Final Report of Project NKFIH-PD17-125261
Gergely Barcza - principal investigator (PI)

The project entitled "Accurate description of multi-reference electronic structures in point-like defects" was Gergely Barcza's postdoctoral program between 01.09.2017-31.08.2020. In this project, the PI attempted to synthesize the main research topics of two groups of the hosting Wigner Institute by applying ab initio computational methods developed in Örs Legeza's group for the description of defected solids investigated in Ádám Gali's group.

The first year focused on method and code development to perform post-DFT calculations, i.e.,
i) the implementation of the canonical density matrix embedding theory (DMET) [Knizia12] and
ii) the construction of ab initio Hamiltonian of plane-wave Kohn-Sham orbitals of Quantum Espresso (QE) [QE] density functional theory (DFT) program package according to the complete active space (CAS) [Jensen07].
In the project, the DFT-CAS problem is solved using the ab initio density matrix renormalization group (DMRG)[White99] code developed in Legeza's group [QC-DMRG-Budapest] which also hosted the PI.
In the second year, the latter computational scheme, dubbed DFT-CAS-DMRG in the following, was preferred to investigate the many-body energy spectrum of defected boron nitride sheet of current experimental and theoretical interest.
In the third year, further applications of the DFT-CAS-DMRG were initiated and a novel approach was worked out to retrieve a portion of dynamical correlation effects.

Comparing the achieved results with the proposed work plan, a rather technical modification is to be noted, i.e., code developments related to periodic DFT calculation were integrated in the QE package (and not in the VASP package as proposed in the plan). Considering the comparable computational performance of the two codes, the open-source QE was preferred due to its comprehensive user-friendly documentation. The developed computational framework was applied on current problems of budding interest.
Furthermore, during the project period the project's interest in point-defected systems broadened to study model Hamiltonians as well in order to understand the screening effect of impurities better.

In the following, the relevant methodological developments are outlined. Results of the published papers are briefly summarized (with open-accessible links) whereas the manuscript of two closing projects is appended.

Method and code developments

a) Implementation and testing of the density matrix embedding theory (DMET)
In the beginning of the project, the canonical DMET method with self-consistency on the total occupation number [Knizia12] was implemented and integrated with the Budapest DMRG program package. Following several numerical tests on molecules which corroborate more recent findings in the literature [Ye18], it was concluded that the control of accuracy of canonical DMET calculations might be nontrivial in correlated problems of interest. Furthermore, the extension of the concept for the accurate description of excited states is challenging in its standard framework.

As a consequence, the PI initiated the application of the more robust complete active space (CAS) protocol to treat the strongly correlated orbital active space of the problem beyond DFT level of theory.

b) Construction of the plane-wave ab initio Hamiltonian for DFT-CAS-DMRG calculations
In order to perform ab initio DMRG spectrum calculations, as input for the algorithm, the matrix elements of the Hamiltonian of the electronic system expanded in the Kohn-Sham orbitals are needed, i.e., the list of the so-called one- and two-electron integrals. Such functionality is available in standard edition of atomic quantum chemical codes, e.g., MOLPRO and ORCA.
On the contrary, in plane-wave codes, e.g., QE and VASP, which are routinely used to study ab initio solid state problems, it was not publicly available.
As part of the project, a routine package was developed in Fortran90 to export the relevant interaction parameters from the Quantum Espresso DFT-SCF ab initio code [QE].

The features of the current version of the in-house developed interface for constructing Hamiltonian matrix elements are summarized as:
- the ab initio Hamiltonian of the active space is constructed according to the canonical complete active space method (CAS) which treats the effect of the inactive electrons on mean-field level of theory. Nevertheless, alternative scheme for constructing Hamiltonian matrix is also available to avoid double counting effects.
- active space orbitals can be selected according to one-electron energy criteria or to spatial localization
- matrix elements are stored in the FCIDUMP format defined by the MOLPRO program package which is readable by standard wave-function based post-SCF computational codes
- the code is partially parallelized
- the code is restricted to gamma-point only bands
- the code is readily extendable to treat screening effects

Various tests and benchmarks confirmed the accuracy of the implementation, in particular tests were performed on closed- and open-shell molecular problems by comparing the vertical excitation spectrum to results obtained using standard atomic codes. Note that in the first year report the CAS method was referred to as frozen core approximation (FCA).
The in-house code is planned to be shared with the QE community following the cleaning and full parallelization of the code.

The developed interface has been applied to perform DMRG calculations on the generated active space of Kohn-Sham orbitals using the Budapest-DMRG code [QC-DMRG-Budapest] in multiple project discussed below.

c) Formulation and implementation of the RAS-FCI method based on many-body states
Following the mathematical derivation of an alternative formulation of the restricted active space full configuration interaction (RAS-FCI) method [Jensen07], a first pilot implementation of the scheme is provided.
The method solves a list of small CAS spaces defined according to the possible RAS excitations. The obtained many-body states provide an optimized many-body basis for the RAS Hamiltonian to be solved compared to the canonical RAS representation.
The method is implemented up to double excitations in the RAS space.
The extension of alternative RAS protocol for DMRG is planned as part of an subsequent project.

d) Analysis of plane-wave orbitals/bands

A code is developed to analyze properties of orbitals, e.g. symmetries, localization, based on their volumetric data stored in Gaussian cube format. The code might be particularly useful for periodic SCF codes where the orbitals are expanded in plane waves and not in atomic orbitals with specific symmetry.

e) Developments related to the Budapest-DMRG

As part of the Budapest-DMRG package, a routine is developed to extract the weight of the leading configurations of arbitrary DMRG many-body wave function, i.e., providing the configuration interaction expansion of the DMRG states, which was used to reveal the spatial symmetry and the multi-reference character of the vertical excited states in the related research projects.

Published papers on ab initio model of defected systems

The potential of the application of the DFT-CAS-DMRG method on defected semiconducting materials of prominent multi-reference character was presented on defected hexagonal boron nitride (hBN). The details of the implementation of ab initio Hamiltonian constructed in plane-wave basis were discussed and the numerical aspects of the computational framework were illustrated on defected hBN nano-flakes and periodic sheet. In particular, the consistency of the many-body DMRG energy spectrum was presented from the perspective of sample size, basis size, and active space selection protocol.

Spectral results for molecules obtained from standard quantum chemical atom-centered basis calculations and plane-wave based counterparts show excellent agreement. The analysis also indicated that rather large hosting environment is needed to reach convergence in spectral properties. Finally, the spectrum for the periodic slab was found in great agreement with extrapolated data for large finite clusters.

The manuscript submitted to Journal of Chemical Theory and Computation (impact factor 5.3) in July of 2020, has just received two positive referee reports suggesting minor changes. It is notable that the paper appeared on arxiv on June of 2020 has already received 2 independent citations. For more details see https://arxiv.org/abs/2006.04557. Following its acceptance, the published version will be deposited in mtak repository.

The magneto-optical properties of the experimentally relevant two-dimensional solid state defect system, i.e., the optically active boron vacancy in hexagonal boron nitride sheet was investigated.

The project was motivated by the growing interest in two-dimensional defected semiconductors.

In particular, the photoluminescence spectrum investigated by DFT ab initio simulations was found well aligned with recent experimental data.

The in-house developed DFT-CAS-DMRG computational framework was applied to describe the system's low-lying vertical many-body excitation spectrum of strongly correlated character. Furthermore, the spatial and spin symmetry properties of the excited states were also obtained.

These results were instrumental in modeling the corresponding decay routes. The optical controllability of spin polarization of the defect based on the proposed model supports recent experimental results [Gottscholl19] and might pave the way towards the manipulation of qubits implemented by point defects in hBN host.

The manuscript [Ivády2019], where Viktor Ivády and the PI shared first authorship, was published at npj Computational Materials (impact factor 8.6) in April of 2020, it has already collected more than 10 independent citations. For more details see http://real.mtak.hu/id/eprint/100627.

Brief summary of projects related to ab initio problems

In a closing project, the in-house developed DFT-CAS-DMRG approach was applied to simulate the energy spectrum of magnesium-vacancy defect hosted in diamond bulk. The energy of the excited high-spin states as well as their optical oscillator strength were investigated on the level of DFT-CASSCF and DFT-CAS-DMRG. In this project, the effect of the applied exchange functional and the sensitivity of the active space selection is also investigated providing a more consistent picture of the system.
Further details, i.e., a short draft of the project entitled "Highly tunable magneto-optical response from MgV color centers in diamond" is appended to the report.

In a just-started project in collaboration with Viktor Ivády, the in-house developed DFT-CAS-DMRG approach was used to investigate the electronic structure of hBN sheet with double boron vacancy which has ground state with prominent multi-reference character.

As an attempt to overcome the limitations of the DFT-CAS-DMRG approach in capturing dynamical correlation effects, the PI proposed an alternative formulation for restricted active space (RAS) computational scheme expanding the effective Hamiltonian in a particularly constructed many-body basis which represents the low-energy states. The advantage of the proposed formalism compared to the canonical RAS formalism is its efficiently compressed basis representation which leads to a significant reduction in required computational efforts.
Details of the method are provided in the appended draft.

Published papers on model systems with localized impurity

The single impurity Anderson and Kondo models were studied, providing the simplest theoretical model for interacting impurity embedded in host environment, which could pave the way toward understanding realistic defected systems as well.

