# Solving the Bethe-Salpeter equation on a subspace: Approximations and consequences for low-dimensional materials

Diana Y. Qiu,<sup>1,2,3</sup> Felipe H. da Jornada,<sup>2,3,4</sup> and Steven G. Louie,<sup>2,3,\*</sup>

<sup>1</sup>Department of Mechanical Engineering and Materials Science, Yale University, New Haven, Connecticut 06520, USA

<sup>2</sup>Department of Physics, University of California, Berkeley, California 94720, USA

<sup>3</sup>Materials Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA

<sup>4</sup>Department of Materials Science and Engineering, Stanford University, Palo Alto, California 94305, USA

(Received 10 July 2020; accepted 21 December 2020; published 15 January 2021)

It is well known that the ambient environment can dramatically renormalize the quasiparticle gap and exciton binding energies in low-dimensional materials, but the effect of the environment on the energy splitting of the spin-singlet and spin-triplet exciton states is less understood. A prominent effect is the renormalization of the exciton binding energy and optical strength (and hence the optical spectrum) through additional screening of the direct Coulomb term describing the attractive electron-hole interaction in the kernel of the Bethe-Salpeter equation. The repulsive exchange interaction responsible for the singlet-triplet splitting, on the other hand, is unscreened within formal many-body perturbation theory. However, Benedict argued that in practical calculations restricted to a subspace of the full Hilbert space, the exchange interaction should be appropriately screened by states outside of the subspace, the so-called S approximation [L. X. Benedict, Phys. Rev. B 66, 193105 (2002)]. Here, we systematically explore the accuracy of the S approximation for different confined systems, including a molecule and heterostructures of semiconducting and metallic layered materials. We show that the S approximation is actually exact in the limit of small exciton binding energies (i.e., small direct term) and can be used to significantly accelerate convergence of the exciton energies with respect to the number of empty states, provided that a particular effective screening consistent with the conventional Tamm-Dancoff approximation is employed. We further find that the singlet-triplet splitting in the energy of the excitons is largely unaffected by the external dielectric environment for most quasi-two-dimensional materials.

DOI: 10.1103/PhysRevB.103.045117

## I. INTRODUCTION

Electron-hole interactions are greatly enhanced in lowdimensional systems as a consequence of both the spatial confinement of the electronic wave functions and the strongly inhomogeneous dielectric environment with a concomitant overall reduction in screening of the Coulomb interaction. For instance, as a consequence of reduced screening in low dimensions, in quasi-one-dimension (quasi-1D), suspended semiconducting carbon nanotubes have strongly bound excitons with binding energies of up to 1 eV [1,2], and in quasi-two-dimensions (quasi-2D), suspended monolayers of layered transition metal dichalcogenides (TMDs) have strongly bound excitons with binding energies on the order of 0.5 eV, orders of magnitude larger than their bulk counterparts [3-9]. In addition, the electronic and optical properties of quasi-low-dimensional systems are easily tuned by changing the ambient dielectric environment through substrate engineering and encapsulation. In few-layer TMDs, for instance, a graphene substrate or capping layer can renormalize the quasiparticle gap and exciton binding energies by 100-300 meV [6,10,11], while encapsulation of few-layer black phosphorus, which displays a weaker intrinsic screening than TMDs, can

Theoretically, excitons can be described with the interacting two-particle Green's function formalism within the ab initio GW plus Bethe-Salpeter equation (BSE), or GW-BSE, approach [14]. However, two significant challenges arise when applying this formalism to low-dimensional systems. First, calculations of low-dimensional systems and nanostructures are typically associated with large unit cells or supercells, so the correlated low-energy neutral excitations often require a large number of band states to be accurately described and are thus computationally demanding. Secondly, while the effect of substrate screening on the quasiparticle (QP) band gaps and exciton binding energies is basically well established, the effect of the ambient dielectric environment on the energy splitting of the spin-singlet and spin-triplet exciton states has not been as thoroughly explored, and as we argue here, much care is needed to account for its effect within the ab initio GW-BSE approach. For example, obtaining accurate theoretical predictions of the singlet-triplet splitting is essential for describing singlet fission processes (the dissociation of one singlet exciton into two triplet excitons) in some organic photovoltaic materials [15–17], exciton lifetimes [18,19], and exciton dispersion and topology [20].

2469-9950/2021/103(4)/045117(9)

qualitatively change the optical response from a system with strongly bound excitons to a system whose optical excitations are nearly free-carrier-like [12,13].

<sup>\*</sup>sglouie@berkeley.edu

The above two theoretical challenges arise as a consequence of the truncation of the Hilbert space for computational efficiency when solving the BSE; that is, only a subspace A associated with low-energy bands close to the Fermi energy is explicitly included in solving the BSE. The remaining subspace B of the Hilbert space-associated with higher-energy bands and/or other physical subsystems such as substrates-is often completely neglected. This approximation comes into two separate parts of the GW-BSE formalism: (a) the construction of the electron-hole interaction terms in the BSE kernel and (b) the construction of the basis functions for the exciton wave function in solving the BSE. In general, different physics and different cutoffs for the partition into subspaces A and B come into these two parts. In this paper, we focus on part (a), which is on the construction of the terms in the BSE kernel. For part (a), in Ref. [21], Benedict argued that instead of completely neglecting subspace B, one could capture its effect by modifying the interaction kernel defined in subspace A. The resulting scheme, known as the S approximation, amounts to simply solving the BSE on the subspace A, but screening the exchange interaction kernel,  $K^{x}$ , in the BSE—which would otherwise involve only the bare Coulomb interaction-using a screening scheme that includes only free electron-hole pair excitations from the subspace B [21].

