Journal of
Materials Chemistry C

Materials for optical, magnetic and electronic devices

Accepted Manuscript

Journal of
Materials Chemistry C

Materials for optical, magnetic and electronic devices

rsc.li/materials-c

Volume 6
Number 5
7 February 2018
Pages 911-1266

ISSN 2050-7526

ROYAL SOCIETY OF CHEMISTRY

PAPER
Nosang V. Myung, Yong-Ho Choe et al.
A noble gas sensor platform: linear dense assemblies of
single-walled carbon nanotubes (LACNTs) in a multi-layered
ceramic/metal electrode system (MLES)

ROYAL SOCIETY OF CHEMISTRY

rsc.li/materials-c

Advancing the hBN Defects Database through Photophysical Characterization of Bulk hBN

Chanaprom Cholsuk,[1, 2, *] Sujin Suwanna,[3] and Tobias Vogl[1, 2, 4, †]

[1]*Department of Computer Engineering, TUM School of Computation,*
Information and Technology, Technical University of Munich, 80333 Munich, Germany
[2]*Munich Center for Quantum Science and Technology (MCQST), 80799 Munich, Germany*
[3]*Optical and Quantum Physics Laboratory, Department of Physics,*
Faculty of Science, Mahidol University, Bangkok 10400, Thailand
[4]*Abbe Center of Photonics, Institute of Applied Physics,*
Friedrich Schiller University Jena, 07745 Jena, Germany

(Dated: September 17, 2025)

Quantum emitters in hexagonal boron nitride (hBN) have gained significant attention due to a wide range of defects that offer high quantum efficiency and single-photon purity at room temperature. Most theoretical studies on hBN defects simulate monolayers, as this is computationally cheaper than calculating bulk structures. However, most experimental studies are carried out on multilayer to bulk hBN, which creates additional possibilities for discrepancies between theory and experiment. In this work, we present an extended database of hBN defects that includes a comprehensive set of bulk hBN defects along with their excited-state photophysical properties. The database features over 120 neutral defects, systematically evaluated across charge states ranging from −2 to +2 (600 defects in total). For each defect, the most stable charge and spin configurations are identified and used to compute the zero-phonon line, photoluminescence spectrum, absorption spectrum, Huang-Rhys (HR) factor, interactive radiative lifetimes, transition dipole moments, and polarization characteristics. Our analysis reveals that the electron–phonon coupling strength is primarily influenced by the presence of vacancies, which tend to induce stronger lattice distortions and broaden phonon sidebands. Additionally, correlation analysis shows that while most properties are independent, the HR factor strongly correlates with the configuration coordinates. All data are publicly available at https://h-bn.info, along with a new application programming interface (API) to facilitate integration with machine learning workflows. This database is therefore designed to bridge the gap between theory and experiment, aid in the reliable identification of quantum emitters, and support the development of machine-learning-driven approaches in quantum materials research.

Keywords: hexagonal boron nitride, defects, density functional theory, database, bulk

1. INTRODUCTION

Solid-state quantum emitters have become essential building blocks for optical quantum technologies, owing to their ability to generate on-demand single photons that serve as flying qubits in applications such as quantum communication [1, 2], quantum memory [3–5], quantum computing [6, 7], and quantum sensing [8, 9]. A variety of platforms, including quantum dots [10, 11], nitrogen-vacancy (NV) centers in diamonds [12, 13], silicon carbide (SiC) [14], two-dimensional transition-metal dichalcogenides (2D TMDs) [15, 16], and defects in hexagonal boron nitride (hBN) [17, 18], have demonstrated single-photon emission under different operating conditions. While epitaxial quantum dots exhibit excellent photon purity, they typically require cryogenic operation [10, 11]. Conversely, bulk crystal defects can function at room temperature but suffer from limited photon out-coupling efficiency due to their three-dimensional geometry [19]. Two-dimensional materials thus offer a compelling compromise. That is, their atomically thin form

reduces the total internal reflection; hence, enhancing the out-coupling efficiency.

Among 2D materials, TMD emitters have shown promise, but to date, most demonstrations remain confined to temperatures below 15 K [15, 16]. In contrast, color centers in hBN operate reliably at room temperature, exhibit high quantum efficiency [18], robustness against aging [20, 21], stability [22], and benefit from high out-coupling due to the atomically thin host lattice [17]. In hBN, a single photon is emitted when an optically excited electron relaxes through a two-level defect state localized within the wide band gap (∼6 eV) of the material. Given the large band gap, hBN has been proven to accommodate a multitude of defect-induced mid-gap states, giving rise to a rich spectrum of optical transitions [23]. Experimentally, quantum emitters in hBN have been observed at photon energies ranging near 1 eV [24], 2 eV [25, 26], 3 eV [27], and 4 eV [28]. This wide range of emission energies makes defect identification a central challenge in hBN research. Currently, only the V_B^{-1} center has a well-established match between theory and experiment [29]. For other proposed emitters such as $C_B V_N$ [30, 31], $C_B C_N$ [28], $C_B C_N C_N$ [32], $C_B C_N C_B C_N$ [33–36], $O_N V_B^{-1}$ [37], a clear consensus of all photophysical properties between experiment and theory remains

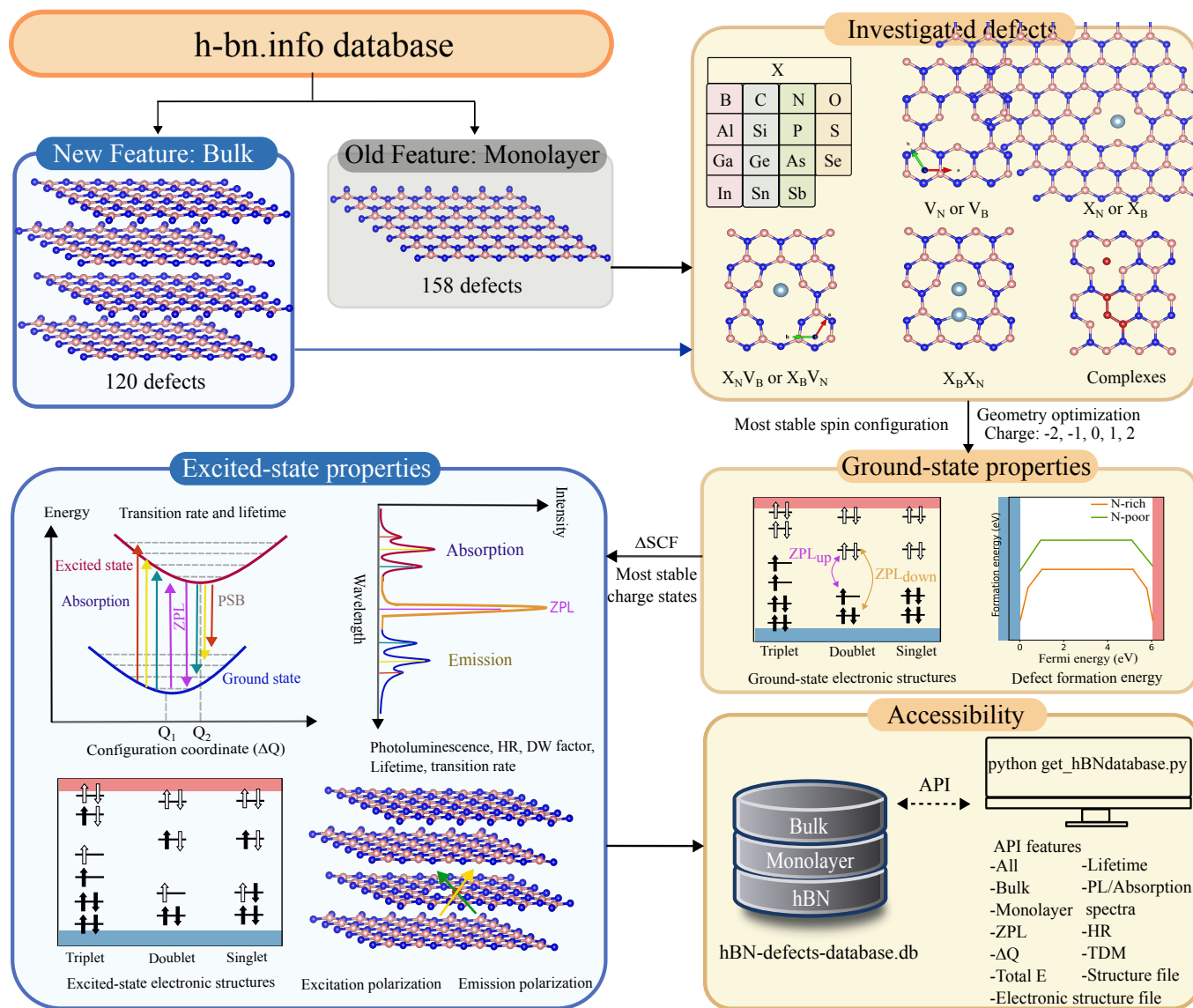* chanaprom.cholsuk@tum.de
† tobias.vogl@tum.de