In particular, various ground-state properties of the single-impurity Anderson model (SIAM) on a tight-binding chain were investigated comparing results obtained by multiple methods (Bethe ansatz, Gutzwiller, magnetic Hartree-Fock, DMRG). The limitation of all approximations was discussed in details. In particular, the DMRG provides excellent data for the ground-state energy and the magnetization for finite external fields compared to exact Bethe ansatz. In strong interaction limit, the amplified finite-size effects challenged the recovery of the exponentially large zero-field susceptibility and the description of the mesoscopically large Kondo screening cloud captured by the Gutzwiller method.
The paper was published in Phys. Rev. B (IF: 3.5).
For more details see http://real.mtak.hu/id/eprint/100627.

The one-dimensional single-impurity Kondo model (SIKM) of magnetic moment impurity interacting with a bath of itinerant electrons was studied in terms of various analytic and numerical methods. Besides the ground-state energy, various correlators, e.g., local spin correlation, impurity-induced magnetization, and corresponding zero-field susceptibilities were investigated. The numerically exact results extrapolated from the DMRG method obtained on finite chains were compared with data from the numerical renormalization group (NRG) method scaled as a function of Wilson cut-off parameter.
The paper published in Phys. Rev. B (IF: 3.5) was honored to be distinguished with the "Editor's suggestion" remark. For more details see http://real.mtak.hu/id/eprint/113911.

Conclusion and perspective

The primary goal of the research project was successfully accomplished yielding publications and manuscripts based on the achieved numerical developments. The application of the in-house developed computational scheme in two currently on-going projects also proves the blossoming collaboration among the local research groups established by the project.

The novel interface for post-DFT calculations is planned to be shared with the QE community as soon as possible.

Owing to the increasing interest in the topic of two-dimensional defected solids, the published papers also induced intensive communication with international research groups.

Furthermore, the PI's recent proposal (FK-20-135496) which is the organic expansion of this closed program was announced to be supported by the funding National Research, Development and Innovation Office in August of 2020.

References

[Knizia12] G. Knizia et al. , Phy. Rev. Lett., 109, 186404 (2012)

[QE] Quantum Espresso. https://www.quantum-espresso.org/

[Jensen07] F. Jensen, "Introduction to Computational Chemistry", 2. ed., publ. John Wiley & Sons. Ltd. (2007)

[White99] White el al., J. Chem. Phys. 110, 4127 (1999)

[QC-DMRG-Budapest] O. Legeza, QC-DMRG-Budapest, a program for quantum chemical DMRG calculations. Copyright 2000–2017, HAS RISSPO Budapest

[Ye18] H.-Z. Ye et al., J. Chem. Phys. 149 194108 (2018)

Highly tunable magneto-optical response from MgV color centers in diamond

Anton Pershin, Gergely Barcza, Ors Legeza, and Adam Gali

Wigner Research Centre for Physics, PO Box 49, H-1525, Budapest, Hungary

The defect quantum bits (qubits) have recently developed into an emerged technology with a broad class of applications ranging from the quantum memory to quantum sensing. Yet, despite their great promise, new types of defects are to be explored when it comes to the large-scale applications (especially, at high temperatures) both in terms of defect types and host materials. In this letter, we examine the potential of magnesium-vacancy in diamond to operate as a qubit by means of evaluating the key electronic and spin properties by applying a robust theoretical analysis. Most shrinkingly, we found that a remarkable electronic structure of MgV enables co-existence of two loosely separated spin-sates with the distinct spectroscopic features, which can be interconverted depending on the temperature, external strain, and defect density. These results show the first example of an unpresidential control over the magneto-optical response from a qubit by modulating the operational conditions.

Control over the formation of defects is vital for all the emerging areas of Physics, Chemistry, and material science, starting with electronics to photocatalysis. Although defects generally tend to deteriorate the material properties, with a rational design, the impurities play a key role in achieving, for instance, high conductivity via doping or in boosting the catalytic activity, as e.g. for water splitting [1] [2]. Moreover, some defects, also known as color centers, develop new states, that are well-separated from the band edges of semi-conductive host materials, and represent a platform for the practical implementation of quantum bits (qubits) [3] [4]. The capability to store and process quantum information with the localized defect spin-states permits in turn realization of various prominent technologies, such as quantum computers, quantum memory, and quantum sensors [5]. Yet, the networks of high-fidelity defect qubits are challenging to achieve in practice, due to the rapid decoherence caused by spin-lattice interactions in the presence of nuclear spins [6] [7]. For the applications in bioimaging and quantum communication, finding an appropriate qubit is even more difficult, since specific emission wavelengths are needed to allow for the optical readout. To this end, the most successful strategy is based on down-converting the signal from the NV(-) center in diamond, even up to the telecommunication wavelengths (1588 nm), but at the expense of the quantum efficiency (which drops to 17 % at best) and a complex optical setup [8]. Therefore, new defect configurations beyond the NV$^-$ center are of high demand and are a subject of extensive current research [9] [10].

For the sake of applications of quantum information theory, photostability of the response is another prerequisite for a successful qubit. Usually, the target charge/spin states, that are suitable for the optical readout, are found within a narrow energy window relative to the position of Fermi level, where the latter is controlled by the amount of doping. However, a new Mg color center is diamond was recently reported by Lühmann et al [11] that by contrast exhibits remarkable photostability. Here, the photoluminescent (PL) measurements show bright and mostly coherent emission with a pronounced zero-phonon line (ZPL) at 557 nm coming from the defect states, and notably, this color center can be implanted with high yield. Most interestingly, the PL signal remains surprisingly stable irrespective of the density

of boron- and phosphorous-dopants. Based on the measured polarization of absorption, the authors assigned the PL feature to a magnesium-vacancy (MgV) complex, however, the charge- and spin-state of the complex was not determined, yet (although likely belonging to either spin-0 or spin-1/2 manifolds, since no optically-detected magnetic resonance was found in the range from 0.5 to 4 GHz [12]).

In this letter, we carry out a robust theoretical investigation to access the role of Mg in the formation of color centers in diamond towards identifying the origin of the photostable experimental PL signal at 557 nm. By going through all the cycles starting with evaluating the formation energies to computing the optical properties, we show that MgV is the most stable simple defect configuration with a photostable -1 charge state and computed PL response in the doublet spin-manifold in an excellent agreement with the experiment. Moreover, we demonstrate that MgV(1-) center possesses an outstanding electronic structure which enables coexistence of several bright transitions, including those at longer wavelengths, that were not experimentally identified, yet. We also show the routes towards stabilization of other emission signals through a rational control of operational and preparation conditions, such as temperature, compressive strain, and defect density. Our findings therefore pave the way towards highly-efficient qubits with the operational wavelengths particularly suitable for the biological applications, that were not reported for color centers in diamond, up to date.

The optimized geometries and electronic structures of four Mg color centers, shown in figure 1, were computed with the HSE06 density functional [13] utilizing the projector-augmented wave method with the kinetic energy cutoff of 400 eV. The calculation for Mg atoms, incorporated into the 512-carbon supercell were performed with VASP package [14]. The zero-phonon line together with the phonon sideband were computed by ΔSCF method. Thermodynamic stability of each defect configuration was evaluated based on the formation energies as a function of Fermi energy (E_F), given as

$$E_{form}^{q}\left(E_F\right)=E_{total}^{q}\left(C512:defect\right)-n_C\mu_C-\mu_{Mg}+q\left(E_F+E_{VBM}\right)+E_{corr}^{q} \quad (1)$$

where E_{total}^{q} is the total energy of the system in a charge state q, μ_C and μ_{Mg} are chemical potentials of C and Mg, respectively, n_C is a number of carbon atoms in the supercell, E_{VBM} is the calculated position of the valence band maximum in diamond. E_{corr}^{q} is the charged defect correction, which was computed by SXDEFECTALIGN code [15]. For the MgV defect, the CASSCF method was further employed to characterize the highly-correlated states based on a cluster of 84 C-atom as implemented in ORCA code [16]. The validity of this model was additionally confirmed by periodic DMRG calculations (see SI for details), that reinforces the robustness of the computed excitation energies.

To provide the theoretical evidences that the experimental PL signal belongs to MgV, we first compared the electronic structures of four simple defect configurations involving Mg, namely, interstitial and substitutional Mg, MgV, and MgV$_2$. From the structural viewpoint, each defect configuration is isomorphic to the respective silicon center: interstitial and substitutional Mg exhibit T$_d$ symmetry, while Mg ion in MgV spans the middle of the cavity formed by two vacancies, hence stabilizing D$_{3d}$ configuration. In turn, MgV$_2$ center essentially mimics the MgV configuration with one additional vacancy attached from side, resulting in the symmetry lowering to C$_1$ instead of (possibly expected) C$_{2v}$. As evident from the HSE electronic structures, shown in figure 1b, in each case, incorporation of Mg into the diamond lattice leads to the development of photoactive localized states inside the band gap. However, with increasing the amount of vacancies, the HOMO-LUMO gap for the most stable charge states (see figure 2) rapidly reduces from 4.46 to 1.5 eV. Note that these values provide a first hint for the lowest transition energy when neglecting the excitonic effects.

