DMRG on top of plane-wave Kohn-Sham orbitals: case study of defected boron nitride

Gergely Barcza,1,2,4 Viktor Ivády,1,3 Tibor Szilvási,4 Márton Vörös,5 Libor Veis,2 Ádám Gali,1,6 and Örs Legeza3

1Wigner Research Centre for Physics, Budapest, PO. Box 49, H-1525, Hungary
2J. Heyrovsky Institute of Physical Chemistry, Czech Academy of Sciences, Prague, CZ-18223, Czechia
3Department of Physics, Chemistry and Biology, Linköping University, Linköping, SE-581 83, Sweden
4Department of Chemical and Biological Engineering, University of Wisconsin-Madison, Madison, WI-53706, USA
5Material Sciences Division, Argonne National Laboratory, Lemont, IL-60439, USA
6Department of Atomic Physics, Budapest University of Technology and Economics, Budapest, H-1111, Hungary

(Dated: June 9, 2020)

In this paper, we analyze the numerical aspects of the inherently multi-reference density matrix renormalization group (DMRG) calculations on top of the periodic Kohn-Sham density functional theory (DFT) using the complete active space (CAS) approach. Following the technical outline related to the computation of the Hamiltonian matrix elements and to the construction of the active space, we illustrate the potential of the framework by studying the vertical many-body energy spectrum of hexagonal boron nitride (hBN) nano-flakes embedding a single boron vacancy point defect with prominent multi-reference character. We investigate the consistency of the DMRG energy spectrum from the perspective of sample size, basis size, and active space selection protocol. Results obtained from standard quantum chemical atom-centered basis calculations and plane-wave based counterparts show excellent agreement. Furthermore, we also discuss the spectrum of the periodic sheet which is in good agreement with extrapolated data of finite clusters. These results pave the way toward applying DMRG method in extended correlated solid state systems, such as point qubit in wide band gap semiconductors.

I. INTRODUCTION

By discovering a plethora of color centres in hexagonal boron nitride (hBN) with various point defect complexes in the last years,1–7 hBN is becoming a focus of interest. In particular, additionally to favourable optical properties, point defects in hBN with spin could possibly implement quantum bits controllable by optically detected magnetic resonance which signal has been recently reported at room temperature.8,9

The standard theoretical approach to study such defects is based on first-principles methods10–25 in particular on the Kohn-Sham density functional theory (KS-DFT),25 which established itself as one of the most prevalent numerical methods both in material science and theoretical chemistry offering balanced compromise between accuracy and computational costs. Despite its immense success attained in the field of weakly correlated systems, it is challenged to describe accurately quasi-degenerate electronic structures exhibiting strong correlations.24

To overcome the limitations of the standard DFT level of theory, various distinct methods have been proposed.25–34 Nevertheless, the common aspect of all these post-DFT approaches is the application of a numerical method which is dedicated to handle the static correlations of the problem at some point of the theory. Besides configuration interaction based methods and variants of the complete active space self-consistent field approach, e.g., Refs. [35–42], density matrix renormalization group (DMRG) procedure is an ideal candidate for treating strong correlations.

DMRG was originally developed to describe one-dimensional quantum models in solid state physics with local interactions.43,44 The method was later generalized to treat long-range interactions found in momentum space representation of lattice models45–47 and in ab initio quantum chemistry.38–50 The success of these developments relies on the efficient factorization of interactions45,48 and the optimization of the DMRG network topologies based on concepts of quantum information theory47,51–53 leading to tremendous reduction in computational costs. The underlying mathematical framework, however, is not restricted to models studied in condensed matter physics54,55 or applications to molecular clusters53,56,57 but among many others it can be also used to study nuclear shell models,58–61 particles in confined potential62–64 or problems in the relativistic domain.65–67 Therefore, mapping a physical problem to a proper model together with an optimal choice of basis could pave the road for DMRG applications in a broad range of disciplines which could surpass conventional methods.

In this work, we present the application of DMRG in a novel direction and use it as a post-DFT approach governed by the complete active space (CAS) protocol.68,69 The discussed approach focuses on the accurate description of the static correlations treating the rest of the electronic structure on effective one-electron level. The method, denoted as DFT-CAS-DMRG in the following, has been applied on top of atomic KS orbitals.70,71 Here, we apply the theoretical framework on top of both atomic and periodic KS orbitals to predict the main features of the low-lying many-body spectrum of molecular and periodic ab initio systems. In fact, quite recently, we have applied this post-DFT approach to the negatively charged
boron vacancy (VB) in single layer hBN confirming the experimental expectations on decay routes.

As prequel of our former study, now we turn our focus to the technical aspects of this post-DFT approach adapted both to KS atomic and periodic orbitals. Particularly, we investigate the vertical many-body energy spectrum of a series of finite VB-hBN samples terminated with hydrogen atoms. As part of the study, to the best of our knowledge, we benchmark ab initio DMRG results on top of plane-wave based Kohn-Sham orbitals against the vertical spectrum computed in localized basis for the first time. Furthermore, tendencies observed on flakes are set against the results obtained for explicit periodic sheet.