Figure 1. Overview of the `h-bn.info` database workflow. The updated database extends our previous work on monolayer hBN by introducing a new feature: bulk hBN defect structures. Investigated defects include native vacancies, substitutional impurities (from groups III–VI), and various complexes, with charge states ranging from –2 to +2. For each defect, the most stable charge and spin configurations are identified. Ground-state properties such as defect formation energies and electronic structures are computed, followed by excited-state properties including ZPL, PL and absorption spectra, configuration coordinates, HR and DW factors, radiative lifetimes, and transition dipole polarizations. All data are compiled into a structured database with an accessible API supporting Python-based queries for integration with ML and data-driven research workflows. The color scheme is defined as follows: blue blocks represent features available only for bulk, while yellow blocks represent features available for both bulk and monolayer.

elusive.

This challenge is based on the fact that many different defects can produce similar signatures, such as overlapping zero-phonon line (ZPL) energies, comparable excited-state lifetimes, similar EPR spectra, and Raman features [38]. A major bottleneck in resolving this issue is the absence of a comprehensive and systematic database covering the photophysical properties of a large set of hBN defects. Therefore, robust defect identification requires a broad and detailed characterization of a

wide range of properties across many possible defects [35]. To mitigate the high computational demand associated with first-principles calculations, several machine learning (ML) approaches have been developed to predict photophysical properties across large defect spaces [39, 40]. However, these models still rely on DFT benchmarks for validation, as the photophysical properties of hBN defects do not follow simple patterns based on chemical type, symmetry, or defect class. The inherent variability and randomness in defect-related optical properties thus

necessitate a comprehensive, DFT-validated repository that includes both ground- and excited-state characteristics to support reliable identification and interpretation of experimental results.

To address this gap, we previously introduced a hBN defect database [41]. However, it was limited to only monolayer hBN and focused primarily on ground-state electronic properties. Since most experimental studies are conducted on bulk hBN samples, there is a pressing need to extend the theoretical framework to suitably accommodate experimental conditions. In this work, we therefore expand the database to include defect properties in bulk hBN and incorporate excited-state photophysical properties. In contrast to our earlier version, as summarized in Fig. 1, the newly developed database includes new defects and an additional analysis of charge stability, where defect charge states ranging from -2 to $+2$ are systematically evaluated. For each defect, the most stable charge and spin configurations are identified and subsequently used for detailed calculations of both ground-state and excited-state properties. All computed data are compiled into a publicly accessible database available at https://h-bn.info. Importantly, the database now provides comprehensive data for both bulk and monolayer hBN, allowing users to select the most relevant system for their research needs. As machine learning models are becoming increasingly popular in materials research, we also provide an application programming interface (API) for automated data retrieval to support broader usability and facilitate integration with ML workflows. As a result, this work serves as a vital resource for guiding experimental findings, supporting defect identification, and facilitating data-driven approaches in quantum materials research.

2. METHOD

This section describes the methodology used to generate the data presented in this work. While each method has been described already in Ref. [41], we still review the details here for the sake of completeness. Each block in Fig. 1 corresponds to a step in the computational pipeline and is detailed below.

A. DFT calculation details

All first-principles calculations in this work were performed using the Vienna Ab initio Simulation Package (VASP) [42, 43], employing the projector augmented-wave (PAW) method for treating core–valence interactions [44, 45]. To accurately capture the electronic structure of wide-bandgap hBN, we employed the screened hybrid functional HSE06, modified with a Hartree–Fock exchange mixing parameter α=0.32, which yields a calculated band gap of 6.09 eV for hBN, in excellent agreement with experiment [46]. A single Γ-point was used for

Brillouin-zone sampling in all defect calculations, and a plane-wave cutoff energy of 500 eV was applied. Spin polarization was included throughout to ensure spin conservation in electronic transitions.

Experimentally, bulk hBN can adopt several stacking sequences, most notably AA' (boron atoms in one layer sit directly above nitrogen atoms in the adjacent layer) and AB (each successive layer is translated by one in-plane bond length) [47]. In this study, we employ the AA' configuration exclusively, as total-energy calculations based on hybrid functional identify it to be slightly more stable than the AB stacking [48, 49]. The van der Waals interactions were included via the D3 correction of Grimme et al. [50].

B. Selection of investigated defects

To systematically explore defects in bulk hBN, we considered native defects, X_N, X_B, $X_N V_B$, $X_B V_N$, $X_B X_N$, and complexes where X is an element from groups III - VI indicated in Fig. 1.

All defect structures were generated within a $6\times6\times2$ supercell (288 atoms), as illustrated in Fig. 1, to minimize spurious defect–defect interactions. Ionic positions were relaxed until residual forces fell below 10^{-2} eV/Å and total energy changes were smaller than 10^{-4} eV. Note that full relaxation of both atomic positions and lattice parameters was also tested and found to cause zero to 20 nm variation in the ZPL, depending on the defect type. Given this minor effect, we chose to relax only the ionic positions. In total, 124 distinct neutral defect configurations were constructed and carried forward for subsequent charge stability analysis.

C. Ground-state properties

1. Ground-state electronic structures

For hBN defects, the most commonly encountered spin configurations include singlet, doublet, and triplet states, each corresponding to a different number of unpaired electrons. In this work, we determine the preferred spin configuration for each defect by comparing the total energies of all relevant spin states. Only the lowest-energy configuration is selected for subsequent calculations.

2. Defect formation energy

Defect formation energy is a key quantity used to determine the most stable charge state of a defect under given thermodynamic conditions. For each defect in this study, charge states ranging from -2 to $+2$ were systematically evaluated by computing their formation energies

using the following expression

$$E^f[X^q] = E_{\text{tot}}[X^q] - E_{\text{tot}}[\text{hBN}] - \sum_i n_i \mu_i$$
$$+ q(\epsilon_{\text{vbm}} + \epsilon_{\text{Fermi}}) + E_{corr}(q), \quad (1)$$

where $E_{\text{tot}}[X^q]$ is the total energy of the supercell containing the defect X in charge state q; $E_{\text{tot}}[\text{hBN}]$ is the total energy of pristine hBN; The term n_i represents the number of atoms added or removed from the system, with μ_i being a chemical potential of atom species i; The Fermi level ϵ_{Fermi} corresponds to the electronic chemical potential, referenced to the valence band maximum (VBM) of pristine hBN, denoted by ϵ_{vbm}. The final term, $E_{corr}(q)$, is a correction for spurious electrostatic interactions arising from periodic boundary conditions in charged supercells [51]. In this work, we employed the Spinney Python code for charge correction [52]. Although charge states q greater than $+3$ or q less than -3 may be possible for certain defects, they were not considered in this work. Future studies may extend the analysis to include these regimes in order to identify additional stable defects and potential quantum emitters.

By comparing the formation energies across charge states, we identify the most stable charge configuration for each defect, which is then used in all subsequent ground- and excited-state property calculations.

