



Cite this: *Phys. Chem. Chem. Phys.*,
2025, 27, 12284

Excitations in lanthanide ions: a systematic evaluation of two-component CAS-CI and GW†

Roman Zielke,  Florian Weigend * and Christof Holzer *

This paper presents a thorough prediction and investigation of ionization energies, atomic levels, and crystal-field splittings in lanthanide ions. We show that a two-component complete active space (CAS) configuration interaction (CI) approach based on two-component density functional theory (DFT) reference states is suitable to yield accurate excitation energies for lower energy terms. DFT references are further shown to be superior to Hartree–Fock (HF) references for predicting both atomic levels and ionization energies. Especially in the Green's function based GW method used to determine ionization energies, the deficiencies of the wave function based HF references are severe, leading to sizable errors. Two-electron contributions to spin–orbit coupling are found to be an important ingredient for obtaining accurate atomic levels. These contributions are taken into account using a screened-nuclear-spin–orbit (SNSO) approach, which is shown to be very accurate. DFT based CAS-CI is further used to calculate crystal-field splittings. The results are well suited to predict the subtle splittings in complexes with unpaired 4f electrons.

Received 27th February 2025,
Accepted 16th May 2025

DOI: 10.1039/d5cp00780a

rsc.li/pccp

1 Introduction

High performance methods are needed to describe the excited states of lanthanide ions and their corresponding complexes. Multiple electronic effects, including electron correlation, spin–orbit coupling (SOC), and spin–other-orbit coupling (SOOC) need to be included to properly describe ground and especially excited states. Standard (time-dependent) density functional theory is often invalid in this regime,¹ as any functional currently available is unable to appropriately describe the multireference character of lanthanides. This even includes specialized functionals designed to describe strong (static) correlation.^{2,3} The ground state term is often already divided into multiple states, many of them slightly or significantly above the ground state level. Although density functional theory (DFT) has recently seen some success in predicting properties of trivalent lanthanide ion complexes,^{4,5} as for example EPR properties, excited states are currently out of reach for DFT. The required time-dependent (TD) DFT description of relativistic open-shell systems is often qualitatively wrong. This hampers the use of TD-DFT at least when aiming to include the f–f transitions in the predicted optical spectra. This is rather unfortunate, as nowadays TD-DFT represents the most important method to predict optical spectra *in silico*.⁶ While recently it was further outlined that also Green's

function based methods can be applied to these highly interesting complexes with unpaired 4f electrons,⁷ the latter approach remains constricted to non-degenerate ground states. This effectively limits its applicability to complexes containing Eu³⁺ and Sm²⁺ ions, and to some extent also Ce³⁺ and Yb³⁺.⁷ It is well known, however, that complete active space (CAS) methods are in principle able to describe the electronic situations prevalent in lanthanide ions,^{8–15} though at steeply increasing costs. Initially, perturbation theory was used to include SOC and other relativistic effects *a posteriori*,^{16,17} CAS+SOC frameworks are broadly available within the OpenMolcas¹⁸ and GAMESS-US¹⁹ packages, and have subsequently been applied to selected states of lanthanide atoms and ions with some success.^{12,20}

In a more recent evaluation,²¹ it has, however, been found that the errors of treating SOC non-variationally are significant. According to this study, the overall errors could be halved for excited states of lanthanide atoms if SOC was instead treated variationally using a relativistic two-component (2c) framework. The necessity to perform fully relativistic configuration interaction simulations significantly adds to the computational effort, requiring enormous computational facilities to tackle even comparably simple problems, as for example presented within the GRASP,²² DIRAC,²³ or Gaussian packages.²⁴ While these are capable of conducting the required relativistic CI calculations, performing simulations for sizable molecular systems is currently out of reach for these programs.^{22–29} Reported relativistic large-scale calculations using CASSCF programs have so far only been reported for the BAGEL program,³⁰ using a relativistic CI applied to the ground states of rhenium, uranium, and dysprosium complexes.^{31,32} Relativistic CASSCF proves to be

Institute for Quantum Materials and Technologies, Karlsruhe Institute of
Technology, Kaiserstrasse 12, D-76131 Karlsruhe, Germany.
E-mail: florian.weigend@kit.edu, holzer@kit.edu

† Electronic supplementary information (ESI) available. See DOI: <https://doi.org/10.1039/d5cp00780a>



numerically very demanding, even with comparably small systems and active space sizes requiring hundreds of cores to be carried out on a feasible time scale.^{32,33}

Therefore, we draw attention to a method developed by Grimme and Marian to ease the computational demands of these kind of calculations: mixing DFT and multireference CI calculations.^{34–37} The resulting DFT-MRCI method tries to incorporate parts of the dynamic correlation from DFT, allowing for a reduction of the active space for which the CI or CAS procedures are carried out. This is considered to be very advantageous as it reduces the size of the active space to the minimum orbital space required to describe the static correlation.³⁷ This statement is even more true for fully relativistic references, where the orthogonality of alpha and beta electrons and choosing a desired multiplicity of the reference wave functions *via* configuration state functions (CSFs) cannot be used to reduce the computational complexity. Instead, Slater determinants have to be used directly, which is conceptually convenient, yet computationally demanding. Still, including SOC directly instead of *via* perturbation theory has proven to be very beneficial, reducing overall errors when SOC is sizable.²¹

Using DFT references for lanthanides has also become more natural recently given the success of Green's function based methods.^{7,38–40} And while certainly lanthanides do put a lot of strain on any method used in theoretical chemistry, the *GW* method has successfully been proven to be applicable even if the underlying DFT reference is not optimal. In this work, we therefore aim at conducting a thorough investigation of DFT references of lanthanide ions, using them to perform subsequent CAS-CI and Green's function based *GW* calculations to determine their spectroscopic properties.