Interestingly, irrespective of the charge state of MgV, the difference between the e_u and e_g Kohn-Sham energies (which are responsible for a ZPL of SiV⁻) is much smaller than 2.23 eV where ZPL was experimentally identified. Noteworthily, in the low spin-0 configurations a small energy separation between the frontier orbitals in MgV(0) promotes strong correlation effects as further verified by the DMRG calculations (see SI); however, the resulting opening of the gap is still not sufficient to explain the PL at 2.23 eV. Our findings suggest that for MgV to be the experimentally observed color center, a higher lying electronic transition, likely involving the deeper a_{2u} localized orbital, should give rise the PL signal, at odds with the Kasha's rule. It also appears that the substitutional Mg is another possible source of the green emission. In the following, however, we show that this defect cannot feature the photostability for a large energy range and is also less thermodynamically favorable as compared to MgV.

The formation energies of the four Mg centers in diamond as a function of Fermi energy are shown in figure 2. Here, none of the four centers exhibits one single charge state within the band gap, that would have explained photostability of the ZPL. Therefore, to be hold for a sufficiently large energy window, photostability should come from photoionization, where the laser pulse of 2.35 eV continuously promotes electrons from (to) valence (conduction) band, thereby stabilizing a unique charge state. In terms of the absolute values of the formation energy, MgV is the most stable defect among the four, with the positions of charge transition levels being favorable for the optical stabilization of MgV(-1) center. Note that the formation energy of MgV decreases with accumulating the negative charge that also contributes into stronger emission, as observed experimentally for the n-type doping condition, due to larger defect density. Considering that the formation energy of a single vacancy in diamond is ~7 eV [17] [18], rather surprising stabilization of MgV over substitutional Mg defect points to a weak (repulsive) interaction between the Mg ion and carbon lattice. This is further confirmed by artificially dispositioning Mg and V shown in figure S1 resulting in a stabilization of the adjacent configuration by ~9 eV for each charge state; therefore, thermal annealing ultimately drives the substitutional Mg into the formation of MgV complexes. In turn, photoionization cannot explain the photostability for substitutional Mg: the non-emissive (due to full occupation of the t levels) -2 charge state appears well-above the ionization threshold by ~0.4 eV, while photostability of a single charge state from the remaining +1, 0, and -1 would involve multi-electron transitions, which is rather unlikely.

To proceed, we further focused on the analysis of the electronic structure and spin properties of MgV complexes. In the neutral charge state, MgV develops five localized levels in a proximity of Fermi level, with two pairs of doubly degenerate (occupied e_u and empty e_g) levels and one deeper lying a_{2u} level as shown in figure 1. In stark contrast to the group-IV elements – vacancy centers, a small energy separation between e_u and e_g permits the appearance of the high-spin states (both triplet and quintet) right above the ground state, as evidenced by both HSE and multireference calculations. In turn, while the quintet state shows D_{3d} symmetry, Jahn-Teller (JT) distortion in the triplet also adds to the stability, yet, not sufficiently to drive the system out of the singlet ground state. This situation translates into the MgV(-1) center with a difference that only one (quartet) high-spin state is possible. Moreover, as computed from the HSE adiabatic potential energy surfaces, the quartet ^4E_u appears as the ground state with the stabilization energy of 37 meV over the doublet ^2E_g at the high symmetry point. Note that both ^4E_u and ^2E_g states are prompt to undergo the JT distortion resulting in a symmetry lowering to C_{2h} upon the 2+4 relaxation scenario (two of six C-Mg bonds become shorter than the remaining four) resulting in the JT energy of 60 and 70 meV, respectfully. Furthermore, due to small energy offset between ^4E_u and ^2E_g, the electron-

phonon interaction also interferes with the alignment of the spin states. Following the methodology from [19], we corrected the adiabatic energies for the zero-point motion; here, we found the true vibronic JT energies of 97 and 84 meV for the doublet- and quartet-state, respectively, that gives the final energy difference of 22 meV. Hence, the thermal energy (~27 meV at 300 K) should be sufficient to enable the spin-conversion to the higher lying doublet state. The further stabilization of the quartet spin-state comes with applying strain: as shown in figure 3a, the uniaxial compressive strain pushes ^4E_u ground state towards lower energies so that ^2E_g cannot be any longer populated at room temperature. Note that such small energy separation between two spin states which ultimately allows for the temperature control and strain engineering of the nature of ground state, is a unique situation for point defects in diamond, to the best of our knowledge.

Next, we addressed the optical properties of the MgV(-1) center. To this end, we first computed the absorption spectra for our cluster model by employing the CASSCF method for a cluster model (where we again observe stabilization of ^4E_u over ^2E_g). As shown in figure 3a, in the energy window below 3 eV, each ^4E_g and ^2E_u state feature two bright transitions which the excitation energies separated by ~1.5 eV. The two lowest ^4E_u and ^2E_g are solely formed by promoting the electrons from e_u to e_g orbitals and are found at ~0.7 and ~1.1 eV, respectively, both being too small to reproduce the experimental PL spectra. As follows from the analysis of the CASSCF wavefunction the $^4A_{2g}$ state at 2.7 eV is governed by a single Slater determinant that justifies the application of ΔSCF method, resulting in turn in the ZPL of ~2.8 eV. Therefore, these findings suggest that despite ^2E_g is higher in energy, $^2E_u^{(2)}$ - ^2E_g transition is responsible for the experimentally observed PL signal. To confirm this, we computed the phonon sideband upon the photoionization scenario from -2 charge state, where, due to the absence of strong correlation, ΔSCF method performs best. The resulting theoretical spectra shown in figure 3c is found to be in an excellent agreement with the experiment, reproducing both the position of ZPL (at 2.2 eV by the HSE calculations) and its vibronic replicas. In turn, the violation of Kasha's rule occurs due the large energy difference between the states (comparable with the $^2E_u^{(2)}$ - ^2E_g excitation energy) and can be also understood from the group theoretical analysis: here, at high-symmetry (D_{3d}) point, the first and second ^2E_u excited states are electronically decoupled, since the transition dipole moment between the states is zero by symmetry. In this case, the nonradiative decay rate, that is a function of the electronic coupling between the initial and final states [20], also vanishes. Hence, once the system is optically initialized in the second excited state, the radiative transition to the ground state becomes the most favorable de-excitation pathway.

Before concluding this letter, we also highlight the notable alignment of the allowed electronic transitions in both spin manifolds for MgV(-1). Since ZPL of ^4E_g -^4E_u state is at half of the $^2E_u^{(2)}$ - ^2E_g energy, a promising down-conversion of the doublet ZPL into two quartets for a pair of coupled MgV(-1) centers may additionally stabilize the quartet state. Here, the down-conversion can spontaneously occur inside the material in the same spin-manifold, in line with the single fission in organic molecules. Moreover, in contrast to the single fission for the acene molecules, the two resulting quartet states are emissive; therefore, the internal quantum efficiency of the entire process should increase, ideally, achieving twice the values of a single defect. The latter potentially expands the application of MgV(-1) to opto-electronics, where the transfer of the (non-emissive) photogenerated triplets to the silicon layer is currently a major challenge for the singlet-fission based technologies. A possible mechanism for the down-conversion at two MgV(0) centers is shown in figure S2 and would require a sizable overlap between the defect wavefunctions, which is controlled by the distance between two defects. Note that such a small separation between two defects was already demonstrated for the NV centers.

To sum up, in this work we performed an extensive theoretical investigation on the role of Mg in the formation of color centers in diamond. Our calculations revealed that MgV is the most stable defect configuration among those considered, with a photostable -1 charge state, achieved in experiment by photoionization with a green laser. Furthermore, we show that this defect features inversion symmetry and emits light from the second excited state at odds with Kasha's rule. It also possesses quasi-degenerate quartet and doublet ground states, each of them exhibiting distinct photophysical properties. While we assigned the experimental PL signal to the $^2E_u^{(2)}$-^2E_g transition, the quartet state (which appeared to be in thermal equilibrium with the doublet ^2E_g) has ZPL of ~1.1 eV, which is right in the near-infrared window in biological tissue. Note that other defects in diamond with similar (prompt) emission wavelength have not been reported, to date. It is also remarkable that the alignment of the stable excited states in both spin manifolds is such that the excitation energy from quartet state is about the half of the $^2E_u^{(2)}$-^2E_g. As both transitions are optically allowed, we anticipate that apart from qubits, a single MgV(-) center can find applications in nonlinear optics, while the combination of two defects might be a source for photon multiplication or up-conversion.

