# Rovibrational energy levels of the hydrogen molecule through nonadiabatic perturbation theory

Jacek Komasa and Mariusz Puchalski
*Faculty of Chemistry, Adam Mickiewicz University, Uniwersytetu Poznańskiego 8, 61-614 Poznań, Poland*

Paweł Czachorowski, Grzegorz Łach, and Krzysztof Pachucki
*Faculty of Physics, University of Warsaw, Pasteura 5, 02-093 Warsaw, Poland*
(Dated: September 30, 2019)

We present an accurate theoretical determination of rovibrational energy levels of the hydrogen molecule and its isotopologues in its electronic ground state. We consider all significant corrections to the Born-Oppenheimer approximation, obtained within nonadiabatic perturbation theory, including the mixed nonadiabatic-relativistic effects. Quantum electrodynamic corrections in the leading $\alpha^5\,m$ and the next-to-leading $\alpha^6\,m$ orders, as well as finite nuclear size effect, are also taken into account but within the Born-Oppenheimer approximation only. Final results for the transition wavelength between rovibrational levels achieve accuracy of the order of $10^{-3}$–$10^{-7}$ cm$^{-1}$, and are provided by simple to use computer code.

## I. INTRODUCTION

The hydrogen molecule is one of the simplest chemical systems. Nevertheless, it has a rich spectrum of rovibrational levels with lifetimes of the order of $10^5$–$10^6$ s. Thanks to these long lifetimes, the contemporary measurements of transition frequencies between rovibrational levels in $H_2$ and its isotopologues has reached an accuracy level of $10^{-8}$ and in unique cases even $10^{-9}$ [1–14]. In parallel, recent advances in a theoretical method—the nonadiabatic perturbation theory (NAPT)—have enabled accurate prediction for an arbitrary vibration and rotation quantum number of an arbitrary hydrogen isotopologue. Thus, theoretical progress and experimental availability make it an attractive candidate for precise tests of quantum electrodynamic (QED) theory and the search for a new physics.

This paper presents the framework of theoretical methods for calculation of nonrelativistic energies and of relativistic and QED corrections up to $\alpha^7\,m$ order, together with the improved calculations of the so called heteronuclear potential. All the recent advances, including the complete $\alpha^6\,m$ and the $\alpha^4\,m^2/M$ corrections, as well as direct four-body calculations of nonrelativistic energies, have been included. Results of these calculations, in the form of pertinent internuclear potentials have been implemented in a publicly available computer code [15]; therefore, this work provides the highest accuracy of all the energy levels and transition frequencies for the hydrogen molecule in the ground electronic state. Depending on the isotopic contents and on the molecular level, the absolute accuracy ranges from $10^{-3}$ to $10^{-4}$ cm$^{-1}$, mostly limited by the yet unknown $\alpha^5\,m^2/M$ correction or the higher-order nonadiabatic nonrelativistic effects.

This work finds application in many areas of physics ranging from astrophysical observations [16], through verification of experimental spectra [17, 18], to measurements of the neutrino mass [19].

## II. NRQED FRAMEWORK

The most convenient theoretical framework for the accurate description of light molecular systems is nonrelativistic quantum electrodynamics (NRQED). It is an effective theory derived from relativistic QED by matching the scattering amplitude up to certain powers in external momenta [20]. The main advantage of NRQED approach is that all the corrections are implemented on the top of a nonrelativistic Hamiltonian, and one can use the standard perturbation theory with the nonrelativistic wave function.

The principal assumption in NRQED, is that the total energy can be expanded in powers of the fine-structure constant $\alpha$ (with $m$ being the mass of the electron)

$$E(\alpha) = \alpha^2 m\, E^{(2)} + \alpha^4 m\, \left( E^{(4)} + E^{(4)}_{\text{FS}} \right) + \alpha^5 m\, E^{(5)}$$
$$+ \alpha^6 m\, E^{(6)} + ..., \tag{1}$$

where $E^{(4)}_{\text{FS}}$ is the finite nuclear size correction. $E^{(n)}$ may include also powers of $\ln\alpha$, which is not shown explicitly. The expansion terms are interpreted subsequently as the nonrelativistic energy $E^{(2)}$, the relativistic correction $E^{(4)}$, the leading QED correction $E^{(5)}$, and the higher-order QED corrections $E^{(i)}$, $i \geq 6$. All these contributions can be expressed as expectation values of certain operators derived within NRQED theory, with a Schrödinger wave function.

