In practical calculations only  $\phi_{n\mathbf{k}}(\mathbf{x})$  are computed and  $\alpha_{n\mathbf{k}s} = 0$  is assumed. However, if a collinear spin-dependent calculation is explicitly performed on a nonmagnetic system, random phases  $\alpha_{n\mathbf{k}\uparrow} - \alpha_{n\mathbf{k}\downarrow}$  will be present in between the two spin channels. Using Eq. (48), i.e., the BSE Hamiltonian, the exciton channel ( $\Delta S_z = 0$ ) can be further blocked into singlets ( $S$ ) with  $(S, S_z) = (0, 0)$  and triplets ( $T$ ) with  $(S, S_z) = (1, 0)$ . The “magnons” channel remains composed of two blocks, which now represent the triplets  $(S, S_z) = (1, +1)$  and  $(S, S_z) = (1, -1)$ . All triplet blocks are degenerate and identical up to the phases  $\alpha_{n\mathbf{k}}$ . Starting from the solutions of the unpolarized BSE  $\tilde{A}_{c\mathbf{v}\mathbf{k}}^{\lambda_s}$  ( $\tilde{A}_{c\mathbf{v}\mathbf{k}}^{\lambda_T}$ ) in the singlet (triplet) block, the full eigenvectors can be reconstructed as

$$A_{c\mathbf{v}\mathbf{k}s}^{\lambda_s} = (-1)^{2s+1} \frac{e^{i(\alpha_{c\mathbf{k}s} - \alpha_{v\mathbf{k}s})}}{\sqrt{2}} \tilde{A}_{c\mathbf{v}\mathbf{k}}^{\lambda_s}, \quad (49a)$$

$$A_{c\mathbf{v}\mathbf{k}s}^{\lambda_T} = \frac{e^{i(\alpha_{c\mathbf{k}s} - \alpha_{v\mathbf{k}s})}}{\sqrt{2}} \tilde{A}_{c\mathbf{v}\mathbf{k}}^{\lambda_T}. \quad (49b)$$

The factor  $\sqrt{2}$  in Eq. (49) ensures that the eigenvectors are normalized to 1.

#### 3. BSE matrix blocking

To summarize, let us explicitly write the general spin structure of the matrix relabelling the states  $n\mathbf{k}$  (with  $n = 1, \dots, N$ ) as  $\bar{n}_s\mathbf{k}$  (with  $\bar{n} = 1, \dots, N/2$ ,  $s = \uparrow, \downarrow$ ). This is just an exact relabeling, which becomes meaningful, i.e.,  $\bar{n}_\uparrow\mathbf{k}$  ( $\bar{n}_\downarrow\mathbf{k}$ ) refers to a “spin-up” (“spin-down”) state if  $\hat{v}^{(\text{SOC})}$  is small and the collinear case notation can be recovered, i.e.,  $\{\bar{n}\uparrow\mathbf{k}\} = \{m\mathbf{k}\uparrow\}$  if  $\hat{v}^{(\text{SOC})} = 0$ . For each set of indexes  $\{\bar{n}\bar{n}'\mathbf{k}, \bar{m}\bar{m}'\mathbf{p}\}$ ,

$$H^{\text{res}} = H^0 + H^{\text{exc}} + H^{e-h\text{-int}} \quad (50)$$

is written in terms of  $4 \times 4$  matrices in the “spin” indexes [31]:

$$H_{s_1 s_2 s_3 s_4}^0 = (\epsilon_{s_1} - \epsilon_{s_2}) \delta_{s_1, s_3} \delta_{s_2, s_4}, \quad (51a)$$

$$H_{s_1 s_2 s_3 s_4}^{\text{exc}} = -iV_{s_1 s_2 s_3 s_4}, \quad (51b)$$

$$H_{s_1 s_2 s_3 s_4}^{e-h\text{-int}} = iW_{s_1 s_2 s_3 s_4}^{\text{st}}. \quad (51c)$$

We have the following three cases:

(i) Noncollinear case. Neither  $S_z$  nor  $S$  is a good quantum number, thus excitons and magnons are mixed.  $H^{\text{exc}}$  is an  $N \times N$  matrix. All the matrix elements of Eq. (51) can be different from zero.

(ii) Collinear magnetic case.  $S_z$  is a good quantum number, while  $S$  is not. The BSE can be blocked in two matrices,  $H^e$  and  $H^m$ , of size  $N/2 \times N/2$  each, which describe separately excitons and magnons. This results from the fact that, for the collinear case,  $V_{s_1 s_2 s_3 s_4} \propto \delta_{s_1, s_2} \delta_{s_3, s_4}$  and  $W_{s_1 s_2 s_3 s_4}^{\text{st}} \propto \delta_{s_1, s_3} \delta_{s_2, s_4}$ , and thus all matrix elements that couple the  $\Delta S_z = 0$  channel to the  $\Delta S_z = \pm 1$  channels are zero.

(iii) Collinear nonmagnetic case. Both  $S_z$  and  $S$  are good quantum numbers. The BSE can be further blocked with four blocks in total, three of which carry the same information. The exciton channel  $H^e$  generates the singlet block ( $\Delta S = 0$ ) and the triplet block ( $\Delta S = 1$ ). The two blocks resulting from the magnon channel  $H^m$  are equivalent to the triplet block. Indeed, since the ground state is nonmagnetic, the distinction between excitons and magnons becomes meaningless. In total, there are two matrices,  $H^{S/T}$ , of size  $N/4 \times N/4$  each. This results from the fact that  $V_{s_1 s_1 s_3 s_3} = V$  for any  $\{s_1, s_3\}$  and  $W_{s_1 s_1 s_2 s_2}^{\text{st}} = W^{\text{st}}$  for any  $\{s_1, s_2\}$ . The blocking in this second step is obtained via the vectors  $1/\sqrt{2}(1, 1)^t$  and  $1/\sqrt{2}(1, -1)^t$ . As a consequence,  $H^{\text{exch}, S} = 2V$  while  $H^{\text{exch}, T} = 0$  and all triplet states are lower in energy.

## V. THE EXCITONIC SPIN POLARIZATION

In the previous section, we discussed the spin structure of  $H^{\text{exc}}$  in different cases and connected the different sized of the matrix to the spin of the exciton in the singlets and triplets channels. We start the present section introducing the matrix that relates the spin of the electron and the hole to the total spin  $S$  and its projection  $S_z$  of the electron-hole pair. We define the triplet and singlet spin states in the usual way,  $|S, S_z\rangle$  [32],

$$|S, S_z\rangle = R_{s_e s_h}^{SS_z} |s_e\rangle \otimes |s_h\rangle. \quad (52)$$

$R_{s_e s_h}^{SS_z}$  are matrices in the basis of the products of the electron and hole spins. These can be written as  $2 \times 2$  matrices as follows:

$$\underline{R}^{1,-1} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad (53a)$$

$$\underline{R}^{1,0} = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \quad (53b)$$

$$\underline{R}^{1,1} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad (53c)$$

$$\underline{R}^{0,0} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \quad (53d)$$

We start by expanding Eq. (45) in the basis of space/spin components, and we write explicitly the  $s_e s_h$  component of the excitonic state:

$$|\lambda\rangle = \sum_{s_e s_h} \int d^3 \mathbf{r}_e d^3 \mathbf{r}_h |s_e s_h\rangle |\mathbf{r}_e \mathbf{r}_h\rangle \langle \mathbf{r}_e \mathbf{r}_h | s_e s_h \lambda \rangle. \quad (54)$$

Equation (54) defines the excitonic wave function projected on the spin state  $|s_e s_h\rangle |\mathbf{r}_e \mathbf{r}_h\rangle$ :

$$\Psi_{s_e s_h}^\lambda(\mathbf{r}_e \mathbf{r}_h) \equiv \langle \mathbf{r}_e \mathbf{r}_h | s_e s_h \lambda \rangle. \quad (55)$$

Thanks to Eq. (52), we can rewrite Eq. (55) in terms of the components on the triplet and singlet spin states,

$$|\lambda\rangle = \sum_{SS_z} \int d^3 \mathbf{r}_e d^3 \mathbf{r}_h |S, S_z\rangle |\mathbf{r}_e \mathbf{r}_h\rangle \langle S, S_z | \lambda \rangle, \quad (56)$$

with

$$\Psi_{S, S_z}^\lambda(\mathbf{r}_e \mathbf{r}_h) \equiv \langle \mathbf{r}_e \mathbf{r}_h | S, S_z \lambda \rangle = \sum_{s_e s_h} R_{s_e s_h}^{S, S_z} \Psi_{s_e s_h}^\lambda(\mathbf{r}_e \mathbf{r}_h). \quad (57)$$

Thanks to Eq. (57), we can introduce several observables that can efficiently describe the noncollinearity of the excitonic state.

We start from the normalization condition:

$$\delta_{\lambda\lambda'} = \langle \lambda | \lambda' \rangle = \sum_{S, S_z} \int d^3 \mathbf{r}_e d^3 \mathbf{r}_h |\Psi_{S, S_z}^\lambda(\mathbf{r}_e \mathbf{r}_h)|^2 = \sum_{S, S_z} N_{S, S_z}^\lambda, \quad (58)$$

with

$$\begin{aligned} N_{S, S_z}^\lambda &\equiv |\langle S, S_z | \lambda \rangle|^2 \\ &= \int d^3 \mathbf{r}_e d^3 \mathbf{r}_h \left| \sum_{c v \mathbf{k} s_e s_h} A_{c v \mathbf{k}}^\lambda R_{s_e s_h}^{S, S_z} \phi_{c \mathbf{k} s_e}(\mathbf{r}_e) \phi_{v \mathbf{k} s_h}^*(\mathbf{r}_h) \right|^2. \end{aligned} \quad (59)$$

It is worth noting that when expanding the square modulus of Eq. (59), only a single summation over  $\mathbf{k}$  survives due to momentum conservation, whereas on all the other indexes double summations remain. From Eq. (59) it follows that we can define an excitonic total average spin and momentum as

$$S_\lambda^2 = \langle \lambda | \hat{S}^2 | \lambda \rangle = \sum_{S_z} S(S+1) N_{S, S_z}^\lambda, \quad (60a)$$

$$S_{z, \lambda} = \langle \lambda | \hat{S}_z | \lambda \rangle = \sum_S S_z N_{S, S_z}^\lambda. \quad (60b)$$

### A. The collinear magnetic case

The meaning of Eq. (59) can be better understood by taking the collinear case. From Eq. (59) it follows that we can define the spin polarization of excitons and magnons as

$$\begin{aligned} N_{S, S_z}^{\lambda e, m} &= \sum_{\substack{c v \\ c' v'}} \mathbf{k} \sum_{s s'} A_{c' v' \mathbf{k} s'}^{\lambda e, m, *} A_{c v \mathbf{k} s}^{\lambda e, m} R_{s \pm s}^{S, S_z, *} R_{s' \pm s'}^{S, S_z} \langle c' \mathbf{k}, s' | c \mathbf{k}, s \rangle \\ &\times \langle v \mathbf{k}, \pm s | v' \mathbf{k}, \pm s' \rangle, \end{aligned} \quad (61)$$

where the  $+$  and  $-$  signs stand for the excitons and magnons channel, respectively. Note that in Eq. (61) the two inner products are between the spatial part of the wave function, which does not impose the spin to be conserved.

To manipulate Eq. (61), we observe, from Eq. (53), that the specific form of the  $R$  matrices imposes that in the  $N_{S, S_z}^{\lambda e}$  case only the  $(S, S_z) = (1, 0)$  and  $(0, 0)$  are nonzero. On the contrary, in the  $N_{S, S_z}^{\lambda m}$  case the nonzero components will be  $(S, S_z) = (1, \pm 1)$ . This marks the distinction between excitons and magnons.

After some simple manipulation, we get

$$N_{S0}^{\lambda e} = \frac{1}{2} \left\{ 1 \pm \sum_{\substack{c v \\ c' v'}} \left[ O_{c v \uparrow \mathbf{k}} + O_{c' v' \downarrow \mathbf{k}} \right] \right\} \quad (62a)$$

$$O_{c v s}^{\lambda e} = A_{c' v' \mathbf{k} s'}^{\lambda e, *} A_{c v \mathbf{k} s}^{\lambda e} \langle c' \mathbf{k}, s' | c \mathbf{k}, s \rangle \langle v \mathbf{k}, s | v' \mathbf{k}, s \rangle, \quad (62b)$$

where the  $+$  ( $-$ ) refers to  $S = 0$  ( $S = 1$ ). Similarly, we get

$$N_{11}^{\lambda_m} = \sum_{c\nu\mathbf{k}} |A_{c\nu\mathbf{k}\uparrow}^{\lambda_m}|^2, \quad (63)$$

$$N_{1-1}^{\lambda_m} = \sum_{c\nu\mathbf{k}} |A_{c\nu\mathbf{k}\downarrow}^{\lambda_m}|^2. \quad (64)$$

From Eqs. (62) and (64) it follows that, in general, also in the collinear case the value of  $S_\lambda$  is not fixed. This is again a manifestation of the fact that  $S$  is not a good quantum number in general. The collinear systems can be characterized by a nonvanishing magnetization that causes the  $\uparrow$  and  $\downarrow$  components of the electronic wave functions to differ by more than a simple phase factor. The consequence is a state-dependent value of the excitonic spin.