Benedict et al. later applied the S approximation to hydrogenated Si clusters [22], and similar approaches to screening the exchange term in the BSE have been used to obtain better agreement with experiment for various nanocluster systems [23,24], as well as to study carbon nanotubes on metallic substrates [25]. Recently, Ref. [26] has explored the effect of applying the S approximation for benzene on a gold surface, using the random-phase approximation (RPA) dielectric function for gold to screen the exchange term when solving the BSE for an isolated benzene molecule, and found that this form of screening results in a large renormalization of the singlet-triplet splitting of the lowest-energy exciton in benzene. These applications of the S approximation further suggest that in general, the screening environment might have a large effect on the singlet-triplet splitting in low-dimensional systems and hence on their exciton lifetimes, dispersion, and dynamics, since the exchange interaction kernel  $K^x$  is responsible for the splitting in energy between spin-singlet and spin-triplet excitonic states. However, there has been to date no systematic analysis of the accuracy of the S approximation against explicitly full-Hilbert-space converged BSE calculations, nor has there been any comparison of the effect of the substrate on the singlet-triplet splitting in the S approximation against an explicit BSE calculation including both substrate and adsorbate. Moreover, the S approximation belongs to a broad class of embedding schemes used to downfold the many-body Hamiltonian across quantum chemistry and condensed matter physics, including in dynamical mean-field theory (DMFT) [27], the constrained RPA [28], self-energy embedding theory (SEET) [29], and configuration interaction calculations [30,31], but it is one of the few schemes where the larger Hilbert space remains computationally accessible thus allowing for numerical validation of the embedding procedure.

In this paper, we rederive the S approximation using the Löwdin partitioning scheme, similarly to Ref. [26], and perform a systematic analysis of its accuracy by comparing the solution of the BSE with and without the S approximation as a function of the size of subspace A with respect to the full Hilbert space. We investigate three prototypical systems: (1) an isolated benzene molecule, where subspaces A (the lowenergy orbitals) and B (the high-energy orbitals) overlap in real space but not energy space; (2) a monolayer of hexagonal boron nitride (h-BN) encapsulated in bulk MoS<sub>2</sub>, where subspaces A (h-BN orbitals) and B ( $MoS_2$  orbitals) do not overlap in real space but subspace B is known to screen subspace A; and (3) a superlattice of monolayer h-BN and graphene, where again subspaces A (h-BN) and B (graphene) do not overlap in real space but graphene can be expected to strongly screen excitations in h-BN. Two major results are obtained. (1) We find that if a carefully designed subspace screening is employed that is consistent with the Tamm-Dancoff approximation (TDA), the S approximation can significantly speed up convergence of the excitation energies with respect to the size of subspace A in molecules in the gas phase where the exchange interaction couples a large number of unoccupied orbitals. We emphasize that consistent with the usage of the TDA, it is essential to only include electron-hole pairs that are forward propagating in time-a subtlety which is either neglected or not made explicit in previous work [21,25,26]. (2) We find that both metallic and semiconducting substrates have a negligible effect on singlet-triplet splitting in layered materials, where the wave functions of the adsorbate and substrate do not significantly hybridize, and we show that this is true in general for any adsorbate where the electron and hole band wave functions can be described by Wannier functions whose extent is smaller than the adsorbate-substrate distance.

This paper is organized as follows. In Sec. II, we rederive the *S* approximation. In Sec. III, we describe the details of our implementation of the *S* approximation. In Sec. IV, we report our results for an isolated benzene molecule. In Sec. V, we report our results for monolayer h-BN encapsulated in bulk  $MoS_2$  and a superlattice of alternating layers of monolayer h-BN and graphene. We conclude in Sec. VI by summarizing our results.

## **II. THE S APPROXIMATION**

Excitons can be described within the two-particle Green's function formalism via the BSE [14]

$$L = L_0 + L_0 K L, \tag{1}$$

where *L* is the electron-hole correlation function,  $L_0$  is the noninteracting counterpart describing the free propagation of an electron and a hole (as quasiparticles), and *K* is an interaction kernel. The BSE is often written, within the TDA [14] and a static approximation for *K*, in terms of an effective eigenvalue problem expressed in a quasiparticle basis,

$$H\psi_{X} = \Omega_{X}\psi_{X},$$

$$H = D + K,$$

$$D_{vc\mathbf{k},v'c'\mathbf{k}'} = \left[\omega - L_{0}^{-1}\right]_{vc\mathbf{k},v'c'\mathbf{k}'}$$

$$= (\varepsilon_{c\mathbf{k}} - \varepsilon_{v\mathbf{k}})\delta_{vv'}\delta_{cc'}\delta_{\mathbf{k}\mathbf{k}'},$$
(2)



FIG. 1. (a) BSE kernel diagrams consisting of an exchange term  $K^x$  and a direct term  $K^d$ . (b) Expansion of the electron-hole correlation function *L* in the BSE in powers of the BSE kernel. (c) Effective screened exchange interaction kernel  $\overline{K}_{AA}^x$ , which contains electron-hole bubbles in the *B* subspace.

where *H* is the effective BSE Hamiltonian;  $\psi_X$  and  $\Omega_X$  are the associated eigenvectors and eigenvalues, which describe electron-hole neutral excitations of the system;  $\varepsilon_n$  are quasiparticle excitation energies; v (*c*) label valence (conduction) bands; and **k** is a **k** point in the Brillouin zone. For simplicity, the spin index is omitted here.

Rigorous conserving approximations for the kernel K can be obtained by expressing it in terms of the functional derivative of the single-particle self-energy  $\Sigma$  with respect to the single-particle Green's function  $G, \tilde{K} = \frac{\delta \Sigma}{\delta G}$ . Within the commonly employed GW approximation for  $\Sigma$  and neglecting the variation of the dynamically screened interaction when computing the functional derivative [14], K can be expressed as the sum of an attractive direct term  $(K^d)$  involving the screened Coulomb interaction and a repulsive exchange term  $(K^{x})$  involving the bare Coulomb interaction. The exchange term,  $K^x$ , is solely responsible for the splitting of spin-singlet and spin-triplet excitations, if spin-orbit interaction is neglected. The two kernel terms are shown diagrammatically in Fig. 1(a), and Fig. 1(b) shows the expansion of the electronhole correlation in Eq. (2) in powers of  $K^x$  and  $K^d$ . An important property that the interaction kernel K must have in Eq. (2) is that it must be a *proper interaction* [32]; that is, no component of K can be written in terms of two interaction kernels  $K^a$  and  $K^b$  connected by a noninteracting correlation function,  $K \neq K^a L_0 K^b$ . This property is automatically satisfied when the kernel is computed in terms of the functional derivative of a conserving electronic self-energy. Following this analysis, one should indeed use the bare Coulomb interaction in  $K^x$  instead of the screened counterpart  $K^W =$  $K^{x} + K^{x}L_{0}K^{x} + \cdots$ , as the latter is improper.