D. Excited-state properties

1. Excited-state electronic structures

To simulate excited-state configurations, we employ the so-called ΔSCF method, a constrained DFT approach in which the occupation of electronic states is manually controlled [53]. In this method, one electron is explicitly promoted to an excited state, allowing us to model the electron excitation.

For spin-polarized systems, such as defects with doublet or triplet ground states, two types of spin-conserving excitations may be possible: spin-up to spin-up or spin-down to spin-down. Both transitions are considered in this work where relevant. In contrast, for singlet defects, only one transition channel is evaluated unless otherwise specified. This ensures consistency with spin-selection rules and captures the dominant optical transitions for each defect configuration.

2. Zero phonon line (ZPL) calculation

In experimental photoluminescence spectra, the sharpest and most prominent peak is typically associated with the ZPL, which corresponds to a purely electronic transition occurring without the involvement of phonons. In this work, the ZPL is calculated as the total energy difference between the fully relaxed excited-state and ground-state geometries as follows:

$$\text{ZPL} = E_{total,excited} - E_{total,ground}. \quad (2)$$

While the zero-point vibrational energies (ZPE) can slightly shift the ZPL, typically in the range of a few meV to 0.1 eV, depending on the defect [54, 55], they are neglected in this study under the assumption that ZPE contributions approximately cancel between the ground and excited states. To ensure spin conservation, we treat spin-up and spin-down transitions separately. For certain defects, this results in two distinct spin-conserving excitation pathways, each yielding its own ZPL.

3. Configuration coordinates

Due to the differences in electronic occupation between the ground and excited states, the atomic geometries of a defect can undergo significant structural relaxation upon excitation. This structural change can be quantified by the so-called configuration coordinate (q_k), which is defined as

$$q_k = \sum_{\alpha, i \in x, y, z} \sqrt{m_\alpha} \left(R_{e,\alpha i} - R_{g,\alpha i} \right) \Delta r_{k,\alpha i}, \quad (3)$$

where α indexes the atoms; i denotes the Cartesian components (x, y, z); and m_α is the mass of atom α. The terms $R_{g,\alpha i}$ and $R_{e,\alpha i}$ correspond to the equilibrium positions of atom α in the ground and excited states, respectively; and $\Delta r_{k,\alpha i}$ is the normalized eigenvector of the phonon mode k, indicating the direction of atomic displacement.

4. HR and DW factors

The Huang-Rhys (HR) factor (S) can be used to quantify the strength of the electron-phonon coupling. This can be computed from the partial HR factor s_k of each phonon mode k, defined as

$$s_k = \frac{\omega_k q_k^2}{2\hbar}, \quad (4)$$

where ω_k is the frequency of the phonon mode k; and q_k is the configuration coordinate displacement corresponding to that mode, as defined previously. Then the HR factor is obtained by summing over all phonon modes, and its spectral distribution can be expressed as

$$S(\hbar\omega) = \sum_k s_k \delta(\hbar\omega - \hbar\omega_k), \quad (5)$$

where δ is the Dirac delta function. In addition to S, the Debye–Waller (DW) factor is also computed to characterize the fraction of emission occurring in the ZPL, which is given by $\text{DW} = e^{-S}$.

5. *Absorption and photoluminescence spectra*

In this work, the photoluminescence $L(\hbar\omega)$ spectrum is calculated from the full-phonon simulation [56], which can be obtained from

$$L(\hbar\omega) = C\omega^3 A(\hbar\omega), \qquad (6)$$

where C is a normalization constant (typically determined by fitting to experimental data), and $A(\hbar\omega)$ is the optical spectral function that captures both ZPL and phonon sidebands, which is computed as

$$A(E_{ZPL} - \hbar\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} G(t)\exp(-i\omega t - \gamma|t|)dt, \quad (7)$$

where γ is a broadening parameter, and $G(t)$ is the generating function defined as

$$G(t) = \exp(S(t) - S(0)). \qquad (8)$$

Here, $S(t)$ is the time-dependent spectral function given by

$$S(t) = \int_0^{\infty} S(\hbar\omega)\exp(-i\omega t)d(\hbar\omega), \qquad (9)$$

where $S(\hbar\omega)$ represents the phonon-resolved Huang–Rhys factor as described in the previous section. The convergence test of the supercell size is shown in Supplementary S1.

6. *Transition rate and lifetime*

A radiative transition rate quantifies how quickly a transition occurs between two defect states with the same spin polarization. It is given by

$$\Gamma_{\mathrm{R}} = \frac{n_D e^2}{3\pi\epsilon_0 \hbar^4 c^3} E_0^3 \mu_{\mathrm{e-h}}^2, \qquad (10)$$

where Γ_{R} is the radiative transition rate; e is the elementary charge; ϵ_0 is vacuum permittivity; E_0 denotes the ZPL energy; $\mu_{\mathrm{e-h}}^2$ is the modulus square of the dipole moment computed by Eq. (12); and n_{D} is 1.85, which is the refractive index of the host hBN in the visible [19].

Since the refractive index of hBN can vary between samples or experimental conditions, our database provides an interactive feature that allows users to input a custom refractive index. The transition rate and corresponding radiative lifetime are automatically updated based on the user-defined value. The radiative lifetime (τ_R) is computed as the inverse of the transition rate

$$\tau_R = \frac{1}{\Gamma_{\mathrm{R}}}. \qquad (11)$$

7. *Transition dipole moment and polarization*

The polarization characteristics of excitation and emission processes can be derived from the transition dipole moments. To compare theoretical predictions with polarization-resolved PL experiments, both excitation and emission dipoles are rotated by 90° and mapped modulo 60° to obtain the nearest alignment relative to the hexagonal crystal axes. Although the hexagonal symmetry is 120°, angles beyond 180° are equivalent due to rotational symmetry. To ensure consistency with experimental conditions, dipole vectors are projected onto the xy-plane of the hBN crystal, allowing the computation of the in-plane polarization visibility.

The transition dipole moments are calculated from the electronic wavefunctions obtained after structural relaxation. These wavefunctions are extracted using the Python package PyVaspwfc [57]. However, since the wavefunctions in the ground and excited states originate from different geometries, two distinct dipole types must be considered: excitation and emission dipoles. The excitation dipole corresponds to the transition from the wavefunction of the optimized ground state to that of the optimized excited state. In contrast, the emission dipole corresponds to the transition from the optimized excited-state wavefunction to the optimized ground-state wavefunction. The transition dipole moment can be obtained from

$$\boldsymbol{\mu} = \frac{i\hbar}{(E_f - E_i)m}\langle\psi_f|\,\mathbf{p}\,|\psi_i\rangle, \qquad (12)$$

where $\psi_{i/f}$ are the initial and final wavefunctions; $E_{i/f}$ are the corresponding eigenvalues; m is the electron mass; and $\mathbf{p}$ is the momentum operator. Because the involved wavefunctions come from geometrically distinct configurations, a modified version of PyVaspwfc [58] is used to correctly evaluate the dipole moments. The resulting dipole vector $\boldsymbol{\mu}$ may have components in all three Cartesian directions and can be expressed as

$$\boldsymbol{\mu} = |\mu_x|\hat{x} + |\mu_y|\hat{y} + |\mu_z|\hat{z}. \qquad (13)$$

If $\mu_z = 0$, the transition dipole lies entirely in the xy-plane, indicating a purely in-plane polarization. This behavior is further analyzed using the linear in-plane polarization visibility, which quantifies the extent to which the emission or absorption is confined within the plane of the hBN crystal.

E. **Database repository**

The hBN defect database is stored as a single SQLite file, `hbn_defects_database.db`, in the GitHub repository `QCS-Theory/hBN-database`. It contains a primary table, `updated_data`, with the following schematics as listed in Tab. I.