References

[1] H. J. Queisser and E. E. Haller, *Defects in Semiconductors: Some Fatal, Some Vital*, Science **281**, 945 (1998).
[2] J. Liu, Z. Wei, and W. Shangguan, *Defects Engineering in Photocatalytic Water Splitting Materials*, ChemCatChem **11**, 6177 (2019).
[3] J. R. Weber, W. F. Koehl, J. B. Varley, A. Janotti, B. B. Buckley, C. G. Van de Walle, and D. D. Awschalom, *Quantum Computing with Defects*, Proceedings of the National Academy of Sciences **107**, 8513 (2010).
[4] D. D. Awschalom, L. C. Bassett, A. S. Dzurak, E. L. Hu, and J. R. Petta, *Quantum Spintronics: Engineering and Manipulating Atom-Like Spins in Semiconductors*, Science **339**, 1174 (2013).
[5] S. Wehner, D. Elkouss, and R. Hanson, *Quantum Internet: A Vision for the Road Ahead*, Science **362**, (2018).
[6] G. Balasubramanian, P. Neumann, D. Twitchen, M. Markham, R. Kolesov, N. Mizuochi, J. Isoya, J. Achard, J. Beck, J. Tissler, V. Jacques, P. R. Hemmer, F. Jelezko, and J. Wrachtrup, *Ultralong Spin Coherence Time in Isotopically Engineered Diamond*, Nature Materials **8**, 5 (2009).
[7] G. Balasubramanian, P. Neumann, D. Twitchen, M. Markham, R. Kolesov, N. Mizuochi, J. Isoya, J. Achard, J. Beck, J. Tissler, V. Jacques, P. R. Hemmer, F. Jelezko, and J. Wrachtrup, *Ultralong Spin Coherence Time in Isotopically Engineered Diamond*, Nature Materials **8**, 5 (2009).
[8] A. Dréau, A. Tcheborateva, A. E. Mahdaoui, C. Bonato, and R. Hanson, *Quantum Frequency Conversion of Single Photons from a Nitrogen-Vacancy Center in Diamond to Telecommunication Wavelengths*, Phys. Rev. Applied **9**, 064031 (2018).
[9] D. J. Christle, P. V. Klimov, C. F. de las Casas, K. Szász, V. Ivády, V. Jokubavicius, J. Ul Hassan, M. Syväjärvi, W. F. Koehl, T. Ohshima, N. T. Son, E. Janzén, Á. Gali, and D. D. Awschalom, *Isolated Spin Qubits in SiC with a High-Fidelity Infrared Spin-to-Photon Interface*, Phys. Rev. X **7**, 021046 (2017).

[10]   B. Pingault, D.-D. Jarausch, C. Hepp, L. Klintberg, J. N. Becker, M. Markham, C. Becher, and M. Atatüre, *Coherent Control of the Silicon-Vacancy Spin in Diamond*, Nature Communications **8**, 1 (2017).

[11]   T. Lühmann, R. John, R. Wunderlich, J. Meijer, and S. Pezzagna, *Coulomb-Driven Single Defect Engineering for Scalable Qubits and Spin Sensors in Diamond*, Nature Communications **10**, 1 (2019).

[12]   T. Lühmann, N. Raatz, R. John, M. Lesik, J. Rödiger, M. Portail, D. Wildanger, F. Kleißler, K. Nordlund, A. Zaitsev, J.-F. Roch, A. Tallaire, J. Meijer, and S. Pezzagna, *Screening and Engineering of Colour Centres in Diamond*, J. Phys. D: Appl. Phys. **51**, 483002 (2018).

[13]   A. V. Krukau, O. A. Vydrov, A. F. Izmaylov, and G. E. Scuseria, *Influence of the Exchange Screening Parameter on the Performance of Screened Hybrid Functionals*, J. Chem. Phys. **125**, 224106 (2006).

[14]   G. Kresse and J. Furthmüller, *Efficient Iterative Schemes for* Ab Initio *Total-Energy Calculations Using a Plane-Wave Basis Set*, Phys. Rev. B **54**, 11169 (1996).

[15]   C. Freysoldt, J. Neugebauer, and C. G. Van de Walle, *Fully Ab Initio Finite-Size Corrections for Charged-Defect Supercell Calculations*, Phys. Rev. Lett. **102**, 016402 (2009).

[16]   F. Neese, *Software Update: The ORCA Program System, Version 4.0*, WIREs Computational Molecular Science **8**, e1327 (2018).

[17]   J. Shim, E.-K. Lee, Y. J. Lee, and R. M. Nieminen, *Density-Functional Calculations of Defect Formation Energies Using Supercell Methods: Defects in Diamond*, Phys. Rev. B **71**, 035206 (2005).

[18]   J. C. Bourgoin, *An Experimental Estimation of the Vacancy Formation Energy in Diamond*, Radiation Effects **79**, 235 (1983).

[19]   G. Thiering and A. Gali, Ab Initio *Magneto-Optical Spectrum of Group-IV Vacancy Color Centers in Diamond*, Phys. Rev. X **8**, 021063 (2018).

[20]   Z. Burshtein, *Radiative, Nonradiative, and Mixed-Decay Transitions of Rare-Earth Ions in Dielectric Media*, OE **49**, 091005 (2010).
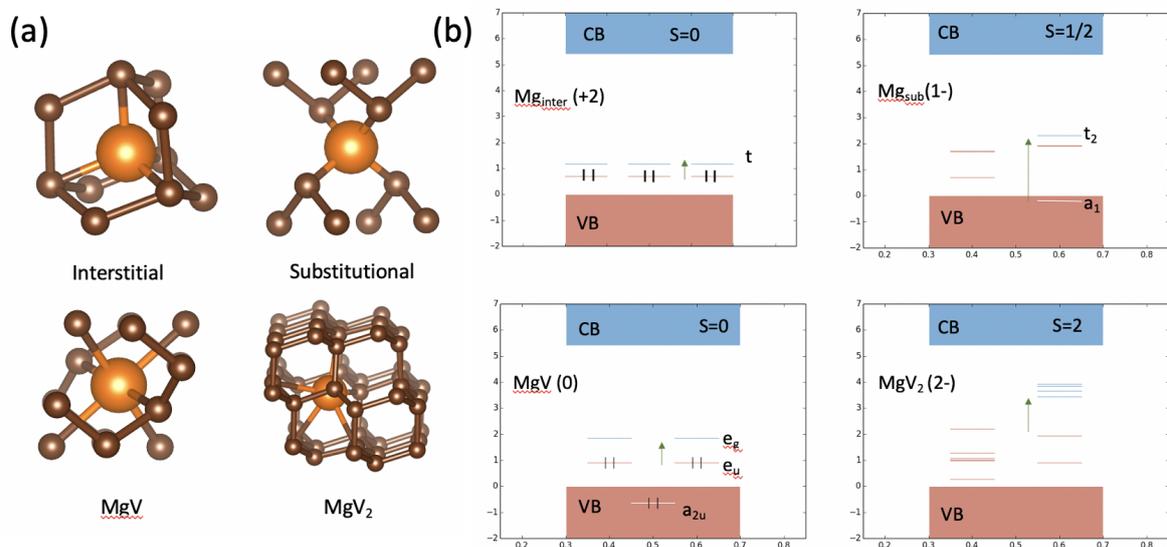
Figure 1. (a) optimized configurations of Mg color centers in diamond and (b) their electronic structure as computed with HSE functional for the representative charge states. The orbitals were classified based on the high symmetry configurations while the green errors indicate lowest symmetry-allowed excitation.
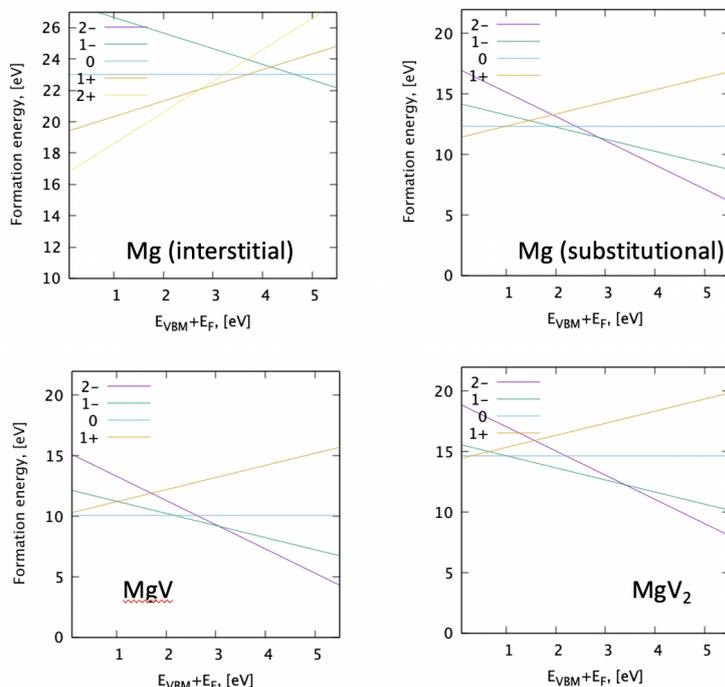
Figure 2. Formation energy diagrams computed for the Mg color centers in diamond from figure 1
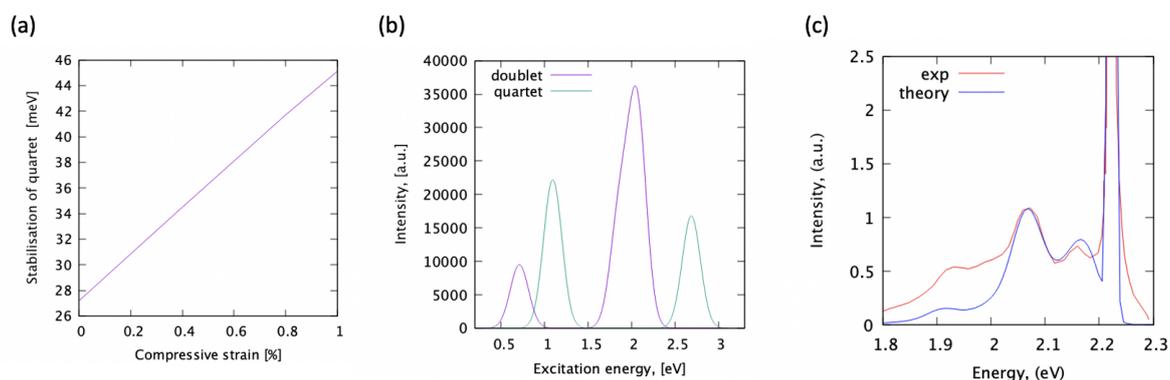
Figure 3. (a) stabilization of the quartet ground state over doublet as a function of applied compressive strain, computed at the HSE adiabatic potential energy surfaces including the relaxation effects, (b) CASSCF absorption spectra from both ground states, and (c) HSE phonon sideband for the $^2E_u^{(2)}$ - ^2E_g transition, overlapped with the experimental PL signal from the Mg color center.