The paper is structured as follows. In section II, we introduce shortly the DMRG method and also discuss the technical details of the Hamiltonian matrix construction and of the active space selection. In section III, we describe the main features of the studied molecular systems. In section IV, the numerical aspects of the simulations are summarized. In section V, the results are discussed. Finally, conclusion is provided in section VI.

II. METHODS

A. Quantum chemical Hamiltonian and molecular orbitals

In this study, the molecular system of \( N \) electrons with coordinate \( \{ r_i \} \) is investigated in the presence of nuclei of atomic number \( \{ Z_\mu \} \) with fixed position \( \{ d_\mu \} \), i.e., in the Born-Oppenheimer approximation. Considering the Coulomb interaction between charged particles, the corresponding time-independent Schrödinger equation,

\[
\hat{H} \Psi = \left[ -\sum_i^N \left( \frac{1}{2} \nabla_i^2 + \sum_\mu Z_\mu \frac{1}{|r_i - d_\mu|} \right) + \sum_{i<j}^N \frac{1}{|r_i - r_j|} + \sum_{\mu<\nu} \frac{Z_\mu Z_\nu}{|d_\mu - d_\nu|} \right] \Psi(r_1, \ldots, r_N) = E \Psi(r_1, \ldots, r_N), \tag{1}
\]

is to be solved to obtain stationary electronic wave function \( \Psi(r_1, \ldots, r_N) \) with energy \( E \). Hamiltonian (1) measured in Hartree atomic units describes the kinetic energy of the electrons, their potential energy induced by the charged nuclei in addition to the electron-electron Coulomb interaction. Owing to the two-body interactions, the many-electron equation, Eq. (1), is not separable to a system of independent single-electron equations. An approximate mean field level solution can be provided by the variational Hartree-Fock method which treats Hartree and exchange interactions exactly but completely neglects correlation effects. Based on the Hartree-Fock solution, various methods were developed to recover the correlation energy which methods may show limitations on accessible system size or accuracy.

The density functional theory (DFT) can provide a viable compromise by mapping the original problem of interacting electrons to a gas of non-interacting particles in an effective potential which also includes the relevant effects of the electron-electron interactions in addition to the external potentials. In standard DFT procedure the energy is minimized with respect to the electron density obtained iteratively from the Kohn-Sham orbitals, \( \{ \psi_j(r) \} \). Whereas the Hartree potential is treated exactly in the DFT method, the accurate description of the many-orbital interactions, encoded in the exchange-correlation energy functional, is the actual challenge of the approach. Various approximations for the functional had been suggested that were optimized in given condensed matter systems but the method became only a standard in computational chemistry with the advent of hybrid functionals which include a portion of the exact exchange energy of the Hartree-Fock theory to give a more precise characterization of exchange effects.

The method has numerous advantages and became extremely popular for treating weakly correlated systems in the last decades. Notwithstanding the enormous success and broad applicability of the approach, one may have to go beyond the pure DFT approach to provide an accurate description of strongly-correlated electronic states. In the following, we summarize the simplest such possible procedure applicable on one-electron level theory based on the complete active space method.

B. Matrix elements of the ab-initio Hamiltonian

Considering large supercells, the electronic bands become flat, i.e., dispersionless in the reduced Brillouin zone. Furthermore, band states of \( \Gamma \)-point, which is the center of the Brillouin zone with zero momentum, are purely real-valued. Consequently, \( \Gamma \)-point states of large supercells can be regarded as molecular orbitals of solid state systems. Hence, similarly to molecular orbitals obtained by atomic codes, electronic states computed by periodic programs can also be employed to construct the corresponding ab initio Hamiltonian. The following discussion summarizes the computation of the Hamiltonian’s matrix elements of spin restricted molecular orbitals and \( \Gamma \)-point only periodic states. In the following, these localized \( \Gamma \)-point periodic states are also referred to as orbitals. For the sake of simplicity, spin-restricted orbital set is considered but the concept can be readily

\[
H \Psi = \left[ -\sum_i^N \left( \frac{1}{2} \nabla_i^2 + \sum_\mu Z_\mu \frac{1}{|r_i - d_\mu|} \right) + \sum_{i<j}^N \frac{1}{|r_i - r_j|} + \sum_{\mu<\nu} \frac{Z_\mu Z_\nu}{|d_\mu - d_\nu|} \right] \Psi(r_1, \ldots, r_N) = E \Psi(r_1, \ldots, r_N), \tag{1}
\]
generalized.

In terms of operators $\hat{a}_{i\sigma}^\dagger$ and $\hat{a}_{i\sigma}$, which creates and annihilates electron with $\sigma$ spin projection in orbital $i$, the second quantized representation of the ab initio Hamiltonian operator, Eq. (1), is parameterized as

$$\hat{H} = \sum_{ij,\sigma} t_{ij} \hat{a}_{i\sigma}^\dagger \hat{a}_{j\sigma} + \frac{1}{2} \sum_{ijkl,\sigma\sigma'} V_{ijkl} \hat{a}_{i\sigma}^\dagger \hat{a}_{j\sigma}^\dagger \hat{a}_{k\sigma'} \hat{a}_{l\sigma'} + E^{\text{muc}}$$

with system specific one- and two-electron integrals, $t_{ij}$, and $V_{ijkl}$, respectively. Assuming all-electron description, the former integrals expand the kinetic contributions and the electron-nuclear attractions in the basis of electronic orbitals, $\{\psi_j(r)\}$,