## III. NONRELATIVISTIC ENERGY $E^{(2)}$

The leading term of Eq. (1) is an eigenvalue of the Coulomb-Schrödinger Hamiltonian (in atomic units)

$$H^{(2)} = -\sum_a \frac{\vec{\nabla}_a^2}{2\,m_a} + \sum_{a>b} \frac{Z_a\,Z_b}{r_{ab}}, \tag{2}$$

where $m_a$ is the mass of the $a$th particle, $Z_a$ its charge, and $a$, $b$ go over all of the particles in the system (two electrons and two nuclei). This eigenvalue is determined using two distinct methods. The first one, called here "direct," relies on

a variational solution of the four-body Schrödinger equation with fully nonadiabatic wave function expanded in the basis of the nonadiabatic James-Coolidge (naJC) functions [21]. By this approach, the energy of a given rovibrational level can be evaluated to a very high accuracy but requires a separate computationally intensive variational calculation. Hence, currently, a limited fraction of the energy levels evaluated with this method are available. The second method relies on the expansion of the energy in a small parameter, being a ratio of the electron-to-nuclei mass. Within this perturbative approach, called the nonadiabatic perturbation theory (NAPT), a set of nuclear interaction potentials is constructed, which, in turn, enable all the energy levels collectively to be determined in a single calculation at insignificant computational cost. Both methods are briefly described in the following subsections.

### A. Direct evaluation of $E^{(2)}$

This method is conceptually very simple. It relies on a variational solution of the Schrödinger equation

$$H^{(2)}\Psi = E^{(2)}\Psi \qquad (3)$$

by expanding the wave function $\Psi$ in a four-particle basis

$$\Psi = \sum_k c_k\, \psi_{\{k\}}(\vec{r}_A, \vec{r}_B, \vec{r}_1, \vec{r}_2)\,. \qquad (4)$$

The obtained matrix form of the Schrödinger equation

$$(\mathbb{H} - E^{(2)}\,\mathbb{S})\,\mathbb{C} = 0 \qquad (5)$$

is then solved using an inverse iteration method with efficiently parallelized linear algebra algorithms. The basis functions employed in this expansion are the exponential functions of the form [21]

$$\psi_{\{k\}} = e^{-\alpha\,R - \beta(\zeta_1+\zeta_2)}\, R^{k_0}\, r_{12}^{k_1}\, \eta_1^{k_2}\, \eta_2^{k_3}\, \zeta_1^{k_4}\, \zeta_2^{k_5}, \qquad (6)$$

called nonadiabatic James-Coolidge (naJC) functions for their resemblance to the classic functions introduced by James and Coolidge in 1933 [22]. The $R \equiv r_{AB}$ variable represents the internuclear distance, $r_{12}$ —the interelectron distance and the meaning of the remaining variables present in this function is the following $\zeta_a = r_{aA} + r_{aB}$ and $\eta_a = r_{aA} - r_{aB}$. The $\alpha$ and $\beta$ in Eq. (6) denote nonlinear variational parameters, common for the whole set of basis functions called "sector," and $k_j$ are non-negative integers collectively denoted as $\{k\}$. If needed, two or more sectors (with different pairs of $\alpha^{(i)}$ and $\beta^{(i)}$) can be used. Technical details concerning the evaluation of the matrix elements of $\mathbb{H}$ and $\mathbb{S}$ can be found in Refs. [21, 23, 24].

From the study of the energy convergence with the increasing size of the basis set, we observe that the nonrelativistic results obtained with this method reach a relative numerical accuracy of $10^{-13} - 10^{-14}$ [21, 23–25]. In principle, this accuracy can be further increased, but in view of the fact that the uncertainties from physical constants appear at the level of $10^{-12}$, there is currently no need for such efforts. All these available energies corresponding to particular molecules and rovibrational levels are supplied to the computer code [15] as external data to be read whenever needed.

### B. NAPT—the nonadiabatic perturbation theory

NAPT was introduced in Refs. [26, 27] and expanded later in Refs. [28, 29]. It relies on a perturbative separation of electronic and nuclear movements. In the zeroth order, molecular energy is obtained from the Born-Oppenheimer (BO) internuclear potential [30]. The next order, which goes with the ratio of the electron mass $m_e$ to the nuclear reduced mass $\mu_n$, is represented as an adiabatic correction to the BO potential [31]. Higher-order corrections, which are quadratic in the mass ratio, give additional $R$-dependent potentials and $R$-dependent nuclear masses in the effective nuclear equation [29]. These higher-order nonadiabatic potentials lead to radically increased accuracy of nonrelativistic levels. This subsection provides a concise description of the NAPT method.

Let the total wave function $\Psi$ be a solution of the stationary Schrödinger equation (3) with the Hamiltonian partitioned into the electronic and nuclear parts

$$H^{(2)} = H_{el} + H_n\,. \qquad (7)$$

The clamped nuclei electronic Hamiltonian

$$H_{el} = -\sum_a \frac{\nabla_a^2}{2\,m_e} + V \qquad (8)$$

consists of the electronic kinetic energy term and the potential $V$, which includes all the Coulomb interactions,

$$V = -\frac{1}{r_{1A}} - \frac{1}{r_{1B}} - \frac{1}{r_{2A}} - \frac{1}{r_{2B}} + \frac{1}{r_{12}} + \frac{1}{R}, \qquad (9)$$

with the fixed positions of the nuclei. After separation of the center of mass motion, the nuclear Hamiltonian in the reference frame fixed at the geometrical center of the nuclei, is

$$\begin{aligned} H_n &= -\frac{\nabla_R^2}{2\,\mu_n} - \frac{\nabla_{el}^2}{2\,\mu_n} - \left(\frac{1}{M_B} - \frac{1}{M_A}\right)\vec{\nabla}_R \cdot \vec{\nabla}_{el} \\ &= H_n' + H_n'', \end{aligned} \qquad (10)$$

where $\vec{\nabla}_{el} = \frac{1}{2}\sum_a \vec{\nabla}_a$ and $\mu_n = (1/M_A + 1/M_B)^{-1}$ is the nuclear reduced mass. The $H_n'$ part is even with respect to the inversion whereas the $H_n''$ is odd and vanishes for a homonuclear molecule.