While it is straightforward to justify the usage of the the bare interaction in  $K^x$  through the aforementioned analytic arguments, the situation is more nuanced in numerical calculations when the BSE is solved in a subspace of the total Hilbert space. Motivated by Benedict [21], we formally separate the Hilbert space into two subspaces. Subspace *A*—the one that we wish to explicitly solve for—may include (i) only a small number of quasiparticle states necessary to describe

correlated neutral excitations and/or (ii) a subset of the physical system of interest, e.g., a molecule or a quasi-2D material but not the substrate supporting it. We denote the remainder of the Hilbert space by subspace B.

We write the full BSE effective Hamiltonian in a block form in terms of the two subspaces as

$$H^{\rm BSE} = \begin{pmatrix} H_{AA} & H_{AB} \\ H_{BA} & H_{BB} \end{pmatrix},\tag{3}$$

where the noninteracting electron-hole correlation function is diagonal in the quasiparticle basis, so any terms involving  $L_0^{AB}$  or  $L_0^{BA}$  [and  $D^{AB}$  or  $D^{BA}$  in Eq. (2)] must be zero. We emphasize that this is not an approximation, unlike what was originally stated in the proposal of the *S* approximation [21]. The form of the energy-dependent effective Hamiltonian projected onto subspace *A* is described by Löwdin partitioning [33] as

$$H_{\rm eff}(\omega) = H_{AA} + H_{AB}L_B(\omega)H_{BA},$$
  
$$\overline{L}_B(\omega) \equiv (\omega - H_{BB})^{-1},$$
 (4)

where  $\overline{L}_B(\omega)$  is the interacting electron-hole correlation function for the subspace *B* alone, i.e., without any cross term with subspace *A*.

The effective Hamiltonian  $H_{\text{eff}}$  in *A* cannot be obtained without explicit knowledge of *B*, so it is typically approximated by truncating the full Hamiltonian to the *A* subspace  $(H_{\text{eff}} \approx H_{AA})$ . Calculations are then converged with respect to the size of subspace *A*. This approach is reasonable as long as the eigenenergies of  $H_{AA}$  are far from those of  $H_{BB}$ at convergence, and the coupling of the subspaces  $H_{AB}$  and  $H_{BA}$  is small. Additionally, even though there are formalisms that solve the BSE without explicitly including unoccupied states [34,35], the Hilbert space is still often separated into subspaces *A* and *B*, for instance, through the definition of the projection operator or by not explicitly including the supporting substrate in the corresponding Hamiltonian.

An intermediate approximation, which is expected to be more accurate than setting  $H_{\text{eff}} \approx H_{AA}$ , is to include only the exchange contribution to the interaction kernel in coupling terms other than  $K_{AA}$ . This is physically motivated by the fact that unlike  $K^x$ , the direct term  $K^d$  depends on the spatial overlap of electron and hole states between initial and final states. Hence  $K_{AB}^d$  and  $K_{BA}^d$  are expected to vanish when subspaces A and B are separated in space, such as when subspace Acontains the states for a molecule and subspace B contains the states for the substrate supporting it. Within this approximation, we obtain

$$H_{\text{eff}}(\omega) = D_A + K_{AA}^d + \overline{K}_{AA}^x(\omega),$$
  

$$\overline{K}_{AA}^x(\omega) = K_{AA}^x + K_{AB}^x \overline{L}_B(\omega) K_{BA}^x(\omega),$$
  

$$\overline{L}_B(\omega) = L_B^0(\omega) + L_B^0(\omega) K_{BB} \overline{L}_B(\omega).$$
(5)

We identify  $\overline{L}_B$  in Eq. (5) as the electron-hole polarization propagator for subspace *B* evaluated within the ring approximation [32], so that  $\overline{K}_{AA}^x$  is simply the exchange contribution to the interaction kernel but with the bare Coulomb interaction modified to include screening arising from positive-energy electron-hole pair excitations from subspace *B*. The effective Hamiltonian in Eq. (5) is equivalent to solving for the electron-hole correlation function in the subspace *A* but using a modified screened exchange interaction kernel, which incorporates the coupling of subspace *A* with subspace *B*. The effective interaction kernel  $\overline{K}_{AA}^x$  is shown diagrammatically in Fig. 1(c). We note that in the limit of small exciton binding energies or vanishing direct interaction, Eq. (5) is exact, as the modified interaction kernel  $\overline{K}_{AA}^x$  is equivalent to the downfolding within Löwdin partitioning [33].

We note that the approximation obtained in Eq. (5) is equivalent to the *S* approximation obtained by Benedict [21] but with an important difference: Consistent with the use of the Tamm-Dancoff approximation in the BSE, the screening from subspace *B* that is used in  $\overline{K}_{AA}^x$  only contains positiveenergy electron-hole pair excitations as opposed to positive and negative excitations. This is an often overlooked detail, which can cause an overscreening within the *S* approximation if neglected.

Finally, we emphasize that even if the interaction kernel  $K^d$  and  $K^x$  are strictly static, the effective interaction kernel  $\overline{K}_{AA}^{x}(\omega)$  picks up a dynamic dependence due to the downfolding procedure. Correspondingly, the screened Coulomb interaction constructed from states from outside the subspace A should rigorously be self-consistently evaluated at the excitation energy  $\omega = \Omega_X$  of the state X that one is interested in solving for. However, we drop the dynamical dependence in  $\overline{K}_{AA}^{x}(\omega)$  in this paper. This is partially justified by the fact that when subspace B is separated enough in energy,  $\overline{K}_{AA}^{x}(\omega)$ will be a smooth function up to an energy scale of the order of the smallest quasiparticle band-to-band transition energy difference  $D^B_{vc\mathbf{k},v'c'\mathbf{k}'}$  in the *B* subspace. This is typically much larger than the excitation energy of the low-energy excitons we are interested in when we use the S approximation to accelerate the convergence with respect to the number of bands included in the A subspace. Yet, this is not strictly the case when the B subspace overlaps in energy with the lowest exciton excitation energy, as in the case when subspace B is a substrate. The static approximation to  $\overline{K}_{AA}^{x}(\omega)$  in this case may cause an additional error, although our calculations show that this error is small, as we show in the next sections.