$$t_{ij} = -\int \psi_i^*(r) \left( \frac{1}{2} \nabla^2 + \sum_{\mu} \frac{Z_\mu}{|r - d_\mu|} \right) \psi_j(r) dr$$

with nuclei of atomic number $Z_\mu$ at fixed coordinate $d_\mu$. Note that applying pseudopotentials, which is optional in atomic codes, but rather mandatory in plane-wave based calculations, the nuclei and the frozen core electrons are described by non-local potential functions. When states either in the band gap or close to it are considered, the correlation effects with core electron are negligible and the pseudopotential approximation is adequate. As a consequence of the approach, i.e., by reducing the number of electrons and making the core potential softer, plane-wave calculations become considerably less demanding.

The interelectronic repulsions yield two-electron integrals,

$$V_{ijkl} = \int \frac{\psi_i^*(r) \psi_j^*(r') \psi_k(r') \psi_l(r)}{|r - r'|} dr dr',$$

whereas the nuclear-nuclear repulsion gives a constant energy shift,

$$E^{\text{muc}} = \sum_{\mu < \nu} \frac{Z_\mu Z_\nu}{|d_\mu - d_\nu|}.$$  

\textit{Evaluation of the two-electron matrix elements}

The definition of the two-electron integrals (4) prescribes an expensive six dimensional real space integral. It is easy to show that the matrix elements can be equivalently evaluated by a three dimensional integral in reciprocal space $q$,

$$V_{ijkl} = \int \frac{\rho_H(q) \rho_{ijkl}(-q)}{|q|^2} dq$$

using auxiliary density operators,

$$\rho_H(q) = \int \psi_i^*(r) \psi_l(r) e^{iqr} dr.$$  

We implemented the computation of Hamiltonian matrix elements within the plane-wave based Quantum Espresso suite\(^{74,75}\) adopting the momentum space representation, Eq. (6), where the evaluation reduces to two wave function multiplications and Fourier transformation besides a threefold integration. Restricting ourselves to the usage of norm-conserving wave functions, we do not face the problem of augmentation charges observed in systems with ultrasoft pseudopotentials. The integrable divergence for $|q| \to 0$ can be analytically treated using Gygi’s approach.\(^{27}\)

The number of two-electron matrix elements to be evaluated scales as quartic of the number of active orbitals thus their effective computation is the most critical part of constructing the Hamiltonian matrix. In order to optimize performance, balancing between memory size and numerical efforts, portions of the computationally expensive auxiliary operators, $\rho_H(q)$, are cached during the two-electron matrix evaluation.

\section{DMRG}

In the following, the basic concept of the DMRG method is summarized, whereas interested reader finds thorough overviews of DMRG and related numerical approaches in the context quantum chemistry (see Refs. \cite{53 and 56}).

Many-body wave function $|\Psi\rangle$, i.e., an eigenstate of Eq. (2), expanded in the space of $L$ spatial-orbitals reads as

$$|\Psi\rangle = \sum_{\{n\}} C^{(n)} \prod_{i=1}^{L} (\hat{a}_{i\uparrow}^\dagger)^{n_{i\uparrow}} (\hat{a}_{i\downarrow}^\dagger)^{n_{i\downarrow}} |0\rangle,$$

with notation $n = (n_{1\uparrow} n_{1\downarrow} n_{2\uparrow} n_{2\downarrow} n_{3\uparrow} n_{3\downarrow} \cdots n_{L\uparrow} n_{L\downarrow})$ where $n_{i\sigma} \in \{0,1\}$. The components of the state specific $C$ tensor increase exponentially with system size $L$ scaling as $2^{2L}$. Hence, the exact solution of problem (2) on current classical machines is generally limited to systems represented with a dozen of orbitals.

It is shown\(^{78}\) that tensor $C$ is factorizable to a set of numerically more manageable matrices using the equivalent matrix product state (MPS) form of it, i.e.,

$$C^{(n)} = \prod_{i=1}^{L} A_i^{(n_{i\uparrow} n_{i\downarrow})}.$$

The dimension of the $A_i$ matrix grows towards the center of the MPS chain as $\dim(A_i) = [4^{i-1}, 4^i]$ for $i \leq L/2$ and $\dim(A_i) = [4^{L-i+1}, 4^{L-i}]$ otherwise (for even $L$). Consequently, recovering exact MPS representation of tensor $C$, the overall computational cost of the MPS matrices also grows exponentially just as the original tensor representation. To overcome this limit, the DMRG approach provides an approximate description of the $C$ tensor in terms of optimized matrices, $A_i^{\text{DMRG}}$, truncated to a fixed manageable bond dimension, $M$, i.e.,
\[ \dim(A_{\text{DMRG}}) \leq [M,M] \]. Increasing \( M \), the precision of the approximation is well controlled approaching variationally the exact solution. In the DMRG protocol the MPS matrices are locally optimized and truncated by minimizing the discarded entanglement between the left and right neighboring blocks of the MPS chain, obtained from the reduced density matrix of the block. The algorithm iterates through the MPS chain in sequential order back and forth until reaching convergence. For a detailed tutorial about the DMRG approach in MPS formalism, we refer to Ref. \[55\].