To provide access, we use the GitHub REST API `contents` endpoint. Users type

Column name	Type and description	Option
Host	TEXT: host material identifier	-
Defect	TEXT: defect chemical formula/code	-
Defect name	TEXT: descriptive name of the defect	-
Charge state	INTEGER: integer charge state	-
Spin multiplicity	TEXT: spin configuration of defects	-
Optical spin transition	TEXT: allowed optical spin transition	-
Excitation properties: dipole_x (Debye)	REAL: x-component of excitation dipole moment	abs_dipole_x
Excitation properties: dipole_y (Debye)	REAL: y-component of excitation dipole moment	abs_dipole_y
Excitation properties: dipole_z (Debye)	REAL: z-component of excitation dipole moment	abs_dipole_z
Excitation properties: linear In-plane Polarization Visibility	REAL: Visibility	abs_visibility
Excitation properties: Intensity (Debye)	REAL: Strength of excitation transition dipole moment	abs_tdm
Excitation properties: Characteristic time (ns)	REAL: Time of excitation transition	abs_lifetime
Excitation properties: Angle of excitation wrt the crystal axis	REAL: excitation polarization angle	abs_angle
Emission properties: dipole_x (Debye)	REAL: x-component of emission dipole moment	ems_dipole_x
Emission properties: dipole_y (Debye)	REAL: y-component of emission dipole moment	ems_dipole_y
Emission properties: dipole_z (Debye)	REAL: z-component of emission dipole moment	ems_dipole_z
Emission properties: linear In-plane Polarization Visibility	REAL: Visibility	ems_visibility
Emission properties: Intensity (Debye)	REAL: Strength of emission transition dipole moment	ems_tdm
Emission properties: ZPL (eV)	REAL: ZPL energy	ZPL
Emission properties: ZPL (nm)	REAL: ZPL wavelength	ZPL_nm
Emission properties: lifetime (ns)	REAL: Radiative lifetime of emission	lifetime
Emission properties: Angle of emission wrt the crystal axis	REAL: emission polarization angle	ems_angle
Emission properties: Polarization misalignment (degree)	REAL: dipole misalignment angle	misalignment
Emission properties: Configuration coordinate ($amu^{1/2}$/Å)	REAL: Q value	Q
Emission properties: HR factor	REAL: S value	HR
Emission properties: DW factor	REAL: DW value	DW
Emission properties: Ground-state total energy (eV)	REAL: total energy	E_ground
Emission properties: Excited-state total energy (eV)	REAL: total energy	E_excited
Ground-state structure	BLOB: atomic structure file (CIF file)	structure_ground
Excited-state structure	BLOB: atomic structure file (CIF file)	structure_excited
Ground-state electronic structure	BLOB: raw OUTCAR file (VASP format) for electronic structure	band_ground
Excited-state electronic structure	BLOB: raw OUTCAR file (VASP format) for electronic structure	band_excited
PL lineshape	BLOB: raw PL lineshape file with the broadening parameter (γ) in Eq. (7) equal to 1	PL
Raman spectrum	BLOB: raw Raman spectrum file	raman

Table I. Structure of the updated_data table in hbn_defects_database.db. The leftmost column provides the database field names, and the central column specifies the corresponding field's data type along with a concise description. The rightmost column lists the retrieval option keys used by get_database(option=["..."]). The first six rows are always returned in the get_database() function. See Sec. 4 for instructions in detail.

```
GET https://raw.githubusercontent.com/
    QCS-Theory/hBN-database/main/
    hbn_defects_structure.db
```

This returns the SQLite binary, which users can save locally and open with any SQLite-compatible tool. Alternatively, in Python, one can call get_database(...) function to load and (optionally) filter the data. See Sec. 4 for step-by-step procedures.

3. RESULTS AND DISCUSSION

In this section, we first analyze correlations among key photophysical parameters. We then investigate the impact of vacancy on PSB broadening. Finally, we present the overall distribution of hBN defects in a histogram.

A. Correlation matrix

The correlation matrix among various photophysical parameters is computed using the Spearman rank correlation method, as shown in Fig. 2. Overall, most parameters exhibit low correlation coefficients, indicating that they are largely independent of one another. However, notable correlations are observed between the configuration coordinate and the HR factor.

These dependencies are expected and can be explained analytically. As shown in Eqs. (4) and (5), the HR factor is proportional to the square of the configuration coordinate, making their correlation inherently strong. This relationship validates the internal consistency of the computed dataset.

B. Relationship between the number of vacancies and PSB broadening

We now investigate the behavior of the HR factor, which is directly related to PSB broadening and thus reflects the strength of electron–phonon coupling. In Fig. 3, defects are categorized into four groups based on their vacancy content: no vacancy with a single impurity, no vacancy with complexes, one vacancy, and two vacancies.

Fig. 3(a) shows the correlation between the Q and the HR factor. As expected from theory, the HR factor increases approximately with the square of Q, confirming the analytical relationship. Notably, defects without vacancies tend to exhibit lower values of both Q and the HR factor. In contrast, defects with one or two vacancies span a wider range of values, with a significant fraction showing strong electron–phonon coupling (i.e., higher Q and HR values). This trend is further supported by the histograms shown in Figs. 3(b) and 3(c), where the distributions of the high HR factor and high Q mostly belong to the defects with vacancies.

PL lineshape comparisons in Fig. 3(d) further reinforce this observation. Defects without vacancies, including isolated point defects and complexes, typically exhibit the HR factors in the range of 0-2 and 0-3.8, respectively. In contrast, defects containing vacancies show significantly broader HR distributions: 0–12 for single-vacancy defects and 3.5–4 for those with two vacancies. As reflected in the PL lineshapes, vacancy-containing defects tend to display broader PSB. These results suggest that having more vacancies tends to induce stronger lattice distortion upon excitation compared to having no vacancy. This enhanced structural relaxation leads to greater geometric variation between the ground and excited states, thereby increasing the Q and, consequently, the HR factor.

C. Distribution of hBN defects

To further analyze trends across different defect classes, Fig. 4(a) presents the distribution of ZPL energies for defects categorized by the number of vacancies. The results suggest that there is no clear correlation between ZPL energy and the number of vacancies: all three classes (no vacancy, one vacancy, two vacancies) span a wide ZPL range. Interestingly, vacancy-free defects with a single impurity appear to populate both the low and high ends of the ZPL spectrum.

At the lowest ZPL bin, the O_N defect is found, with a ZPL energy of 0.24 eV. However, when taking defect formation energies into account, this configuration is energetically less favorable compared to its positively charged counterpart (O_N^{+1}), suggesting that such a low-energy transition may not be experimentally observable under typical conditions.

On the opposite end, the Ga_B defect exhibits a ZPL of 5.82 eV, which is the highest in the dataset. This configuration is thermodynamically stable based on its defect formation energy. However, its electronic structure does not feature well-defined two-level mid-gap states. This behavior can be attributed to the fact that Ga and B belong to the same group in the periodic table, resulting in no difference in the number of valence electrons. Consequently, the electronic transition involved in Ga_B closely resembles a near-band-edge transition, yielding a ZPL value approaching the intrinsic bandgap of hBN.

For defects containing complexes with one or more vacancies, the ZPL energies consistently fall within the 1–4 eV range. This trend can be attributed to the nature of vacancies, which typically introduce localized electronic states within the mid-gap region of hBN. Given the intrinsic bandgap of approximately 6 eV, these vacancy-induced states often lie near the center of the gap. As a result, optical transitions between such defect levels commonly span energy differences of 1–4 eV, leading to ZPL values confined within this range.

Fig. 4(b) presents the distribution of misalignment between excitation and emission polarizations. The results show that there is no correlation between the misalignment and the number of vacancies. For the majority of defects, excitation and emission dipoles are nearly parallel (minimal misalignment). Since the polarization angles are mapped modulo 60° to account for the hexagonal crystal symmetry, most values cluster near 0°. However, one obvious outlier with a misalignment of 51° is identified as the $O_N V_B^{-1}$ defect. Further analysis reveals that this large offset arises from a strong difference in polarization visibility: the excitation dipole is primarily in-plane and linearly polarized, while the emission dipole has an out-of-plane component. As such, the polarization misalignment is substantial.