In the next section, we describe how to implement the *S* approximation within an *ab initio* GW-BSE framework.

## III. NUMERIC IMPLEMENTATION OF THE S APPROXIMATION

In this section, we detail the modifications in the BERKE-LEYGW software package [36–38] to implement the *S* approximation. For all systems, we solve for the electronhole excitations within the *ab initio* GW plus Bethe-Salpeter equation (GW-BSE) approach in the Tamm-Dancoff approximation (TDA). We modify the kernel of the BSE to implement the *S* approximation by screening the exchange matrix elements in the following way:

$$\langle vc\mathbf{k}|K^{x}|v'c'\mathbf{k}'\rangle = \sum_{\mathbf{G}\mathbf{G}'} M^{*}_{vc\mathbf{k}}(\mathbf{Q},\mathbf{G})\overline{W}_{\mathbf{G}\mathbf{G}'}(\omega=0,\mathbf{Q})M_{v'c'\mathbf{k}'}(\mathbf{Q},\mathbf{G}').$$
<sup>(6)</sup>

Here, the BSE matrix elements are written in the basis of electron-hole pairs vc at a **k**-point **k**. On the right-hand side,

 $M_{vck}(\mathbf{Q},\mathbf{G}) = \langle v\mathbf{k} + \mathbf{Q} | e^{i(\mathbf{Q}+\mathbf{G})\cdot\mathbf{r}} | c\mathbf{k} \rangle$ , where **G** is a reciprocal lattice vector and  $\mathbf{Q}$  is the center-of-mass momentum of the electron-hole pair (for simplicity, the numerical examples we show will only consider excitations where  $\mathbf{Q} = 0$ ).  $\overline{W}_{\mathbf{G}\mathbf{G}'}(\omega=0,\mathbf{Q})$  is the screened Coulomb interaction in the static limit. In the absence of the S approximation,  $\overline{W}$  is the usual bare Coulomb interaction,  $\overline{W}_{\mathbf{GG}'}(\mathbf{Q}) = v(\mathbf{Q} + \mathbf{G}) \,\delta_{\mathbf{GG}'}$ , where  $v(\mathbf{Q} + \mathbf{G})$  are the Fourier components of the Coulomb interaction. Since we solve the BSE within the TDA, solutions with positive and negative excitation energies do not mix, and hence the virtual electron-hole pairs introduced by the exchange diagram in the BSE also do not couple positive and negative excitation energies. Consequently, within the S approximation,  $\overline{W}$  is screened by the polarizability of the B subspace created only by positive electron-hole excitations, which, at  $\omega = 0$ , is equal to half of the noninteracting polarizability that is typically computed within the random-phase approximation (RPA). That is,

$$\overline{W}_{\mathbf{G}\mathbf{G}'}(\omega=0,\mathbf{Q}) = \overline{\epsilon}_{\mathbf{G}\mathbf{G}'}^{-1}(\omega=0,\mathbf{Q})v(\mathbf{Q}+\mathbf{G}), \quad (7)$$

where  $\overline{\epsilon}$  is a dielectric matrix of the form

$$\overline{\epsilon}_{\mathbf{G}\mathbf{G}'}(\omega=0,\mathbf{Q}) = \delta_{\mathbf{G}\mathbf{G}'} - v(\mathbf{Q}+\mathbf{G})\frac{1}{2}\overline{\chi}^{0}_{\mathbf{G}\mathbf{G}'}(\omega=0,\mathbf{Q}).$$
(8)

The factor of  $\frac{1}{2}$  removes the electron-hole pairs that are backward propagating in time, and  $\overline{\chi}^0$  is the static, noninteracting RPA polarizability due to all electron-hole pairs not in subspace *A*,

$$\overline{\chi}_{\mathbf{G}\mathbf{G}'}^{0}(\mathbf{Q}) = \sum_{n}^{\mathrm{occ}} \sum_{n'}^{\mathrm{unocc}} \sum_{\mathbf{k}}^{\mathbf{k}} \frac{M_{nn'\mathbf{k}}^{*}(\mathbf{Q},\mathbf{G})M_{nn'\mathbf{k}}(\mathbf{Q},\mathbf{G}')}{E_{n\mathbf{k}+\mathbf{Q}} - E_{n'\mathbf{k}}} - [\chi^{AA}]_{\mathbf{G}\mathbf{G}'}^{0}(\mathbf{Q}), \qquad (9)$$

while  $[\chi^{AA}]^0$  is the noninteracting RPA polarizability formed by transitions from occupied to unoccupied states both belonging to the *A* subspace,

$$[\chi^{AA}]^{0}_{\mathbf{G}\mathbf{G}'}(\mathbf{Q}) = \sum_{n \in A}^{\operatorname{occ}} \sum_{\mathbf{k}' \in A}^{\operatorname{unocc}} \sum_{\mathbf{k}} \frac{M^{*}_{nn'\mathbf{k}}(\mathbf{Q}, \mathbf{G})M_{nn'\mathbf{k}}(\mathbf{Q}, \mathbf{G}')}{E_{n\mathbf{k}+\mathbf{Q}} - E_{n'\mathbf{k}}}.$$
(10)

In Eqs. (9) and (10), we have dropped the  $\omega = 0$  variable for compactness.