### B. The collinear nonmagnetic case

Let us conclude this section by considering the case of a collinear and nonmagnetic system. When Eqs. (48) and (49) are plugged into Eq. (61), it turns out that

$$N_{SS_z}^{\lambda_e(S)} = \delta_{S,0} \delta_{S_z,0}, \quad (65a)$$

$$N_{SS_z}^{\lambda_e(T_0)} = \delta_{S,1} \delta_{S_z,0}, \quad (65b)$$

$$N_{SS_z}^{\lambda_m(T_1)} = \delta_{S,1} \delta_{S_z,-1}, \quad (65c)$$

$$N_{SS_z}^{\lambda_m(T_{-1})} = \delta_{S,1} \delta_{S_z,1}, \quad (65d)$$

as it should from the blocking of the BSE matrix.

## VI. A PERTURBATIVE APPROACH TO THE EFFECT OF SPIN-ORBIT COUPLING

We now proceed to consider the SOC as a perturbation. To first order, the perturbation just gives a correction to the DFT eigenvalues and does not touch the wave functions. Numerically, the perturbation needs to be applied by mapping the solution of the DFT calculation without SOC into the solution of the DFT calculation with SOC. In the first part of this section, we discuss such mapping.

Once a mapping is defined, the SOC perturbation could be directly applied to the KS energies before solving both the GW scheme and the BSE or after. In the second part of this section, we discuss such a distinction.

### A. Mapping procedure

Now, in order to define a mapping procedure in practice we distinguish between the KS Hamiltonian without SOC,  $\hat{h}^{\text{KS},0}$ , and the standard KS Hamiltonian  $\hat{h}^{\text{KS}}$ , with  $\hat{h}^{\text{KS}} - \hat{h}^{\text{KS},0} = \hat{v}^{\text{(SOC)}}$ . Moreover, let us focus here on the situation in which the ground state is nonmagnetic, since it is the case for which calculations are actually performed in the present paper and also because it is the more complex case. The generalization to magnetic systems is straightforward.  $\hat{h}^{\text{KS},0}$  is then collinear and spin-independent. It has eigenvectors  $\phi_{n\mathbf{k}}^0(\mathbf{x})$  and energies  $\epsilon_{n\mathbf{k}}^{\text{KS},0}$ .

The key passage now is how to connect the eigenvectors of the SOC-free case to the full spinorial case. To this end, we define a map based on the overlap between  $\phi_{n\mathbf{k}}^0(\mathbf{x})$  and  $\vec{\phi}_{n\mathbf{k}}(\mathbf{x})$ . Since both represent a complete basis set for each

$k$ -point, we just need to expand  $\vec{\phi}_{n\mathbf{k}}(\mathbf{x})$  in terms of the spinor defined extending  $\phi_{n\mathbf{k}}^0(\mathbf{x})$ , first including the spin index [i.e., using Eq. (48)] and then constructing the spinors [i.e., using Eq. (46)]. Notice that in doing so, the gauge with  $e^{i\alpha_{n\mathbf{k}}} = 1$  is assumed. Moreover, in the degenerate spaces, the ‘‘up’’ and ‘‘down’’ spinors are always chosen among all possible random combinations. Let us call  $\vec{\phi}_{n\mathbf{k}}^0(\mathbf{x}) = \langle \mathbf{x} | \vec{m}\mathbf{k}s \rangle$  the result of such an expansion. Then it follows that

$$|\vec{n}\mathbf{k}\rangle = \sum_{ms} |\vec{m}\mathbf{k}s\rangle \langle \vec{m}\mathbf{k}s | \vec{n}\mathbf{k} \rangle = \sum_{ms} \Delta_{n,ms}^{\mathbf{k}} |\vec{m}\mathbf{k}s\rangle. \quad (66)$$

We now want compute the first-order perturbation theory correction to the energy with respect to the perturbation  $\hat{V} = \hat{h}^{\text{KS}} - \hat{h}^{\text{KS},0}$ . Using Eq. (66), it can be expressed as

$$\begin{aligned} \epsilon_{m\mathbf{k}s} &= \epsilon_{m\mathbf{k}}^0 + \langle \vec{m}\mathbf{k}s | (\hat{h}^{\text{KS}} - \hat{h}^{\text{KS},0}) | \vec{m}\mathbf{k}s \rangle \\ &= \epsilon_{m\mathbf{k}}^0 + (\langle \vec{m}\mathbf{k}s | \hat{h}^{\text{KS}} | \vec{m}\mathbf{k}s \rangle - \epsilon_{m\mathbf{k}}^0) \\ &= \epsilon_{m\mathbf{k}}^0 + \sum_n |\Delta_{n,ms}^{\mathbf{k}}|^2 \epsilon_{n\mathbf{k}} - \epsilon_{m\mathbf{k}}^0, \end{aligned} \quad (67)$$

where in the third line we inserted two completeness relations over  $|\vec{n}\mathbf{k}\rangle$ , we used the fact that  $|\vec{n}\mathbf{k}\rangle$  are eigenstates of  $\hat{h}^{\text{KS}}$ , and we used the definition of  $\Delta_{n,ms}^{\mathbf{k}}$ . Equation (67) is exact. We now define a generic mapping function,  $f_{\mathbf{k}}$ , which maps every eigenstate of  $\hat{h}^{\text{KS}}$  to the eigenstates of  $\hat{h}^{\text{KS},0}$ . In case neither  $|\vec{m}\mathbf{k}s\rangle$  nor  $|\vec{n}\mathbf{k}\rangle$  is degenerate, the mapping function can be defined as

$$f_{\mathbf{k}}(ms) = n \quad \text{if} \quad |\Delta_{n,ms}^{\mathbf{k}}|^2 = D_n^{\mathbf{k}}, \quad (68)$$

with  $D_n^{\mathbf{k}} = \max_{ms} |\Delta_{n,ms}^{\mathbf{k}}|^2$ . The extension to the degenerate case is discussed in Appendix F. Using the mapping procedure, we can approximate Eq. (67) as

$$\begin{aligned} \epsilon_{m\mathbf{k}s} &\approx \epsilon_{m\mathbf{k}}^0 + (\epsilon_{f_{\mathbf{k}}(ms)\mathbf{k}} - \epsilon_{m\mathbf{k}}^0) \\ &\approx \epsilon_{m\mathbf{k}}^0 + \Delta \epsilon_{m\mathbf{k}s}^{\text{SOC}}. \end{aligned} \quad (69)$$

Note also that for each eigenvalue  $\epsilon_{m\mathbf{k}}^0$ , a pair of spinorial states is selected, one for  $s = \uparrow$  and one for  $s = \downarrow$ , since to each  $m\mathbf{k}$  are associated two spinors via Eqs. (48) and (46). This property will be crucial to defining spin-conserving ( $\Delta S_z \approx 0$ ) and spin-flip ( $\Delta S_z \approx \pm 1$ ) excitations.

We also define the quality of the mapping as

$$Q = \min_{n\mathbf{k}} (D_n^{\mathbf{k}}), \quad (70)$$

and we monitor its value.  $0 < Q \leq 1$ , and  $Q = 1$  means the mapping is exact.

### B. Perturbative SOC within GW and BSE

We now want to use the mapping to apply the corrections in the GW-BSE scheme. Formally, one should first apply the SOC corrections and then solve the GW-BSE scheme. Applying them before would significantly increase the computational load, thus reducing the advantages of using a perturbative approach compared to the full approach. We thus want to apply the corrections afterward. In the next section, we will check the quality of the scheme against full SOC calculations.

Both in GW and in BSE, the KS energies enter in the RPA screening. In  $W^{\text{RPA}}$  one can reasonably expect no significant changes due to SOC. Then, as far as GW is concerned, the KS energies enter in two more points: the definition of the self-energy and the Dyson equation for  $G$ . The Dyson equation for  $G$  is recast in terms of Eq. (33) for the QP energies, and it gives

$$\epsilon_{mks}^{\text{GW}} \approx \epsilon_{mk}^{\text{KS}} + [\Sigma_{mmk}(\epsilon_{mk}^{\text{GW}}) - v_{mmk}^{\text{xc}}] + \Delta\epsilon_{mks}^{\text{SOC}}. \quad (71)$$

Since the perturbative approach assumes that only the energies change, all matrix elements remain in the collinear basis set. Only the self-energy depends explicitly on the energies. Thus small differences between applying SOC corrections before or after can be due to the self-energy. Further differences with the full SOC approach are expected to be higher order.

In the BSE scheme instead there is no straightforward way to apply the corrections afterward, and a new procedure needs to be defined. To this end, we follow the ‘‘perturbative-BSE’’ (pBSE) approach of Ref. [11], which, however, we critically discuss and refine. In the pBSE, the SOC potential extended to the  $e$ - $h$  basis set is defined as

$$V_{ss'}^{(\text{SOC})}(\mathbf{x}_c, \mathbf{x}_v) = [v_{ss',c}^{(\text{SOC})}(\mathbf{x}_c) \otimes \mathbb{1}_v] - [\mathbb{1}_c \otimes v_{ss',v}^{(\text{SOC})}(\mathbf{x}_v)]. \quad (72)$$

This potential is then used to define the SOC corrections  $\Delta\omega_{\lambda s}$  to the BSE eigenvalues without SOC  $\omega_{\lambda}^0$  via its expectation value on the excitonic state.

For the spin-conserving, optically active channel, we use the mapping procedure proposed in the literature [11]. Introducing

$$\Psi_{ss'}^{\lambda S,0}(\mathbf{x}_c, \mathbf{x}_v) = \delta_{s,s'} \sum_{cvks} A_{cvks}^{\lambda S} \phi_{cks}^{0,*}(\mathbf{x}_c) \phi_{vks}^0(\mathbf{x}_v),$$

where  $A_{cvks}^{\lambda S}$  is obtained from Eqs. (49), the SOC corrections are defined as

$$\begin{aligned} \Delta\omega_{\lambda}^{S,0} &= \langle \Psi^{0,\lambda} | \hat{V}^{(\text{SOC})} | \Psi^{0,\lambda} \rangle \\ &= \sum_{cvks} \sum_{c'v'k's'} A_{cvks}^{\lambda S,*} A_{c'v'k's'}^{\lambda S} \\ &\quad \times [ \langle \psi_{cks}^0 | \hat{v}^{(\text{SOC})} | \psi_{c'k's'}^0 \rangle \\ &\quad + \langle \psi_{vks}^0 | \hat{v}^{(\text{SOC})} | \psi_{v'k'±s}^0 \rangle ] \\ &\approx \sum_{cvks} |A_{cvks}^{\lambda S}|^2 (\Delta\epsilon_{cks}^{\text{SOC}} - \Delta\epsilon_{vks}^{\text{SOC}}). \end{aligned} \quad (73)$$

We used the mapping approximation and we neglect the terms with  $c \neq c'$ ,  $v \neq v'$ , and  $s \neq s'$ .  $\delta(\mathbf{k} - \mathbf{k}')$  is instead imposed by the Bloch Hamiltonians. This is a simplification, and one should carefully check how the SOC splitting compares with the exchange splitting (see Appendix E for more details).

For the spin flipping (or magnons) channel instead, the exchange interaction is always  $\approx 0$  and the mapping procedure is unique, with

$$\Psi_{ss'}^{\lambda T,\pm 1}(\mathbf{x}_c, \mathbf{x}_v) = \delta_{s,-s'} \sum_{cvk} A_{cvks}^{\lambda T} \phi_{ck±s}^{0,*}(\mathbf{x}_c) \phi_{v\mathbf{k}\mp s}^0(\mathbf{x}_v)$$

and

$$\Delta\omega_{\lambda}^{T,\pm 1} \approx \sum_{cvk} |A_{cvks}^{\lambda T}|^2 (\Delta\epsilon_{ck±s}^{\text{SOC}} - \Delta\epsilon_{v\mathbf{k}\mp s}^{\text{SOC}}). \quad (74)$$

Equations (73) and (74) are used in the present paper to compute the optical properties of TMDs.

## VII. RESULTS FOR PARADIGMATIC MATERIALS

Bulk TMDs are indirect gap semiconductors, but when going to a single layer, their gap becomes direct (or quasidirect), making them suitable for applications in the fields of electronics, optoelectronics, and sensing. The  $K^+$  and  $K^-$  points of the hexagonal Brillouin zone are the location of the band extrema of the TMDs. At these inequivalent points, linked by time-inversion symmetry, the spin-split valence-band maximum (VBM) and conduction-band minimum (CBM) are almost completely spin-polarized in an opposite way in the two valleys, allowing for a selective valley excitation by  $\sigma^+$ - and  $\sigma^-$ -polarized light. All the results that will be discussed from this point on were obtained using the computational methods and parameters presented in Appendix D.