2 Theory

In the fully relativistic CI ansatz, the wave function $|\Psi\rangle$ is expanded in terms of Slater determinants $|\psi\rangle$,²⁸

$$|\Psi\rangle = \sum_p^{N_{\text{det}}} C_p |\psi\rangle \quad (1)$$

where the number of determinants is directly obtained from the number of spinors N_s and electrons n_e *via* the combinatorial n choose k approach, *i.e.*

$$N_{\text{det}} = \binom{N_s}{n_e} \quad (2)$$

The Slater determinants ψ can be obtained as a properly antisymmetrized product of molecular orbitals or spinors ϕ , which are themselves constructed from chosen basis functions χ and the spinor expansion coefficients. Within the exact-2-component (X2C) theory,^{41–43} they are given as

$$\phi_p = \sum_{\mu} \left(c_{\mu,p}^{\alpha} \chi_{\mu} + c_{\mu,p}^{\beta} \chi_{\mu} \right) \quad (3)$$

As outlined before, a construction of CSFs is superficial, as no computational savings can be achieved. Instead, in a fully

relativistic ansatz we can directly aim at optimizing the CI coefficients C minimizing the CAS-CI energy

$$E = e^{\text{HF}} + \langle \Psi | \hat{H}^{\text{CAS-CI}} | \Psi \rangle \quad (4)$$

The CAS-CI Hamiltonian is given by²⁴

$$\hat{H}^{\text{CAS-CI}} = \sum_{tu} h_{tu}^c \hat{E}_{tu} + \frac{1}{2} \sum_{tuvw} (tu|vw) (\hat{E}_{tu} \hat{E}_{vw} - \delta_{uv} \hat{E}_{tw}) \quad (5)$$

where h^c is the inactive contribution of the Fock matrix as outlined by eqn (20) of ref. 24. The first term of eqn (5) can be evaluated as

$$h_{tu}^c = h_{tu} + \sum_i [(tu|ii) - (ti|iu)] \quad (6)$$

We note that within the CAS-CI procedure, eqn (5) is evaluated using the spinor coefficients obtained from either DFT or HF, with no further differentiation being necessary. $(tu|vw)$ denotes a standard two-electron Coulomb integral, and $\hat{E}_{tu}$ are the generators of the unitary group⁴⁴

$$\hat{E}_{tu} = \hat{a}_t^{\dagger} \hat{a}_u \quad (7)$$

composed of creation ($\hat{a}^{\dagger}$) and annihilation ($\hat{a}$) operators. The indices $t, u, \dots$ run over all active orbitals while i includes all inactive, closed shell orbitals. Within the CASSCF procedure, it is now required to subsequently optimize the CI coefficients C as well as the expansion coefficients c iteratively, which quickly becomes very tedious. Effectively, relativistic CASSCF calculations therefore become untractable for even moderately sized molecular systems if the space spanned by the active orbitals exceeds more than a few spinors. The non-linear dependence further increases the amount of computational resources required for 2c CASSCF calculations by increasing the number of SCF iterations needed to converge both CI and expansion coefficients. It would therefore be of great help to simply use a good initial set of expansion coefficients c , only needing to determine the CI coefficients C once. HF theory, which is commonly used as starting point, lacks any correlation, and the resulting references are often neither qualitatively nor quantitatively correct. In contrast, Kohn–Sham (KS) references have been shown to be more capable of incorporating dynamic electron correlation, providing better starting points for CI methods.^{34–37}

Within DFT, only the question about the most suitable starting reference arises. The single determinant ansatz used in Kohn–Sham theory is, however, not able to describe the (nearly) degenerate orbital situation in many open-shell systems containing heavy elements. We therefore suggest instead using a fractional occupation approach that can yield qualitatively correct density distributions for atomic and molecular systems containing lanthanides or other heavy elements. The fractional occupation is chosen in a way that restores the spherical symmetry of the reference density of an atom, automatically leading to the correct degeneracy patterns. This avoids conceptual issues of introducing multiple Slater determinants in density functional theory, while still yielding suitable starting points for further CAS-CI treatments. Contrary, as outlined by McWeeny,⁴⁵ and Thyssen,⁴⁶ the average-of-configuration (AOC) ansatz, merging restricted open-



shell HF (ROHF) and CASSCF is well suited to provide good initial guesses for wave function theory. This ansatz is therefore preferred by wave function theory, yielding reliable values also for electronically complicated systems.^{47,48}

This is, however, improper for DFT, because in the latter, only density dependent variables can formally arise, not wave function based determinants from a restricted open shell scheme. We note that the CI energy is computed explicitly within the chosen active space and is not simply added to the DFT correlation energy. As such, the DFT correlation contributions are not double-counted within the active space. This avoids the typical double counting issue that arises in more empirical DFT-CI methods, where correlation corrections are directly added to the DFT total energy.

3 Computational details

A two-component Slater determinant based CAS-CI code was implemented in the TURBOMOLE package,^{49,50} allowing us to perform CAS-CI calculations using various HF and DFT starting points. The hybrid PBE0^{51,52} and local hybrid CHYF⁵ density functional approximations have been used. The x2c-QZVPPall-2c basis set⁵³ was employed for lanthanide ions, while x2c-TZVPPall-2c was used for the remaining lighter elements if applicable. HF and KS reference states were converged to at least $10^{-8}E_h$ and $10^{-7}E_h$ for differences in energy and difference density, respectively. Fine integration grids (gridsize 3a) were used for DFT.⁵⁴ Eigenvalue self-consistent evGW was used as described in ref. 7, further making use of a frequency sampling algorithm.⁴⁰ As an HF or KS reference in the evGW@DFT method, the canonical reference state with integer occupation numbers was used. For CAS-CI calculations, the respective fractional occupied reference density was employed,

$$D_{\mu\nu}^{\eta\zeta} = \sum_p n_p c_{\mu,p}^{\eta} c_{\nu,p}^{\zeta} \quad (8)$$

with occupation numbers n_p being 1.0 for spinors below the 4f level, $n/14$ for the 4fⁿ states, and 0.0 for the virtual spinors. Atomic densities from this fractional occupied states are spherically symmetric, with the spin and magnetization densities vanishing in each point in space. The excitation energies were subsequently obtained by diagonalizing the CAS-CI Hamiltonian. All-electron relativistic X2C theory was taken into account to include spin-orbit coupling.^{55,56} A row-dependent screened-nuclear-spin-orbit (SNSO) approach was chosen to account for two-electron spin-orbit effects in a semiempirical manner.^{57–59}

The structure of the aquatic $[\text{Nd}(\text{H}_2\text{O})_9]^{3+}$ complex has been taken from ref. 60. For the $[\text{Nd}(\text{H}_2\text{O})_9](\text{BrO}_3)_3$ complex, the crystal structure was taken from ref. 61 and the position of the hydrogen was optimized using PBE0 and a ECP49MWB basis set for Nd^{3+} , with the f-electrons being absorbed into the effective core potential (ECP).⁶² The def2-TZVP basis set was used for the remaining atoms.^{58,63,64}

4 Results

Before diving into the prediction of atomic levels of lanthanide atoms, we will perform an evGW@HF and evGW@DFT investigation of trivalent lanthanides to mainly detect qualitative differences in the underlying HF and DFT references. While GW yields ionization energies, *i.e.* charged excitations, those are also closely connected to the excited states *via* the Bethe–Salpeter equation.⁶⁵ Intrinsic information about the quality of excited states is therefore tightly linked to the ionization energies recovered from a given reference.