The unperturbed (zeroth-order) wave function is assumed in the form of the product

$$\Psi_a(\vec{r}, \vec{R}) = \phi_{el}(\vec{r}; \vec{R})\,\chi(\vec{R}) \qquad (11)$$

of the nuclear wave function $\chi$ and the electronic wave function $\phi_{el}$ which implicitly depends on the nuclear coordinates $\vec{R}$. The function $\phi_{el}$ fulfills the electronic Schrödinger equation

$$H_{el}\,\phi_{el} = \mathcal{E}_{el}(R)\,\phi_{el}\,, \qquad (12)$$

while $\chi$ satisfies the nuclear equation with $\mathcal{E}_{el}(R)$ as an interaction potential.

Having this in mind, the total wave function can be expressed as a sum of terms parallel to and orthogonal to $\phi_{el}$

$$\Psi = \phi_{el}\,\chi + \delta\phi_{na}\,, \qquad (13)$$

where the latter means that the electronic matrix element

$$\langle \delta\phi_{\text{na}}|\phi_{\text{el}}\rangle_{\text{el}} = 0 \qquad (14)$$

vanishes. The symbol $\langle \ldots \rangle_{\text{el}}$ used henceforth represents an integration over electronic coordinates only. With such a representation of wave function, the Schrödinger equation (3) can also be decomposed into parallel and orthogonal parts

$$\left[ (H_{\text{el}} - \mathcal{E}_{\text{el}}) + (\mathcal{E}_{\text{el}} + H_{\text{n}} - E^{(2)}) \right] |\phi_{\text{el}}\,\chi + \delta\phi_{\text{na}}\rangle = 0 \qquad (15)$$

and rearranged further to

$$(\mathcal{E}_{\text{el}} - H_{\text{el}})|\delta\phi_{\text{na}}\rangle = (\mathcal{E}_{\text{el}} + H_{\text{n}} - E^{(2)})|\phi_{\text{el}}\,\chi + \delta\phi_{\text{na}}\rangle. \qquad (16)$$

Since $\delta\phi_{\text{na}}$ is orthogonal to $\phi_{\text{el}}$, the formal solution into the above equation can be expressed in the following recursive form employing the reduced resolvent $\frac{1}{(\mathcal{E}_{\text{el}} - H_{\text{el}})'}$:

$$|\delta\phi_{\text{na}}\rangle = \frac{1}{(\mathcal{E}_{\text{el}} - H_{\text{el}})'}\left[ H_{\text{n}}|\phi_{\text{el}}\,\chi\rangle + (\mathcal{E}_{\text{el}} + H_{\text{n}} - E^{(2)})|\delta\phi_{\text{na}}\rangle \right]. \qquad (17)$$

The left multiplication of Eq. (15) by $\langle\phi_{\text{el}}|$ gives

$$(\mathcal{E}_{\text{el}} + \mathcal{E}_{\text{a}} + H_{\text{n}} - E^{(2)})|\chi\rangle = -\langle\phi_{\text{el}}|H_{\text{n}}|\delta\phi_{\text{na}}\rangle_{\text{el}}, \qquad (18)$$

with $\mathcal{E}_{\text{a}}(R) \equiv \langle\phi_{\text{el}}|H_{\text{n}}|\phi_{\text{el}}\rangle_{\text{el}}$ being the adiabatic correction potential. Finally, insertion of (17) to the above equation forms a perturbative expansion for the effective nuclear Hamiltonian

$$(\mathcal{E}_{\text{el}} + \mathcal{E}_{\text{a}} + H_{\text{n}} - E^{(2)})|\chi\rangle = -(H_{\text{n}}^{(2)} + H_{\text{n}}^{(3)} + H_{\text{n}}^{(4)} + \ldots)|\chi\rangle. \qquad (19)$$

The leading terms of this series have the following explicit form:

$$H_{\text{n}}^{(2)} = \left\langle \phi_{\text{el}} \middle| H_{\text{n}} \frac{1}{(\mathcal{E}_{\text{el}} - H_{\text{el}})'} H_{\text{n}} \middle| \phi_{\text{el}} \right\rangle_{\text{el}}, \qquad (20)$$

$$H_{\text{n}}^{(3)} = \left\langle \phi_{\text{el}} \middle| H_{\text{n}} \frac{1}{(\mathcal{E}_{\text{el}} - H_{\text{el}})'} (H_{\text{n}} + \mathcal{E}_{\text{el}} - E^{(2)}) \right.$$
$$\left. \times \frac{1}{(\mathcal{E}_{\text{el}} - H_{\text{el}})'} H_{\text{n}} \middle| \phi_{\text{el}} \right\rangle_{\text{el}}, \qquad (21)$$