Highly tunable magneto-optical response from MgV color centers in diamond
Supporting Information

i \ S	GS 0	GS 1
1	0.000	0.163
2	1.935	0.357
3	2.518	1.153
4	2.519	1.154
5	2.814	1.517
6	3.460	1.517
7	3.460	2.776
8	3.748	2.925

TABLE I. Many-body energy spectrum obtained of MgV^0 system for ground state geometry (GS) are measured relative to the ground state in eV. Excitation level and the total spin of the state are denoted by i and S, respectively. DFT calculations are performed using PBE functional. DMRG spectrum is calculated on CAS(16,22) correlating 22 electrons on 16 orbitals.

i \ S	GS $\frac{1}{2}$	GS $\frac{3}{2}$	D$_3$d $\frac{1}{2}$	D$_3$d $\frac{3}{2}$	C$_1$h $\frac{1}{2}$	C$_1$h $\frac{3}{2}$
1	0.000	0.079	0.000	0.063	0.000	0.021
2	0.001	0.080	0.001	0.064	0.268	0.383
3	0.530	1.075	0.508	1.025	0.522	1.034
4	0.531	1.076	0.509	1.026	0.707	1.297
5	0.850	2.546	0.850	2.638	1.001	2.781
6	1.402	3.210	1.388	3.268	1.486	3.606
7	1.603	3.424	1.584	3.476	1.657	3.816
8	1.603	3.424	1.584	3.477	1.717	4.078

TABLE II. Many-body energy spectrum obtained of MgV$^-$ system for the GS geometry and excited geometries with D$_3$d and C$_1$h spatial symmetry. DFT calculations are performed using PBE functional. DMRG spectrum is calculated on CAS(26,43).

i \ S	CAS(16,23) $\frac{1}{2}$	CAS(16,23) $\frac{3}{2}$	CAS(26,23) $\frac{1}{2}$	CAS(26,23) $\frac{3}{2}$	CAS(26,43) $\frac{1}{2}$	CAS(26,43) $\frac{3}{2}$
1	0.000	0.099	0.000	0.100	0.000	0.079
2	0.001	0.100	0.001	0.101	0.001	0.080
3	0.555	1.077	0.555	1.077	0.530	1.075
4	0.556	1.076	0.556	1.078	0.531	1.076
5	0.898	2.591	0.895	2.583	0.850	2.546
6	1.521	3.196	1.517	3.193	1.402	3.210
7	1.711	3.417	1.707	3.412	1.603	3.424
8	1.713	3.417	1.710	3.412	1.603	3.424

TABLE III. Many-body energy spectrum obtained of MgV$^-$ system for ground state (GS) geometry. DFT-CAS-DMRG calculations are performed using three distinct active spaces, i.e. CAS(16,23), CAS(26,23), CAS(26,43).

i \ S	PBE $\frac{1}{2}$	PBE $\frac{3}{2}$	PBE0 $\frac{1}{2}$	PBE0 $\frac{3}{2}$	HSE06 $\frac{1}{2}$	HSE06 $\frac{3}{2}$
1	0.000	0.079	0.000	0.081	0.000	0.124
2	0.001	0.080	0.001	0.082	0.001	0.125
3	0.530	1.075	0.539	1.093	0.596	1.169
4	0.531	1.076	0.540	1.094	0.597	1.171
5	0.850	2.546	0.859	2.553	0.982	2.886
6	1.402	3.210	1.417	3.223	1.465	3.551
7	1.603	3.424	1.623	3.438	1.734	3.635
8	1.603	3.424	1.623	3.438	1.734	3.729

TABLE IV. Many-body energy spectrum obtained of MgV$^-$ system for GS geometry. DFT-CAS-DMRG calculations are performed on CAS(26,43) using PBE, PBE0 and HS06 functionals. (In QE calculations, PBE denotes: (1 4 3 4 0 0) EXX 0; PBE0 denotes: (6 4 8 4 0 0) EXX 0.25; HSE06 denotes: (0 0 12 0 0 0) EXX 0.25.

A. Discussion

The convergence of the density functional theory- complete active space- density matrix renormalization group (DFT-CAS-DMRG) calculations is tested in terms of the CAS size. Spectra obtained on CAS(16,23), CAS(26,23), CAS(26,43) active systems are compiled in Table III. Comparing results on CAS(16,23) and many-body spectrum obtained on larger virtual space, i.e., CAS(26,23), we find excellent agreement suggesting that the virtual orbitals play insignificant role in the description of the low-energy excitations. Extending the active space towards the valence space, i.e., CAS(26,43), the DMRG spectrum also varies marginally. Nevertheless, considering the variational nature of the CAS-DMRG calculations, the obtained lower absolute energies suggest the eminence of the valence orbitals compared to the virtuals in the characterization of the spectrum.

The effect of the choice of the exchange-correlation functional on the DMRG spectrum is also investigated in Table IV. The many-body spectrum for PBE and for hybrid PBE0 are in excellent agreement. HSE06 functional provides similar excitation structure with increased gaps, i.e., typically of 10%.

The presented DFT-CAS-DMRG analysis provide an insight into the overall features of the many-body spectrum by capturing strong correlation effects, nevertheless the description does neither incorporate dynamical correlations nor screening effects. The direct comparison with experimental data is also hindered by neglecting geometric relaxations.

B. Methods

Spin-restricted Kohn-Sham orbitals are used for generating CAS. Depending on the total electron occupation of the system, frontier orbitals are optimized by smearing the electron pair with highest energy and the unpaired electron evenly on the partially occupied degenerate e orbitals for MgV^0 and for MgV^-, respectively. Orbitals around the valence band maximum level are selected to form the active space.

Many-body electronic excitation spectrum is computed using the density matrix renormalization group (DMRG)

approach[1] applied on top of periodic Kohn-Sham orbitals.[2] DFT and DMRG simulations are performed using the QE[3,4] and the Budapest DMRG package,[5] respectively. The two program suites are interfaced using our in-house developed code which constructs the Hamiltonian of the active orbitals corresponding to the complete active space (CAS) method.[6,7] The accuracy of the DMRG calculations is controlled by the dynamic block state selection approach and the computational efficiency is improved by initialization procedures which are inspired by quantum information theoretical considerations.[8] Investigating multiple excitations, distinct DMRG calculations are performed for target states with different total spin.

[1] S. R. White, Phys. Rev. Lett. **69**, 2863 (1992).

[2] G. Barcza, V. Ivády, S. Tibor, V. Márton, V. Libor, Á. Gali, and Ö. Legeza, "DMRG on plane-wave KS orbitals: case study of negatively charged boron vacancy point defect in hexagonal boron nitride," (2020), arXiv:2006.04557.

[3] 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, N. Marzari, F. Mauri, R. Mazzarello, S. Paolini, A. Pasquarello, L. Paulatto, C. Sbraccia, S. Scandolo, G. Sclauzero, A. P. Seitsonen, A. Smogunov, P. Umari, and R. M. Wentzcovitch, Journal of Physics: Condensed Matter **21**, 395502 (2009).

[4] P. Giannozzi, O. Andreussi, T. Brumme, O. Bunau, M. B. Nardelli, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, M. Cococcioni, N. Colonna, I. Carnimeo, A. D. Corso, S. de Gironcoli, P. Delugas, R. A. D. Jr, A. Ferretti, A. Floris, G. Fratesi, G. Fugallo, R. Gebauer, U. Gerstmann, F. Giustino, T. Gorni, J. Jia, M. Kawamura, H.-Y. Ko, A. Kokalj, E. Küçükbenli, M. Lazzeri, M. Marsili, N. Marzari, F. Mauri, N. L. Nguyen, H.-V. Nguyen, A. O. de-la Roza, L. Paulatto, S. Poncé, D. Rocca, R. Sabatini, B. Santra, M. Schlipf, A. P. Seitsonen, A. Smogunov, I. Timrov, T. Thonhauser, P. Umari, N. Vast, X. Wu, and S. Baroni, Journal of Physics: Condensed Matter **29**, 465901 (2017).

[5] Ö. Legeza, L. Veis, and T. Mosoni, "QC-DMRG-Budapest, a program for quantum chemical DMRG calculations," (2020).

[6] C. Cramer, *Essentials of Computational Chemistry: Theories and Models* (Wiley, 2005).

[7] F. Jensen, *Introduction to Computational Chemistry* (John Wiley and Sons, Inc., Hoboken, NJ, USA, 2006).