4.1 Ionization energies of trivalent lanthanide ions

As outlined in ref. 7, the GW method is a formidable tool to tackle ionization energies of lanthanide ions. Table 1 and Fig. 1 show that modern local hybrid functionals such as CHYF are an excellent basis for GW, in contrast to Hartree–Fock, which yields large errors, possibly related to its inability to correctly describe the ground state of lanthanide ions.

A first remarkable trend is observed for HF. As the nuclear charge increases, the inability of HF to describe the ground states sufficiently well becomes obvious, and the error of the obtained evGW@HF ionization energies amounts to several eV. As seen in Table 1, the error of evGW@HF increases monotonically with higher nuclear charge Z . A population analysis of the HF wave function reveals that for Ce^{3+} , instead of the correct 4f¹ configuration, a mixture of 4f¹ and 5d¹ reference states is found, while a population analysis of the CHYF density reveals the correct 4f¹ occupation. Due to error cancellation, the obtained ionization energy of HF is nevertheless remarkably close to the experimental value, while the error is highest for Ce^{3+} for CHYF.

In contrast, ionization energies predicted by evGW@DFT improve as the nuclear charge increases, with the CHYF local hybrid functional outperforming all previously tested functionals.⁷ As shown in Table 1, the deviations from experimental ionization energies monotonically decrease with increasing nuclear charge Z ,

Table 1 Mean absolute deviation (MAD) and standard deviation (STD) for the lanthanide ion ionization energies from the evGW@HF and evGW@DFT approaches with respect to the NIST database values.⁶⁶ The x2c-QZVPPall-2c basis set was used for all ions. All values in eV

	HF	PBE0		CHYF	
		SNSO	No SNSO	SNSO	No SNSO
Ce ³⁺	0.07	0.85	0.64	0.77	0.57
Pr ³⁺	0.34	0.90	0.72	0.88	0.78
Nd ³⁺	0.43	0.63	0.45	0.59	0.53
Pm ³⁺	0.63	0.61	0.68	0.39	0.52
Sm ³⁺	0.90	0.55	0.80	0.27	0.41
Eu ³⁺	0.99	0.53	0.68	0.19	0.38
Gd ³⁺	1.24	0.37	0.73	0.05	0.40
Tb ³⁺	1.44	0.73	0.72	0.32	0.33
Dy ³⁺	1.44	0.86	0.52	0.54	0.46
Ho ³⁺	1.71	0.63	0.65	0.22	0.29
Er ³⁺	2.00	0.60	0.83	0.00	0.33
Tm ³⁺	2.27	0.40	0.69	−0.24	0.04
Yb ³⁺	2.12	0.24	0.64	−0.40	0.01
MAD	1.38	0.64	0.68	0.45	0.44
STD	0.71	0.20	0.10	0.37	0.21



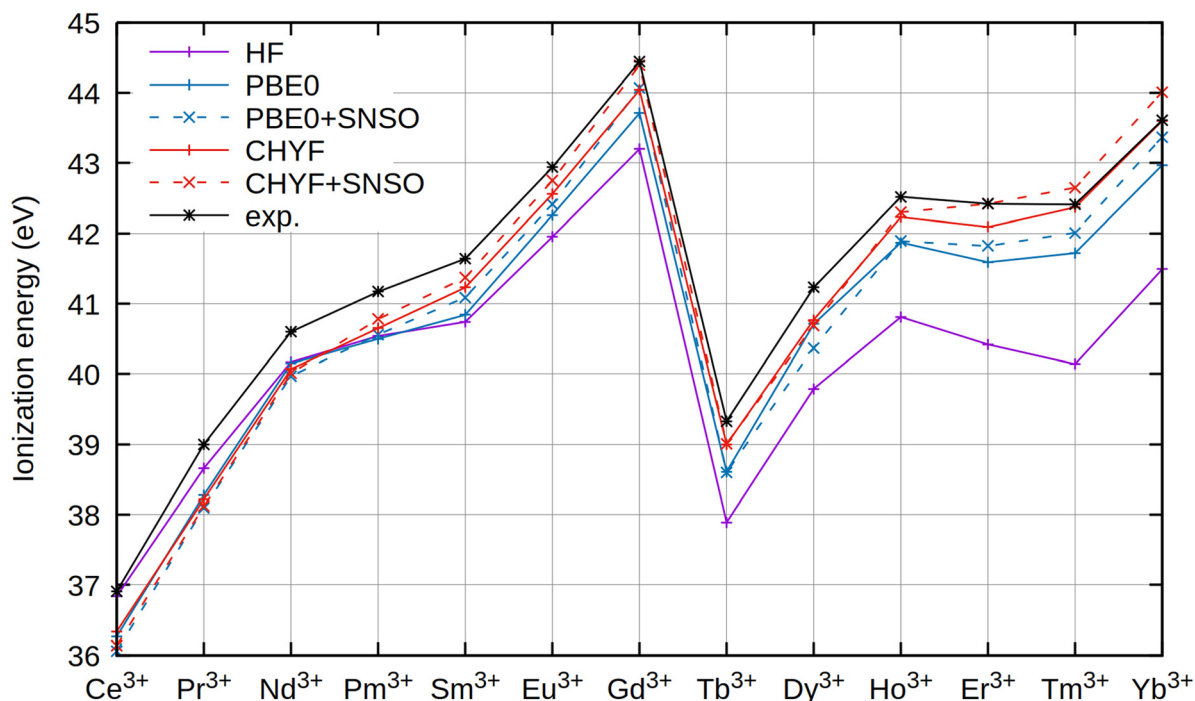


Fig. 1 Trend of ionization energies (in eV) for the trivalent lanthanoid ions Ln^{3+} , calculated using evGW@HF and evGW@DFT approaches. The x2c-QZVPall-2c basis set was used for all ions. For the dashed lines, the SNSO approach was used. Experimental references have been taken from the NIST database.⁶⁶ All values in eV.

though at an overall significantly smaller magnitude compared to the HF reference. Especially from Sm^{3+} on, evGW@CHYF yields accurate ionization energies and the correct $4f^n$ occupation in every case. evGW@PBE0 finally yields a more uniform distribution of errors as outlined by both Fig. 1 and Table 1, with the overall error being approximately independent from Z , though at a significantly elevated level of errors compared to CHYF.