### A. Effect of SOC in the band structure

As shown in Table I, the valence-band spin splitting at K is of several hundreds of meV, whereas the conduction band is split only a few tens of meV. Because the lowest SOC split conduction band can either have the same as or the opposite spin character of the top valence band, it is convenient to define the conduction-band splitting as  $\text{CBMS} = \epsilon_{\text{CBM}}^{\sigma} - \epsilon_{\text{CBM}}^{-\sigma}$ , where  $\sigma$  is the spin of the top of the valence band at K. In this way, the two different situations are marked by the sign of the splitting, i.e., if the bottom of the conduction band has the same spin character as the top of the valence band, the resulting CBMS will be negative.

Similarly to previous calculations, and shown in Figs. 2 and 3, we observe that for Mo-based systems, the bottom of the conduction at K has the same character of the top of the valence, whereas the opposite is true for W-based systems. We find a reduction (increase) of the absolute value of the CBM splitting for the  $\text{WX}_2$  ( $\text{MoX}_2$ ) systems, as the conduction state with the same spin character of the VBM is less corrected than the one with the opposite character (if  $\sigma$  is the spin of the top of the valence band at K,  $\Delta\epsilon_{\sigma,c}^{\text{GW}} < \Delta\epsilon_{-\sigma,c}^{\text{GW}}$ ). At the GW level, the sign of the CB splitting is consistent with the literature; the observed dispersion in the computed numerical values is possibly due to the use of the GdW approach in one case [37] and to different calculation parameters or numerical implementations in the remaining cases [33–36].

As a consequence of the positive/negative CBM splitting, based only on the analysis of the GW electronic states, it could be argued that for W-based systems, the optically active and spin-allowed transition should have higher energy with respect to the dark spin-forbidden transition, while the opposite should be true for Mo-based systems. However, it is not possible to predict whether the lowest optical excitation is spin-forbidden (dark) or spin-allowed (bright) only on the basis of the electronic band structure. The dark or bright character of the lowest optical excitation is indeed the result of a delicate interplay between the different contributions to the excitation energy, and the full treatment of SOC at a nonperturbative level may change qualitatively the picture.

TABLE I. SOC splitting of the highest valence and lowest conduction band (VBM, CBM) at K. The conduction-band splitting is defined as  $\text{CBMS} = \epsilon_{\text{CBM}}^{\sigma} - \epsilon_{\text{CBM}}^{-\sigma}$ , where  $\sigma$  is the spin of the top of the valence band at K. All energies are in meV.

	SOC splitting					
	Present work				Literature	
	VBM		CBM		CBM	
DFT	GW	DFT	GW	DFT	GW	
WSe <sub>2</sub>	458	516	46	25	37 [33] 40 [34]	7 [35], 10 [36], 6 [33]
WS <sub>2</sub>	408	394	33	12	33 [33]	10 [35], 12 [36], 5 [33]
MoSe <sub>2</sub>	186	191	-23	-29	-21 [33]	-42 [35], -41 [36], -14 [33]
MoS <sub>2</sub>	145	151	-3	-9	-3 [33]	-15 [35], -15 [36], -31 [33]

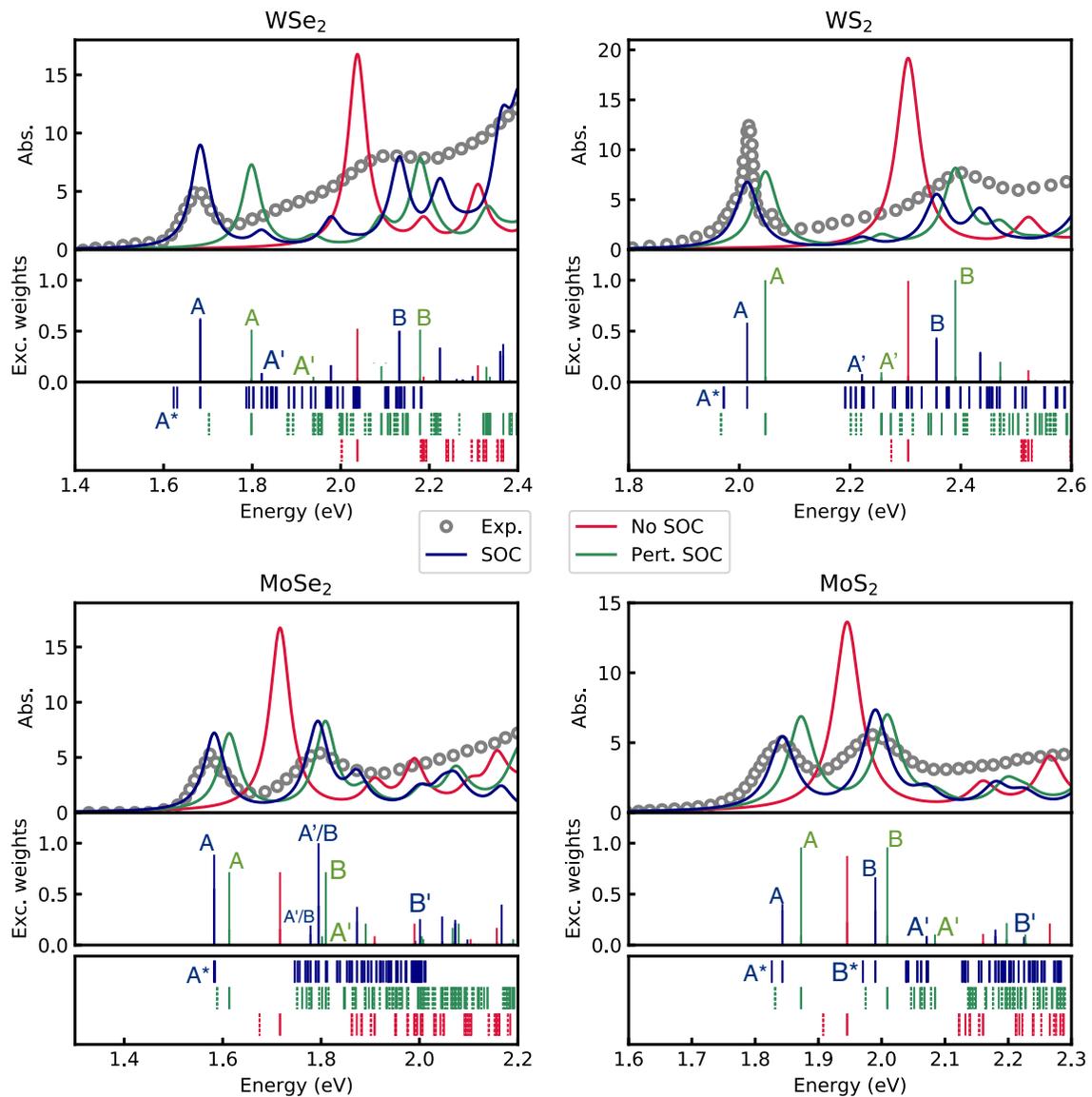


FIG. 1. Top panel: Bethe-Salpeter spectra of single-layer WSe<sub>2</sub>, WS<sub>2</sub>, MoS<sub>2</sub>, and MoSe<sub>2</sub>. Middle panel: oscillator strengths (normalized to 1). Bottom panel: each bar represents an excitation regardless of its oscillator strength. No SOC, full SOC, and perturbative SOC results are presented in red, blue, and green, respectively. All the theoretical curves have been rigidly redshifted  $\sim 0.1$ – $0.2$  eV to align the energetic position of peak A to the experimental one. Experimental data are taken from Ref. [38] for MoS<sub>2</sub>, and from Ref. [39] for the remaining systems.

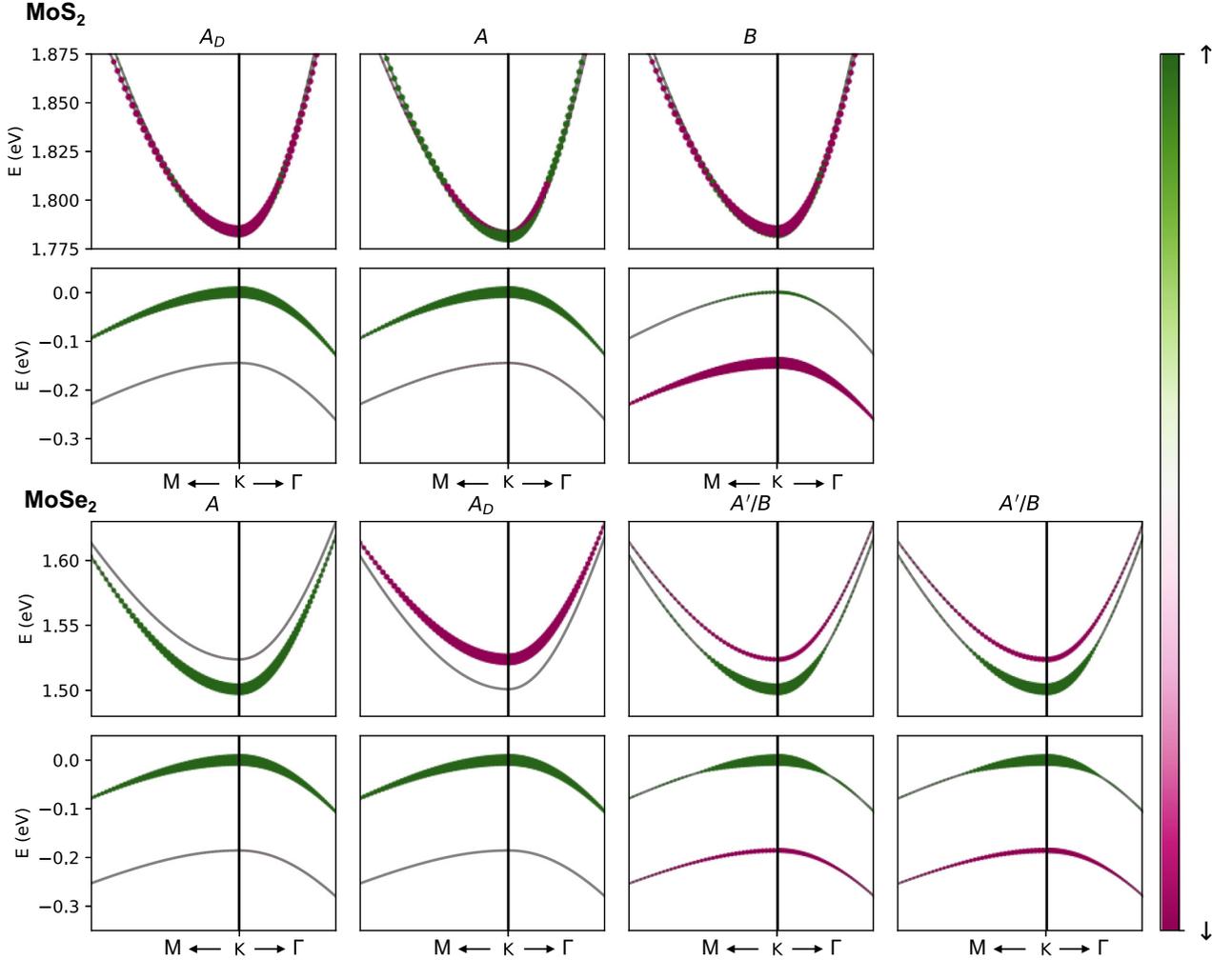


FIG. 2. Zoom of the spin-split VBM and CBM around K showing the single-particle contributions to the main optical excitations: the width of the points is proportional to the contribution of the state to the excitation, while the color refers to the spin character. Top panels: MoS<sub>2</sub>; bottom panels: MoSe<sub>2</sub>.