[8] Sz. Szalay, M. Pfeffer, V. Murg, G. Barcza, F. Verstraete, R. Schneider, and Ö. Legeza, Int. J. Quant. Chem. **115**, 1342 (2015).

Combining the n-electron valence state formalism with the restricted active space configuration interaction method

Gergely Barcza[1,2,*] and Tibor Szilvási[3]

[1] *Wigner Research Centre for Physics, H-1525, Budapest, Hungary*
[2] *J. Heyrovský Institute of Physical Chemistry, Czech Academy of Sciences, CZ-18223, Prague, Czechia*
[3] *Department of Chemical and Biological Engineering,*
The University of Alabama, Tuscaloosa, AL-35487, USA

We introduce the n-electron valence states formalism for the restricted active space configuration interaction method. The proposed variational formalism provides an accurate description not only the ground state but also the low-lying vertical excitation spectrum. Benchmark tests are performed on both molecules and periodic systems to compare with state-of-the-art methods.

I. INTRODUCTION

The accurate theoretical description of static electronic structures is provided by the solution of their time-independent Schrödinger equation. In Born–Oppenheimer approximation, the second quantized representation of the corresponding ab initio Hamiltonian,[1] Eq. (1), is written

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

with molecule specific one- and two-electron integrals, t_{ij}, and V_{ijkl}, respectively. Here, operators $\hat{a}^{\dagger}_{i\sigma}$ and $\hat{a}_{i\sigma}$ creates and annihilates electron with σ spin projection in molecular orbital i. The rank of the associated computational basis, $\mathbf{\Phi}$, increases exponentially with spin-orbital size L, i.e., 4^L neglecting particle and spin conservation. Therefore, the exact solution on current classical machines is restricted to a dozen of orbitals. In order to provide accurate description of larger systems various approaches have been developed taking advantage of the internal structure of the Hamiltonian and its relevant, i.e., low-energy, eigenstates.

One of the most important protocols which aims to keep the problem computationally tractable is the complete active space (CAS) scheme[2–4] which restricts the large configuration space to an effective subset spanning given $L_{\text{CAS}} \leq L$ orbitals. The protocol classifies the set of canonical molecular orbitals to three categories, i.e., the so-called core and virtual orbitals are kept on Hartree-Fock level of theory 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. The method recovers the correlation effects among the active orbitals but it completely neglects correlations encoded outside the active space. In practical simulation of large systems, due to scaling of the method, the active space of $L_{\text{CAS}} \ll L$ orbitals is selected to focus on the description of states with static correlations, i.e., wave functions which differ strikingly from the reference Slater determinant of the Hartree-Fock solution. Even though the principal

features of the electronic states are characterized by the static correlations, the dynamical effects, i.e., contribution of intractable number of excited configurations with small weights, can be also crucial to provide theoretical description which is comparable with experimental data.[5,6] Such effects are well described as determinental excitations using coupled cluster theory (CC)[7] or on the level of orbitals itself, i.e., Kohn-Sham orbitals of density-functional theory (DFT).[5,8,9]

Besides various approaches based on perturbation theory,[10–16] many-body expansion theory,[17–20] and tailored methods,[21,22] the restricted active space (RAS) protocol[3,4,23] is one computational scheme which tries to combine the two worlds of correlation effects in order to provide a balanced description of electronic correlations. Extending the horizon of the CAS protocol, the simplest variant of the RAS method distinguishes additional two sets of orbitals, the excitable valence (Val) and virtual (Virt) spaces where restricted R level of electron excitation can be considered in between them and the active orbitals. In practice, this simple protocol of limited excitation scheme, in the following denoted RAS(R), is already suitable for capturing some dynamical correlation effects in addition to static correlations.

Standard wave-function based numerical methods, e.g., variants of configuration interaction (CI) method,[1,3,4,24,25] which are based on expansions of configuration space relative to a reference state, are naturally able to implement both CAS and RAS protocols. In the standard CAS approximation of a system with n_e electrons, n_c electrons occupy the inactive orbitals according to their HF level filling and $n_a = n_e - n_c$ electrons are correlated in the active space. In the RAS approach the CAS determinental space is enlarged such a way that restricted $|r| \leq R$ number of particles (for $r > 0$ electrons and for $R < 0$ holes) are promoted from the reference configuration. Most importantly, the corresponding Hamiltonian matrix is expanded in the basis of all the possible determinental states considering the possible restricted electronic promotions. Formally, the effective configurational space is decomposed as $\mathbf{\Phi}_{\text{eff}} = \mathbf{\Phi}_{\text{core}} \mathbf{\Phi}_{\text{RAS}} \mathbf{\Phi}_{\text{CAS}}$. Here, $\mathbf{\Phi}_{\text{core}}$ denotes the string of modes corresponding to the inactive frozen cores while $\mathbf{\Phi}_{\text{RAS}}$ and $\mathbf{\Phi}_{\text{CAS}}$ the set of

possible configurations restricted to the RAS and the CAS orbital subspaces. In the following, frozen cores are taken implicit and notation Φ_{core} is omitted for the sake of brevity. Describing L_{CAS} orbitals on full-CI level of theory and limiting excitations in RAS(R) space of a large number of L_{RAS} (exclusively core or virtual) orbitals up to R particles i.e., $R \ll L_{\mathrm{RAS}}$, the dimension of the Hamiltonian expanded in the configuration basis scales as $\mathcal{O}(L_{\mathrm{RAS}}^{R} 4^{L_{\mathrm{CAS}}})$.

Contrary to this demanding solution, to the best of our knowledge, we propose a novel implementation for the RAS concept applicable not only for CI methods but also for arbitrary wave function approaches, e.g., tensor-network state algorithms for ab initio problems,[26] in particular density matrix renormalization group (DMRG).[27,28] In our approach, the standard RAS result is approximated by the combination of specially constructed many-body wave functions, i.e., n-electronic valence states (NEV),[15] obtained on a series of CAS spaces of L_{CAS} orbitals. Deploying this alternative formulation of RAS(R)-FCI, denoted in the following NEV-RAS(R)-FCI, an effective Hamiltonian of rank $\mathcal{O}(L_{\mathrm{RAS}}^{R})$ expanded in the basis of many-body states is to be solved. In the proposed RAS representation expressing the accessible configurational space in terms of relevant many-body states, significantly larger active space might be reachable on single and double excitation level, i.e., $R = 1 - 2$, compared to the canonical RAS-FCI. The method converges variationally to the exact solution increasing R.

The method is outlined in three key steps:

- 1) a list of CAS spaces is defined by the possible $r = \{0, \ldots, R\}$ electron excitations in the Val-Virt spaces

- 2) full CI calculation are performed on this series of CAS spaces

- 3) the obtained full CI results are combined together to minimize the energy of the actual many-body state of the defining problem (1) by diagonalization

II. THEORY

A. n-electron valence states

In the following, we adopt the notation applied in paper[15] which introduced the so called n-electron valence states (NEV) in perturbation theory. The possible determinants of a given CAS space, characterized by s inactive electron configuration part with r excitations relative to the reference state, are selected from the determinental space as the orbital product of the inactive $n(\Phi_s^{-r}) = n_c - r$ electrons, $\Phi_s^{-r} \in \mathbf{\Phi}_{\mathrm{RAS}}$, and the possible configurations in the active orbitals filled with $n_a + r$

electrons using projector $\hat{P}_s^r$. Formally, the set of the resulted basis states reads

$$\mathbf{\Phi}_s^r \equiv \hat{P}_s^r \mathbf{\Phi}_{\mathrm{eff}} = \{|\Phi_s^{-r}\Phi^r\rangle \in \mathbf{\Phi}_{\mathrm{eff}} \mid n(\Phi^r) = n_a + r\}. \quad (2)$$

The eigenvalue-problem of the corresponding Hamiltonian projected from Eq. (1) reads

$$\hat{H}_s^r |\Phi_s^{-r}\Psi_{s,\alpha}^r\rangle = \hat{P}_s^r \hat{H} \hat{P}_s^r |\Phi_s^{-r}\Psi_{s,\alpha}^r\rangle = E_{s,\alpha}^r |\Phi_s^{-r}\Psi_{s,\alpha}^r\rangle \quad (3)$$

where $|\Psi_{s,\alpha}^r\rangle$, referred to as fully uncontracted valence state in NEVPT theory, is the αth many-body eigenstate of Hamiltonian $\hat{H}_s^r$ and expands formally in the $\mathbf{\Phi}_s^r$ basis as

$$|\Phi_s^{-r}\Psi_{s,\alpha}^r\rangle = \sum_{x \in \mathbf{\Phi}_s^r} c_{s\alpha,x}^r |x\rangle. \quad (4)$$

The projected Hamiltonian (3) reads

$$\hat{H}_s^r = \sum_{ij,\sigma} t_{s;ij\sigma}^r \hat{a}_{i\sigma}^\dagger \hat{a}_{j\sigma} + \frac{1}{2} \sum_{ijkl,\sigma\sigma'} V_{ijkl} \hat{a}_{i\sigma}^\dagger \hat{a}_{j\sigma'}^\dagger \hat{a}_{k\sigma'} \hat{a}_{l\sigma} \quad (5)$$
$$+ E_{\mathrm{core};s}^r + E_{\mathrm{nuc}},$$

where V_{ijkl} integrals are restricted to the active orbital set. The one-electron integrals of the CAS space $\mathbf{\Phi}_s^r$, $t_{ij;\sigma}^{s,r}$, describes not only the kinetic energy of the active electrons and their attraction to nuclei but also their interaction with the mean-field electrons associated to $\mathbf{\Phi}_s^r$. The one-electron interactions are written as

$$t_{s;ij\sigma}^r = t_{ij} + \frac{1}{2}\left(\sum_{c \in C_{s\sigma}^r} \left(2V_{iccj} - V_{icjc}\right) + \sum_{c \in C_{s\sigma'}^r} 2V_{iccj}\right) \quad (6)$$

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 in distinct spin channels, $C_{s\sigma}^r$, corresponding for a given Φ_s^{-r} NEV configuration. Finally, the additional energy contribution of the core electrons is summed up in term $E_{\mathrm{core};s}^r$.