Compared to wave functions obtained from typical quantum chemical approaches, the DMRG method parameterizes the eigenstates in terms of local variational objects instead of excitations of a reference configuration. Consequently, DMRG can precisely describe active spaces up to 40-80 orbitals due to the implicitly polynomial scaling of computational costs for gapped non-critical systems. Most typically the method is applied to obtain the ground state properties but it can be used to describe not only the vertical but also the relaxed low-lying electronic excitations as well.\[79,80\]

In the following, we compute the vertical excitation spectrum on the relaxed ground state geometry. In the DMRG truncation procedure, the reduced density matrix of the blocks is formed of the equally weighted linear combination of all target states.

### D. Complete active space (CAS) method

The investigated systems, consisting of dozens of atoms, is described by hundreds of Kohn-Sham orbitals which cannot be directly treated by the DMRG owing to the computational costs. Therefore, an optimal selection of orbitals with tractable size is needed which are responsible for the strong static correlations. The complete active space (CAS) scheme\[68,69\] classifies the set of orbitals to three categories, i.e., the so called core and virtual orbitals are frozen to the mean field level and filled with two and zero electrons, respectively. The third class comprises of the so called active orbitals which are populated with the rest of electrons minimizing the energy.

Accordingly, the virtual orbitals does not play any role in the corresponding CAS Hamiltonian whereas the core electrons affect the electrons of the active space through the Coulomb interactions, i.e., the Hamiltonian of the active space reads

\[ \hat{H}^{\text{CAS}} = E^{\text{num}} + E^{\text{core}} + \sum_{ij,\sigma} t_{ij}^{\text{CAS}} \hat{a}_{i\sigma}^{\dagger} \hat{a}_{j\sigma} + \frac{1}{2} \sum_{ijkl,\sigma\sigma'} V_{ijkl} \hat{a}_{i\sigma}^{\dagger} \hat{a}_{j\sigma'}^{\dagger} \hat{a}_{k\sigma'} \hat{a}_{l\sigma} \]  

(10)

with the previously defined \( V_{ijkl} \) integrals restricted to the active orbital set. The one-electron integrals of the CAS space, \( t_{ij}^{\text{CAS}} \), describes not only the kinetic energy of the active electrons and their attraction to nuclei but also their interaction with the core electrons. Describing the active electrons with the DMRG method, which treats exactly the electron exchange, the one-electron interactions are written as

\[ t_{ij}^{\text{CAS}} = t_{ij} + \frac{1}{2} \sum_{c} (2V_{iccj} - V_{icjc}) \]  

(11)

to treat the Coulombic effects of the frozen electrons on the active orbitals. Here, the summation runs only on the indices of the core orbitals. Finally, the additional energy contribution of the core electrons is summed up in term \( E^{\text{core}} \), i.e.,

\[ E^{\text{core}} = 2 \sum_{c} t_{cc} + \sum_{cc'} (2V_{cc'cc'c} - V_{cc'cc'}). \]  

(12)

Including more and more orbitals in the active space, the corresponding CAS ground state energy gets lower and lower consistently, i.e., enlarging the active space class the energy of the CAS approaches variationally the ground state energy of the original problem.

In practice, the active space is restricted to the most important orbitals with fractional occupational numbers. Even though the method has limitations to provide correct description of dynamical correlations using relatively small active space, it captures static correlations with high accuracy providing valuable insight into the low-lying energy spectrum and the essential structure and symmetry properties of the corresponding electronic eigenstates.

Note also that, contrary to alternative post-DFT methods,\[81\] the CAS Hamiltonian (11) to be investigated does not include the Kohn Sham energies explicitly but only the Kohn Sham orbitals by construction. Also, the absolute energies of the states computed from the CAS Hamiltonian are not trivially comparable with counterparts obtained on the DFT level of theory due to the different description of the exchange and correlation effects.

### E. Protocols for selecting the active space orbitals

The studied systems are described by hundreds of orbitals, which limits the applicability of sophisticated CAS selection schemes based on large-scale post-self-consistent field calculations.\[88,82\] Nevertheless, in the studied defect systems, the low-lying many-body energy spectrum is expected to be conceptually described by the localized defect orbitals (see Fig. 1) and their interaction with the hosting environment, i.e., orbitals with large lobes around the central atoms can potentially interfere with the in-gap orbitals. Therefore, in the following, we apply two approaches to define the active space: i) we select canonical orbitals with prominent localization at the core of the defect, ii) we construct the active space based on energy window around the valence band maximum (VBM) level to verify the applicability of the CAS selection based on orbital localization.
It is to be noted that we do not localize the canonical orbital set in order to preserve their point group symmetry thus providing the possibility of studying the symmetry properties of the many-body states. In practice, such orbitals are taken into account whose degree of localization on the central part of the defect yield a critical value according to their orbital volumetric data or (in case of availability) to their projection on atomic basis.

Comparing the many-body excited state spectrum yielded from the two CAS selection protocols, we find that the structure of the spectra is essentially identical up to some 0.001–0.2 eV shifts in energy as discussed in section V B, i.e., the most important orbitals are the localized in-gap orbitals incorporated in both of the applied protocols. Nevertheless, obtaining slightly lower absolute energies with the applied variational computational procedure indicates that the orbital selection based on localization is preferable to the one based on KS energy. Correspondingly, results discussed in section V C are based on DMRG calculations performed on CAS which is selected according to orbital localization.