and

$$H_{\text{n}}^{(4)} = \left\langle \phi_{\text{el}} \middle| H_{\text{n}} \frac{1}{(\mathcal{E}_{\text{el}} - H_{\text{el}})'} (H_{\text{n}} + \mathcal{E}_{\text{el}} - E^{(2)}) \frac{1}{(\mathcal{E}_{\text{el}} - H_{\text{el}})'} \right.$$
$$\left. \times (H_{\text{n}} + \mathcal{E}_{\text{el}} - E^{(2)}) \frac{1}{(\mathcal{E}_{\text{el}} - H_{\text{el}})'} H_{\text{n}} \middle| \phi_{\text{el}} \right\rangle_{\text{el}}. \qquad (22)$$

Let us concentrate for a while on homonuclear molecules ($H_{\text{n}}'' = 0$). Taking into account the form (10) of the nuclear Hamiltonian, we can transform the above formulas further, e.g.,

$$H_{\text{n}}^{(2)} = \left\langle H_{\text{n}}\phi_{\text{el}} \middle| \frac{1}{(\mathcal{E}_{\text{el}} - H_{\text{el}})'} \middle| H_{\text{n}}\phi_{\text{el}} \right\rangle_{\text{el}} \qquad (23)$$

$$+ \frac{1}{\mu_{\text{n}}} \vec{\nabla}_R \left\langle \vec{\nabla}_R\phi_{\text{el}} \middle| \frac{1}{(\mathcal{E}_{\text{el}} - H_{\text{el}})'} \middle| H_{\text{n}}\phi_{\text{el}} \right\rangle_{\text{el}}$$
$$- \frac{1}{\mu_{\text{n}}} \left\langle H_{\text{n}}\phi_{\text{el}} \middle| \frac{1}{(\mathcal{E}_{\text{el}} - H_{\text{el}})'} \middle| \vec{\nabla}_R\phi_{\text{el}} \right\rangle_{\text{el}} \vec{\nabla}_R$$
$$- \frac{1}{\mu_{\text{n}}^2} \vec{\nabla}_R \left\langle \vec{\nabla}_R\phi_{\text{el}} \middle| \frac{1}{(\mathcal{E}_{\text{el}} - H_{\text{el}})'} \middle| \vec{\nabla}_R\phi_{\text{el}} \right\rangle_{\text{el}} \vec{\nabla}_R.$$

The nuclear function $\chi(\vec{R})$ can be factorized into the product of radial and angular parts,

$$\chi_{JM}(\vec{R}) = \frac{\chi_J(R)}{R} Y_{JM}(\vec{n}), \qquad (24)$$

with spherical harmonic $Y_{JM}$ and $\vec{n} = \vec{R}/R$. Such a factorization followed by integration over the angular variables leads to the radial form of the nuclear Hamiltonian

$$H_{\text{n}}^{(2)} = \mathcal{U}(R) + \left( \frac{2}{R} + \frac{\partial}{\partial R} \right) \mathcal{V}(R) \qquad (25)$$
$$- \frac{1}{R^2} \frac{\partial}{\partial R} R^2 \mathcal{W}_\parallel(R) \frac{\partial}{\partial R} + \frac{J(J+1)}{R^2} \mathcal{W}_\perp(R),$$

in which

$$\mathcal{U}(R) = \left\langle H_{\text{n}}\phi_{\text{el}} \middle| \frac{1}{(\mathcal{E}_{\text{el}} - H_{\text{el}})'} \middle| H_{\text{n}}\phi_{\text{el}} \right\rangle, \qquad (26)$$

$$\mathcal{V}(R) = \frac{1}{\mu_{\text{n}}} \left\langle \partial_R\phi_{\text{el}} \middle| \frac{1}{(\mathcal{E}_{\text{el}} - H_{\text{el}})'} \middle| H_{\text{n}}\phi_{\text{el}} \right\rangle, \qquad (27)$$

$$\mathcal{W}_\parallel(R) = \frac{1}{\mu_{\text{n}}^2} \left\langle \partial_R\phi_{\text{el}} \middle| \frac{1}{(\mathcal{E}_{\text{el}} - H_{\text{el}})'} \middle| \partial_R\phi_{\text{el}} \right\rangle, \qquad (28)$$

$$\mathcal{W}_\perp(R) = \frac{1}{2\mu_{\text{n}}^2 R^2} \left\langle \phi_{\text{el}} \middle| \vec{L}_{\text{el}} \frac{1}{(\mathcal{E}_{\text{el}} - H_{\text{el}})'} \vec{L}_{\text{el}} \middle| \phi_{\text{el}} \right\rangle. \qquad (29)$$

In the last equation, it is assumed that the electronic wave function $\phi_{\text{el}}$ in a ground molecular $\Sigma$ state is rotationally invariant, which implies $\vec{L}_n = -\vec{L}_{\text{el}}$ with $\vec{L}_n = -iR \times \vec{\nabla}_R$ and $\vec{L}_{\text{el}} = -i\sum_a \vec{r}_a \times \vec{\nabla}_a$.