## **IV. BENZENE MOLECULE**

In this section, we calculate the neutral excitations of a single benzene molecule in a  $16 \times 16 \times 16$ -Åsupercell. First, we use density functional theory (DFT) [39,40] in the local-density approximation (LDA) [41], as implemented in QUANTUM ESPRESSO [42], to obtain a mean-field starting point for our GW and GW-BSE calculation. Plane-wave components up to 80 Ry are included in the wave function. We then performed a GW-BSE calculation with the BERKE-LEYGW code [37,38,43]. In the GW calculation, the dynamical screening effects were accounted for within the Hybertsen-Louie generalized plasmon-pole (HL-GPP) model [43]. The dielectric matrix includes plane-wave components up to 20 Ry, and the summation over unoccupied states includes 2000 states. The static remainder technique was used to speed up convergence with respect to the number of empty states [44]. We truncate the Coulomb potential in a sphere with a radius of 8 Å to prevent spurious interactions between periodic images of the supercell. For these parameters, we find a highest-occupied-molecular-orbital– lowest-unoccupied-molecular-orbital (HOMO-LUMO) gap of 10.4 eV. For the BSE calculation, we include plane-wave components up to 6 Ry in the dielectric matrix. Then, we solve the BSE with and without the *S* approximation and examine the convergence of the singlet and triplet excitation energies with respect to the number of unoccupied bands explicitly included in the BSE Hamiltonian. The number of occupied states included in the BSE Hamiltonian is fixed at 10.

We focus on two low-energy exciton states in analyzing the results of the BSE: (1) the lowest-energy excited state, which we label  $E_{d,1}$ , whose singlet exciton is dark for polarizations of light in the plane of the benzene ring perpendicular to the carbon-carbon bond, and (2) the lowest-energy exciton state for which the singlet exciton is bright when the polarization of the external electric field is in the plane of the benzene ring perpendicular to the carbon-carbon bond, which we label  $E_{h_1}$ . To analyze the effect of the exchange interaction, we compare the convergence of the spin-singlet exciton  $E_{d,1}^{S}(E_{b,1}^{S})$  for state "d"(state "b"), which has a contribution from the exchange interaction, with the convergence of the spin-triplet exciton  $E_{d1}^{T}$  ( $E_{b1}^{T}$ ) for state d (state b), which has no contribution from the exchange interaction. We note here that both d and b states can be either spin singlet or spin triplet. Only the b spin-singlet state is optically active. For the d states, the spin-singlet state is optically inactive due to the mirror plane symmetry of the molecular orbitals composing the exciton. The convergence of the singlet states and triplet states with respect to the number of unoccupied states and without the S approximation is shown by the red circles and green squares in Fig. 2, respectively. We see that the excitation energy of the  $E_{d,1}$  singlet and triplet states converge rapidly with respect to the number of unoccupied bands, while the excitation energy of the  $E_{b,1}$  singlet converges very slowly, requiring more than 200 bands to converge the error in the excitation energy to less than 100 meV. Including dynamical effects in the screening changes the eigenvalues by less than 50 meV. The triplet state associated with the  $E_{b,1}$  singlet state, however, converges much more rapidly with the number of unoccupied bands, suggesting that the slow convergence can be primarily attributed to the exchange interaction, making the  $E_{b,1}$  exciton an ideal test case of the S approximation.

Next, we apply the *S* approximation to the benzene molecule by screening the exchange interaction. The results are shown by the blue triangles in Fig. 2. We see that the *S* approximation does indeed speed up convergence with respect to the number of unoccupied states. To understand where this improvement is coming from, we then solve the BSE with the direct Kernel term, which is responsible for the finite binding energy of bound excitons, set to zero—the limit where the *S* approximation becomes exact. The convergence of the BSE with only the exchange interaction is shown in Fig. 3. In the restricted subspace, the convergence of the eigenvalues for  $E_{b,1}$  is slow, but when the exchange is screened within the *S* approximation, the eigenvalue immediately jumps to the converged value after a small number of unoccupied states



FIG. 2. Convergence of exciton excitation energies for a benzene molecule with respect to the number of unoccupied bands ( $N_{unocc}$ ) used to diagonalize the BSE Hamiltonian of (a) the singlet excitation energy for the lowest-energy dark exciton  $E_{d,1}^S$  (which is dark due to a separate selection rule), (b) the excitation energy for the triplet state [spin partner to the state in (a)] associated with the lowestenergy dark exciton  $E_{d,1}^T$ , (c) the singlet excitation energy for the lowest-energy bright exciton  $E_{b,1}^S$ , and (d) the excitation energy for the triplet state [spin partner to the state in (c)] associated with the lowest-energy bright exciton  $E_{b,1}^T$ . Red circles give singlet excitation energies calculated in the A subspace only. Blue triangles are singlet excitation energies calculated with the S approximation. The red dashed line is the linear extrapolation of the red circles to  $1/N_{unocc} \rightarrow 0$ . The green squares show the convergence of the triplet state, which does not include any exchange contribution.



FIG. 3. Convergence of singlet excitation energy in a benzene molecule for  $E_{d,1}$  (top panel) and  $E_{b,1}$  (bottom panel) with respect to the number of unoccupied bands ( $N_{\text{unocc}}$ ) used to diagonalize the BSE Hamiltonian with only the exchange interaction included in the interaction kernel (i.e.,  $K^d = 0$ ). Red circles are excitation energies in the restricted subspace only. Blue triangles are excitation energies calculated with the *S* approximation. The dashed line is the linear extrapolation of the red circles to  $1/N_{\text{unocc}} \rightarrow 0$ .



FIG. 4. (a) Illustration of supercell with strained h-BN encapsulated in  $MoS_2$ . The dashed rectangle indicates the boundaries of the supercell used in the calculation, which is periodic in all directions. (b) Band structure at LDA level of the  $MoS_2$  and h-BN supercell (red lines) superimposed on the band structure of an isolated  $MoS_2$  slab (blue circles).

have been included. These results suggest that the *S* approximation can be used to moderately speed up convergence of the excitation energies when the exchange interaction couples a large number of unoccupied orbitals.