### B. Optical absorption spectra

In the top panels of Fig. 1, we show the GW+BSE absorbance spectra calculated without SOC and with perturbative and nonperturbative SOC, and we compare them with available experimental data [39]. In the middle panel, bars represent oscillator strengths for each excitation, while in the bottom panel they signal the presence of an excitation, independently from its oscillator strength. The optical spectra are characterized by the presence of spin-split strongly bound excitons and their corresponding series. The lowest bright excitation (the so-called A exciton) is mainly composed of transitions at the six-equivalent K-points of the BZ from the VBM to the first unoccupied state with the same spin-character, which is the CBM for MoX<sub>2</sub> and the CBM + 1 for WX<sub>2</sub>; see also Figs. 2 and 3. In Table II, the A exciton binding energies, defined as the difference between the electronic and the optical gap, are shown for the different levels of calculation together with the difference, between perturbative and full SOC, of the absolute position of the A peak,  $\Delta E_A$ , and of the energy of the single-particle transition that mainly gives rise

to it,  $\Delta E_{IP}$ . In general agreement with previous calculations [40,41], in free-standing MLs the binding energy of exciton A is large. By looking at the A exciton binding energy, it could be simply argued that perturbative SOC systematically underbinds the A exciton by  $\sim 20$  meV. However, by taking a closer look at the absolute positions of the peak, the effect of applying perturbative SOC seems far from systematic, and thus predictable. Indeed, when looking at  $\Delta E_A$  in Table II and at the spectra in Fig. 1, the absolute value of the A peak position may vary by as much as 116 meV in the case of WSe<sub>2</sub>. Moreover, the blueshift of the A exciton in the perturbative SOC scheme is almost entirely due to larger independent-particle transition energies for W-based materials, as shown by the  $\Delta E_{IP}$  column of Table II, but not for Mo-based ones. This furthermore means that, while for W-based materials in the perturbative and full SOC schemes the  $e$ - $h$  interaction that renormalizes the independent-particle (IP) transition energies is very similar, for Mo-based MLs, perturbative SOC underestimates the strength of  $e$ - $h$  attraction by  $\sim 20$  meV. Finally, it is worth mentioning that experimental binding energies depend on the substrate and are generally smaller than what

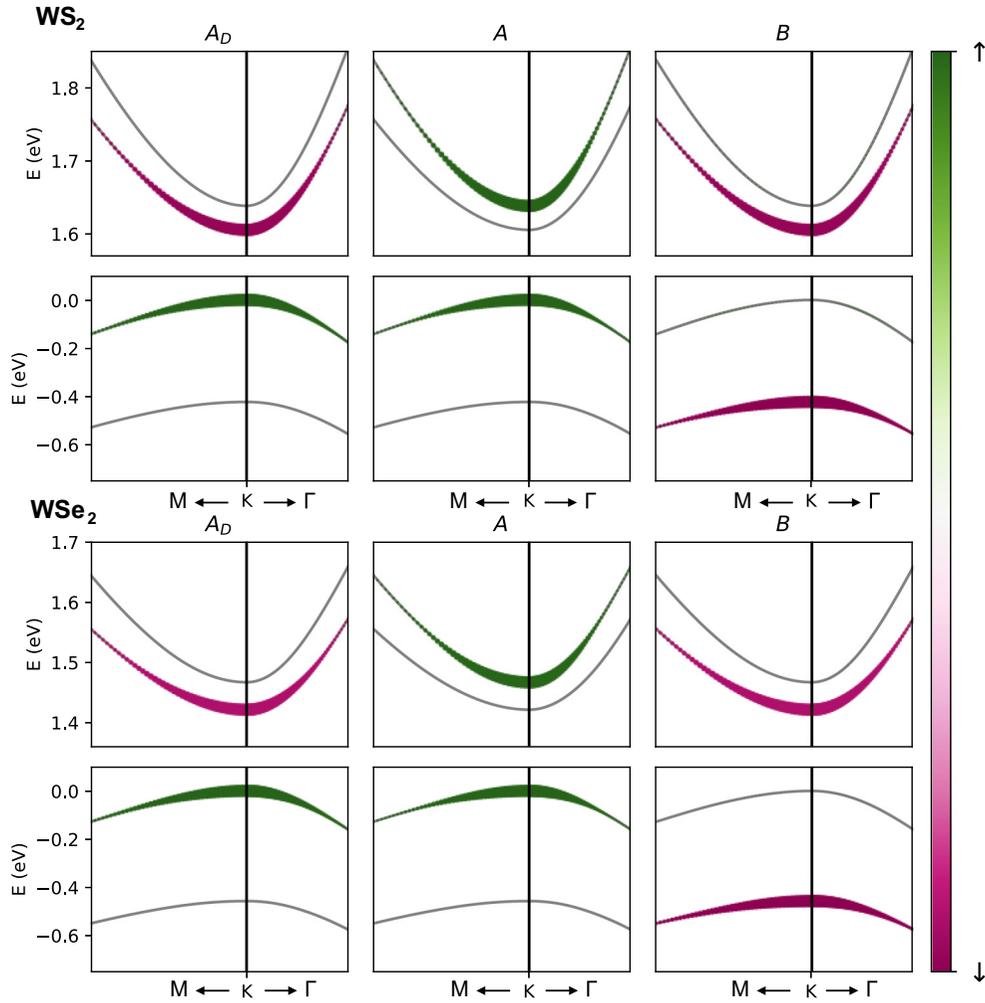


FIG. 3. Zoom of the spin-split VBM and CBM around K showing the single-particle contributions to the main optical excitations: the width of the points is proportional to the contribution of the state to the excitation, while the color refers to the spin character. Top panels:  $\text{WS}_2$ ; bottom panels:  $\text{WSe}_2$ .

has been found theoretically for free-standing layers [42–50]. Therefore, the direct comparison with theoretical data is not meaningful due to the presence of the substrate that modifies the electronic screening.

As shown in Figs. 2 and 3, the excitonic peak B is mainly composed of single-particle transitions located at K from the

TABLE II. A exciton binding energy, difference between the perturbative and full SOC schemes of the absolute position of the A peak,  $\Delta E_A$ , and of the independent-particle transition that mainly contributes to the A peak,  $\Delta E_{\text{IP}}$ . All energies are in (meV).

	A exciton binding energy (meV)			$\Delta E_A$ (meV)	$\Delta E_{\text{IP}}$ (meV)
	SOC	Pert. SOC	No SOC		
$\text{WSe}_2$	550	533	596	116	119
$\text{WS}_2$	631	610	658	33	32
$\text{MoSe}_2$	666	648	547	31	13
$\text{MoS}_2$	691	673	679	29	11

spin-split VBM-1 to the first conduction band with its same spin character, namely the CBM for W-based systems and the CBM + 1 for Mo-based ones. The IP B-A splitting, presented in Table III, arises from the DFT splitting of the conduction and valence bands in the case of perturbative SOC, and from the spinorial GW calculation for the full SOC scheme. Within both approaches, the mixing of transitions reduces such IP splitting, and the perturbative SOC reproduces quite well the full SOC B-A splitting with the exception of  $\text{WSe}_2$ , where there is a difference of about  $\sim 70$  meV. The origin of such a large deviation in the case of  $\text{WSe}_2$  can be ascribed to large differences already at the IP level: the IP B-A splitting difference between full and perturbative SOC [ $\Delta(E_B - E_A)_{\text{IP}}$ ] is  $\sim 80$  meV in the case of  $\text{WSe}_2$ , while it ranges from 7 to 12 meV for the remaining cases. It is easy to show that  $\Delta(E_B - E_A)_{\text{IP}} = \Delta_{\text{GW}}(\text{VBMS}) - \Delta_{\text{GW}}(\text{CBMS})$ , where  $\Delta_{\text{GW}}(\text{VBMS})$  [ $\Delta_{\text{GW}}(\text{CBMS})$ ] is the GW correction to the top (bottom) valence- (conduction-) band splitting. Thus, in the end, in  $\text{WSe}_2$  the large difference between full and perturbative SOC B-A splittings is due to large GW corrections

TABLE III.  $A'$ - $A$  and  $B$ - $A$  splittings within full and perturbative SOC schemes. BSE and independent-particle (IP) results are compared for the case of  $B$ - $A$  splitting. The  $B$  exciton in the full SOC  $\text{MoSe}_2$  case is not univocally defined; see the discussion in the text. Experiments were either performed on a  $\text{SiO}_2/\text{Si}$  substrate [42,45,47,49,54,55], or encapsulating the monolayer in hexagonal BN [44,46,53,56–58]. All energies are in meV.

	B-A splitting (meV)					A'-A splitting (meV)		
	Pert. SOC		SOC		Expt.	Pert. SOC	SOC	Expt.
	IP	BSE	IP	BSE				
$\text{WSe}_2$	410	380	491	450	410–430 [42,44,47,54]	140	140	130–160 [44,46,47,49,56,57]
$\text{WS}_2$	380	340	382	340	371–395 [42,45,54]	210	210	160 [45]
$\text{MoSe}_2$	210	200	221	200;210	155 [53] 190–220 [55]	190	200;210	155 [53]
$\text{MoS}_2$	150	140	160	150	124–150 [45,54,58]	210	230	175 [58]

to valence- and conduction-band splittings, which are, moreover, of opposite sign and thus sum up. On the contrary, for  $\text{WS}_2$ , the GW corrections to valence- and conduction-band splittings have the same sign and tend to cancel each other. Finally, for Mo-based systems the two contributions add up, as in  $\text{WSe}_2$ , but to a lower value, being one order of magnitude smaller.

In the perturbative SOC scheme, excitons  $A$  and  $B$  are compelled to have the same oscillator strength, which should be half of the oscillator strength of the NO SOC calculation they stem from, as shown in the middle panel of Fig. 1. On the contrary, when using a full spinorial formulation, the oscillator strengths of the two structures may be different, yielding a better agreement with experiment [51]. In particular, we find that for  $W$ -based materials the oscillator strength of the  $A$  exciton is larger than that of the  $B$  exciton, while the opposite is true for Mo-based ones, although the case of  $\text{MoSe}_2$  must be handled with care, as better detailed in the following.

The  $A'$  structure is identified as the first structure following  $A$  that has similar single-particle composition. From Table III and Fig. 1 it can be seen that the sequence of  $A, B, A'$  peaks of the full SOC calculations is well reproduced by the perturbative SOC approach, with the  $W$ -based materials featuring an  $A'$  peak clearly below the  $B$  one. The comparison with experiments shows that the  $B$ - $A$  splittings are well reproduced, but  $A'$ - $A$  splittings are in most of the cases overestimating the experimental values. This different behavior may be due to the different origin of the two splittings and to the corresponding sensitivity to the dielectric environment. While the  $B$ - $A$  splitting is connected to the spin-orbit splitting of the conduction and valence bands, but is not affected by the changes in the electronic screening due to substrates and/or sample encapsulations, the  $A'$ - $A$  splitting due the hydrogenoid series is instead strongly affected, just like the binding energy. However, it is worth noting that the  $A'$ - $A$  splitting may require finer  $k$ -point grids to converge its absolute value and make a meaningful comparison with experiments, therefore we performed full SOC calculations on a denser  $39 \times 39 \times 1$   $k$ -point grid for  $\text{MoS}_2$ ,  $\text{MoSe}_2$ , and  $\text{WSe}_2$ , finding  $A$ - $A'$  splitting values of 270, 231, and 215 meV, respectively. The converged values enhance the overestimation with respect to experiment, suggesting that the substrate effect on the  $A$ - $A'$  splitting needs further investigations, but this is beyond the scope of this paper.

### C. Exchange-driven intravalley mixing

In the perturbative SOC approach, the identification of the  $A$  and  $B$  structures is straightforward as the  $A$  and  $B$  excitons are derived from the same excitation of the underlying SOC-free BSE calculation to which different corrections are applied. Thus, by construction, in the perturbative SOC scheme the single-particle transitions that contribute to the  $A$  structure are completely disjoint from those that contribute to the  $B$  structure. However, it was recently pointed out for  $\text{MoS}_2$  that, when using a full spinorial formalism, an exchange-driven intravalley mixing was found, meaning that transitions pertaining to the  $B$  exciton contributed (minimally) to the  $A$  exciton and vice versa [51]. We confirm this finding for  $\text{MoS}_2$ , and we found a similar mixing of contributions for the  $W$ -based materials. The mixing of transitions can be easily recognized looking at Fig. 2, where for the  $B$  excitons we see that all bands contribute. But while for  $\text{MoS}_2$  this mixing is minimal, we find a huge effect for the  $B$  exciton of  $\text{MoSe}_2$ . In this case, the structures that we labeled  $A'/B$  arise from four almost degenerate excitons, which are linear combinations of single-particle transitions from the VBM to the CBM (characterizing the  $A$  series) and from the VBM-1 and the CBM + 1 (characterizing the  $B$  series). The degree of mixing is one order of magnitude larger than in the other cases, meaning the exchange interaction, leading to the intravalley mixing between VBM  $\rightarrow$  CBM and (VBM-1) $\rightarrow$ (CBM + 1) transitions, is playing a prominent role. To understand this enhancement, we must look back at the structure of the BSE Hamiltonian  $\Delta S_z = 0$  channel in the collinear magnetic case; see Sec. IV A 3. The magnitude of the exchange term should be compared to the splitting of the diagonal part: the smaller such splitting is, the more the off-diagonal exchange term will mix the two transitions.