III. INVESTIGATED SYSTEMS

Defect-free hexagonal boron nitride monolayer has been investigated numerically on the level of ab initio dynamical mean-field theory using crystal atomic orbitals. The computed direct and indirect band gaps are of around 10 eV. 83

Various defects embedded in hBN with potential in technology have been recently investigated, for a current review see Ref. [84]. In this work, we study hBN nanoflakes and sheet hosting a negatively charged boron vacancy (VB-hBN). The defect system of D3h symmetry exhibits spin triplet ground state which can be understood as following. The in-plane dangling bonds and the out-of-plane p_z orbitals of the three neighbor nitrogen atoms, located at the core of the defect, provide six single-particle defect states. The two non-degenerate a and the two degenerate e defect states plotted in Fig. 1 are occupied with ten electrons in the negatively charged system. The highest-lying e dangling bond becomes half occupied yielding spin-triplet ground state. Analysing the single-electron spectrum of VB-hBN also reveals that these localized in-gap orbitals are close in energy. 19

Therefore, the accurate description of the strongly-correlated excited states necessitates the application of multi-reference methods. In our recent work, 72 we provided a detailed description of magneto-optical properties of the system and also studied its many-body electronic spectrum applying the DFT-CAS-DMRG method. In this work, we turn our focus on the computational aspects of the DFT-CAS-DMRG approach illustrated on VB-hBN samples.

IV. COMPUTATIONAL DETAILS

We study two-dimensional planar VB-hBN flakes of various sizes, i.e., investigating clusters of increasing size is necessary to capture long-range correlations and to minimize the finite-size effects of the terminating hydrogen atoms. In particular, we study molecules B4N3H9, B18N15H15, B36N30H31, B46N50H27, B50N50H31. In addition, we investigate a periodic sheet of a 182 atom supercell which is large enough to minimize the interference between neighboring impurity sites and to reach convergence even restricting to Γ-point-only description.

The hosting molecular flakes embed the vacancy in the center of the model so that the relaxed lattice structures exhibit D3h symmetry. The relaxed molecular geometries of the charged defects are obtained on DFT-PBE level of theory with a homogeneous compensating background charge using VASP package. 86 The finite flakes are embedded in a cubic supercell of 30 Å, corresponding, 30 Å vacuum size is applied in the perpendicular direction of the periodic sheet.

The electronic structure of the molecules is described in terms of spin restricted Kohn-Sham DFT orbitals using PBE functional. Self-consistent field (SCF) calculations are performed using both atomic basis based quantum chemical program suite ORCA 87 and plane-wave based Quantum Espresso (QE) 75 with norm-conserving pseudo potentials. The actual capabilities of the periodic code...
FIG. 2. Snippet of KS energy spectrum close to the VBM, for VB-hBN model with 18 boron atoms. DFT calculations on the DFT-PBE level of theory are performed using plane-wave basis with kinetic energy cutoff of the wave function at $\epsilon = 400, 500, 600, 700, 800$ and $900$ eV. Valence, virtual, doubly/partially occupied in-gap orbitals are visualized by color black, blue, gray and red, respectively.

FIG. 3. Vertical energy spectrum relative to the ground state energy for model $B_{18}N_{18}H_{15}$. The active space is selected according to KS energy. Orbitals are expanded in PW basis with various energy cut-off $\epsilon$. For better visibility spin-triplet and spin-singlet energy levels are contrasted using solid green and dashed purple line, respectively.

To demonstrate the robustness of the plane-wave based calculations, we investigate both the Kohn-Sham single-electron and the resulted vertical many-body excitation spectrum of the model with 18 boron atoms in terms of the energy cutoff, which are plotted in Figs. 2 and 3, respectively. The kinetic energy cutoff of the wave function is varied in the $\epsilon = 400 \rightarrow 900$ eV range whereas the cutoff for charge density and potential is fixed as $4 \epsilon$ in agreement with the default QE parametrization. The details of the KS spectrum around the VBM are presented in Fig. 2 where the energy is measured relative to the lowest lying defect state. We find that the energy spectrum is rather independent of the applied energy cut-off which also illustrates the quick convergence of the DFT calculations. The observed small fluctuations are acknowledged to the fact that $\epsilon$ is increased in steps of 100 and not in steps of the lowest $\epsilon = 400$ eV value. Correspondingly, due to the convergence of the KS orbitals observed already for surprisingly low cut-offs, the orbital-dependent DMRG results found in Fig. 3 vary marginally for $\epsilon > 400$ eV. Therefore, in the following, according to the observed quick convergence in cutoff energy, PW results for $\epsilon = 700$ eV are used as reference data.