Lastly, turning to the radiative lifetime distribution shown in Fig. 4(c), the result suggests that most hBN defects exhibit lifetimes ranging from approximately 1 ns to 100 μs. Furthermore, there is no clear correlation
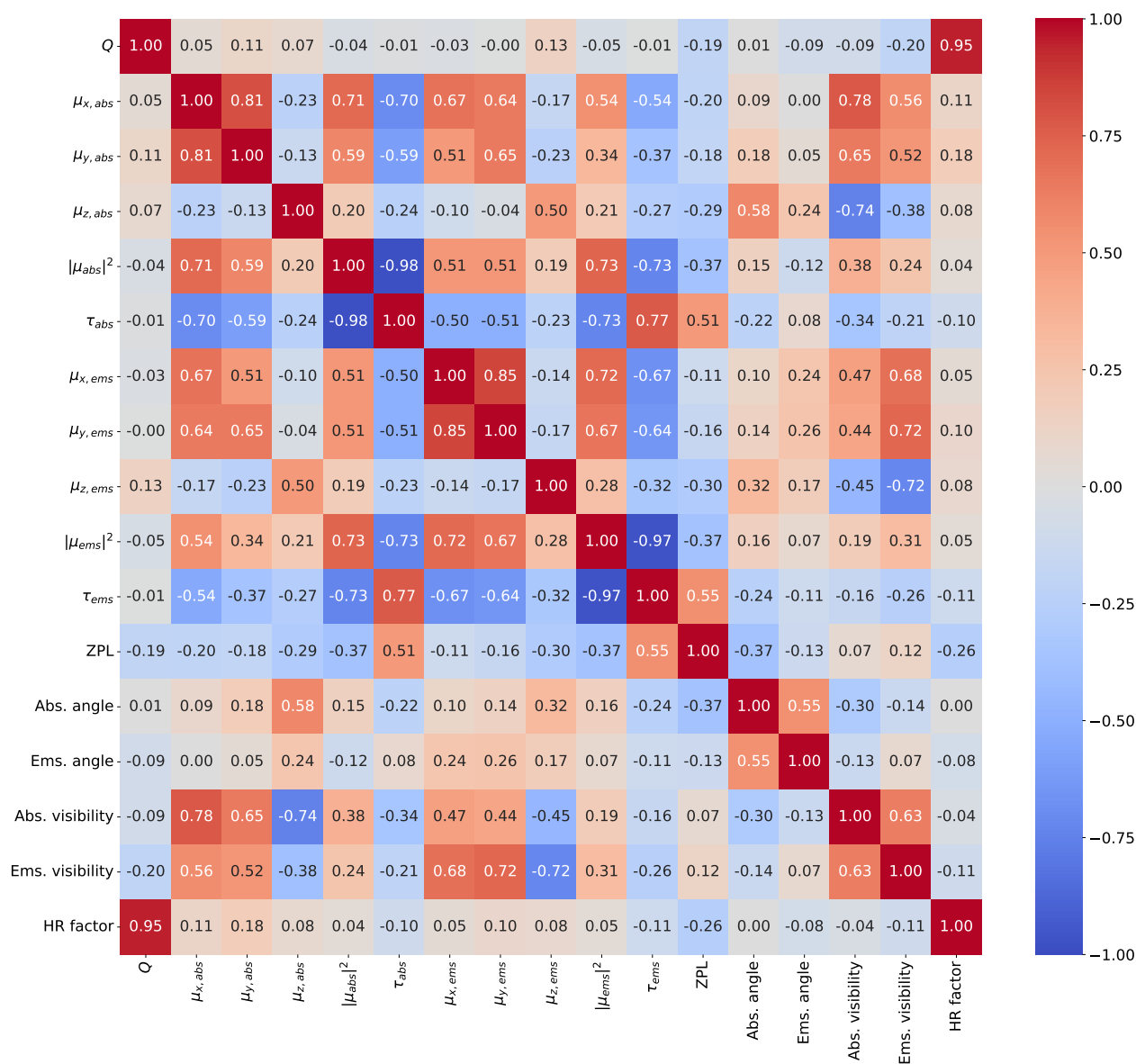
Figure 2. Correlation matrix based on the Spearman rank correlation method among 17 properties. High absolute values indicate strong correlations, while values close to zero imply weak or no correlation. Positive values represent increasing relationships between property pairs, whereas negative values indicate inverse relationships.

between the lifetime and the number of vacancies. More broadly, no systematic relationship is observed between the number of vacancies and other key defect properties, as further illustrated in Supplementary S2. Full example calculations of defect properties in the database are given in Supplementary S3.

4. PROGRAMMATIC RETRIEVAL OF THE HBN DEFECT DATABASE VIA THE GITHUB API

In this section, we describe a procedure for accessing the hBN defect database directly from our GitHub repository using the GitHub REST API, apart from accessing our website https://h-bn.info. We provide a Python module, get_hBN_defects_database.py, which automates (i) accessing the raw SQLite database and (ii) extracting a user-specified subset of entries. This functionality enables seamless retrieval of the complete dataset, thereby facilitating integration with machine
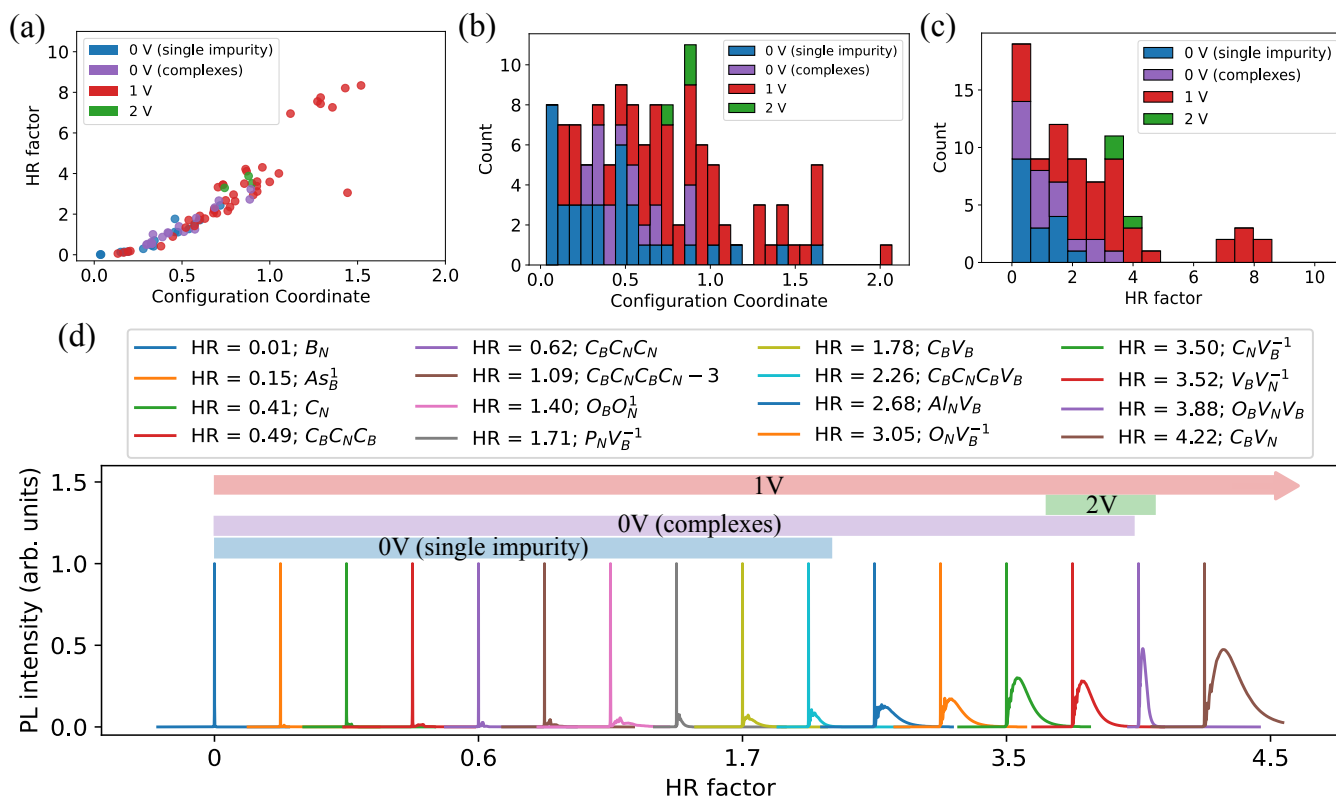
Figure 3. (a) Correlation between configuration coordinate and HR factor. (b) and (c) show histograms of the configuration coordinate between 0 and 2.5 amu$^{1/2}$Å and HR factor between 0 and 10, respectively. Defect types are color-coded as follows: blue for single-impurity defects without vacancies; purple for vacancy-free defect complexes (comprising multiple impurities); red for defects with a single vacancy; and green for defects with two vacancies. (d) displays selected but representative PL lineshapes categorized by the number of vacancies in each defect. The color stripe indicates the corresponding HR factor range for each defect category.

learning workflows. The step-by-step procedure is as follows:

1. **Obtain the retrieval script.**
   Download `get_hBN_defects_database.py` and place it in your working directory.