Furthermore, in Fig. 1 and Table 1 the impact of relativistic effects on two-electron integrals was assessed by also incorporating an SNSO approach.^{57,58} For ionization energies, this effect leads only to minor corrections in the resulting GW values. However, the effect is rather dependent on the actual lanthanide with, for example, Gd^{3+} and Yb^{3+} seeing rather large corrections from the inclusion of SNSO, while Pr^{3+} and Tb^{3+} only see a minor change in the overall obtained ionization energy. Nevertheless, this effect should not be overemphasized, as parts of these deviations can also arise from the full neglect of changes in the underlying spinors from the evGW update, as in the latter still only the quasiparticle eigenvalues are updated in each iteration. The relativistic effects induced by the modified two-electron interactions are significant, though rather unsystematic. Their magnitude has been described by Boettger,⁵⁷ and is generally below the intrinsic error of the evGW method, essentially producing the same mean average deviation as clearly shown in Table 1. In general, DFT references can be assumed to be superior to HF references. Concerning the choice of the basis set, x2c-QZVPall-2c was found to yield converged values. We note that significant deviations can be found for smaller triple- ζ basis sets as shown in ref. 7, and currently we recommend using quadruple- ζ basis

sets for investigations of excited and ionized states of lanthanide ions.

We therefore also assume that starting CI procedures from DFT references can be advantageous, as the complete negligibility of correlation can lead to qualitatively wrong reference states in HF theory, as do pure DFT functionals for lanthanide ions.

We note that for any system an optimal specific amount of exact exchange can be obtained by tuning the density functional. However, modern local hybrid functionals like CHYF have shown to be rather robust in this respect,^{4,5} being able to intrinsically adapt the amount of exact exchange incorporated.

4.2 Atomic levels of lanthanide ions

In terms of excited states of lanthanide ions, the GW-BSE method has severe drawbacks when trying to describe excitations from systems with (nearly) degenerate ground states.⁷ CAS-CI and also CASSCF resolve this issue, leading to well defined atomic levels, though with the restriction of CAS-CI requiring a reference state of viable symmetry. Fig. 2–4 outline a set of CAS-CI calculations at three different references, namely PBE0, CHYF and HF.

The results obtained for all three references are rather similar, pointing out that for atomic levels both HF and DFT can achieve good starting points if an initial wave function or density with the correct spherical symmetry is generated. Especially for the low energy states below 7000 to 8000 cm^{-1} , usually belonging to the ground state term symbol, only very small deviations are obtained for all references. This is strictly tied to the inclusion of at least approximate two-electron spin-



orbit contributions (2e SOC). For example, in this work as well as in ref. 21 the SNSO approach is used to ensure this. The exact details of the chosen SNSO approach are only of minor importance for the investigated lanthanide ions, and testing of modern SNSO approaches yields only small overall differences well below the error of the initial CAS-CI method, which is very likely carried over to CASSCF.⁵⁹ We also expect other atomic mean field approaches, as for example the AMFI approach,^{16,67} to yield similar results. If no correction for 2e SOC effects is included, the splitting of the ground state term symbols can be off by several thousand wave numbers. This is further emphasized through Fig. 5, where the deterioration of calculated atomic levels is obvious when compared to the HF variant including SNSO correction as shown in Fig. 4. The same is also observed for the DFT cases, as seen in the ESI.[†] Involving these 2e SOC effects is therefore absolutely crucial in the prediction of the atomic levels of all lanthanide ions if two-component or four-component Hamiltonians are used. In the case of perturbative spin-orbit corrections, the results presented by Freidzon and co-workers however outline that this issue is unproblematic.¹² This is consistent with recent calculations of lanthanide complexes using the same perturbative CASSCF-SO approach for Eu^{3+} , Sm^{3+} , Tb^{3+} , and Yb^{3+} ions.^{68–71} In case of perturbation theory, 2e SOC effects are recovered by the Breit–Pauli spin-orbit operator if the one- and averaged two-electron parts are both included in the perturbation operator used to calculate the spin-orbit coupling matrix elements (SOCMEs).^{17,72} The above-mentioned studies on lanthanide ions and lanthanide complexes also show that a CAS treatment that only considers the 4f shell is often sufficient to obtain qualitatively correct results for the lowest terms.^{12,71} Statistical analysis of the resulting data, presented in Table 2, already indicates that CAS-CI@DFT is sufficient to reliably yield atomic levels with less than 4% error across all tested lanthanide ions. In absolute numbers, this translates to an average error of less than 200 cm^{-1} . CAS-CI@HF yields higher errors for the ground state terms,

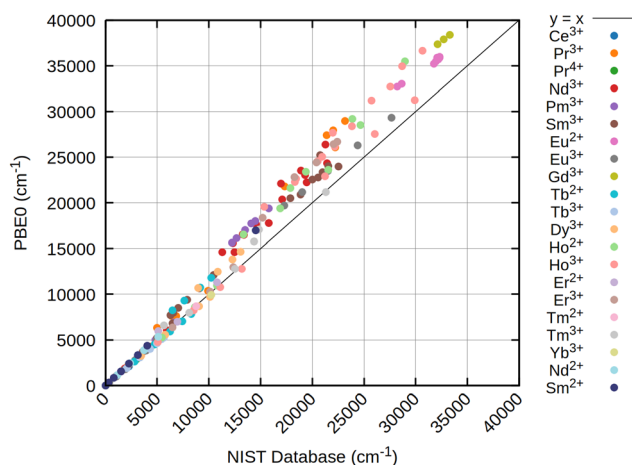


Fig. 2 Scatter plot of calculated atomic levels of 21 lanthanide ions at the CAS-CI@PBE0/x2c-QZVPPall-2c level of theory with an SNSO 2e SOC correction. As reference on the the x-axis, atomic levels from the NIST database⁶⁶ are taken. All values in cm^{-1} .