We also perform a series of DFT calculations for the model with 18 borons applying various complexity of atomic basis sets, i.e., STO-3G, 6-31G, cc-pVDZ, cc-pVTZ, cc-pVQZ, aug-cc-pVDZ. The corresponding Kohn-Sham one-electron spectra and, as reference, the result obtained in the plane-wave basis (PW) with energy cutoff of $\epsilon = 700$ eV are presented in Fig. 4. Interestingly, whereas the details of the spectrum show some sensitivity to the applied basis, the energy gaps of the localized state are rather independent of the chosen atomic basis to be described, DMRG calculations are performed for target states with total spin 0 and 1. In the DMRG truncation procedure, the density matrix is constructed of the equally weighted linear combination of all target states. As discussed in Ref. [53], the convergence is sped up by initialization procedures inspired by quantum information theoretical considerations and the accuracy of the simulations is controlled by the dynamic block state selection approach. In the calculations, the quantum information loss is kept below threshold value $\chi = 10^{-5}$.

Note that the direct comparison with experimental data is not straightforward as the important structural relaxation effects are not incorporated in the current level of theory. Furthermore, it is also to be kept in mind that the neglected orbital relaxation effects as well as the accurate treatment of dynamical screening effects decrease the presented vertical excitation energies.

V. RESULTS
A. Energy spectrum of $B_{18}N_{18}H_{15}$ obtained in localized and in plane-wave basis sets

To demonstrate the robustness of the plane-wave based calculations, we investigate both the Kohn-Sham single-electron and the resulted vertical many-body excitation spectrum of the model with 18 boron atoms in terms of the energy cutoff, which are plotted in Figs. 2 and 3, respectively. The kinetic energy cutoff of the wave function is varied in the $\epsilon = 400 \rightarrow 900$ eV range whereas the cutoff for charge density and potential is fixed as $4 \epsilon$ in agreement with the default QE parametrization. The details of the KS spectrum around the VBM are presented in Fig. 2 where the energy is measured relative to the lowest lying defect state. We find that the energy spectrum is rather independent of the applied energy cut-off which also illustrates the quick convergence of the DFT calculations. The observed small fluctuations are acknowledged to the fact that $\epsilon$ is increased in steps of 100 and not in steps of the lowest $\epsilon = 400$ eV value. Correspondingly, due to the convergence of the KS orbitals observed already for surprisingly low cut-offs, the orbital-dependent DMRG results found in Fig. 3 vary marginally for $\epsilon > 400$ eV. Therefore, in the following, according to the observed quick convergence in cutoff energy, PW results for $\epsilon = 700$ eV are used as reference data.

We also perform a series of DFT calculations for the model with 18 borons applying various complexity of atomic basis sets, i.e., STO-3G, 6-31G, cc-pVDZ, cc-pVTZ, cc-pVQZ, aug-cc-pVDZ. The corresponding Kohn-Sham one-electron spectra and, as reference, the result obtained in the plane-wave basis (PW) with energy cutoff of $\epsilon = 700$ eV are presented in Fig. 4. Interestingly, whereas the details of the spectrum show some sensitivity to the applied basis, the energy gaps of the localized state are rather independent of the chosen atomic basis...
In particular, the ORCA program describes the spin the atomic and by the plane-wave basis computations. The distinct treatment of unpaired electrons applied by the PW basis applying energy cutoff at \( \epsilon = 700 \text{ eV} \). Color key as in Fig. 2.

The many-electron DMRG energy spectrum resulted in the CAS constructed of the series of KS-DFT-SCF orbitals is plotted in Fig. 5. The active space for all the testing CAS calculations were selected according to the one-electron energies correlating 32 electrons in 21 orbitals. The overall structure of the many-electron spectrum obtained in reasonably large basis sets, i.e., cc-pVDZ and beyond, is rather independent of the applied basis even for excitations of \( 4 - 5 \text{ eV} \). Using more and more sophisticated atomic basis, the absolute energy of each excitation lowers but with varying tendency. Nevertheless, approaching the completeness of the basis, i.e., using correlation consistent and PW basis with large cutoff \( \epsilon \), the spectrum becomes consistent with slight variations yielding \( \sim 90\% \) agreement on the many-body excitation energies. It is also notable that the singlet-triplet gap, which involves predominantly spin excitation of the partially filled defect \( e'' \) level, is predicted surprisingly well, i.e., with 10\% accuracy, even on the minimal STO-3G level of theory. Based on these results, cc-pVDZ atomic basis is an optimal choice for reasonably accurate non-periodic DFT-CAS-DMRG calculations.

The observed discrepancies in the spectrum obtained for various orbital sets reflect not only the distinct character of the underlying incomplete basis sets but also the limitations of the CAS based description. In particular, as valence orbitals can be very close in energy with accidental ordering as observed in Fig. 4, the selection of a dozen of valence orbitals for the CAS can be biased which could be cured with the study of a much larger and optimally constructed active space.

### B. Sensitivity of active space selection

Having a CAS based description in our hands, careful selection of active orbitals is crucial. In the following CAS selection is investigated from the perspective of active space size and selection protocols on the example of \( \text{B}_{36}\text{N}_{36}\text{H}_{21} \) on cc-pVDZ level of theory.

First, the large gap, \( \Delta \approx 5 \text{ eV} \), also observed in Fig. 4 and the partial occupation of the defect \( e'' \) level suggest that virtual orbitals does not play crucial role in the de-
scription of low-lying electronic states excited with energy of a few eV. Similar conclusion is drawn from the molecular orbitals obtained in PW SCF calculations, i.e., some of the lowest-lying virtuals are out of the plane of the sheet and expand in the vacuum, i.e., the system can be ionized with around energy $\Delta$.