2. **Import the retrieval function.**
   In your Python script or interactive session, import the function

   ```
   from get_hBN_defects_database import get_database
   ```

3. **Fetch and filter the database.**
   Invoke `get_database()` with the desired filtering criteria

   ```
   data = get_database(
       option=["ZPL"],
       host=["monolayer", "bulk"],
       spin_multiplicity=["singlet", "doublet",
       "triplet"],
       charge_state=[-2, -1, 0, 1, 2],
       optical_spin_transition=["up", "down"],
   ```

   ```
       value_range=(2.0, 4.0),
       download_db=False
   )
   ```

The keyword arguments perform the following functions:

option: Specifies which database columns to return. The complete set of valid keys is listed in Tab. I. To retrieve all columns, use

   ```
   option = ["all"]
   ```

host: selects between the monolayer and bulk hBN datasets. By default, both are returned.

spin_multiplicity: filters defects by their spin multiplicity. If omitted, all multiplicities are included.

charge_state: filters defects by charge state. Defaults to all if unspecified.

optical_spin_transition: filters by optical spin transition (i.e. "up" refers up to up or "down" refers to down to down). Both are returned if this argument is not provided.
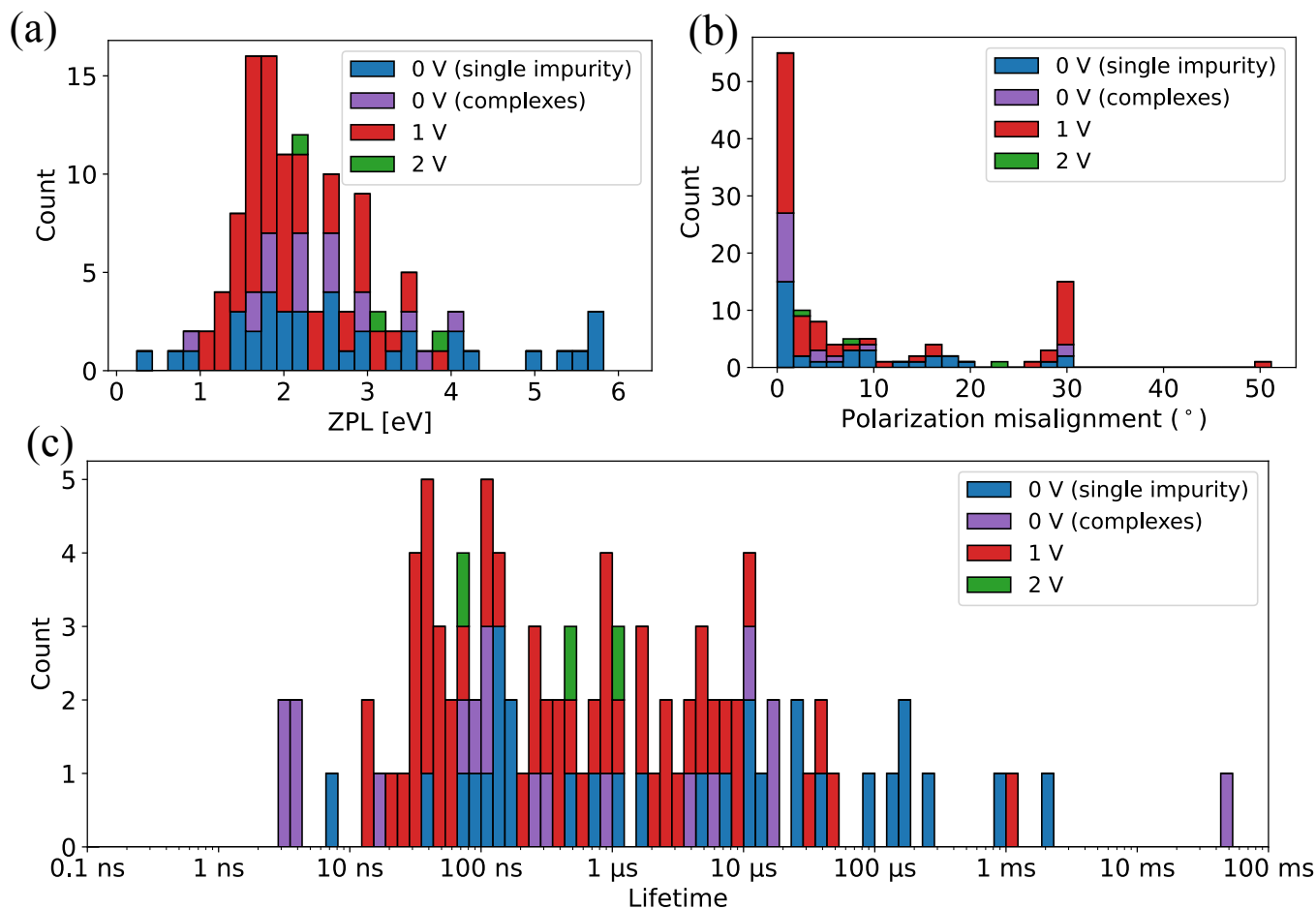
Figure 4. (a) Distribution of ZPL energies. (b) Distribution of polarization misalignment angles. (c) Distribution of radiative lifetimes. Blue bins (0V) represent single-impurity defects without vacancies; purple bins (0V) represent vacancy-free defect complexes (containing more than two impurities); red bins (1V) correspond to defects with one vacancy; and green bins (2V) represent defects containing two vacancies.

value_range: restricts the numeric range of the selected property. When omitted, no range filtering is applied.

download_db: If set to `True`, one can download the raw SQLite database file (named `hbn_defects_<options>.db`) to the working directory.

4. **Inspect the results.**
   The object `data` is a `pandas.DataFrame` containing the filtered subset, ready for downstream analysis or visualization. If `download_db=True`, the corresponding `.db` file will also be saved locally.

The complete code examples are available in Supplementary S4.

5. FURTHER UPDATES TO THE DATABASE

In addition to the completely new database of both ground and excited state properties in bulk hBN presented in this work, our hBN defects database has also undergone other updates since its initial release. This includes

1. **Version 1.1**: Incorporation of configuration coordinate data for 277 monolayer hBN defects.

2. **Version 1.2**: Inclusion of 554 total energy values for each structure, covering both ground- and excited-state configurations for monolayer hBN defects.

3. **Version 1.3**: Inclusion of Raman spectra for 100 defects in hBN monolayers obtained from Ref. [38]

6. CONCLUSION

In summary, we have established an extended database of hBN defects that includes a comprehensive set of defects in bulk hBN, along with their excited-state photophysical properties. In addition to our previous work focusing on monolayer hBN and ground-state characteristics, this new database incorporates additional defect types in bulk hBN, charge-state analysis, and excited-state properties. Key properties such as ZPL, PL lineshape, absorption lineshape, HR factor, interactive radiative lifetimes, transition dipole moments, and polarization characteristics have been systematically computed.