B. Hamiltonian expanded in the basis of n-electron valence states

In the $R = 0$ limit, i.e., all restricted active orbitals are frozen to their HF occupation, the theory is equivalent to the standard CAS solution for L_{CAS}. In the complete set limit, reaching $\tilde{R} \equiv R = 2\max(L_c, L_v)$, the theory is equivalent to the CAS solution for $n_a + 2L_c$ electrons correlated on $L_{\mathrm{CAS}} + L_c + L_v$ orbitals. In practice, $L_c \ll L_v$ for standalone molecules expanded in localized orbital set and $L_c \gg L_v$ for large gaped periodic systems. More importantly, it is to be noted that, by construction, all the computed eigenstates are orthonormal, i.e.,

$$\langle \Psi_{s',\alpha'}^{r'} \Phi_{s'}^{-r'} |\Phi_s^{-r}\Psi_{s,\alpha}^r\rangle = \delta_{\alpha,\alpha'}\delta_{r,r'}\delta_{s,s'}, \quad (7)$$

for $|r| \leq R$, $1 \leq s \leq d^r$, $1 \leq \alpha \leq m_s^r$, where d_r denotes the number of the possible r-particle-NEV configurations which scales as $\mathcal{O}(L_{\mathrm{RAS}}^r)$ and m_s^r the number of the computed many-body states for NEV configuration Φ_s^{-r}. The orthonormality is the consequence of orthogonality of the NEV configurations and the orthornormality of the many-body states belonging to a particular NEV configuration.

Therefore, the collection of the obtained eigenfunctions provides an incomplete set of basis many-body states spanning the CI excitation RAS space, which is shortly denoted as

$$\Psi_m^R = \left\{ |\Phi_s^{-r}\Psi_{s,\alpha}^r\rangle \mid |r| \leq R,\ 1 \leq s \leq d^r,\ 1 \leq \alpha \leq m \right\}, \tag{8}$$

where we presumed that m low-energy many-body states are generated for all Φ_s^{-r} NEV configuration, for the sake of brevity.

Accordingly, the defining Hamiltonian of the molecule, Eq.(1), is expanded in basis Ψ_m^R which is denoted in the following as

$$H_{r',s',\alpha';r,s,\alpha}^{\mathrm{NEV-RAS}} = \langle \Psi_{s',\alpha'}^{r'}\Phi_{s'}^{-r'}|\hat{H}|\Phi_s^{-r}\Psi_{s,\alpha}^r\rangle. \tag{9}$$

The efficient evaluation of the matrix elements, Eq.(9), is provided in the appendix. Solving the eigenvalue problem, i.e.,

$$\hat{H}^{\mathrm{NEV-RAS}}|\Psi_\alpha^{\mathrm{NEV-RAS}}\rangle = E_\alpha^{\mathrm{NEV-RAS}}|\Psi_\alpha\rangle, \tag{10}$$

the many-body RAS energies, E_α and the corresponding many-body wave functions, $|\Psi_\alpha\rangle$, are obtained. Using a single iterative index over Ψ_m^R, the wave functions reads

$$|\Psi_\alpha^{\mathrm{NEV-RAS}}\rangle = \sum_{x\in\Psi} c_{\alpha,x}|x\rangle \tag{11}$$

revealing that the construction of arbitrary expectation value of the orbitals, e.g., occupation/spin density, correlators, is straightforward.

C. Assessing computational costs

A key computational bottleneck of a "naive" implementation is the memory-footprint of the quasi-full matrix of Hamiltonian (10), i.e., expanding it up to single and double RAS excitations scales as $\mathcal{O}(mL_{\mathrm{RAS}} \times mL_{\mathrm{RAS}})$ and $\mathcal{O}((mL_{\mathrm{RAS}})^2 \times (mL_{\mathrm{RAS}})^2)$, respectively. Furthermore, the cost of obtaining the low-energy spectrum by iterative diagonalization routine[29] scales polynomially with the rank of the Hamiltonian.

Therefore, as of 2020, a typical computational node equipped with 100 cores and 100 GB RAM, could treat around 100k NEV many-body states, Eq. 8 presuming dense matrix structure for the NEV Hamiltonian. Fixing $m = 10$ and keeping 4-10 orbitals in the CAS spaces Ψ_s^r, where each FCI calculation takes around 0.5-30 min using a single computational core, 10k completely independent FCI calculations takes 1-50 hours utilizing 100 cores. The actual construction of (10) is performed on comparable time scale whereas its partial diagonalization takes around 10 hours using the Lánczos method.[29] It is also notable that the generation of matrix elements of the defining Hamiltonian (1), i.e., FCIDUMP, scaling with quartic of the number of orbitals, also becomes critical for hundreds of orbitals.

In conclusion, the total computational time sums up to a 0.5-5 days on a regular computational machine for an effective Hamiltonian expressed in the basis of 10k NEV configurations which translates around to 5k or 50 orbitals treated on RAS(1)-CI or RAS(2)-CI, respectively.

D. Optimizing the eigendecomposition of the NEV-RAS Hamiltonian

In the following, we shortly study the internal structure of the dense Hamiltonian matrix (9) in order to reduce numerical efforts. Owing to the fact that the Hamiltonian only contains one- and two-electron scatterings, the resulting off-diagonal matrix elements can be numerically negligible yielding sparse representation for the matrix. In practice, matrix elements could be potentially discarded having marginal influence on the solution, i.e. $|\langle j|H^{\mathrm{NEV-RAS}}|i\rangle| < 10^{-10}$. Analyzing the pattern of matrix elements, typically we observe that up to 25-75% of the matrix elements have negligible weight, however there are no segregated NEV states, i.e., all states are connected to some others.

In addition, applying the reverse Cuthill-McKee algorithm[30] on the NEV-RAS Hamiltonian, its potential blockdiagonal structure corresponding to the spatial symmetry of the many-body states is revealed. Considering this feature, the eigendecomposition problem of the large Hamiltonian matrix (9) is more efficiently performed on the separate subblocks of smaller rank.

Furthermore, taking advantage of the inner product structure of the NEV states Ψ, similarly to bipartite operator decomposition in DMRG, the memory footprint of the algorithm can be significantly reduced.

III. BENCHMARK AND DISCUSSION

In this paper, the FCI solutions to generate NEV states, Eq. (8), are obtained by exact diagonalization implemented as part of the Budapest-DMRG program package.[31] The NEV-RAS-FCI concept, i.e., Eq. (9), has been worked out up to double excitations.

In the following, as proof of principle, some preliminary calculations are given for solids and molecules, in table I and II, respectively. Reference results are provided by DMRG/CC, NEV-RAS-FCI results are also set against CAS-FCI results obtained on small active spaces formed of the frontier orbitals.

S $\backslash$ i	DMRG(40,70)		NEV-RAS(1)-FCI(4,6,TS=4)			CAS-FCI(4,6)		
	E_i	Δ_i^{ref}	E_i	Δ_i	$\Delta_i/\Delta_i^{\mathrm{ref}}$	E_i	Δ_i	$\Delta_i/\Delta_i^{\mathrm{ref}}$
$^3A_2'$	-1939.0080	0.0000	-1939.0061	0.0000	NaN	-1938.9971	0.0000	NaN
$^1E'$	-1938.9235	0.0845	-1938.9200	0.0862	1.0194	-1938.7670	0.2301	2.7216
$^1A_2''$	-1938.8361	0.1719	-1938.8326	0.1735	1.0097	-1938.7597	0.2374	1.3812

TABLE I. Snippet of the low-lying vertical many-body excitation spectrum of the periodic the VB-hBN sheet measured relative to the ground state energy in Ry. Reference calculation is provided by DMRG where 70 electrons are correlated on 40 orbitals. The standard CAS-FCI(4,6), i.e., exact diagonalization, is performed on active space of 4 orbitals with 6 active electrons. The NEV-RAS(1)-FCI solution (up to the single excitation level on the restricted orbital space) is composed of results obtained on a series of constructed CAS spaces of 4 orbitals which reproduces the reference energies with great accuracy. Notable that the reference DMRG calculation took 250 min on a cluster, the NEV-RAS-CIS calculation took 15 min on a laptop.