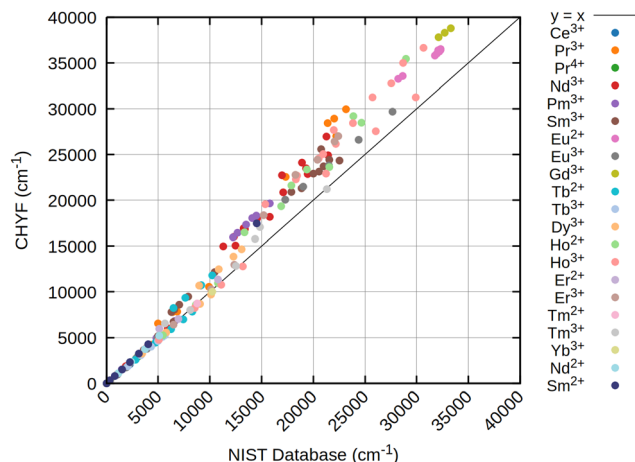


Fig. 3 Scatter plot of calculated atomic levels of 21 lanthanide ions at the CAS-CI@CHYF/x2c-QZVPPall-2c level of theory with an SNSO 2e SOC correction. As reference on the the x-axis, atomic levels from the NIST database⁶⁶ are taken. All values in cm^{-1} .

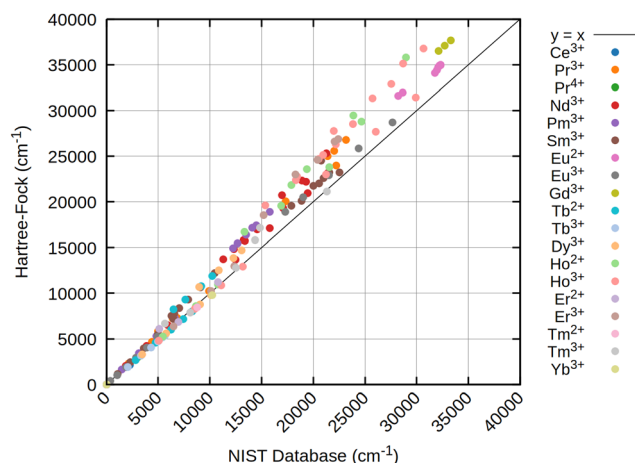


Fig. 4 Scatter plot of calculated atomic levels of 19 lanthanide ions at the CAS-CI@HF/x2c-QZVPPall-2c level of theory with an SNSO 2e SOC correction. As reference on the the x-axis, atomic levels from the NIST database⁶⁶ are taken. All values in cm^{-1} .

amounting to above 5% and absolute errors of up to 300 cm^{-1} , which is nevertheless still acceptable. However, HF completely fails to converge to the correct electronic ground state for Ce^{3+} , Nd^{2+} , and Sm^{2+} , which therefore have been omitted from the statistical analysis. While for the Ce^{3+} ion, simply including the 5d orbital into the active space cures the otherwise incorrect results, no simple solution is found for Nd^{2+} and Sm^{2+} . With even simple atoms being problematic, this points at molecular complexes being even more troublesome. Contrary, neither CHYF nor PBE0 have deficiencies when trying to converge to the correct state of Ce^{3+} , Nd^{2+} , and Sm^{2+} , yielding the correct 4fⁿ configuration with no artificial 5d admixtures. A special case is Pr^{2+} , which also did not converge to the correct state with PBE0, whereas solely CHYF managed to correctly predict the ground state occupation of this ion. CAS-CI@DFT can therefore have a

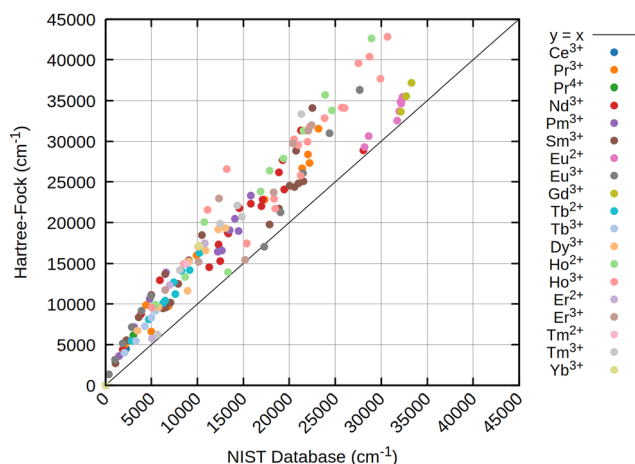


Fig. 5 Scatter plot of calculated atomic levels of 19 lanthanide ions at the CAS-CI@HF/x2c-QZVPAll-2c level of theory without 2e SOC correction. As reference on the x-axis, atomic levels from the NIST database⁶⁶ are taken. All values in cm^{-1} .

Table 2 Mean deviation (MD) and standard deviation (STD) for the lanthanide ion test set using the CAS-CI@HF and CAS-CI@DFT approaches. The x2c-QZVPAll-2c basis set was employed for all ions in all configurations. Values are given in absolute units (cm^{-1}) and relative units (%). The SNSO correction has been used in all calculations. Only ground state terms are considered here

	HF		PBE0		CHYF	
	STD	MD	STD	MD	STD	MD
Rel. error [%]	5.1	5.7	3.2	3.8	3.2	3.7
Abs. error [cm^{-1}]	242	289	164	201	156	190

double advantage in real applications: not only does it lead to a slightly better description of the excited states, but it also has a higher chance of finding the right starting point to enable the calculation of excited states.

For excited states belonging to higher terms, the deviations found with respect to data from the NIST atomic level database⁶⁶ increase with increasing excitation energy. Observed deviations are roughly proportional to the energy of the atomic level. Deviations further are monotonically growing, being too large in nearly all cases. These effects are nearly independent from the chosen reference functions, with CAS-CI@HF and CAS-CI@DFT yielding similar deviations for all lanthanide ions as described in Fig. 2–4. It can be attributed to dynamic correlation becoming increasingly important for excited state terms. However, as shown in earlier work,⁴⁷ simply increasing the active space only slowly leads to a lowering of the energies of these atomic states. Increasing the active space size is therefore inefficient, and future developments that improve higher excited states of lanthanides should therefore focus on more efficient ways to recover dynamic correlation. The detailed data presented in these figures show that the error patterns are rather intuitive. For example, for Sm^{2+} and Eu^{3+} , both featuring a $4f^6$ electron configuration, nearly identical deviations are found. For the

lowest F manifold, deviations of only a few hundred wave numbers are observed, while the first D manifold is off by 2000–3000 cm^{-1} , depending on the reference. Consequently, if higher lying terms are taken into account, the subsequent statistical deviations from the NIST atomic levels increase as shown in Table 3. Notably, the order of error is now turned upside down, with CAS-CI@HF being superior to CAS-CI@DFT, at least as long as 2e SOC is taken into account using the SNSO approximation. Table 3 also shows that neglecting the 2e SOC leads to a drastic increase in overall errors. The obtained results are, as outlined previously, not meaningful for any lanthanide ion and can deviate by tens of thousands of wave numbers from the NIST values.