# V. SUBSPACE PARTITIONING IN LOW-DIMENSIONAL MATERIALS ON A SUBSTRATE

#### A. Boron nitride encapsulated in MoS<sub>2</sub>

In this section, we study how the singlet-triplet splitting may be affected by the presence of substrates and whether the S approximation can accurately describe this effect. We first consider the case of a monolayer of h-BN encapsulated between bulklike slabs (four layers) of MoS<sub>2</sub>. In order to provide a computationally feasible test of the S approximation, we set up an artificial system where the BN is strained by 20% in order to match the lattice of  $MoS_2$ . The  $MoS_2$  and BN slabs are separated by 6 Å to prevent hybridization of the wave functions. The supercell setup is shown in Fig. 4(a). In Fig. 4(b), we compare the band structure of the supercell with the band structure of the  $MoS_2$  slab alone. The overlap of the two band structures indicates that there is minimal hybridization of the wave functions near the Fermi level. We emphasize that this supercell setup is not meant to describe realistic experimental conditions but rather to provide a computationally tractable test case, where the wave functions of subspaces A and B do not hybridize but the screening environment is dramatically changed from quasi-2D (for a freestanding atomic layer of monolayer h-BN) to bulklike (for h-BN encapsulated in MoS<sub>2</sub>). Here, we focus on the effect of screening from MoS<sub>2</sub> on excitations in h-BN because MoS<sub>2</sub> is more polarizable and h-BN has a large singlet-triplet splitting on the order of 100 meV compared with 20 meV in MoS<sub>2</sub>, so any screening effect on the exchange interaction is expected to be more pronounced in h-BN.

In solving the BSE, the dielectric matrix is calculated on a  $24 \times 24 \times 1$  **k**-point grid with plane-wave components up to 5 Ry included in the dielectric matrix and band states with

TABLE I. Singlet-triplet splitting for the lowest-energy exciton in monolayer h-BN encapsulated in  $MoS_2$  ( $MoS_2 + h$ -BN) and in a superlattice with alternating layers of monolayer h-BN and graphene (graphene + h-BN) calculated in three ways: (1) with the basis of the BSE matrix restricted to the subspace of h-BN (*A* subspace), (2) with the *S* approximation, and (3) with the full Hamiltonian including states from both h-BN and MoS<sub>2</sub> or graphene.

	Singlet-triplet splitting (eV)		
	A subspace	S approximation	Full Hamiltonian
$MoS_2 + h-BN$	0.07	0.07	0.06
Graphene + h-BN	0.1	0.1	0.09

energies up to 5 Ry included in the sum over unoccupied states. We calculate two dielectric matrices: one that includes the full supercell of monolayer h-BN encapsulated in a slab of MoS<sub>2</sub> ( $\epsilon_{BN+MoS_2}$ ) and another that includes only the MoS<sub>2</sub> slab ( $\epsilon_{MoS_2}$ ). Then, we solve the BSE for three cases: (1) the conventional treatment considering a freestanding monolayer of h-BN, where the direct kernel  $K^d$  is screened by the full dielectric matrix  $\epsilon_{BN+MoS_2}$  and the exchange kernel  $K^x$  is unscreened; (2) the S-approximation case, where  $K^d$  is screened by  $\epsilon_{\text{BN+MoS}}$ , and  $K^x$  is screened by  $\epsilon_{\text{MoS}}$ ; and (3) the explicit BSE calculation on the full supercell containing both h-BN and MoS<sub>2</sub>. We compare the singlet-triplet splitting of the lowest-energy exciton in h-BN for the three cases described above. In case (3), when the BSE is solved for the full system including MoS<sub>2</sub> and h-BN, the calculation includes many excitons arising from MoS<sub>2</sub>. To identify the spin-singlet and spin-triplet excitons arising primarily from h-BN, we calculated the charge density of the electron contribution to each exciton state when the hole is fixed at the center of the h-BN slab. We identify excitons with maximum electron charge density that lies within 3 Å of the center of the h-BN slab as excitons in h-BN.

The singlet-triplet splitting of the lowest-energy h-BN exciton for the three cases is shown in the first row of Table I. We see that encapsulation causes a small reduction in the singlet-triplet splitting. Explicitly including the MoS<sub>2</sub> states in the BSE Hamiltonian reduces the singlet-triplet splitting by 5 meV or about 7%. The S approximation gives the same result as the freestanding h-BN. Here, the small error in the S approximation, compared with the full BSE Hamiltonian, arises from neglecting wave-function hybridization between subspaces A and B and the assumption that the screening from subspace B is static. We also perform an additional calculation at a reduced distance of 3.5 Å to isolate the effect of substrate distance. At the reduced distance, hybridization effects are larger due to the artificial periodicity of the model system, but we can qualitatively understand the role of the Bsubspace in the exciton singlet-triplet splitting. When the substrate adsorbate distance is reduced to 3.5 Å, including the substrate screening within the S approximation reduces the singlet-triplet splitting by only 2 meV.

## B. Boron nitride/graphene superlattice

We next consider what happens when the external screening environment is metallic rather than semiconducting by



FIG. 5. (a) Illustration of supercell with a superlattice of alternating layers of graphene and monolayer h-BN. The dashed rectangle indicates the boundaries of the periodic supercell used in the calculation, which is periodic in all directions. (b) Band structure at LDA level of the graphene and h-BN supercell (red lines) superimposed on the band structure of the isolated h-BN monolayer (blue circles).

constructing a bulklike superlattice of alternating monolayers of h-BN and monolayers of graphene. Again, in order to make the calculation computationally tractable and reduce wave-function hybridization, we construct an artificial supercell where the h-BN and graphene monolayers are separated by 5 Å and the lattice of the h-BN is strained to match that of graphene [Fig. 5(a)]. Figure 5(b) shows the DFT band structure of the freestanding h-BN slab overlaid on the band structure of the h-BN/graphene superlattice to show that hybridization of the wave functions is minimal. We repeat the three kinds of calculations in Sec. V A for the h-BN/graphene superlattice.