S $\backslash$ i	DMRG(20,6)		NEV-RAS(1)-FCI(4,6,TS)			NEV-RAS(1)-FCI(4,6,TS=6)			CAS-FCI(4,6)		
	E_i	Δ_i^{ref}	E_i	Δ_i	$\Delta_i/\Delta_i^{\mathrm{ref}}$	E_i	Δ_i	$\Delta_i/\Delta_i^{\mathrm{ref}}$	E_i	Δ_i	$\Delta_i/\Delta_i^{\mathrm{ref}}$
1	-39.0319	0	-38.9384	0	NaN	-38.9232	0	NaN	-38.9211	0	NaN
2	-38.9900	0.0420	-38.8894	0.0489	1.1663	-38.8787	0.0445	1.0598	-38.8684	0.0527	1.2565
3	-38.9699	0.0621	-38.8709	0.0675	1.0880	-38.8545	0.0686	1.1060	-38.8460	0.0752	1.2114
4	-38.9051	0.1269	-38.7946	0.1438	1.1333	-38.7893	0.1339	1.0557	-38.7812	0.1399	1.1028
5	-38.7593	0.2726	-38.6555	0.2828	1.0376	-38.6385	0.2846	1.0442	-38.6032	0.3180	1.1666
6	-38.7565	0.2754	-38.6489	0.2894	1.0508	-38.6320	0.2912	1.0573	-38.5959	0.3252	1.1807
7	-38.7375	0.2944	-38.6331	0.3053	1.0371	-38.6105	0.3127	1.0622	-38.5608	0.3603	1.2240
8	-38.7358	0.2962	-38.6161	0.3222	1.0880	-38.6087	0.3145	1.0619	-38.5478	0.3734	1.2606
9	-38.7234	0.3085	-38.6071	0.3313	1.0739	-38.5992	0.3240	1.0503	-38.3873	0.5338	1.7305
10	-38.7188	0.3131	-38.6057	0.3327	1.0626	-38.5717	0.3515	1.1227	-38.3797	0.5414	1.7292

TABLE II. Low-lying vertical energy spectrum of CH2 measured in Hartree. DMRG is applied as reference method correlating 6 electrons on 20 orbitals. The performance of NEV-RAS-CI method is set against the CAS-FCI. The average error of the energy gaps is 8%, 7% and 31% for NEV-RAS(1)-FCI(4,6,TS=NaN), NEV-RAS(1)-FCI(4,6,TS=6)and CAS-FCI(4,6), respectively. TS=6/NaN denotes that up to only 6/all possible NEV states are kept to construct the effective Hamiltonian. Analyzing the structure of the reference wave functions obtained from DMRG suggests that NEV-RAS(2)-FCI should provide much more accurate estimate for absolute energies as well.

IV. CONCLUSION AND PERSPECTIVE

We proposed a novel formalism of restricted active space protocol based on n-electron valance states to extend the capabilities of methodology to multiple dozens of orbitals. The approach, denoted NEV-RAS, is expected to be particularly well-suited to describe ab initio systems with orbital space separable according to its correlation pattern, e.g. localized defects in solids, see table I, to provide a balanced description of static and dynamical correlations already in terms of NEV-RAS(1). The main advantages of the method, compared to alternatives, e.g., tailored coupled cluster, that it is variational and it is capable of describing not only the ground state but the vertical excitation states as well.

The presented RAS scheme is straightforward to generalize as a multilayered protocol,i.e., most RAS orbitals are treated only up to single excitation level while the more intricate ones are on higher level of theory.

Furthermore, the NEV states based RAS protocol also trivially applicable to the reference-free DMRG method which is notorious for its CAS-only description neglecting dynamical correlations of external orbital space.

* barcza.gergely@wigner.mta.hu

[1] Szabo, A.; Ostlund, N. S. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*, 1st ed.; Dover Publications, Inc.: Mineola, 1996.

[2] Roos, B. O. *Advances in Chemical Physics*; John Wiley & Sons, Ltd, 2007; pp 399–445.

[3] Cramer, C. *Essentials of Computational Chemistry: Theories and Models*; Wiley, 2005.

[4] Jensen, F. *Introduction to Computational Chemistry*; John Wiley and Sons, Inc.: Hoboken, NJ, USA, 2006.

[5] Becke, A. D. Density functionals for static, dynamical, and strong correlation. *The Journal of Chemical Physics* **2013**, *138*, 074109.

[6] Benavides-Riveros, C. L.; Lathiotakis, N. N.; Marques, M. A. L. Towards a formal definition of static and dynamic electronic correlations. *Phys. Chem. Chem. Phys.* **2017**, *19*, 12655–12664.

[7] Cizek, J. On the Correlation Problem in Atomic and Molecular Systems. Calculation of Wavefunction Components in UrsellType Expansion Using QuantumField Theoretical Methods. *The Journal of Chemical Physics* **1966**, *45*, 4256–4266.

[8] Kohn, W.; Sham, L. J. Self-Consistent Equations Including Exchange and Correlation Effects. *Phys. Rev.* **1965**, *140*, A1133–A1138.

[9] Sham, L. J.; Schlüter, M. Density-Functional Theory of the Energy Gap. *Phys. Rev. Lett.* **1983**, *51*, 1888–1891.

[10] Møller, C.; Plesset, M. S. Note on an Approximation Treatment for Many-Electron Systems. *Phys. Rev.* **1934**, *46*, 618–622.

[11] Epstein, P. S. The Stark Effect from the Point of View of Schroedinger's Quantum Theory. *Phys. Rev.* **1926**, *28*, 695–710.

[12] Nesbet, R. K.; Hartree, D. R. Configuration interaction in orbital theories. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* **1955**, *230*, 312–321.

[13] Andersson, K.; Malmqvist, P. A.; Roos, B. O.; Sadlej, A. J.; Wolinski, K. Second-order perturbation theory with a CASSCF reference function. *The Journal of Physical Chemistry* **1990**, *94*, 5483–5488.

[14] Malrieu, J.-P.; Heully, J.-L.; Zaitsevskii, A. Multiconfigurational second-order perturbative methods: Overview and comparison of basic properties. *Theoretica chimica acta* **1995**, *90*, 167–187.

[15] Angeli, C.; Cimiraglia, R.; Evangelisti, S.; Leininger, T.; Malrieu, J.-P. Introduction of n-electron valence states for multireference perturbation theory. *The Journal of Chemical Physics* **2001**, *114*, 10252–10264.

[16] Malmqvist, P. .; Pierloot, K.; Shahi, A. R. M.; Cramer, C. J.; Gagliardi, L. The restricted active space followed by second-order perturbation theory method: Theory and application to the study of CuO2 and Cu2O2 systems. *The Journal of Chemical Physics* **2008**, *128*, 204109.

[17] Richard, R. M.; Lao, K. U.; Herbert, J. M. Understanding the many-body expansion for large systems. I. Precision considerations. *The Journal of Chemical Physics* **2014**, *141*, 014108.

[18] Zimmerman, P. M. Strong correlation in incremental full configuration interaction. *The Journal of Chemical Physics* **2017**, *146*, 224104.

[19] Zimmerman, P. M. Incremental full configuration interaction. *The Journal of Chemical Physics* **2017**, *146*, 104102.

[20] Eriksen, J. J.; Gauss, J. Generalized Many-Body Expanded Full Configuration Interaction Theory. *The Journal of Physical Chemistry Letters* **2019**, *10*, 7910–7915.

[21] Kinoshita, T.; Hino, O.; Bartlett, R. J. Coupled-cluster method tailored by configuration interaction. *The Journal of Chemical Physics* **2005**, *123*, 074106.

[22] Veis, L.; Antalk, A.; Brabec, J.; Neese, F.; Legeza, .; Pittner, J. Coupled Cluster Method with Single and Double Excitations Tailored by Matrix Product State Wave Functions. *The Journal of Physical Chemistry Letters* **2016**, *7*, 4072–4078.

[23] Olsen, J.; Roos, B. O.; Jörgensen, P.; Jensen, H. J. A. Determinant based configuration interaction algorithms for complete and restricted configuration interaction spaces. *The Journal of Chemical Physics* **1988**, *89*, 2185–2192.

[24] Siegbahn, P. E. M. *Chem. Phys. Lett.* **1984**, *109*, 417.

[25] Knowles, P. J.; Handy, N. C. A new determinant-based full configuration interaction method. *Chem. Phys. Lett.* **1984**, *111*, 315.

[26] Szalay, Sz.; Pfeffer, M.; Murg, V.; Barcza, G.; Verstraete, F.; Schneider, R.; Legeza, Ö. Tensor product methods and entanglement optimization for ab initio quantum chemistry. *Int. J. Quantum Chem.* **2015**, *115*, 1342–1391.

[27] White, S. R.; Martin, R. L. Ab initio quantum chemistry using the density matrix renormalization group. *The Journal of Chemical Physics* **1999**, *110*, 4127–4130.

[28] Baiardi, A.; Reiher, M. The density matrix renormalization group in chemistry and molecular physics: Recent developments and new challenges. *The Journal of Chemical Physics* **2020**, *152*, 040903.

[29] Lánczos, K. An iteration method for the solution of the eigenvalue problem of linear differential and integral operators. *Journal of Research of the National Bureau of Standards* **195**, *45*, 255–282.

[30] George, A.; Liu, J. W. *Computer Solution of Large Sparse Positive Definite*; Prentice Hall Professional Technical Reference, 1981.

[31] Legeza, Ö.; Veis, L.; Mosoni, T. QC-DMRG-Budapest, a program for quantum chemical DMRG calculations.