In Table 4, detailed values of the 7F_J and 5D_J states of the Eu^{3+} ion are listed to give a concise comparison of various feasible approaches for calculating excited states of this ion. Throughout the literature, the optical properties of the Eu^{3+} ion are the most thoroughly theoretically investigated, with various studies also focusing on the atomic ion.^{7,12,13,47,73,74}

The overview given in Table 4 indicates that it is still challenging to extract accurate excitation energies for even a single lanthanide ion. Eu^{3+} , featuring a 7F_0 ground state is one of the simplest examples, as this state is non-degenerate. Even the simple complete open shell CI (COSCI) variants¹³ are capable of well describing the 7F_J manifold, especially if 2e SOC corrections are taken into account. However, there are no significant advantages of using the 4c COSCI formalism compared to our simple 2c CAS-CI ansatz, with overall errors being basically identical. In fact, the 1c ansatz that includes SOC perturbatively yields excellent values,¹² being nearly indistinguishable even from the 4c results. The only outlier is the evGW-BSE based calculation on the bare ion,⁷ which, however, acts on the true (symmetry broken) KS ground state instead of a CAS or a spherically averaged reference state. evGW-BSE is therefore not able to capture any significant multireference character, leading to substantial deviations in the 7F_J manifold.

For the 5D_J states, deviations are elevated as demonstrated in Fig. 2–4. All CI based approaches yield too high energies for this state. The 4c COSCI and the 2c CAS-CI approach deliver values of roughly 20 000 cm^{-1} , being approximately 2000–3000 cm^{-1} larger than the experimental reference values. The CAS-CI@DFT calculations are slightly more accurate, indicating that the inclusion of dynamic correlation is more important than a full 4c treatment of the underlying reference state. Furthermore, even extensive Kramers-restricted CI spaces as used in the KRCI and MCDF-CI approaches are merely able to lower

Table 3 Mean relative deviation (MRD) and standard deviation (STD) for the lanthanide ion excitation energies using the CAS-CI@HF and CAS-CI@DFT approach. The x2c-QZVPAll-2c basis set was used for all ions in all configurations. All values in %

	HF		PBE0		CHYF	
	STD	MRD	STD	MRD	STD	MRD
SNSO	7.7	9.2	9.8	11.2	10.7	12.2
No SNSO	31.3	42.0	27.0	35.5	25.8	35.7



Table 4 Atomic levels determined for the 7F_J and 5D_J states of Eu^{3+} . 2c CAS-CI results were determined within this work, while the remaining results have been taken from literature. All values in cm^{-1}

	$\text{Eu}^{3+}:\text{LaCl}_3^{66}$ Exp.	COSCI^{13} 4c + ΔSOO	COSCI^{47}	KRCl^{47} 4c + ΔSOO	MCDF-CI^{73} 4c + Breit	CAS-CI 2c HF	CAS-CI 2c CHYF	CAS-CI + d 2c CHYF	evGW-BSE ⁷ 2c TPSSH	SO-CASSCF ¹² 1c + XMCQPD2
7F_J										
$J = 0$	0	0	0	0	0	0	0	0	0	0
$J = 1$	370	375	340	374	347	366	329	332	690	380
$J = 2$	1040	1058	964	1041	964	1026	933	939	732	1040
$J = 3$	1890	1962	1801	1916	1773	1899	1743	1753	801	1880
$J = 4$	2860	3022	2793	2931	2711	2920	2702	2716	2370	2830
$J = 5$	3910	4192	3893	4040	3734	4042	3770	3784	3443	3860
$J = 6$	4940	5429	5069	5210	4810	5233	4912	4926	5229	4970
5D_J										
$J = 0$	17 270	20 565	20 706	18 476	18 969	18 892	20 046	19 796	17 298	17 830
$J = 1$	19 030	22 213	22 202	20 138	20 607	20 487	21 498	21 260	17 987	19 450
$J = 2$	21 510	24 746	24 556	22 588	22 991	22 911	23 776	23 546	21 017	22 140
$J = 3$	24 390	27 478		25 536	25 823	25 864	26 598	26 372		25 370
$J = 4$	27 640	30 795			29 239	28 719	29 678	29 502		28 960

this to 18 000–19 000 cm^{-1} ,^{47,73} still yielding deviations of 1000–1500 cm^{-1} with respect to the NIST database. While it can be shown that including more and more correlation by expanding the active space indeed recovers the dynamic correlation needed to yield improved atomic levels, the convergence with respect to included states is slow. This is confirmed by our CAS-CI+d calculation, which sets the active space to the 4f+5d orbitals. The latter only lowers the 5D_0 state by 250 cm^{-1} , and has no significant effect on the already excellent 7F states at all. 1c SO-CASSCF used in ref. 12 delivers lower deviations than the 2c and 4c approaches, even though they also only use an active space composed of the 4f orbitals and electrons. This can largely be amounted to the “extended multi-configuration quasi-degenerate second order perturbation theory” (XMCQDPT2) correction that is used to include dynamic electron correlation.⁷⁵ Surprisingly, the 2c evGW-BSE approach yields values closest to those of the experiment, even though it is unable to capture important static correlation effects. Yet, evGW-BSE likely captures a large amount of the dynamic electron correlation, yielding good results overall – although the results likely are also artificially enhanced by error cancellation.