Hence, we typically include only a couple of virtual orbitals in the active space and select active orbitals mostly from the valence band besides the six localized orbitals presented in Fig. 1. As an illustration, we show many-electron energy spectrum obtained for KS-energy-selected active space of various sizes in Fig. 6. Increasing CAS size as more and more correlation effect is possible to be retrieved, consequently every many-electron energy decreases but with a particular rate depending on the structure of the corresponding many-body state. We find that the main features of the spectrum are already captured using the six orbital CAS, nevertheless the energy of some many-body states lower with more than 1 eV including more dynamical correlations by increasing CAS size. The results also show that a CAS of around 30 orbitals is already large enough to describe the low-lying application-relevant excitations.

It is notable that besides the ground state there are some excitations of prominent single-reference character whose energy decreases less dramatically with increasing active space. In particular, the lowest-lying spin-singlet spatial-doublet state is mainly characterized by spin excitations within the degenerate defect $e''$ level as we have shown in Ref. [72]. For smaller active space sizes, i.e. $<30$, the spin-triplet doublet state with single-reference character found around 4 eV describes excitation from the localized $a'_1$, similarly to the singlet state detected around 4.5 eV. Nevertheless, extending the active space further with lower-energy orbitals, which exhibit significant localization around the defect, even these states are shifted by 0.5 eV for active spaces of $\geq 30$ orbitals.

This observation also underlines the detailed analysis of the eigenstates, which indicates that valence orbitals showing larger overlap with the defect orbitals become more correlated in the post-DFT calculations than orbitals far from the center confirming the chemical intuition of CAS selection based on localization. In Fig. 7, we demonstrate the results obtained for 25 active orbitals selected by the two CAS construction schemes, i.e., selecting orbitals based on their KS energy and on their localization. In the first case the occupied orbitals are kept in the active space based solely on their proximity to the Fermi surface. In the latter case the localization filters further the orbitals, i.e., only those orbitals are considered whose atomic basis projection on the first neighbor three nitrogen atoms reaches at least 0.005-0.01. Comparing DMRG spectrum obtained in the two differently constructed active spaces of 25 orbitals, corroborates the expectations, i.e. all the many-electron states get lower in energy with $0.0001 - 0.2$ eV using localization based CAS. In fact, the latter calculation recovers essentially the spectrum obtained for the larger CAS of 30 orbitals selected according to KS energy (see Fig. 6). In larger systems, where the corresponding orbitals are more distributed in space, the CAS selection scheme has even more drastic effect.

C. Large model and single sheet results

The KS-PBE energy levels close to the gap are depicted in Fig. 8 for various model sizes and for periodic sheet with 80 boron atoms. Having observed good convergence on the relatively cheap cc-pVDZ basis set for $B_{18}N_{18}H_{15}$ in Sect. VA, all molecules are treated on the cc-pVDZ level of theory whereas the periodic sample is computed.
FIG. 8. Snippet of the KS energy spectrum of orbitals around the gap on the DFT-PBE level of theory for models with 6, 18, 36, 60, 90 boron atoms and for the periodic sheet with 80 boron atoms computed using cc-pVDZ atomic and plane-wave basis with cut-off at 700 eV, respectively. Color key as in Fig. 2.

FIG. 9. Relative vertical energy spectrum of models with 6, 18, 36, 60, 90 boron atoms and of single layer with 80 borons obtained in localization-selected CAS of 40 orbitals. The models and the sheet are described in cc-pVDZ and PW basis, respectively. Color code as in Fig. 3.

using PW basis with cut-off at $\epsilon = 700$ eV. We find that the structure of the in-gap spectrum is similar for models with 18, 36, 60 and 90 boron atoms, however, the energy gaps are enlarged in smaller models due to the quantum confinement. It is also clearly observed that the confinement affects strongly the typical extended molecular orbitals but hardly the six localized defect states which are coded essentially with color red and gray in Fig. 8. Furthermore, we also find that the doubly occupied energy spectrum of the larger molecules is in good agreement with result obtained on the periodic sheet (also noting differences of partial-filled and virtual energies owing to technical reasons discussed in Sect. V.A.

The yielded vertical many-electron DMRG spectrum obtained for CAS with 40 localized orbitals in the active space is presented in Fig. 9 for the distinct systems. The ground state has single reference nature in each cases, i.e. > 0.95% weight for the leading configuration with half-filled $e$ orbitals is observed, confirming the applicability of DFT as ground-state calculation on VB-hBN systems. Disregarding the smallest model with 6 borons, we find that the overall structure of the vertical many-body spectrum is largely independent of system size. Nevertheless, the actual excitation energies generally increase with system size yielding convergence for model with 60-90 boron atoms for lower-lying excitations. Interestingly, the minimal excitation energy, i.e., singlet-triplet energy gap, is less dependent on the model size. Its stability is the consequence of the corresponding many-body state, which is dominated by the spin flip on the half-filled twofold degenerate $e$ orbitals.