The Hamiltonian $H_{\text{n}}^{(2)}$ contains only the terms proportional to $(m_{\text{e}}/\mu_{\text{n}})^2$, but $H_{\text{n}}^{(3)}$ and $H_{\text{n}}^{(4)}$ may have the terms with the second and higher powers of the electron-to-nucleus mass ratio. Because we are interested in the leading nonadiabatic correction, i.e. in the terms proportional to $(m_{\text{e}}/\mu_{\text{n}})^2$, all the $\mathcal{O}((m_{\text{e}}/\mu_{\text{n}})^3)$ corrections are neglected. However, some terms from $H_{\text{n}}^{(3)}$ and $H_{\text{n}}^{(4)}$ can be represented as $(m_{\text{e}}/\mu_{\text{n}})^2$ correction to the potential, or $(m_{\text{e}}/\mu_{\text{n}})^3$ to the nuclear kinetic energy. Their representation is not unique and may have different but equivalent forms, which differ by a commutator $[H_{\text{n}} + \mathcal{E}_{\text{el}} - E^{(2)}, Q]$ whose expectation value vanishes for an arbitrary $Q$. These terms are also neglected for consistency reasons. Hence we include only a term from the $H_{\text{n}}^{(3)}$, which has a unique representation and is the $(m_{\text{e}}/\mu_{\text{n}})^2$ correction to the potential [27], namely

$$\delta\mathcal{V}(R) = -\frac{1}{2\mu_{\text{n}}^2} \partial_R\mathcal{E}_{\text{el}} \left\langle \partial_R\phi_{\text{el}} \middle| \left[ \frac{1}{(\mathcal{E}_{\text{el}} - H_{\text{el}})'} \right]^2 \middle| \partial_R\phi_{\text{el}} \right\rangle \qquad (30)$$

and this correction is added to $\mathcal{V}(R)$ in Eq (25). The omitted components of $H_{\mathrm{n}}^{(3)}$ and of the higher-order Hamiltonians remain the main source of the uncertainty of the nonrelativistic results obtained within NAPT. The magnitude of this uncertainty can be estimated for each level separately by the value of the second-order NAPT correction to this level scaled by the $m_{\mathrm{e}}/\mu_{\mathrm{n}}$ factor. This estimation has been validated by the direct variational computations described in the preceding subsection. Because the direct calculations give an energy that corresponds to the perturbative series summed up to infinite order, the difference between the NAPT and the direct results accounts for all the omitted higher-order corrections [23, 25]—their values turn out to be smaller than the simple uncertainty estimation by the scaling. It has also been found that the missing contribution grows proportionally to $J(J+1)$ with a slope depending on the vibrational quantum number. We should also mention that the higher-order nonadiabatic corrections are singular at $R = 0$. It means that the nonadiabatic expansion does not work properly at a very small distance where the nuclear kinetic energy becomes comparable to the electronic one. Nevertheless, at $R = 0$ the exact wave function is as small as $10^{-25}$ (in a.u.), which makes these singular terms numerically negligible, but their existence indicates possible limitations of NAPT.

Equation (19), after reduction to one-dimensional form and the neglect of the $\mathcal{O}((m_{\mathrm{e}}/\mu_{\mathrm{n}})^3)$ terms, can be explicitly written as [27]

$$\left[ -\frac{1}{R^2} \frac{\partial}{\partial R} \frac{R^2}{2\,\mu_{\parallel}(R)} \frac{\partial}{\partial R} + \frac{J(J+1)}{2\,\mu_{\perp}(R)R^2} + \mathcal{Y}(R) \right] \chi_J(R)$$
$$= E^{(2)} \chi_J(R), \tag{31}$$

where the functions

$$\frac{1}{2\,\mu_{\parallel}(R)} \equiv \frac{1}{2\,\mu_{\mathrm{n}}} + \mathcal{W}_{\parallel}(R) \tag{32}$$

and

$$\frac{1}{2\,\mu_{\perp}(R)} \equiv \frac{1}{2\,\mu_{\mathrm{n}}} + \mathcal{W}_{\perp}(R) \tag{33}$$

can be interpreted as $R$-dependent vibrational and rotational masses, and where the potential $\mathcal{Y}(R)$ for the movement of the nuclei consists of the BO potential $\mathcal{E}_{\mathrm{el}}(R)$ [30], the adiabatic correction $\mathcal{E}_{\mathrm{a}}(R)$ [31], and the nonadiabatic correction $\delta\mathcal{E}_{\mathrm{na}}(R)$ [29] potentials. The latter correction is expressed in terms of the functions defined above:

$$\delta\mathcal{E}_{\mathrm{na}}(R) = \mathcal{U}(R) + \left( \frac{2}{R} + \frac{\partial}{\partial R} \right) [\mathcal{V}(R) + \delta\mathcal{V}(R)]. \tag{34}$$