In the perturbative SOC calculations, where the  $A$  and  $B$  series cannot mix by construction, the  $A'$  and  $B$  peaks happen to be in close proximity to one another, as also reported in [52]. The proximity in energy between these two states explains the enhanced, exchange-driven, intravalley mixing, and this finding is robust with respect to denser  $k$ -point sampling. A recent up-conversion experiment on a hexagonal boron nitride encapsulated sample [53] is able to resolve two structures 150 and 155 meV above the  $A$  exciton, which could not be previously resolved [54,55]. While it is not possible to experimentally assign to such structures a pure  $A'$  and  $B$  or

TABLE IV. A-A\* splitting (i.e., bright-dark splitting of exciton A). All energies are in meV. A-A\* splittings from Ref. [33] are taken from  $G_0W_0$ -PBE + BSE calculations for Mo-based materials, while they are taken from  $G_0W_0$ -HSE + BSE calculations for W-based ones. The experimental value of the A-A\* splitting for MoS<sub>2</sub> from Ref. [64] was estimated from the digitalization of Fig. 6 of the corresponding paper. Experiments were either carried out on Si/SiO<sub>2</sub> substrate [34,60,64,65], on sapphire [64], or on hBN encapsulated samples [62,63,66,67].

	A-A* splitting (meV)					
	Theory				Literature (Other)	Expt.
	Present work		No SOC			
SOC (GW)	SOC (scissor)	Pert. SOC	No SOC			
WSe <sub>2</sub>	56	77	96	36	80 [36], 16 [33]	55 [63], $\simeq$ 47 [60], 57 [34], 47 [62]
WS <sub>2</sub>	42	72	80	30	80 [36], 11 [33]	40 [63], 47 [60]
MoSe <sub>2</sub>	-1	7	25	42	10 [36], -11 [33]	-1.5 [66], -1.3 [67], $\simeq$ 0 [63] $\simeq$ 0 [60] -30 [65]
MoS <sub>2</sub>	17	19	41	38	25 [36], 20 [68], 5 [33]	98 [60], < 20 [64], 14 [67]

a mixed A'/B character, this finding confirms that the ‘‘accidental’’ proximity of the A' and B excitations is present both in our theoretical calculations on free-standing MLs and in experiments including a dielectric environment.

#### D. Dark-bright splitting

Several recent experimental and theoretical works have shown that dark (spin-forbidden and finite-momentum) excitons are present near the first bright A exciton [59–62]. The knowledge of their energetic position is crucial in order to understand the exciton dynamics in view of the possible use of TMD-MLs in optoelectronic devices. Indeed, at the experimental level, spin-forbidden dark excitons could be revealed by photoluminescence [63], ellipsometry measurements with out-of-plane light polarization [64], by using photocurrent spectroscopy [65] or brightening with magnetic field [34,60,66], or near-field coupling to surface plasmon polaritons [62]. In particular, here we focus on the lowest-energy spin-forbidden dark excitons (which we label A\*) in order to see how the perturbative and full SOC schemes compare.

In Table IV, the theoretical A-A\* splittings are shown and compared with the available theoretical and experimental literature. The no-SOC scheme corresponds to the non-spin-polarized calculation, where the triplet, spin-forbidden excitons are obtained simply by switching off the exchange term in the BSE kernel. A full SOC BSE calculation employing a rigid shift of the conduction bands (scissor operator), opening the single-particle DFT gap to the corresponding GW value, was also performed. In this last case, the SOC-related splittings of the single-particle energies are the DFT ones. Within our results, we note that different theoretical schemes produce different splitting values, and that only when full-SOC is used is a better agreement with experimental data reached.

The case of MoS<sub>2</sub> is more controversial and needs further discussion. Magnetophotoluminescence experiments carried out on a Si/SiO<sub>2</sub> substrate show an extremely large A-A\* splitting of  $\sim$ 100 meV [60], whereas upon hBN encapsulation, which is known to enhance the optical quality of the samples, an A-A\* splitting of 14 meV is reported [67], and spectroscopic ellipsometry shows an A-A\* splitting smaller than 20 meV [64]. These last experimental results are more consistent with theoretical predictions, which, within

all schemes and implementations, range from 5 to 40 meV. Indeed, the A-A\* splitting can be thought of in terms of two main contributions: the first due to the exchange term of the BSE kernel, and the second due to the CBM splitting. The first contribution is always of positive sign and is the only one present in the no-SOC scheme. The second contribution is positive for W-based systems and negative for the Mo-based ones. In the case of the perturbative SOC and full SOC with scissor schemes, this contribution arises from the DFT CBM splitting, whereas for the full SOC calculation it arises from the GW one. While the exchange contribution, estimated from the no-SOC scheme, is approximatively the same for all materials, analysis of the atomic origins of the spin splitting of the conduction bands shows that the CMB splitting is expected to be small for Mo-based TMDs and in particular for MoS<sub>2</sub> [69]. On this basis, the overall A-A\* splitting is expected (and theoretically predicted) to be smaller for Mo-based TMDs.

The scattering of the theoretical estimate of the A-A\* splitting can again be explained by the different approach (GdW approach for [36]) and/or different calculation parameters or numerical implementations [33]. However, within each calculation the picture is consistent with a more or less material-independent exchange contribution plus material-specific CBM, CBM + 1 splittings (see Table I and Ref. [33]).

#### E. Excitonic spin polarization

In a full spinorial treatment, it is possible to determine the exciton's spin expectation value, namely  $\langle S^2 \rangle$ , given by Eq. (60), and its projections  $N_{SM}^\lambda$  along the singlet state and along the three components of the triplet state, given by Eq. (59). It is worth noting that such an analysis is not possible in a perturbative SOC approach where only the excitation energies are corrected and the single-particle wave functions are kept fixed to the nonmagnetic case. In Fig. 4, the dots represent  $\langle S \rangle$  for each excitation, the color of the dot is determined by its oscillator strength, while the absorption spectrum is plotted as a guide for the eye in a full line. There are two main classes of excitations: those whose  $\langle S \rangle \sim 1$  and those whose  $\langle S \rangle \sim 0.6$ . All the  $\langle S \rangle \sim 1$  excitons are dark, as they are mainly composed of linear combinations of the  $|S = 1, M = \pm 1\rangle$  states and are thus spin-forbidden excitations. The situation for  $\langle S \rangle \sim 0.6$  is less homogeneous; some

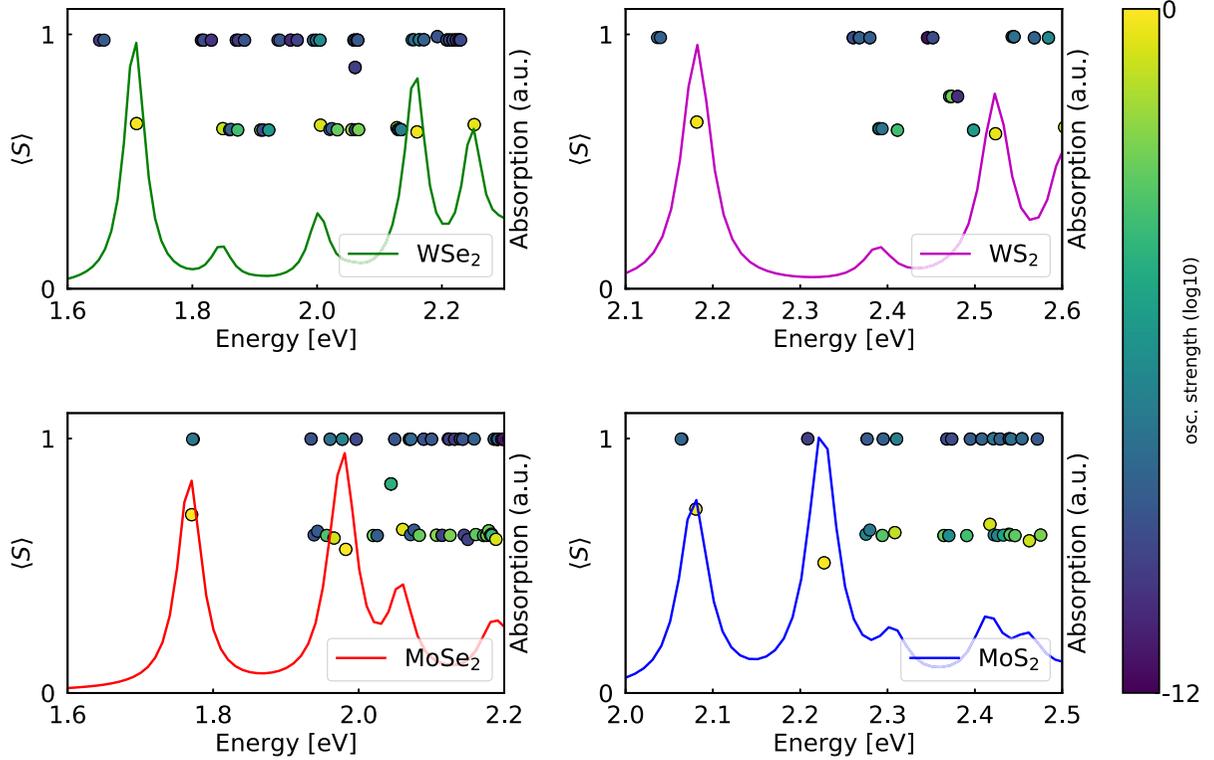


FIG. 4. Dots: excitonic total spin expectation value. The dot color refers to the oscillator strength of that specific excitation (normalized to 1). Full lines, as a guide for the eye, show the absorption spectra. Top panel: W-based materials: WSe<sub>2</sub> left, WS<sub>2</sub> right. Bottom panel: Mo-based materials: MoSe<sub>2</sub> left, MoS<sub>2</sub> right.

excitations are bright and build up the optical spectrum, while others are dark. In this case, the origin of the low oscillator strength cannot be attributed to spin conservation but to dipole symmetry rules. The fact that  $\langle S \rangle \sim 0.6$ , significantly lower than 1, is evidence of a strong component along the singlet state.

Indeed, in Table V we report the decomposition of the main structures identified in the absorption spectrum, along with that of the dark A excitation. The A and B excitons are linear combinations of the  $|S = 0, M = 0\rangle$  and  $|S = 1, M = 0\rangle$  states, whereas the dark A peak is made mainly of  $|S = 1, M = \pm 1\rangle$  contributions.

First of all, it can be seen that  $\langle S_z \rangle = 0$  for all the excitations, including the dark ones related to the magnon channel; this is a consequence of the symmetry between the K and K' valley, which contribute equally. Secondly, in principle in a full spinorial, noncollinear, approach, all the  $|S, M\rangle$  components would be allowed to mix, while, as discussed in Sec. IV A and shown in Ref. [5], in a collinear but spin-polarized case, only the singlet and the  $M = 0$  component of the triplet are allowed to mix. The spin analysis of the low-energy excitations shows that here we are somewhat close to this case: the spinor states contributing to the low-energy excitations are strongly spin-polarized, and thus the mixing of the singlet and  $M = \pm 1$  triplet components is very small. Moreover, SOC removes the degeneracy between the single-particle contributions pertaining to the A and B excitons, while at the same time this SOC splitting is much larger than the exchange contribution which couples such transitions. As

a consequence, the  $|S = 0, M = 0\rangle$  and  $|S = 1, M = 0\rangle$  states are strongly mixed, and the triplet or singlet character of the excitations is destroyed (see also Appendix E). Finally, the fact that for the bright excitations the weights of the  $|S = 0, M = 0\rangle$  and  $|S = 1, M = 0\rangle$  components are not exactly equal, as shown in Table V, is an effect of the small but finite exchange term. As shown in [51] for MoS<sub>2</sub>, and found here also for the remaining cases, exchange is responsible for the intravalley mixing of the A and B single-particle contributions.

## VIII. CONCLUSIONS

In this paper, a detailed derivation of the GW and BSE equations by including the full spinorial nature of the wave functions is illustrated. This formulation allows us to obtain the collinear, spin, and non-spin-polarized cases, as well as the noncollinear cases, in a natural way. The spin-orbit interaction can then be included in a nonperturbative way from the ground-state up to the excited-state simulations. Beyond the formal derivation of all the main equations, we carry out a systematic analysis of electronic and optical properties of most representative group VI TMD monolayers, comparing at the same level of numerical implementation the results without SOC and with SOC at a perturbative and a nonperturbative level. While in most of the observables considered the perturbative and nonperturbative approaches for SOC give very similar results, the dark-bright splittings are generally improved when the SOC is included in a nonperturbative

TABLE V. Excitonic spin analysis:  $\langle S \rangle$  and  $\langle S_z \rangle$  expectation values, and the projections of the excitonic wave functions along the singlet state and along the three components of the triplet state.