4.3 Nonaqua complexes of neodymium(III)

Crystal-field splittings of lanthanide ions are still difficult to grasp, yet can also be obtained from the approach outlined above for the bare ions. The difficulty of receiving crystal-field splittings is then redirected to calculating a suitable reference for the molecular ground state. It can still be assumed that the 4f shells are mainly atomic, with no significant mixing with the valence orbitals of the surrounding ligands. While this can be troublesome for many DFT functionals, modern local hybrid functionals constructed from first principles as CHYF can handle the required situations, providing enough flexibility and the rigor to yield qualitatively correct starting points for a subsequent fully relativistic CAS treatment.

First, we investigate the aquatic $[\text{Nd}(\text{H}_2\text{O})_9]^{3+}$ complex, as for this complex experimental and theoretical investigations of the crystal-field splitting have recently been carried out.^{60,76} Nielsen and co-workers found that in solution, a capped square

antiprism (cSAP) is prevalent, rather than the commonly assumed tricapped trigonal prism (TTP).^{60,77} Table 5 compares the splitting of the $I_{9/2}$ state of Nd^{3+} for these two configurationally different variants using the structures provided by ref. 60. Starting from a PBE0 or CHYF reference using the fractional occupation ansatz described for atoms in the previous section, we find that the experimentally observed crystal-field splitting of the $I_{9/2}$ ground state is too large for a tricapped trigonal prism. The latter only amounts to 148–222 cm^{-1} for TTP, while cSAP produces significantly larger splittings of up to 256–289 cm^{-1} , better aligning with the experimental findings of 350–374 cm^{-1} .⁷⁶ Employing a perturbative spin-orbit correction *a posteriori* to a CASSCF reference wave function leads to generally larger splittings, yielding 304 cm^{-1} for the highest crystal-field splitted state of the $I_{9/2}$ manifold. Although this is in slightly better agreement with the experiment than the CAS-CI approach, the latter is much closer for the first excited state of the $I_{9/2}$ manifold, predicting 33–35 cm^{-1} , being in line with experimental measurements of 23–39 cm^{-1} , while CASSCF-SO overestimates this splitting by a factor of 4.

Given the assumption that the actual structure of Nd^{3+} is highly dynamic, which leads to even larger crystal-field splittings,⁶⁰ our CAS-CI results are well suited to predict subtle splittings in this complicated lanthanide ion. As determined by Satten, the splitting of the $I_{9/2}$ manifold is even larger in the $[\text{Nd}(\text{H}_2\text{O})_9](\text{BrO}_3)_3$ crystal.⁶¹ Table 6 illustrates the values calculated for this structure. Fully relativistic CAS-CI calculations confirm that the splitting found in this crystal is overall slightly larger than for the aquatic solution. For the CHYF reference, this effect is predicted to be moderate, while the PBE0 reference anticipates a more pronounced shift, which fits the overall observed splittings better. However, the differences between the splittings found in solution and in the crystal are overestimated by the PBE0 reference, which especially estimates state 4 to shift by 150 cm^{-1} , though only an actual change of 8–32 cm^{-1} is found experimentally. The latter predicted change is more in line with the CAS-CI calculated in the CHYF reference, which anticipates a change of roughly 43 cm^{-1} . Furthermore, CAS-CI@CHYF exhibits a nearly constant shift with respect to the experimental values of ref. 61,



Table 5 Crystal-field splittings simulated for the $I_{9/2}$ state of $[\text{Nd}(\text{H}_2\text{O})_9]^{3+}$ for the capped square antiprism (cSAP) and tricapped trigonal prism (TTP) using the CAS-CI@DFT approach outlined in this work. The x2c-QZVPall-2c (Nd) and x2c-TZVPPall-2c (H,O) basis sets were used. Experimental values extracted from ref. 76. CASSCF-SO values taken from ref. 60. All values in cm^{-1}

TTP		cSAP		Ref. 60 (cSAP)	Exp.	
PBE0	CHYF	PBE0	CHYF	CASSCF-SO	Abs.	Em.
0	0	0	0	0	0	0
69	5	35	33	123	23	39
129	95	175	124	146	—	—
180	142	220	188	292	153	164
222	148	289	256	304	350	374

while CAS-CI@PBE0 yields less systematic deviations in that specific case. We note in passing that the HF reference state leads to no meaningful results in the case of the $[\text{Nd}(\text{H}_2\text{O})_9](\text{BrO}_3)_3$ crystal, due to converging to the wrong state. Therefore, it is omitted from the discussion.

Conclusions

We have investigated the optical spectra of all relevant lanthanide ions using a newly implemented 2c CAS-CI approach based on KS determinants. Taking the 4f shell and electrons as active space on top of the Kohn–Sham reference yields very good agreement for the ground state terms. For higher lying terms, still significant errors are observed, though they are slightly lower than their 4c CAS based counterparts. We therefore conclude that for ground state terms, the static and dynamic correlation captured by including the 4f shell in the CAS treatment is sufficient. Our CAS-CI@DFT is therefore especially well suited to predict crystal-field splittings in lanthanide ions, as long as excitations are solely described by the 4f manifold. In ref. 7 it has been demonstrated for Eu^{3+} , that the excited $^5\text{D}_j$ state can participate in charge-transfer excitations, exceeding the pure 4f orbital space. This has recently been shown experimentally using time-resolved emission spectroscopy.⁷⁸ In this case, only including the 4f shell in the active space may lead to acceptable excitation energies, although the transition properties will be off. The correlation between the 4f shell and the surrounding ligands needs to be carefully taken into account to correctly predict these properties.

Table 6 Crystal-field splittings simulated for $I_{9/2}$ state of the $[\text{Nd}(\text{H}_2\text{O})_9](\text{BrO}_3)_3$ crystal for the capped square antiprism (cSAP) and tricapped trigonal prism (TTP) using the CAS-CI@DFT approach as outlined in this work. The x2c-QZVPall-2c (Nd) and x2c-TZVPPall-2c (H,O) basis sets were used. Experimental values taken from ref. 61. Δ values denote differences between the calculated and measured splittings. State labels assigned according to ref. 77. All values in cm^{-1}

State m_j	PBE0	CHYF	ΔPBE0	ΔCHYF	Exp.
$\pm 9/2$	0	0	0	0	0
$\pm 7/2$	83	46	−34	−71	117
$\pm 5/2$	136	85	−48	−99	184
$\pm 3/2$	421	281	58	−82	363
$\pm 1/2$	442	299	60	−83	382

Author contributions

Roman Zielke: conceptualization (equal); investigation (lead); methodology (equal); software (lead); validation (equal); visualization (lead); writing – original draft (supporting); writing – review & editing (equal). Florian Weigend: conceptualization (equal); investigation (supporting); methodology (equal); validation (supporting); writing – original draft (supporting); writing – review & editing (equal). Christof Holzer: conceptualization (equal); investigation (supporting); methodology (equal); software (supporting); validation (equal); visualization (supporting); writing – original draft (lead); writing – review & editing (equal).