Let us now return to the heteronuclear case, i.e. to the $H_{\mathrm{n}}$ Hamiltonian (10) in its full form. The unitary transformation from Ref. [28]

$$\tilde{H} = \left( e^{\lambda\,\vec{r}\cdot\vec{\nabla}_R} \right)^+ H\, e^{\lambda\,\vec{r}\cdot\vec{\nabla}_R} \tag{35}$$

with $\vec{r} = \sum_a \vec{r}_a$ and the nuclear mass asymmetry parameter

$$\lambda = -\frac{m_{\mathrm{e}}}{2} \left( \frac{1}{M_B} - \frac{1}{M_A} \right) \tag{36}$$

enables the heteronuclear part of the potential to be expressed as an additional correction to the potential only,

$$\delta\mathcal{E}'_{\mathrm{na}}(R) = \lambda^2 \left[ \left\langle \phi_{\mathrm{el}} \left| \frac{1}{m_{\mathrm{e}}} \nabla_R^2 + \frac{1}{2} r^i r^j \nabla_R^i \nabla_R^j(V) \right| \phi_{\mathrm{el}} \right\rangle_{\mathrm{el}} \right.$$
$$\left. + \left\langle \phi_{\mathrm{el}} \left| \vec{r}\cdot\vec{\nabla}_R(V) \frac{1}{(\mathcal{E}_{\mathrm{el}} - H_{\mathrm{el}})'} \vec{r}\cdot\vec{\nabla}_R(V) \right| \phi_{\mathrm{el}} \right\rangle_{\mathrm{el}} \right], \tag{37}$$

so that, finally,

$$\mathcal{Y}(R) = \mathcal{E}_{\mathrm{el}}(R) + \mathcal{E}_{\mathrm{a}}(R) + \delta\mathcal{E}_{\mathrm{na}}(R) + \delta\mathcal{E}'_{\mathrm{na}}(R). \tag{38}$$

There is, however, one problem with the heteronuclear correction $\delta\mathcal{E}'_{\mathrm{na}}$, because it behaves for small $R$ as $1/R^4$ and thus is singular. As mentioned earlier in this paragraph, NAPT does not work at very small nuclear distances, but this region is numerically insignificant. In practice, one can modify this potential $\delta\mathcal{E}'_{\mathrm{na}}(R) = \delta\mathcal{E}'_{\mathrm{na}}(R')$ for $R < R'$ and check that for a small $R'$ the results do not depend on its choice at the aimed precision.

Often it is the dissociation energy $D_{vJ}$ of rovibrational levels which is of interest. For this reason, we fix the origin of the energy scale to the separated atoms limit and make all the potentials vanish at infinity:

$$\tilde{\mathcal{Y}}(R) = \mathcal{Y}(R) - \mathcal{Y}(\infty). \tag{39}$$

Similarly, we subtract the asymptotic value from $\mathcal{W}$ potentials,

$$\tilde{\mathcal{W}}(R) = \mathcal{W}(R) - \mathcal{W}(\infty), \tag{40}$$

so that the $R$-dependent mass functions (32) and (33) correctly tend to the reduced atomic mass. In this convention, the eigenvalue $E^{(2)}$ of the Hamiltonian in Eq. (31) corresponds to the negative of $D_{vJ}$.

Equation (31) has been solved using two distinct numerical methods. One based on the Numerov integration method [32] and the other on the discrete variable representation (DVR) method, with mutual agreement between both of them. In this work we use the DVR method, due to its great efficiency and simplicity of implementation.

## C. Discrete Variable Representation

The DVR method is a pseudospectral method, making use of both a discrete grid and an associated basis set. There are many different flavors of DVR, using various basis sets and crafted for different integration ranges. The variant employed here [15] rests on the Fourier-basis version proposed by Colbert and Miller in Ref. [33]. It assumes the following expansion of the radial nuclear wave function:

$$\chi(R) = \sum_{n=1}^{N} f_n \phi_n(R), \tag{41}$$

where $\phi_n(R)$ are particle-in-a-box functions

$$\phi_n(R) = \left(\frac{2}{b-a}\right)^{1/2} \sin\left[\frac{n\pi(R-a)}{b-a}\right], \qquad (42)$$

where $R \in [a,b]$. The coefficients $f_n$ can be expressed via a numerical quadrature with weights $w_n$

$$f_n = \sum_{m=1}^{N} w_n \phi_n(R_m)\chi(R_m). \qquad (43)$$

The position $R$ is discretized on $N$ points—equal to the number of basis functions $\phi_n$

$$R_m = a + \frac{m(b-a)}{N+1}, \quad \Delta R = \frac{b-a}{N+1}, \qquad (44)$$

for $m = 1, ..., N$ (which means that neither $a$ nor $b$ are grid points themselves). The weights $w_n$ are all equal to the grid separation $\Delta R$ in this type of DVR [34]. Combining Eqs. (41) and (43), one gets

$$\chi(R) = \sum_{n=1}^{N}\sum_{m=1}^{N} \Delta R\, \phi_n(R_m)\chi(R_m)\phi_n(R)$$
$$= \sum_{m=1}^{N} \chi_m\, \varphi_m(R), \qquad (45)$$

where $\varphi_m(R)$ is a DVR orthonormal position basis [not to be confused with the auxiliary basis of Eq. (42)], and $\chi_m$ is proportional to the value of the wave function on the $R_m$ grid point