	$\langle S^2 \rangle$	$\langle S_z \rangle$	$ S = 0, M = 0\rangle$	$ S = 1, M = -1\rangle$	$ S = 1, M = 0\rangle$	$ S = 1, M = 1\rangle$
WSe <sub>2</sub>						
A*	1.899	0	0.034	0.466	0.034	0.466
A	0.823	0	0.464	0.007	0.522	0.007
A'	0.776	0	0.487	0.007	0.499	0.007
B	0.748	0	0.502	0.037	0.425	0.037
WS <sub>2</sub>						
A*	1.946	0	0.019	0.481	0.02	0.481
A	0.838	0	0.458	0.006	0.531	0.006
A'	0.776	0	0.488	0.006	0.501	0.006
B	0.730	0	0.51	0.022	0.446	0.022
MoSe <sub>2</sub>						
A	0.951	0	0.404	0.001	0.594	0.001
A*	1.988	0	0.004	0.496	0.004	0.496
A'/B	0.730	0	0.51	0.003	0.484	0.003
A'/B	0.635	0	0.559	0.003	0.435	0.003
B'	0.718	0	0.516	0.005	0.474	0.005
MoS <sub>2</sub>						
A*	1.991	0	0.003	0.497	0.003	0.497
A	1.011	0	0.377	0.001	0.621	0.001
B*	1.997	0	0.001	0.499	0.001	0.499
B	0.535	0	0.614	0.003	0.38	0.003
A'	0.778	0	0.486	0.001	0.512	0.001
B'	0.705	0	0.522	0.003	0.471	0.003

way. The exchange-driven intravalley mixing, absent by construction within the perturbative approach, is found to hugely impact the nature of the B exciton for the case of MoSe<sub>2</sub> that is found to be strongly mixed with the A' exciton. Furthermore, the spin character of all the excitons in the IR-vis region for the four TMDs analyzed, which is impossible in a perturbative SOC approach, and is not yet available, has been obtained and discussed.

#### ACKNOWLEDGMENTS

M.P. acknowledges funding from INFN20-TIME2QUEST project. A.M.-S. thanks the Marie-Curie-COFUND program Nano TRAIN for Growth II (Grant Agreement 713640) and the Ramón y Cajal programme (Grant No. RYC2018-024024-I, MINECO, Spain). The computations were performed on the Tirant III cluster of the Servei d'Informàtica of the University of Valencia (project vlc82) and on the Mare Nostrum cluster of the Barcelona Supercomputing Center (project FI-2020-2-033 and FI-2020-3-0021). D.S. and A.M. acknowledge the funding received from MIUR PRIN Grant No. 20173B72NB and from the European Union projects: MaX *Materials design at the eXascale* H2020-EINFRA-2015-1, Grant agreement No. 676598, and H2020-INFRAEDI-2018-2020/H2020-INFRAEDI-2018-1, Grant agreement No. 824143; *Nanoscience Foundries and Fine Analysis - Europe* H2020-INFRAIA-2014-2015, Grant agreement No. 654360. M.M. acknowledges that part of the computing resources and the related technical support used for this work have been

provided by CRESCO/ENEAGRID High Performance Computing infrastructure and its staff [70]. CRESCO/ENEAGRID High Performance Computing infrastructure is funded by ENEA, the Italian National Agency for New Technologies, Energy and Sustainable Economic Development, and by Italian and European research programmes (see Ref. [71] for more information).

#### APPENDIX A: SPIN-DEPENDENT EXCHANGE CORRELATION IN A SPINORIAL BASIS

The density matrix and the xc potential can be written in terms of the density ( $n$ ) and the magnetization ( $\mathbf{m}$ ) as

$$\underline{\rho}(\mathbf{x}) = n(\mathbf{x})\underline{\sigma}_0 + \mathbf{m}(\mathbf{x}) \cdot \underline{\sigma}, \quad (\text{A1a})$$

$$\underline{v}^{\text{xc}}(\mathbf{x}) = \phi^{\text{xc}}(\mathbf{x})\underline{\sigma}_0 + \mathbf{B}^{\text{xc}}(\mathbf{x}) \cdot \underline{\sigma}, \quad (\text{A1b})$$

where we also introduced the exchange-correlation magnetic field,  $\bar{\mathbf{B}}^{\text{xc}}$ , and the density-potential,  $\phi^{\text{xc}}$ .

Although known xc functionals are (local) functions of the modulus of  $\mathbf{m}$  alone, spin-dependent KS equations can be solved also in the LDA for a noncollinear system. This is obtained by calculating  $\underline{v}^{\text{KS}}$ , which depends on the LDA only from the magnitude of the local magnetization, by rotating the magnetization vector in each point in space into the local frame of spin-quantization and evaluating the potential  $v_{\text{xc}}(n(r), m(r))$  and then rotating back to the global reference frame.

## APPENDIX B: HEDIN'S EQUATIONS IN SPIN, SPACE, AND TIME REPRESENTATIONS

The main ingredient of the approach is the electronic Green's function (GF):

$$G(1, 2) = -i\langle \mathcal{T}\{\hat{\Psi}(1)\hat{\Psi}^\dagger(2)\} \rangle, \quad (\text{B1})$$

where  $\langle \dots \rangle$  is the trace evaluated with the exact density matrix,  $1 = (\mathbf{x}_1, s_1, t_1)$  includes space, spin, and time, and operators are in the Heisenberg representation.

From Eq. (13) it follows that the noninteracting Green's function is the KS one:

$$G^{(0)}(1, 2) \equiv G^{\text{KS}}(1, 2). \quad (\text{B2})$$

Hedin's equations can be derived by using the functional derivative Schwinger approach [72], where  $H$  is perturbed with a spin-collinear *time-dependent auxiliary field*  $\eta(1)$ ,

$$\hat{H}_\eta(t_1) = \hat{H} + \int d\mathbf{x}_1 \eta(1)\hat{\Psi}^\dagger(1)\hat{\Psi}(1). \quad (\text{B3})$$

We can safely use a collinear perturbation because, as explained in Ref. [21], this is consistent with the fact that the Coulomb interaction is spin-independent. More elaborate auxiliary fields must be introduced in the case of noncollinear electron-electron mediated interactions.

It can be easily proved [21,72] that  $G$  solves a set of self-consistent, integrodifferential equations. The first equation is the usual Dyson equation,

$$G(1, 2) = G^{(0)}(1, 2) + G^{(0)}(1, 3)\Sigma^{Hxc}(3, 4)G(4, 2), \quad (\text{B4})$$

with repeated subscripts summed and repeated arguments integrated, if not explicitly written.

In Eq. (B4),  $\Sigma^{Hxc}$  is composed of two terms:

$$\Sigma^{Hxc}(3, 4) = \Sigma(3, 4) + v^H(3)\delta(3, 4), \quad (\text{B5})$$

with  $v^H$  the Hartree potential,

$$v^H(3) = v^H(\mathbf{x}_3) = -iv(\mathbf{x}_3 - \mathbf{x}_5)G(5, 5^+), \quad (\text{B6})$$

and  $\Sigma$  the exchange and correlation self-energy, also known as the Mass operator.  $\Sigma$  can be rewritten in terms of an irreducible vertex function  $\tilde{\Gamma}$ :

$$\Sigma(1, 2) = -iG(1, 3)\tilde{\Gamma}(3, 2; 4)W(4, 1), \quad (\text{B7})$$

with

$$\begin{aligned} \tilde{\Gamma}(1, 2; 3) &= -\frac{\delta G^{-1}(1, 2)}{\delta \eta(3)} = \delta(1, 3)\delta(2, 3) \\ &+ \frac{\delta \Sigma(1, 2)}{\delta G(4, 5)}G(4, 6)\tilde{\Gamma}(6, 7; 3)G(7, 5). \end{aligned} \quad (\text{B8})$$

$W(4, 1)$  is the electronic screened interaction

$$W(4, 1) = v(4, 1) + v(4, 5)\tilde{\chi}(5, 6)W(6, 1), \quad (\text{B9})$$

which is written in terms of  $\tilde{\chi}$ , the irreducible electronic response function:

$$\tilde{\chi}(5, 6) = G(5, 7)\tilde{\Gamma}(7, 8; 6)G(8, 5). \quad (\text{B10})$$

Equations (B7)–(B10) represent the spin Hedin equations and completely solve the many-body problem.

Starting from the equation for the response function and the vertex, a Dyson-like equation for a response function

appears. However, due to the structure of the variables, such an equation cannot be directly cast in terms of the two-point response function [not even with the approximation defined in Eq. (35)]. It must be cast in terms of the four-point function  $L(1, 3; 2, 4)$ . Defining  $K = \partial \Sigma / \partial G$  and inserting Eq. (B8) into Eq. (B10), one obtains

$$L(13, 24) = L_0(13, 24) + L_0(13, 1'3')K(1'3', 2'4')L(2'4', 24), \quad (\text{B11})$$

where  $L_0 = GG$  and  $L(11, 22) = \chi(1, 2)$ .

## APPENDIX C: ROTATION IN THE SPINORIAL BASIS OF THE DIFFERENT COMPONENTS OF HEDIN'S EQUATIONS

The two maps needed to rotate Hedin's equations in the spinorial basis are defined in Eq. (17). The goal of the different sections of this Appendix is to demonstrate how the two maps follow from the manipulation of Hedin's equations.

### 1. The Dyson equation

The transformation of the self-energy operator follows easily by taking Eq. (B4) and expanding both  $G$  and  $G^{(0)}$  using  $M_2$ . It follows that

$$\begin{aligned} \Sigma_{\mathbf{I}_1\mathbf{I}_2}^{Hxc}(t_1, t_2) &= \phi_{\mathbf{I}_1s_1}^*(\mathbf{x}_1)\Sigma^{Hxc}(1, 2)\phi_{\mathbf{I}_2s_2}(\mathbf{x}_2) \\ &= \Sigma_{\mathbf{I}_1\mathbf{I}_2}(t_1, t_2) + v_{\mathbf{I}_1\mathbf{I}_2}^H(t_1)\delta(t_1 - t_2). \end{aligned} \quad (\text{C1})$$

$v_{\mathbf{I}_3\mathbf{I}_2}^H$  is defined in Eq. (C3), while  $\Sigma_{\mathbf{I}_1\mathbf{I}_2}$  is defined in Eq. (22b).

### 2. The Hartree potential

From Eq. (C1), it follows that

$$v_{\mathbf{I}_3\mathbf{I}_4}^H(t_3) = \phi_{\mathbf{I}_3s_3}^*(\mathbf{x}_3)v^H(3)\phi_{\mathbf{I}_4s_4}(\mathbf{x}_3). \quad (\text{C2})$$

By using Eq. (B6), we see that

$$v_{\mathbf{I}_3\mathbf{I}_4}^H(t_3) = -iV_{\mathbf{I}_3\mathbf{I}_4} G_{\mathbf{I}_5\mathbf{I}_5}, \quad (\text{C3})$$

with

$$V_{\mathbf{I}_3\mathbf{I}_4} = \phi_{\mathbf{I}_1s_1}^*(\mathbf{x}_1)\phi_{\mathbf{I}_1s_1}(\mathbf{x}_1)v(\mathbf{x}_1 - \mathbf{x}_3)\phi_{\mathbf{I}_3s_3}^*(\mathbf{x}_3)\phi_{\mathbf{I}_4s_3}(\mathbf{x}_3). \quad (\text{C4})$$

Equation (C4) is the proof of Eq. (19c).

### 3. The vertex function

To rotate the vertex function, we observe that from Eq. (20) it follows that  $G_{\mathbf{I}_1\mathbf{I}_2}^{-1}$  rotates like  $\Sigma_{\mathbf{I}_1\mathbf{I}_2}$ . This implies that we can rewrite

$$\tilde{\Gamma}(3, 2; 4) = -\phi_{\mathbf{I}_3s_3}^*(\mathbf{x}_3)\phi_{\mathbf{I}_2s_2}^*(\mathbf{x}_2)\frac{\delta G_{\mathbf{I}_3\mathbf{I}_2}^{-1}(t_3, t_2)}{\delta \eta_{\mathbf{I}_4\mathbf{I}_5}(t_5)}\frac{\delta \eta_{\mathbf{I}_4\mathbf{I}_5}(t_5)}{\delta \eta(4)}, \quad (\text{C5})$$

where we have introduced

$$\eta_{\mathbf{I}_4\mathbf{I}_5}(t_5) = \phi_{\mathbf{I}_4s_5}(\mathbf{x}_5)\eta(5)\phi_{\mathbf{I}_5s_5}^*(\mathbf{x}_5) \quad (\text{C6})$$

and

$$\tilde{\Gamma}_{\mathbf{I}_3\mathbf{I}_2}^{-1}(t_3, t_2; t_4) \equiv \frac{\delta G_{\mathbf{I}_3\mathbf{I}_2}^{-1}(t_3, t_2)}{\delta \eta_{\mathbf{I}_4\mathbf{I}_5}(t_4)}. \quad (\text{C7})$$

From Eq. (C6), the functional derivative appearing on the right-hand side of Eq. (C5) can be easily calculated to give

$$\begin{aligned} & \tilde{\Gamma}(3, 2; 4) \\ &= -\phi_{\mathbf{I}_3 s_3}(\mathbf{x}_3) \phi_{\mathbf{I}_2 s_2}^*(\mathbf{x}_2) \phi_{\mathbf{I}_4 s_4}(\mathbf{x}_4) \phi_{\mathbf{I}_5 s_4}^*(\mathbf{x}_4) \frac{\delta G_{\mathbf{I}_3 \mathbf{I}_2}^{-1}(t_3, t_2)}{\delta \eta_{\mathbf{I}_4 \mathbf{I}_5}(t_4)}. \end{aligned} \quad (\text{C8})$$

Equation (C4) is the proof of Eq. (19e).