Data availability

The data supporting this article have been included as part of the ESI.†

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

Roman Zielke thanks the Turbomole GmbH for financial support. Florian Weigend gratefully acknowledges financial support from the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through the Collaborative Research Centre “4f for Future” (CRC 1573, project number 471424360), project Q.

References

- 1 L. Petit, A. Borel, C. Daul, P. Maldivi and C. Adamo, *Inorg. Chem.*, 2006, **45**, 7382–7388.
- 2 A. D. Becke, *J. Chem. Phys.*, 2003, **119**, 2972–2977.
- 3 A. D. Becke, *J. Chem. Phys.*, 2005, **122**, 064101.
- 4 C. Holzer and Y. J. Franzke, *J. Chem. Phys.*, 2022, **157**, 034108.
- 5 C. Holzer and Y. J. Franzke, *J. Chem. Theory Comput.*, 2025, **21**, 202–217.
- 6 M. Marques and E. Gross, *Annu. Rev. Phys. Chem.*, 2004, **55**, 427–455.
- 7 C. Holzer, *J. Chem. Theory Comput.*, 2023, **19**, 3131–3145.
- 8 H. Moriyama, Y. Watanabe, H. Nakano, S. Yamamoto and H. Tatewaki, *J. Chem. Phys.*, 2010, **132**, 124310.
- 9 S. Yamamoto and H. Tatewaki, *J. Chem. Phys.*, 2011, **134**, 164310.
- 10 M. Sekiya, F. Sasaki and H. Tatewaki, *Phys. Rev. A: At., Mol., Opt. Phys.*, 1997, **56**, 2731–2740.
- 11 P. A. Tanner, Y.-Y. Yeung and L. Ning, *J. Phys. Chem. A*, 2014, **118**, 8745–8752.
- 12 A. Y. Freidzon, I. A. Kurbatov and V. I. Vovna, *Phys. Chem. Chem. Phys.*, 2018, **20**, 14564–14577.
- 13 O. Visser, L. Visscher, P. J. C. Aerts and W. C. Nieuwpoort, *J. Chem. Phys.*, 1992, **96**, 2910–2919.
- 14 W. Van den Heuvel, S. Calvillo and A. Soncini, *Phys. Chem. Chem. Phys.*, 2016, **18**, 15807–15814.



- 58 M. Filatov, W. Zou and D. Cremer, *J. Chem. Phys.*, 2013, **139**, 014106.
- 59 J. Ehrman, E. Martinez-Baez, A. J. Jenkins and X. Li, *J. Chem. Theory Comput.*, 2023, **19**, 5785–5790.
- 60 V. R. M. Nielsen, M. Grasser, B. Le Guennic and T. J. Sørensens, *Inorg. Chem.*, 2025, **64**, 3463–3475.
- 61 R. A. Satten, *J. Chem. Phys.*, 1953, **21**, 637–648.
- 62 M. Dolg, H. Stoll and H. Preuss, *J. Chem. Phys.*, 1989, **90**, 1730–1734.
- 63 F. Weigend and R. Ahlrichs, *Phys. Chem. Chem. Phys.*, 2005, **7**, 3297–3305.
- 64 D. Cremer, W. Zou and M. Filatov, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 2014, **4**, 436–467.
- 65 X. Blase, I. Duchemin, D. Jacquemin and P.-F. Loos, *J. Chem. Phys. Lett.*, 2020, **11**, 7371–7382.
- 66 A. Kramida, Yu. Ralchenko, J. Reader and NIST ASD Team, *NIST Atomic Spectra Database (ver. 5.12)*, [Online], Available: <https://physics.nist.gov/asd> [2025, February 24], National Institute of Standards and Technology, Gaithersburg, MD, 2024.
- 67 M. Iliaš, V. Kellö, L. Visscher and B. Schimmelpfennig, *J. Chem. Phys.*, 2001, **115**, 9667–9674.
- 68 F. Gendron, B. Moore II, O. Cador, F. Pointillart, J. Autschbach and B. Le Guennic, *J. Chem. Theory Comput.*, 2019, **15**, 4140–4155.
- 69 K. A. Romanova, A. Y. Freidzon, A. A. Bagaturyants and Y. G. Galyametdinov, *J. Phys. Chem. A*, 2014, **118**, 11244–11252.
- 70 F. Gendron, M. Grasser and B. Le Guennic, *Phys. Chem. Chem. Phys.*, 2022, **24**, 5404–5410.
- 71 M. Grasser and B. Le Guennic, *Phys. Chem. Chem. Phys.*, 2024, **26**, 7203–7210.
- 72 O. Vahtras, H. Ågren, P. Jørgensen, H. Jørgen Aa. Jensen, T. Helgaker and J. Olsen, *J. Chem. Phys.*, 1992, **96**, 2118–2126.
- 73 M. Seth, K. G. Dyall, R. Shepard and A. Wagner, *J. Phys. B: At., Mol. Opt. Phys.*, 2001, **34**, 2383.
- 74 C. Naleway, M. Seth, R. Shepard, A. F. Wagner, J. L. Tilson, W. C. Ermler and S. R. Brozell, *J. Chem. Phys.*, 2002, **116**, 5481–5493.
- 75 A. A. Granovsky, *J. Chem. Phys.*, 2011, **134**, 214113.
- 76 V. R. M. Nielsen, P. R. Nawrocki and T. J. Sørensens, *J. Phys. Chem. A*, 2023, **127**, 3577–3590.
- 77 P. R. Nawrocki and T. J. Sørensens, *Phys. Chem. Chem. Phys.*, 2023, **25**, 19300–19336.
- 78 Y. Kitagawa, T. Tomikawa, K. Aikawa, S. Miyazaki, T. Akama, M. Kobayashi, M. Wang, S. Shoji, K. Fushimi, K. Miyata, Y. Hirai, T. Nakanishi, K. Onda, T. Taketsugu and Y. Hasegawa, *Commun. Chem.*, 2025, **8**, 24.