Furthermore, comparing spectra obtained on the periodic sheet with 80 boron atoms and larger flake models in Fig. 9, rather good agreement is observed. In particular, for excitations below 3 eV, which are actually the most relevant for possible quantum bit applications, the agreement is quantitative. Therefore, the results indicate that clusters with 60-90 borons can already reproduce the low-energy features of large periodic sheets, whereas the smaller flakes, especially $B_6N_6H_9$, have limitations to capture the physics of the single layer due to the non-negligible perturbative effect of the terminating hydrogen atoms.

The effect of the active space selection on the electronic spectrum, i.e., the relevance of dynamical correlation is investigated as well. First, in addition to the presented 40 orbital calculations, we performed further DMRG computations on 25 orbitals selected according to localization concluding very similar many-body energy spectrum. In fact, for models with 6-18 borons the variation is insignificant for all many-body states whereas for larger systems the energies shift with up to 5% for the higher excitations. Investigating the spectrum resulted of 10−15 relevant orbitals, more significant deviations of 10−15% are observed in the excitation regime above 3 eV for larger systems. These results suggest that active space of 30-40 orbitals might be already sufficient to provide reasonable description for the low-energy excitations. Nevertheless, the accurate description of the missed portion of dynamical correlations could potentially lower the gaps of larger systems to some extent.

VI. CONCLUSION

We report the first detailed implementation and discussion of DFT-CAS-DMRG method on plane-wave basis. As application of the approach, we investigate the electronic structure of VB-hBN defect systems of current interest, i.e., the sheet with a single defect and its molecular models. Applying our novel computational scheme to construct the CAS Hamiltonian of QE-DFT orbitals, we
find great agreement on the many-electron spectrum with results obtained using standard atomic quantum chemical program suite. We show that flakes with around 60-90 boron atoms are large enough to reflect faithfully the low-lying energy spectrum of the sheet with 80 borons. In agreement with expectations, we confirm that static correlation effects can be attributed mainly to the six localized defect orbitals and using CAS of 30-40 orbitals can provide a reasonable description of the main electronic features of the system. The simulation of the periodic sheet also serves as validation and benchmark of the PW based approach whose actual potential advantage is to be realized in bulk calculations. Given the underlying theoretical background of the presented DFT-CAS-DMRG method, we hypothesize that it might be well-suited for predicting the main features of the low-lying many-body method, we hypothesize that it might be well-suited for predicting the main features of the low-lying many-body spectrum of wide-band-gap semi-conducting defect materials with localized ingap orbitals which class of materials show emerging potentials in various quantum technology applications.

ACKNOWLEDGEMENTS

GB acknowledges the support from the NKFIH PD-17-125261 and the Bolyai Research Scholarship of HAS. VI is supported by the MTA Premium Postdoctoral Research Program and by the Knut and Alice Wallenberg Foundation through WBSQD2 project (Grant No. 2018.0071). LV acknowledges the support from Czech Science Foundation (grant no. 18-18940Y). ØL and AG acknowledge support of the NKFIH through the National Quantum Technology Program (Grant No. 2017-1.2.1-NKP-2017-00001). ØL acknowledges grant NKFIH K120569 and the support of the Alexander von Humboldt Foundation. AG acknowledges the Hungarian NKFIH grants No. KKP129866 of the National Excellence Program of Quantum-coherent materials project. The computations were performed on resources provided by the Wigner RCP and by the Swedish National Infrastructure for Computing (SNIC 2018/3-625 and SNIC 2019/1-11) at the National Supercomputer Centre (NSC).
79. G. Barcza, W. Barford, F. Gebhard, and O. Legeza, Phys. Rev. B 87, 245116 (2013).
80. W. Hu and G. K.-L. Chan, Journal of Chemical Theory and Computation 11, 3000 (2015).
81. M. Bockstedte, F. Schütz, T. Garratt, V. Ivády, and A. Gali, npj Quantum Materials 3, 31 (2018).
82. C. J. Stein and M. Reiher, Journal of Chemical Theory and Computation 12, 1760 (2016).
83. T. Zhu, Z.-H. Cui, and G. K.-L. Chan, Journal of Chemical Theory and Computation 16, 141 (2020).
84. A. Sajid, M. J. Ford, and J. R. Reimers, Reports on Progress in Physics 83, 044501 (2020).
85. J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. 77, 3865 (1996).
86. G. Kresse and J. Hafner, Phys. Rev. B 49, 14251 (1994).
87. F. Neese, WIREs Computational Molecular Science 2, 73 (2012).
88. M. Kallay, P. R. Nagy, D. Mester, Z. Rolik, G. Samu, J. Csonotos, J. Csoka, B. P. Szabo, L. Gyevi-Nagy, B. Hegely, I. Ladjanszki, L. Szegedy, B. Ladoczki, K. Petrov, M. Farkas, P. D. Mezei, and A. Ganyecz, “Mrcc, a quantum chemical program suite,” (2020).
89. M. Kallay, P. R. Nagy, D. Mester, Z. Rolik, G. Samu, J. Csonotos, J. Csoka, B. Szabo, L. Gyevi-Nagy, B. Hegely, I. Ladjanszki, L. Szegedy, B. Ladoczki, K. Petrov, M. Farkas, P. Mezei, and A. Ganyecz, The Journal of Chemical Physics 152, 074107 (2020).
90. Ö. Legeza, L. Veis, and T. Mosoni, “QC-DMRG-Budapest, a program for quantum chemical DMRG calculations,” (2020).