#### 4. The response function

The rotation of the response function follows from Eq. (B10), after using Eqs. (17) and (C8). It follows that

$$\begin{aligned} & \tilde{\chi}(1, 2) \\ &= \phi_{\mathbf{I}_1 s_1}(\mathbf{x}_1) \phi_{\mathbf{I}'_1 s_1}^*(\mathbf{x}_1) \tilde{L}_{\mathbf{I}_1 \mathbf{I}'_1}(\mathbf{I}_1, \mathbf{I}'_1, t_1, t_2) \phi_{\mathbf{I}_2 s_2}(\mathbf{x}_2) \phi_{\mathbf{I}'_2 s_2}^*(\mathbf{x}_2), \end{aligned} \quad (\text{C9})$$

which, using Eq. (B9), demonstrates Eq. (24). The response function is a particular case. Indeed,  $\tilde{L}_{\mathbf{I}_1 \mathbf{I}'_1}(\mathbf{I}_1, \mathbf{I}'_1, t_1, t_2) \neq \tilde{\chi}_{\mathbf{I}_1 \mathbf{I}'_1}(\mathbf{I}_1, \mathbf{I}'_1, t_1, t_2)$ , since  $\tilde{\chi}_{\mathbf{I}_1 \mathbf{I}'_1}$ , which results from  $M_4$ :  $\tilde{\chi}$ , can only be used to reconstruct  $\tilde{\chi}(1, 2)$  via the inversion of  $M_4$ , while  $\tilde{L}_{\mathbf{I}_1 \mathbf{I}'_1}(\mathbf{I}_1, \mathbf{I}'_1, t_1, t_2)$  contains enough

information to reconstruct both  $\tilde{\chi}(1, 2)$  and  $\tilde{L}(1, 2)$ . However, when the Dyson equation for the response function is written in the spinorial representation (or more generally in a wave-function basis set), the contraction is not present anymore and the matrix element of  $L$  appears.

#### 5. The exchange-correlation self-energy operator

If now we use Eqs. (20) and (C8) to expand  $G$  and  $\tilde{\Gamma}$  in Eq. (B7) for  $\Sigma$ , we get

$$\begin{aligned} \Sigma_{\mathbf{I}_1, \mathbf{I}_2}(t_1, t_2) &= -iG_{\mathbf{I}_1 \mathbf{I}_3}(t_1, t_3) \tilde{\Gamma}_{\mathbf{I}_3 \mathbf{I}_2}(\mathbf{I}_3, \mathbf{I}_2, t_3, t_2; t_4) \\ &\quad \times [\phi_{\mathbf{I}_1 s_1}^*(\mathbf{x}_1) \phi_{\mathbf{I}'_1 s_1}(\mathbf{x}_1) W(4, 1) \phi_{\mathbf{I}_4 s_4}^*(\mathbf{x}_4) \phi_{\mathbf{I}_4 s_4}(\mathbf{x}_4)], \end{aligned} \quad (\text{C10})$$

which implies Eq. (22b).

#### APPENDIX D: COMPUTATIONAL DETAILS

The density functional theory (DFT) simulations have been performed using the plane-wave QUANTUM ESPRESSO code [73]. A Perdew-Burke-Ernzerhof (PBE) exchange-correlation functional [74] and optimized norm-conserving [75] pseudopotentials have been used. A  $16 \times 16 \times 1$  Monkhorst-Pack grid [76] of  $k$ -points is used to sample the Brillouin zone for structural optimization runs. A kinetic energy cutoff of 140 Ry has been used. Structure relaxation is assumed at convergence when the maximum component of the residual forces on the ions is smaller than  $10^{-5}$  Ry/Bohr. The lattice parameters of the four systems are in very good agreement with existing literature, being  $a = 3.12$  Å (for MoS<sub>2</sub>), 3.25 Å (for MoSe<sub>2</sub>), 3.12 Å (for WS<sub>2</sub>), and 3.25 Å (for WSe<sub>2</sub>). Once the optimized atomic structures have been obtained, self- and non-self-consistent DFT calculations have been performed to

obtain Kohn-Sham (KS) eigenvalues and eigenfunctions to be used in the many-body simulations done by using the many-body code YAMBO [6,77]. Specifically, we calculated the quasiparticle (QP) energies by using the GW perturbative one-shot method, and the optical excitation energies and the optical spectra by solving the Bethe-Salpeter equation (BSE) [1,72,78–80]. For GW simulations, a plasmon-pole approximation for the inverse dielectric matrix has been applied [81]; a kinetic energy cutoff of 10 Ry (60 Ry) is used for the correlation, i.e., the  $\Sigma_c$  (exchange,  $\Sigma_x$ ) part of the self-energy; and the sum over the unoccupied states for  $\Sigma_c$  and the dielectric matrix is performed up to  $\sim 30$  eV above the VBM. To speed up the convergence with respect to empty states, we adopted the technique described in Ref. [82]. The Bethe-Salpeter equation to obtain the optical spectrum and exciton spatial localization has been solved within the Tamm-Dancoff [1,83] approximation (which is generally valid for bulk compounds to describe neutral excitations well below the plasma frequency of the material). Two occupied and two unoccupied states have been used to build up the excitonic Hamiltonian. For both the GW and BSE simulations, we used a  $k$ -grid of  $33 \times 33 \times 1$ , which is enough to provide a good convergence in the position of the A-B excitons, which are the main focus of this paper.

#### APPENDIX E: PERTURBATIVE SOC AND EXCHANGE

It is important to notice that there are two different cases to be considered for spin-conserving excitations (i.e.,  $\Delta S_z = 0$ ):

(a) The singlet-triplet splitting, i.e., the exchange splitting  $\Delta E_{\lambda}^{T-S}$  due to  $H^{\text{exch}}$ , is bigger than the SOC correction.

(b) SOC is, on the one hand, small enough to be considered a perturbation, and on the other hand, significantly bigger than the  $\Delta E_{\lambda}^{T-S}$  splitting.

In case (a), we expect the singlet and triplet excitation to remain far apart in energy. The SOC splitting just corrects such energies. Thus it makes sense to define the pSOC as an energy shift in the  $S$  (and  $T$ ) channel. This is the scheme reported in the literature and in the main text.

In case (b) instead, we expect the singlet and triplet structure to be destroyed. Neither the singlet Hamiltonian, with  $H^{\text{exch},S} = 2V$ , nor the triplet one, with  $H^{\text{exch},T} = 0$ , should be used, rather an average Hamiltonian with  $H^{\text{exch}} = V$  should be used. The resulting eigenvectors  $\tilde{A}_{c\mathbf{k}}^{\lambda}$  define

$$\Psi_{s's'}^{\lambda,s}(\mathbf{x}_c, \mathbf{x}_v) = \delta_{s,s'} \sum_{c\mathbf{k}} \tilde{A}_{c\mathbf{k}}^{\lambda} \phi_{c\mathbf{k}s}^{0,*}(\mathbf{x}_c) \phi_{v\mathbf{k}s}^0(\mathbf{x}_v), \quad (\text{E1})$$

where now  $s$  is considered to be an approximate quantum number. The two associated transitions would then be

$$\Delta\omega_{\lambda,s}^0 \approx \sum_{c\mathbf{k}} |\tilde{A}_{c\mathbf{k}}^{\lambda}|^2 (\Delta\epsilon_{c\mathbf{k}s}^{\text{SOC}} - \Delta\epsilon_{v\mathbf{k}s}^{\text{SOC}}). \quad (\text{E2})$$

The final result is that each peak is split into two peaks with equal intensity both belonging to the  $\Delta S_z = 0$  channel, one for  $\uparrow - \uparrow$  transitions and the other for  $\downarrow - \downarrow$  transitions. Accordingly,  $S$  is no longer a good quantum number for the exciton, and one would expect  $\langle \hat{S} \rangle \neq 0, 1$ .

## APPENDIX F: MAPPING AND DEGENERATE STATES

In the case of degeneracy, we define the subgroups of degenerate spaces as  $G_{ms}^k$  and  $G_n^k$ . We observe that the states without SOC,  $|m\mathbf{k}s\rangle$ , are more degenerate than the states with SOC,  $|n\mathbf{k}\rangle$ , such that to each  $G_{ms}^k$  may correspond multiple  $G_n^k$ . Then we define, for each  $n\mathbf{k}$ ,

$$|\Delta_{n,G_{ms}^k}^k|^2 = \sum_{ms \in G_{ms}^k} |\Delta_{n,ms}^k|^2, \quad (\text{F1})$$

$$D_n^k = \max_{G_{ms}^k} |\Delta_{n,G_{ms}^k}^k|^2. \quad (\text{F2})$$

The mapping function is then defined as

$$f_{\mathbf{k}}(ms) = n \quad \text{if} \quad |\Delta_{n,G_{ms}^k}^k|^2 = D_n^k, \quad (\text{F3})$$

where the mapping is done recursively, doing a loop over the  $n$  index, and we assign  $n$  randomly to one of the states  $ms \in G_{ms}^k$ , which has no other  $n$  assigned. Moreover, if  $n \in G_n^k$ , the procedure picks up the  $n \in G_n^k$  with maximum projection in  $G_n^k$ , which is again random. Since we are dealing with degenerate energies, these choices do not affect the final result. In the end, there is one  $|\vec{n}\mathbf{k}\rangle$  state assigned to each  $|m\mathbf{k}s\rangle$  state.