We find that similar to the case of h-BN encapsulated in  $MoS_2$ , graphene has a very small effect on the singlet-triplet splitting of excitons in h-BN. Inclusion of the graphene in the full BSE Hamiltonian reduces the singlet-triplet splitting of the lowest-energy exciton in h-BN by 5 meV or about 5%. Once again, the *S* approximation gives the same result as the the conventional treatment using a freestanding h-BN with the direct term screened.

## C. Analysis of singlet-triplet splitting

Here, we connect the effect of the substrate screening on the exciton singlet-triplet splitting to the character of the exciton states in the adsorbate in order to understand why the effect of the external screening is small. The screenedexchange contribution to the BSE kernel matrix elements in the *S* approximation is given by

$$\langle vc\mathbf{k}, \mathbf{Q} | K^{SX} | v'c'\mathbf{k}', \mathbf{Q}' \rangle = \int d^3r d^3r' \psi_{v\mathbf{k}}^*(\mathbf{r}) \psi_{c\mathbf{k}+\mathbf{Q}}(\mathbf{r}) \\ \times \overline{W}(\mathbf{r} - \mathbf{r}') \psi_{v'\mathbf{k}}(\mathbf{r}') \psi_{c'\mathbf{k}'+\mathbf{Q}'}^*(\mathbf{r}'),$$
(11)

where  $\overline{W}$  is the Coulomb interaction screened within the *B* subspace of the BSE and evaluated within the TDA, as discussed in the previous sections.

The matrix elements can also be transformed into a Wannier basis

$$\langle v \mathbf{R}_{v}; c \mathbf{R}_{c} | K^{SX} | v' \mathbf{R}_{v'}; c' \mathbf{R}_{c'} \rangle$$

$$= \int d^{3}r d^{3}r' \mathcal{W}_{v}^{*}(\mathbf{r} - \mathbf{R}_{v}) \mathcal{W}_{c}(\mathbf{r} - \mathbf{R}_{c}) \overline{\mathcal{W}}(\mathbf{r} - \mathbf{r}')$$

$$\times \mathcal{W}_{v'}(\mathbf{r}' - \mathbf{R}_{v'}) \mathcal{W}_{c'}^{*}(\mathbf{r}' - \mathbf{R}_{c'}),$$
(12)

where  $W_n(\mathbf{r} - \mathbf{R})$  is a Wannier function for band *n* centered at **R**, and *v* and *c* label valence and conduction bands, respectively.

Since Wannier functions are well localized in real space, the singlet-triplet splitting is dominated by matrix elements in Eq. (12) that correspond to having the electron and the hole in the same position, and so, for a two-band Wannier exciton, it is proportional to

$$\Delta_{ST} \sim \langle v \mathbf{R}; c \mathbf{R} | K^{SX} | v \mathbf{R}; c \mathbf{R} \rangle$$
  
=  $\int d^3 r d^3 r' \mathcal{W}_v^*(\mathbf{r}) \mathcal{W}_c(\mathbf{r}) \overline{\mathcal{W}}(\mathbf{r} - \mathbf{r}') \mathcal{W}_v(\mathbf{r}') \mathcal{W}_c^*(\mathbf{r}').$  (13)

Because of the localization of the Wannier functions, the integral above is given by the screened Coulomb interaction with separations constrained to the dimension of the unit cell, a short length scale at which the screening due to the substrate  $\overline{\epsilon}^{-1}(\mathbf{r},\mathbf{r}')$  is not effective. Hence the screening due to the substrate will not lead to a modification in the singlet-triplet exciton energies, unless the separation between the substrate and the sample of interest is smaller than the typical size of the Wannier function. This is not the case in TMD monolayers, where the Wannier functions are smaller than the unit cell [12,45], and more generally, we predict that the singlet-triplet splitting in most quasi-2D systems will not be influenced by the screening due to the substrate. However, the substrate can still affect the singlet-triplet splitting in other systems, such as adsorbates that are metals or molecules, where the extent of the Wannier functions is larger than the substrate-adsorbate distance. This is consistent with previous calculations [26] that show that the exciton singlet-triplet splitting in a benzene molecule can be reduced by over 10% when the moleculesubstrate separation is smaller than the typical size of the Wannier or Boys function of the frontier orbitals in a benzene molecule ( $\leq 5$  Å).

## VI. SUMMARY

In summary, we rederive the *S* approximation using the Löwdin partitioning—obtaining a new expression with an additional factor of  $\frac{1}{2}$ —and perform a systematic analysis of its accuracy by comparing the solution of the BSE with and without the *S* approximation as a function of the size of subspace *A* of the Hilbert space. We find that screening the exchange interaction in the BSE with a screening carefully designed to be consistent with the TDA can moderately speed up convergence with respect to subspace size in confined systems, such as molecules, where the exchange interaction couples a large number of occupied and unoccupied states in the same spatial region. We find that both metallic and semiconducting substrates have a negligible effect on singlet-triplet splitting in layered materials, where the wave functions of the system of interest and substrate do not significantly

hybridize. In general, the screening from a substrate will not affect the exciton singlet-triplet splitting in the sample unless the extent of the maximally localized Wannier function for the single-particle states is larger than the adsorbate-substrate distance.

## ACKNOWLEDGMENTS

This work was supported by the Center for Computational Study of Excited State Phenomena in Energy Materials (C2SEPEM) at the Lawrence Berkeley National Laboratory, which is funded by the U.S. Department of Energy, Office of Science, Basic Energy Sciences, Materials Sciences and Engi-