$$\varphi_m(R) = \sum_{n=1}^{N} \phi_n(R_m)\phi_n(R)\sqrt{\Delta R}, \qquad (46)$$

$$\chi_m = \chi(R_m)\sqrt{\Delta R}. \qquad (47)$$

It can be shown [33] that for $\phi_n$ of Eq. (42) the DVR basis function $\varphi_n$ exhibits asymptotically [i.e. for $(a-b) \to \infty$

and $N \to \infty$] the following property:

$$\varphi_n(R_m) = \frac{\delta_{nm}}{\sqrt{\Delta R}}. \qquad (48)$$

That is why the potential-energy matrices in the $\varphi_n$ DVR basis are diagonal:

$$V_{ij} = \sum_{n=1}^{N} \Delta R\, \varphi_i(R_n)V(R_n)\varphi_j(R_n) = \delta_{ij}V(R_j) \qquad (49)$$

and

$$\langle\chi|V|\chi\rangle \approx \sum_{n=1}^{N}\sum_{i=1}^{N}\sum_{j=1}^{N} \Delta R\, \chi_i\chi_j V(R_n)\varphi_i(R_n)\varphi_j(R_n)$$
$$= \sum_{n=1}^{N} \chi_n\chi_n V(R_n). \qquad (50)$$

The interval of $R$ in our problem is $[0,\infty)$, so $a = 0$. In practical applications $b$ and $N$ cannot be infinite, but already values as small as $N = 200$ and $R_N = 10.0$ are usually sufficient for most of our purposes. They only need to be increased when investigating highly excited vibrational states.

The matrix elements of differential operators are nondiagonal but still can be expressed by simple formulas. Namely, the Hamiltonian matrix elements in the BO approximation are given by

$$H_{ij} = \frac{1}{\mu_n\Delta R^2}(-1)^{i-j}\left(\frac{1}{(i-j)^2} - \frac{1}{(i+j)^2}\right), \qquad (51)$$

$$H_{ii} = \frac{1}{2\mu_n\Delta R^2}\left(\frac{\pi^2}{3} - \frac{1}{2i^2}\right) + \frac{J(J+1)}{2\mu_n R_i^2} + \mathcal{E}_{\text{el}}(R_i). \qquad (52)$$

The nonadiabatic Schrödinger equation (31) leads to a more elaborate formula—not only because of the "distance-dependent masses" present, but also because $\mathcal{W}_\parallel(R)$ is subjected to differentiation:

$$H_{ij} = \begin{cases} \dfrac{1}{\Delta R^2}\left(\dfrac{1}{2\mu_a} + \tilde{\mathcal{W}}_\parallel(R_i)\right)\left(\dfrac{\pi^2}{3} - \dfrac{1}{2i^2}\right) \\ \quad + \dfrac{\tilde{\mathcal{W}}_\parallel'(R_i)}{R_i} + \dfrac{1}{2}\tilde{\mathcal{W}}_\parallel''(R_i) + \left(\dfrac{1}{2\mu_a} + \tilde{\mathcal{W}}_\perp(R_i)\right)\dfrac{J(J+1)}{R_i^2} + \tilde{\mathcal{Y}}(R_i) \ \text{ for } i = j, \\ \dfrac{(-1)^{i-j}}{\Delta R^2}\left(\dfrac{1}{\mu_a} + \tilde{\mathcal{W}}_\parallel(R_i) + \tilde{\mathcal{W}}_\parallel(R_j)\right)\left(\dfrac{1}{(i-j)^2} - \dfrac{1}{(i+j)^2}\right) \qquad \text{ for } i \neq j, \end{cases} \qquad (53)$$

where $\tilde{\mathcal{W}}_\parallel'$ and $\tilde{\mathcal{W}}_\parallel''$ are the first and second derivatives of $\tilde{\mathcal{W}}_\parallel$ with respect to $R$, and $\tilde{\mathcal{Y}}(R)$ has been defined in Eqs. (38) and (39). Note that in the above formula the reduced nuclear mass

$\mu_n$ is replaced with the reduced atomic mass $\mu_a$

$$\frac{1}{\mu_a} = \frac{1}{m_A + m_e} + \frac{1}{m_B + m_e}. \qquad (54)$$

This is because Eq. (53) is written with respect to the disso-

ciation limit and, as discussed in Refs. [27, 28], $\mu_\perp(R)$ and $\mu_\parallel(R)$ tend to $\mu_a$ for $R \to \infty$.

The accuracy of $\mathcal{W}_\parallel$, $\mathcal{W}_\perp$, and $\delta\mathcal{E}_{na}$ is considered to be high enough to not contribute to the total nonrelativistic uncertainty. This uncertainty is dominated by the estimate of the neglected NAPT term, which is $m_e/\mu_n$ times the leading nonadiabatic correction. This missing-term uncertainty tends to be the largest source of the total theoretical error if the direct nonadiabatic results are not available.