- 
- [1] G. Onida, L. Reining, and A. Rubio, *Rev. Mod. Phys.* **74**, 601 (2002).
- [2] M. Bode, M. Heide, K. von Bergmann, P. Ferriani, S. Heinze, G. Bihlmayer, A. Kubetzka, O. Pietzsch, S. Blügel, and R. Wiesendanger, *Nature (London)* **447**, 190 (2007).
- [3] F. Kuemmeth, S. Ilani, D. C. Ralph, and P. L. McEuen, *Nature (London)* **452**, 448 (2008).
- [4] S. Picozzi, A. Continenza, M. Kim, and A. J. Freeman, *Phys. Rev. B* **73**, 235207 (2006).
- [5] C. Rödl, F. Fuchs, J. Furthmüller, and F. Bechstedt, *Phys. Rev. B* **77**, 184408 (2008).
- [6] D. Sangalli, A. Ferretti, H. Miranda, C. Attaccalite, I. Marri, E. Cannuccia, P. M. Melo, M. Marsili, F. Paleari, A. Marrazzo, G. Prandini, P. Bonfà, M. O. Atambo, F. Affinito, M. Palummo, A. M. Sanchez, C. Hogan, M. Grüning, D. Varsano, and A. Marini, *J. Phys.: Condens. Matter* **31**, 325902 (2019).
- [7] M. Wu, Z. Li, T. Cao, and S. G. Louie, *Nat. Commun.* **10**, 2371 (2019).
- [8] C. E. M. Nielsen, L. Merring-Mikkelsen, and M. O. Sauer, Master's thesis, Aalborg University, 2020; M. O. Sauer, C. E. M. Nielsen, L. Merring-Mikkelsen, and T. G. Pedersen, [arXiv:2102.12215](https://arxiv.org/abs/2102.12215).
- [9] J. Hafner, *J. Comput. Chem.* **29**, 2044 (2008).
- [10] J. Deslippe, G. Samsonidze, D. A. Strubbe, M. Jain, M. L. Cohen, and S. G. Louie, *Comput. Phys. Commun.* **183**, 1269 (2012).
- [11] D. Y. Qiu, F. H. da Jornada, and S. G. Louie, *Phys. Rev. Lett.* **111**, 216805 (2013).
- [12] A. Molina-Sánchez, M. Palummo, A. Marini, and L. Wirtz, *Phys. Rev. B* **93**, 155435 (2016).
- [13] A. Molina-Sánchez, D. Sangalli, K. Hummer, A. Marini, and L. Wirtz, *Phys. Rev. B* **88**, 045412 (2013).
- [14] G. Giorgi, K. Yamashita, and M. Palummo, *J. Phys. Chem. Lett.* **9**, 5891 (2018).
- [15] M. Palummo, M. Bernardi, and J. C. Grossman, *Nano Lett.* **15**, 2794 (2015).
- [16] I. Guilhon, M. Marques, L. K. Teles, M. Palummo, O. Pulci, S. Botti, and F. Bechstedt, *Phys. Rev. B* **99**, 161201(R) (2019).
- [17] T. Deilmann, P. Krüger, and M. Rohlfing, *Phys. Rev. Lett.* **124**, 226402 (2020).
- [18] G. Wang, A. Chernikov, M. M. Glazov, T. F. Heinz, X. Marie, T. Amand, and B. Urbaszek, *Rev. Mod. Phys.* **90**, 021001 (2018).
- [19] M. Koperski, M. R. Molas, A. Arora, K. Nogajewski, A. O. Slobodeniuk, C. Faugeras, and M. Potemski, *Nanophotonics* **6**, 1289 (2017).
- [20] S. Manzeli, D. Ovchinnikov, D. Pasquier, O. V. Yazyev, and A. Kis, *Nat. Rev. Mater.* **2**, 17033 (2017).
- [21] F. Aryasetiawan and S. Biermann, *Phys. Rev. Lett.* **100**, 116402 (2008).
- [22] We do not consider here external vector potentials.
- [23] F. Schwabl, R. Hilton, and A. Lahee, *Advanced Quantum Mechanics*, Advanced Texts in Physics (Springer, Berlin, Heidelberg, 2005).
- [24] A. Van Yperen-De Deyne, E. Pauwels, V. Van Speybroeck, and M. Waroquier, *Phys. Chem. Chem. Phys.* **14**, 10690 (2012).
- [25] A. Dal Corso and A. Mosca Conte, *Phys. Rev. B* **71**, 115106 (2005).
- [26] Relativistic pseudopotentials are often created directly using the Dirac equation, which includes relativistic corrections to all orders [84,85].
- [27] Note that one could still define a collinear functional of the sole density [86]. However it is more convenient to work within noncollinear DFT, which gives not only the exact density but also the exact spin-magnetization.
- [28] U. v. Barth and L. Hedin, *J. Phys. C* **5**, 1629 (1972).
- [29] G. Stefanucci and R. van Leeuwen, *Nonequilibrium Many-Body Theory of Quantum Systems: A Modern Introduction* (Cambridge University Press, Cambridge, 2013).
- [30] M. C. T. D. Müller, C. Friedrich, and S. Blügel, *Phys. Rev. B* **94**, 064433 (2016).
- [31] An alternative and equivalent way would be to directly write the BSE in the basis set of the calculation without SOC. In such a case, the matrix elements of the SOC potential and of the difference  $v^{Hxc}[\rho^{SOC}] - v^{Hxc}[\rho^0]$  would explicitly appear in  $H^0$ .
- [32] J. Sakurai, *Modern Quantum Mechanics, Revised Edition*, edited by S. Tuan (Addison-Wesley, Reading, MA, 1994).
- [33] J. P. Echeverry, B. Urbaszek, T. Amand, X. Marie, and I. C. Gerber, *Phys. Rev. B* **93**, 121107(R) (2016).
- [34] X.-X. Zhang, T. Cao, Z. Lu, Y.-C. Lin, F. Zhang, Y. Wang, Z. Li, J. C. Hone, J. A. Robinson, D. Smirnov, S. G. Louie, and T. F. Heinz, *Nat. Nanotechnol.* **12**, 883 (2017).
- [35] M. Drüppel, T. Deilmann, J. Noky, P. Marauhn, P. Krüger, and M. Rohlfing, *Phys. Rev. B* **98**, 155433 (2018).
- [36] T. Deilmann and K. S. Thygesen, *Phys. Rev. B* **96**, 201113(R) (2017).
- [37] M. Rohlfing, *Phys. Rev. B* **82**, 205127 (2010).
- [38] K. F. Mak, C. Lee, J. Hone, J. Shan, and T. F. Heinz, *Phys. Rev. Lett.* **105**, 136805 (2010).

- [39] D. Kozawa, R. Kumar, A. Carvalho, K. Kumar Amara, W. Zhao, S. Wang, M. Toh, R. M. Ribeiro, A. H. Castro Neto, K. Matsuda, and G. Eda, *Nat. Commun.* **5**, 4543 (2014).
- [40] H.-P. Komsa and A. V. Krasheninnikov, *Phys. Rev. B* **86**, 241201(R) (2012).
- [41] D. Y. Qiu, F. H. da Jornada, and S. G. Louie, *Phys. Rev. B* **93**, 235435 (2016).
- [42] A. Hanbicki, M. Currie, G. Kioseoglou, A. Friedman, and B. Jonker, *Solid State Commun.* **203**, 16 (2015).
- [43] A. Arora, K. Nogajewski, M. Molas, M. Koperski, and M. Potemski, *Nanoscale* **7**, 20769 (2015).
- [44] M. Manca, M. M. Glazov, C. Robert, F. Cadiz, T. Taniguchi, K. Watanabe, E. Courtade, T. Amand, P. Renucci, X. Marie, G. Wang, and B. Urbaszek, *Nat. Commun.* **8**, 14927 (2017).
- [45] H. M. Hill, A. F. Rigosi, C. Roquelet, A. Chernikov, T. C. Berkelbach, D. R. Reichman, M. S. Hybertsen, L. E. Brus, and T. F. Heinz, *Nano Lett.* **15**, 2992 (2015).
- [46] E. Liu, J. van Baren, T. Taniguchi, K. Watanabe, Y.-C. Chang, and C. H. Lui, *Phys. Rev. B* **99**, 205420 (2019).
- [47] K. He, N. Kumar, L. Zhao, Z. Wang, K. F. Mak, H. Zhao, and J. Shan, *Phys. Rev. Lett.* **113**, 026803 (2014).
- [48] A. Chernikov, T. C. Berkelbach, H. M. Hill, A. Rigosi, Y. Li, O. B. Aslan, D. R. Reichman, M. S. Hybertsen, and T. F. Heinz, *Phys. Rev. Lett.* **113**, 076802 (2014).
- [49] G. Wang, X. Marie, I. Gerber, T. Amand, D. Lagarde, L. Bouet, M. Vidal, A. Balocchi, and B. Urbaszek, *Phys. Rev. Lett.* **114**, 097403 (2015).
- [50] A. Chernikov, A. M. van der Zande, H. M. Hill, A. F. Rigosi, A. Velauthapillai, J. Hone, and T. F. Heinz, *Phys. Rev. Lett.* **115**, 126802 (2015).
- [51] L. Guo, M. Wu, T. Cao, D. M. Monahan, Y.-H. Lee, S. G. Louie, and G. R. Fleming, *Nat. Phys.* **15**, 228 (2019).
- [52] M. M. Ugeda, A. J. Bradley, S.-F. Shi, F. H. da Jornada, Y. Zhang, D. Y. Qiu, W. Ruan, S.-K. Mo, Z. Hussain, Z.-X. Shen, F. Wang, S. G. Louie, and M. F. Crommie, *Nat. Mater.* **13**, 1091 (2014).
- [53] B. Han, C. Robert, E. Courtade, M. Manca, S. Shree, T. Amand, P. Renucci, T. Taniguchi, K. Watanabe, X. Marie, L. E. Golub, M. M. Glazov, and B. Urbaszek, *Phys. Rev. X* **8**, 031073 (2018).
- [54] Y. Kikuchi, Y. Miyauchi, R. Takaoka, T. Suzuki, M. Tanaka, and S. Ohno, *Phys. Rev. B* **100**, 075301 (2019).
- [55] G. Wang, I. C. Gerber, L. Bouet, D. Lagarde, A. Balocchi, M. Vidal, T. Amand, X. Marie, and B. Urbaszek, *2D Mater.* **2**, 045005 (2015).
- [56] A. V. Stier, N. P. Wilson, K. A. Velizhanin, J. Kono, X. Xu, and S. A. Crooker, *Phys. Rev. Lett.* **120**, 057405 (2018).
- [57] S.-Y. Chen, T. Goldstein, J. Tong, T. Taniguchi, K. Watanabe, and J. Yan, *Phys. Rev. Lett.* **120**, 046402 (2018).
- [58] C. Robert, M. A. Semina, F. Cadiz, M. Manca, E. Courtade, T. Taniguchi, K. Watanabe, H. Cai, S. Tongay, B. Lassagne, P. Renucci, T. Amand, X. Marie, M. M. Glazov, and B. Urbaszek, *Phys. Rev. Materials* **2**, 011001(R) (2018).
- [59] E. Malic, M. Selig, M. Feierabend, S. Brem, D. Christiansen, F. Wendler, A. Knorr, and G. Berghäuser, *Phys. Rev. Materials* **2**, 014002 (2018).
- [60] M. R. Molas, C. Faugeras, A. O. Slobodeniuk, K. Nogajewski, M. Bartos, D. M. Basko, and M. Potemski, *2D Mater.* **4**, 021003 (2017).
- [61] Y. Zhang, T.-R. Chang, B. Zhou, Y.-T. Cui, H. Yan, Z. Liu, F. Schmitt, J. Lee, R. Moore, Y. Chen, H. Lin, H.-T. Jeng, S.-K. Mo, Z. Hussain, A. Bansil, and Z.-X. Shen, *Nat. Nanotechnol.* **9**, 111 (2014).
- [62] Y. Zhou, G. Scuri, D. S. Wild, A. A. High, A. Dibos, L. A. Jauregui, C. Shu, K. De Greve, K. Pistunova, A. Y. Joe, T. Taniguchi, K. Watanabe, P. Kim, M. D. Lukin, and H. Park, *Nat. Nanotechnol.* **12**, 856 (2017).
- [63] G. Wang, C. Robert, M. M. Glazov, F. Cadiz, E. Courtade, T. Amand, D. Lagarde, T. Taniguchi, K. Watanabe, B. Urbaszek, and X. Marie, *Phys. Rev. Lett.* **119**, 047401 (2017).
- [64] S. Funke, B. Miller, E. Parzinger, P. Thiesen, A. W. Holleitner, and U. Wurstbauer, *J. Phys.: Condens. Matter* **28**, 385301 (2016).
- [65] J. Quereda, T. S. Ghiasi, F. A. van Zwol, C. H. van der Wal, and B. J. van Wees, *2D Mater.* **5**, 015004 (2017).
- [66] Z. Lu, D. Rhodes, Z. Li, D. V. Tuan, Y. Jiang, J. Ludwig, Z. Jiang, Z. Lian, S.-F. Shi, J. Hone, H. Dery, and D. Smirnov, *2D Mater.* **7**, 015017 (2019).
- [67] C. Robert, B. Han, P. Kapuscinski, A. Delhomme, C. Faugeras, T. Amand, M. R. Molas, M. Bartos, K. Watanabe, T. Taniguchi, B. Urbaszek, M. Potemski, and X. Marie, *Nat. Commun.* **11**, 4037 (2020).
- [68] D. Y. Qiu, T. Cao, and S. G. Louie, *Phys. Rev. Lett.* **115**, 176801 (2015).
- [69] K. Kosmider, J. W. González, and J. Fernández-Rossier, *Phys. Rev. B* **88**, 245436 (2013).
- [70] F. Iannone, F. Ambrosino, G. Bracco, M. De Rosa, A. Funel, G. Guarnieri, S. Migliori, F. Palombi, G. Ponti, G. Santomauro, and P. Procacci, CRESCO ENEA HPC clusters: a working example of a multifabric GPFS Spectrum Scale layout, in *2019 International Conference on High Performance Computing Simulation (HPCS)* (IEEE, Dublin, Ireland, 2019), pp. 1051–1052.
- [71] <http://www.cresco.enea.it/english>
- [72] G. Strinati, *Riv. Nuovo Cimento* (1978–1999) **11**, 1 (1988).
- [73] P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, C. Davide, G. L. Chiarotti, M. Cococcioni, I. Dabo, A. D. 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, *J. Phys.: Condens. Matter* **21**, 395502 (2009).
- [74] J. P. Perdew, K. Burke, and M. Ernzerhof, *Phys. Rev. Lett.* **77**, 3865 (1996).
- [75] D. R. Hamann, *Phys. Rev. B* **88**, 085117 (2013).
- [76] H. J. Monkhorst and J. D. Pack, *Phys. Rev. B* **13**, 5188 (1976).
- [77] A. Marini, C. Hogan, M. Grüning, and D. Varsano, *Comput. Phys. Commun.* **180**, 1392 (2009).
- [78] G. Strinati, H. J. Mattausch, and W. Hanke, *Phys. Rev. B* **25**, 2867 (1982).
- [79] W. Hanke and L. J. Sham, *Phys. Rev. Lett.* **33**, 582 (1974).
- [80] W. Hanke and L. J. Sham, *Phys. Rev. B* **21**, 4656 (1980).
- [81] R. W. Godby and R. J. Needs, *Phys. Rev. Lett.* **62**, 1169 (1989).
- [82] F. Bruneval and X. Gonze, *Phys. Rev. B* **78**, 085125 (2008).
- [83] S. M. Dancoff, *Phys. Rev.* **78**, 382 (1950).
- [84] M. J. Oliveira and F. Nogueira, *Comput. Phys. Commun.* **178**, 524 (2008).
- [85] G. B. Bachelet and M. Schlüter, *Phys. Rev. B* **25**, 2103 (1982).
- [86] R. M. Dreizler and E. K. U. Gross, *Density Functional Theory* (Springer-Verlag, Berlin, Heidelberg, 1990).