- C. D. Spataru, S. Ismail-Beigi, L. X. Benedict, and S. G. Louie, Phys. Rev. Lett. 92, 077402 (2004).
- [2] J. Maultzsch, R. Pomraenke, S. Reich, E. Chang, D. Prezzi, A. Ruini, E. Molinari, M. S. Strano, C. Thomsen, and C. Lienau, Phys. Rev. B 72, 241402(R) (2005).
- [3] D. Y. Qiu, F. H. da Jornada, and S. G. Louie, Phys. Rev. Lett. 111, 216805 (2013); 115, 119901(E) (2015).
- [4] K. F. Mak, K. He, C. Lee, G. H. Lee, J. Hone, T. F. Heinz, and J. Shan, Nat. Mater. 12, 207 (2013).
- [5] 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).
- [6] 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).
- [7] B. Zhu, X. Chen, and X. Cui, Sci. Rep. 5, 9218 (2015).
- [8] Z. Ye, T. Cao, K. O'Brien, H. Zhu, X. Yin, Y. Wang, S. G. Louie, and X. Zhang, Nature (London) 513, 214 (2014).
- [9] 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).
- [10] A. J. Bradley, M. M. Ugeda, F. H. da Jornada, D. Y. Qiu, W. Ruan, Y. Zhang, S. Wickenburg, A. Riss, J. Lu, S.-K. Mo, Z. Hussain, Z.-X. Shen, S. G. Louie, and M. F. Crommie, Nano Lett. 15, 2594 (2015).
- [11] F. A. Rasmussen, P. S. Schmidt, K. T. Winther, and K. S. Thygesen, Phys. Rev. B 94, 155406 (2016).
- [12] X. Li, S. Zhang, and Q. Wang, Nanoscale 9, 562 (2017).
- [13] D. Y. Qiu, F. H. da Jornada, and S. G. Louie, Nano Lett. 17, 4706 (2017).
- [14] G. Strinati, Riv. Nuovo Cimento Soc. Ital. Fis. 11, 1 (1988).
- [15] M. B. Smith and J. Michl, Chem. Rev. (Washington, DC) 110, 6891 (2010).
- [16] T. Rangel, K. Berland, S. Sharifzadeh, F. Brown-Altvater, K. Lee, P. Hyldgaard, L. Kronik, and J. B. Neaton, Phys. Rev. B 93, 115206 (2016).
- [17] S. Refaely-Abramson, F. H. da Jornada, S. G. Louie, and J. B. Neaton, Phys. Rev. Lett. **119**, 267401 (2017).
- [18] C. D. Spataru, S. Ismail-Beigi, R. B. Capaz, and S. G. Louie, Phys. Rev. Lett. 95, 247402 (2005).

neering Division, under Contract No. DE-AC02-05CH11231, as part of the Computational Materials Sciences Program, which provided theoretical formulation, algorithm, and code development. D.Y.Q. acknowledges resources from a user project at the Molecular Foundry supported by the Office of Science, Office of Basic Energy Sciences, of the U.S. Department of Energy under the same contract number, which provided computational resources for the calculation on layered materials. This research used resources of the National Energy Research Scientific Computing Center (NERSC), a DOE Office of Science User Facility supported by the Office of Science of the U.S. Department of Energy under Contract

[19] M. Palummo, M. Bernardi, and J. C. Grossman, Nano Lett. 15, 2794 (2015).

No. DE-AC02-05CH11231, which provided computational

- [20] D. Y. Qiu, T. Cao, and S. G. Louie, Phys. Rev. Lett. 115, 176801 (2015).
- [21] L. X. Benedict, Phys. Rev. B 66, 193105 (2002).

resources for the calculation on benzene.

- [22] L. X. Benedict, A. Puzder, A. J. Williamson, J. C. Grossman, G. Galli, J. E. Klepeis, J.-Y. Raty, and O. Pankratov, Phys. Rev. B 68, 085310 (2003).
- [23] A. Franceschetti, L. W. Wang, H. Fu, and A. Zunger, Phys. Rev. B 58, R13367(R) (1998).
- [24] D. O. Demchenko and L.-W. Wang, Phys. Rev. B 73, 155326 (2006).
- [25] C. D. Spataru, Phys. Rev. B 88, 125412 (2013).
- [26] T. Deilmann and K. S. Thygesen, Phys. Rev. B 99, 045133 (2019).
- [27] G. Kotliar, S. Y. Savrasov, K. Haule, V. S. Oudovenko, O. Parcollet, and C. A. Marianetti, Rev. Mod. Phys. 78, 865 (2006).
- [28] F. Aryasetiawan, K. Karlsson, O. Jepsen, and U. Schönberger, Phys. Rev. B 74, 125106 (2006).
- [29] D. Zgid and E. Gull, New J. Phys. 19, 023047 (2017).
- [30] M. Dvorak, D. Golze, and P. Rinke, Phys. Rev. Materals 3, 070801 (2019).
- [31] M. Dvorak and P. Rinke, Phys. Rev. B 99, 115134 (2019).
- [32] A. Fetter and J. Walecka, *Quantum Theory of Many-Particle Systems*, Dover Books on Physics (Dover, Mineola, NY, 2003).
- [33] P.-O. Löwdin, Int. J. Quantum Chem. **21**, 69 (1982).
- [34] D. Rocca, D. Lu, and G. Galli, J. Chem. Phys. 133, 164109 (2010).
- [35] D. Rocca, Y. Ping, R. Gebauer, and G. Galli, Phys. Rev. B 85, 045116 (2012).
- [36] M. S. Hybertsen and S. G. Louie, Phys. Rev. B 35, 5585 (1987).
- [37] M. Rohlfing and S. G. Louie, Phys. Rev. B 62, 4927 (2000).
- [38] J. Deslippe, G. Samsonidze, D. Strubbe, M. Jain, M. L. Cohen, and S. G. Louie, Comput. Phys. Commun. 183, 1269 (2012).
- [39] P. Hohenberg and W. Kohn, Phys. Rev. [Sect.] B 136, B864 (1964).
- [40] W. Kohn and L. J. Sham, Phys. Rev. [Sect.] A 140, A1133 (1965).
- [41] J. P. Perdew and A. Zunger, Phys. Rev. B 23, 5048 (1981).
- [42] P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, 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 *et al.*, J. Phys.: Condens. Matter **21**, 395502 (2009).

- [43] M. S. Hybertsen and S. G. Louie, Phys. Rev. B 34, 5390 (1986).
- [44] J. Deslippe, G. Samsonidze, M. Jain, M. L. Cohen, and S. G. Louie, Phys. Rev. B. 87, 165124 (2013).
- [45] C. Sgiarovello, M. Peressi, and R. Resta, Phys. Rev. B 64, 115202 (2001).