From the current dataset, we computed a correlation matrix and found that only the configuration coordinate and HR factor are strongly correlated, consistent with their analytical relationship. Further analysis showed that defects with vacancies tend to exhibit larger lattice distortions, leading to higher configuration coordinates and HR factors, resulting in broader PSBs. This suggests that broadened PSBs in low-temperature PL may serve as a possible indicator of vacancy-containing defects. Moreover, we analyzed the distribution of all parameters and found no additional strong correlations.

We stress that the primary goal of our work is to construct a large database covering comprehensive properties of many different defects. This is to serve as a screening tool that helps rule out unlikely defects and narrow down promising candidates. This work does not take into account symmetry reduction effects such as the Jahn–Teller distortion; therefore, the specific values of the ZPL may be shifted. Such effects should be carefully considered when determining a defect as a promising candidate. For instance, to use our database, if an experiment reports a ZPL at 2.2 eV, one should focus on defects with calculated ZPLs between 2.0 and 2.4 eV, taking into account the finite accuracy of DFT calculations. This approach effectively eliminates many candidates and suggests the most plausible ones.

Finally, all data are made publicly available through an online platform https://h-bn.info, alongside the GitHub REST API. This enables programmatic retrieval of the complete SQLite database or user-specified subsets of defect properties. The API returns raw or filtered `.db` files on demand, and can be seamlessly integrated into automated pipelines and ML workflows. This resource is therefore intended to assist researchers in experimental interpretation, accelerate defect identification, and facilitate data-driven discovery in the field of solid-state quantum emitters.

DATA AVAILABILITY

All raw data from this work is available from the authors upon reasonable request. The database with browsing interfaces is freely accessible at https://h-bn.info. The SQLite database file can be downloaded via the GitHub REST API.

NOTES

The authors declare no competing financial interest. We acknowledge the assistance of OpenAI's ChatGPT in managing code snippets. All final decisions on integration and adaptation of those suggestions were made by the authors.

ACKNOWLEDGMENTS

[1] N. Gisin, G. Ribordy, W. Tittel, and H. Zbinden, Quantum cryptography, Rev. Mod. Phys. **74**, 145 (2002).

[2] N. Ahmadi, S. Schwertfeger, P. Werner, L. Wiese, J. Lester, E. Da Ros, J. Krause, S. Ritter, M. Abasifard, C. Cholsuk, and et al., QUICK³ - design of a satellite-based quantum light source for quantum communication and extended physical theory tests in space, Adv. Quantum Technol. **7**, 2300343 (2024).

[3] A. Reiserer, N. Kalb, M. S. Blok, K. J. Van Bemmelen, T. H. Taminiau, R. Hanson, D. J. Twitchen, and M. Markham, Robust Quantum-Network Memory Using Decoherence-Protected Subspaces of Nuclear Spins, Phys. Rev. X **6**, 021040 (2016).

[4] C. Cholsuk, A. Çakan, S. Suwanna, and T. Vogl, Identifying electronic transitions of defects in hexagonal boron nitride for quantum memories, Adv. Opt. Mater. **12**, 2302760 (2024).

[5] T. Nateeboon, C. Cholsuk, T. Vogl, and S. Suwanna, Modeling the performance and bandwidth of single-atom adiabatic quantum memories, APL Quantum **1**, 026107

(2024).

[6] T. D. Ladd, F. Jelezko, R. Laflamme, Y. Nakamura, C. Monroe, and J. L. O'Brien, Quantum computers, Nature **464**, 45 (2010).

[7] S. Pezzagna and J. Meijer, Quantum computer based on color centers in diamond, Applied Physics Reviews **8**, 011308 (2021).

[8] H. L. Stern, C. M. Gilardoni, Q. Gu, S. Eizagirre Barker, O. F. J. Powell, X. Deng, S. A. Fraser, L. Follet, C. Li, A. J. Ramsay, H. H. Tan, I. Aharonovich, and M. Atatüre, A quantum coherent spin in hexagonal boron nitride at ambient conditions, Nat. Mater. **23**, 1379 (2024).

[9] A. Gottscholl, M. Diez, V. Soltamov, C. Kasper, D. Krauße, A. Sperlich, M. Kianinia, C. Bradac, I. Aharonovich, and V. Dyakonov, Spin defects in hBN as promising temperature, pressure and magnetic field quantum sensors, Nat Commun **12**, 4480 (2021).

[10] P. Senellart, G. Solomon, and A. White, High-performance semiconductor quantum-dot single-photon sources, Nat. Nanotechnol. **12**, 1026 (2017).

[11] H. Mantynen, N. Anttu, Z. Sun, and H. Lipsanen, Single-photon sources with quantum dots in iii-v nanowires, Nanophotonics **8**, 747 (2019).

[12] M. W. Doherty, N. B. Manson, P. Delaney, F. Jelezko, J. Wrachtrup, and L. C. Hollenberg, The nitrogen-vacancy colour centre in diamond, Physics Reports **528**, 1 (2013), the nitrogen-vacancy colour centre in diamond.

[13] M. V. G. Dutt, L. Childress, L. Jiang, E. Togan, J. Maze, F. Jelezko, A. S. Zibrov, P. R. Hemmer, and M. D. Lukin, Quantum Register Based on Individual Electronic and Nuclear Spin Qubits in Diamond, Science **316**, 1312 (2007).

[14] B. Lienhard, T. Schröder, S. Mouradian, F. Dolde, T. T. Tran, I. Aharonovich, and D. Englund, Bright and photostable single-photon emitter in silicon carbide, Optica **3**, 768 (2016).

[15] J. Klein, M. Lorke, M. Florian, F. Sigger, L. Sigl, S. Rey, J. Wierzbowski, J. Cerne, K. Müller, E. Mitterreiter, P. Zimmermann, T. Taniguchi, K. Watanabe, U. Wurstbauer, M. Kaniber, M. Knap, R. Schmidt, J. J. Finley, and A. W. Holleitner, Site-selectively generated photon emitters in monolayer MoS2 via local helium ion irradiation, Nature Communications **10**, 2755 (2019).

[16] Y. Lei, Q. Lin, S. Xiao, J. Li, and H. Fang, Optically active telecom defects in mote2 fewlayers at room temperature, Nanomaterials **13**, 10.3390/nano13091501 (2023).

[17] T. T. Tran, K. Bray, M. J. Ford, M. Toth, and I. Aharonovich, Quantum emission from hexagonal boron nitride monolayers, Nat. Nanotechnol. **11**, 37 (2016).

[18] N. Nikolay, N. Mendelson, E. Özelci, B. Sontheimer, F. Böhm, G. Kewes, M. Toth, I. Aharonovich, and O. Benson, Direct measurement of quantum efficiency of single-photon emitters in hexagonal boron nitride, Optica **6**, 1084 (2019).

[19] T. Vogl, M. W. Doherty, B. C. Buchler, Y. Lu, and P. K. Lam, Atomic localization of quantum emitters in multilayer hexagonal boron nitride, Nanoscale **11**, 14362 (2019).

[20] T. Vogl, G. Campbell, B. C. Buchler, Y. Lu, and P. K. Lam, Fabrication and deterministic transfer of high-quality quantum emitters in hexagonal boron nitride, ACS Photonics **5**, 2305 (2018).

[21] W. Liu, Y.-T. Wang, Z.-P. Li, S. Yu, Z.-J. Ke, Y. Meng, J.-S. Tang, C.-F. Li, and G.-C. Guo, An ultrastable and robust single-photon emitter in hexagonal boron nitride, Physica E Low Dimens. Syst. Nanostruct. **124**, 114251 (2020).

[22] M. Kianinia, B. Regan, S. A. Tawfik, T. T. Tran, M. J. Ford, I. Aharonovich, and M. Toth, Robust solid-state quantum system operating at 800 k, ACS Photonics **4**, 768 (2017).

[23] C. Cholsuk, S. Suwanna, and T. Vogl, Tailoring the emission wavelength of color centers in hexagonal boron nitride for quantum applications, Nanomaterials **12**, 2427 (2022).

[24] R. Camphausen, L. Marini, S. A. Tawfik, T. T. Tran, M. J. Ford, and S. Palomba, Observation of near-infrared sub-poissonian photon emission in hexagonal boron nitride at room temperature, APL Photonics **5**, 076103 (2020).

[25] A. Dietrich, M. Bürk, E. S. Steiger, L. Antoniuk, T. T. Tran, M. Nguyen, I. Aharonovich, F. Jelezko, and A. Kubanek, Observation of fourier transform limited lines in hexagonal boron nitride, Phys. Rev. B. **98**, 081414 (2018).

[26] M. Hazra, M. Rieger, A. Kumar, M. N. Mishuk, C. Cholsuk, K. Sripathy, V. Villafañe, K. Müller, J. J. Finley, and T. Vogl, Temperature-dependent emission spectroscopy of quantum emitters in hexagonal boron nitride (2025), arXiv:2507.06783 [physics.optics].

[27] A. Gale, C. Li, Y. Chen, K. Watanabe, T. Taniguchi, I. Aharonovich, and M. Toth, Site-Specific Fabrication of Blue Quantum Emitters in Hexagonal Boron Nitride, ACS Photonics **9**, 2170 (2022).

[28] M. Mackoit-Sinkevičienė, M. Maciaszek, C. G. V. de Walle, and A. Alkauskas, Carbon dimer defect as a source of the 4.1 ev luminescence in hexagonal boron nitride, Appl. Phys. Lett. **115**, 212101 (2019).

[29] A. Gottscholl, M. Kianinia, V. Soltamov, S. Orlinskii, G. Mamin, C. Bradac, C. Kasper, K. Krambrock, A. Sperlich, M. Toth, I. Aharonovich, and V. Dyakonov, Initialization and read-out of intrinsic spin defects in a van der waals crystal at room temperature, Nature Materials **19**, 540 (2020).

[30] A. Sajid and K. S. Thygesen, $V_N C_B$ defect as source of single photon emission from hexagonal boron nitride, 2D Mater. **7**, 031007 (2020).

[31] S. Li and A. Gali, Bistable carbon-vacancy defects in h-BN, Front. Quantum Sci. Technol. **1**, 1007756 (2022).

[32] K. Li, T. J. Smart, and Y. Ping, Carbon trimer as a 2 ev single-photon emitter candidate in hexagonal boron nitride: A first-principles study, Phys. Rev. Mater. **6**, L042201 (2022).

[33] A. Kumar, C. Cholsuk, A. Zand, M. N. Mishuk, T. Matthes, F. Eilenberger, S. Suwanna, and T. Vogl, Localized creation of yellow single photon emitting carbon complexes in hexagonal boron nitride, APL Mater. **11**, 071108 (2023).

[34] A. Kumar, Ç. Samaner, C. Cholsuk, T. Matthes, S. Paçal, Y. Oyun, A. Zand, R. J. Chapman, G. Saerens, R. Grange, S. Suwanna, S. Ateş, and T. Vogl, Polarization dynamics of solid-state quantum emitters, ACS Nano **18**, 5270 (2024).

[35] C. Cholsuk, S. Suwanna, and T. Vogl, Comprehensive scheme for identifying defects in solid-state quantum systems, J. Phys. Chem. Lett. **14**, 6564 (2023).

[36] M. Maciaszek and L. Razinkovas, Blue Quantum Emitter in Hexagonal Boron Nitride and a Carbon Chain Tetramer: a First-Principles Study, ACS Appl. Nano Mater. **7**, 18979 (2024).

[37] S. Li and A. Gali, Identification of an oxygen defect in hexagonal boron nitride, J. Phys. Chem. Lett. **13**, 9544–9551 (2022).

[38] C. Cholsuk, A. Cakan, V. Deckert, S. Suwanna, and T. Vogl, Raman signatures of single point defects in hexagonal boron nitride quantum emitters, arXiv.org, e-Print Arch., Phys. , arxiv:2502.21118 (2025), arXiv:2502.21118 [cond-mat.mtrl-sci].

[39] V. B. Shenoy, N. C. Frey, D. Akinwande, and D. Jariwala, Machine learning-enabled design of point defects in 2D materials for quantum and neuromorphic information processing, ACS Nano **14**, 13406 (2020).

[40] F. L. Thiemann, P. Rowe, E. A. Müller, and A. Michaelides, Machine Learning Potential for Hexagonal Boron Nitride Applied to Thermally and Mechanically Induced Rippling, J. Phys. Chem. C **124**, 22278 (2020).

[41] C. Cholsuk, A. Zand, A. Çakan, and T. Vogl, The hBN Defects Database: A Theoretical Compilation of Color Centers in Hexagonal Boron Nitride, J. Phys. Chem. C **128**, 12716 (2024).

[42] G. Kresse and J. Furthmüller, Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set, Comput. Mater. Sci. **6**, 15 (1996).

[43] 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).

[44] P. E. Blöchl, Projector augmented-wave method, Phys. Rev. B **50**, 17953 (1994).

[45] G. Kresse and D. Joubert, From ultrasoft pseudopotentials to the projector augmented-wave method, Phys. Rev. B **59**, 1758 (1999).

[46] C. Elias, P. Valvin, T. Pelini, A. Summerfield, C. J. Mellor, T. S. Cheng, L. Eaves, C. T. Foxon, P. H. Beton, S. V. Novikov, B. Gil, and G. Cassabois, Direct bandgap crossover in epitaxial monolayer boron nitride, Nat. Commun. **10**, 2639 (2019).

[47] S. M. Gilbert, T. Pham, M. Dogan, S. Oh, B. Shevitski, G. Schumm, S. Liu, P. Ercius, S. Aloni, M. L. Cohen, and A. Zettl, Alternative stacking sequences in hexagonal boron nitride, 2D Mater. **6**, 021006 (2019).

[48] D. Wickramaratne, L. Weston, and C. G. Van De Walle, Monolayer to Bulk Properties of Hexagonal Boron Nitride, J. Phys. Chem. C **122**, 25524 (2018).

[49] G. Constantinescu, A. Kuc, and T. Heine, Stacking in bulk and bilayer hexagonal boron nitride, Phys. Rev. Lett. **111**, 036104 (2013).

[50] S. Grimme, J. Antony, S. Ehrlich, and H. Krieg, A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu, The Journal of Chemical Physics **132**, 154104 (2010).

[51] 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).

[52] M. Arrigoni and G. K. Madsen, Spinney: Post-processing of first-principles calculations of point defects in semiconductors with Python, Computer Physics Communications **264**, 107946 (2021).

[53] R. O. Jones and O. Gunnarsson, The density functional formalism, its applications and prospects, Rev. Mod. Phys. **61**, 689 (1989).

[54] Z. Benedek, R. Babar, Á. Ganyecz, T. Szilvási, Ö. Legeza, G. Barcza, and V. Ivády, Symmetric carbon tetramers forming spin qubits in hexagonal boron nitride, npj Computational Materials **9**, 187 (2023).

[55] H. Mishra and S. Bhattacharya, Giant exciton-phonon coupling and zero-point renormalization in hexagonal monolayer boron nitride, Physical Review B **99**, 165201 (2019).

[56] A. Alkauskas, B. B. Buckley, D. D. Awschalom, and C. G. Van De Walle, First-principles theory of the luminescence lineshape for the triplet transition in diamond NV centres, New J. Phys. **16**, 073026 (2014).

[57] L. Liu, Pyvaspwfc. https://github.com/liming-liu/pyvaspwfc (accessed 2022-11-01).

[58] J. Davidsson, Theoretical polarization of zero phonon lines in point defects, J. Phys.: Condens. Matter **32**, 385502 (2020)

Data availability

All raw data from this work is available from the authors upon reasonable request. The database with browsing interfaces is freely accessible at https://h-bn.info. The SQLite database file can be downloaded via the GitHub REST API.