In the case of the heteronuclear-specific correction $\delta\mathcal{E}'_{na}$, the former calculations [28] have been significantly improved here using the previously optimized ECG functions from Ref. [35]. As a result, the numerical uncertainty of $\delta\mathcal{E}'_{na}$ is

also negligible.

## IV. RELATIVISTIC CORRECTION $E^{(4)}$

The second term in the $\alpha$ expansion (1) is the leading relativistic correction of the order $\alpha^4 m$. It is the expectation value of the Breit-Pauli Hamiltonian

$$E^{(4)} = \langle \Psi | H^{(4)} | \Psi \rangle. \tag{55}$$

For the hydrogen molecule in the $^1\Sigma_g^+$ state (in which all the electron spin-dependent terms vanish) this Hamiltonian takes the following form [36]:

$$H^{(4)} = -\sum_a \frac{p_a^4}{8} - \sum_X \frac{p_X^4}{8m_X^3} + \frac{1}{2}\sum_{a,X}\frac{1}{m_X}p_a^i\left(\frac{\delta^{ij}}{r_{aX}} + \frac{r_{aX}^i r_{aX}^j}{r_{aX}^3}\right)p_X^j - \frac{1}{2}p_1^i\left(\frac{\delta^{ij}}{r_{12}} + \frac{r_{12}^i r_{12}^j}{r_{12}^3}\right)p_2^j$$
$$- \frac{1}{2m_A m_B}p_A^i\left(\frac{\delta^{ij}}{r_{AB}} + \frac{r_{AB}^i r_{AB}^j}{r_{AB}^3}\right)p_B^j + \frac{\pi}{2}\sum_{a,X}\left(1 + \frac{\delta_{s_X}}{m_X^2}\right)\delta^3(r_{aX}) + \pi\delta^3(r_{12}), \tag{56}$$

where $a$ goes over the electrons (1 and 2) and $X$ – over the nuclei ($A$ and $B$), and $\delta_s$ depends on the nuclear spin $s$: $\delta_s = 0$ for $s = 0$ or 1, and $\delta_s = 1$ for $s = 1/2$. Its first two terms account for the relativistic correction to the kinetic energy. The third, fourth, and fifth terms are called the Breit corrections (or "orbit-orbit coupling" terms) and can be attributed to the relativistic retardation of the Coulomb potential [37]. The remaining contributions are represented by the so-called contact terms and are proportional to the 3D Dirac delta functions. In practice, even if the above Hamiltonian is used in the fully nonadiabatic approach (treating electrons and nuclei on an equal footing), the second term is neglected—being proportional to the very small $m_e^3/m_X^3$ factor.

For the dissociation energy $D$, one subtracts $E^{(4)}$ from the relativistic correction for separated atoms $E_A^{(4)} + E_B^{(4)}$, where

$$E_X^{(4)} = -\frac{1}{8} + \frac{1}{4}\left(\frac{1}{m_X}\right)^2 + O\left(\frac{1}{m_X}\right)^3. \tag{57}$$

Note that the term proportional to $1/m_X$ is not present in the above formula; consequently, the relativistic recoil correction for separated atoms is very small.

The direct calculations of the relativistic correction with the nonadiabatic wave function have been performed only for the ground molecular state $v = 0, J = 0$ [36, 38]. The results for arbitrary vibrationally and rotationally excited states have been obtained within the NAPT approach,

$$E^{(4)} = E^{(4,0)} + E^{(4,1)} + \ldots, \tag{58}$$

described in the following subsections.

### A. Leading-order relativistic correction, $E^{(4,0)}$

The leading relativistic contribution in the BO approximation consists of the nuclear-mass-independent terms from the Breit-Pauli Hamiltonian (56):

$$H^{(4,0)} = -\frac{p_1^4 + p_2^4}{8} - \frac{1}{2}p_1^i\left(\frac{\delta^{ij}}{r_{12}} + \frac{r_{12}^i r_{12}^j}{r_{12}^3}\right)p_2^j + \pi\delta^3(r_{12})$$
$$+ \frac{\pi}{2}\left(\delta^3(r_{1A}) + \delta^3(r_{2A}) + \delta^3(r_{1B}) + \delta^3(r_{2B})\right). \tag{59}$$

The correction to the BO potential energy is the expectation value with the electronic wave function

$$\mathcal{E}^{(4,0)}(R) = \langle \phi_{el} | H^{(4,0)} | \phi_{el} \rangle. \tag{60}$$

It was observed in [39] that the numerical convergence is significantly improved when the electronic wave function $\phi_{el}$ satisfies the electron-electron cusp condition. The final value of the leading-order relativistic correction is evaluated as the expectation value with the nuclear wave function

$$E^{(4,0)} = \langle \chi | \mathcal{E}^{(4,0)}(R) | \chi \rangle \tag{61}$$

and this is calculated using the DVR function from the nonrelativistic BO approximation. On the basis of the results of Ref. [39], a numerical uncertainty $\delta E^{(4,0)}$ is estimated by $2 \times 10^{-6} E^{(4,0)}$.

### B. Finite nuclear mass relativistic correction $E^{(4,1)}$

The leading finite-nuclear-mass relativistic correction can also be expressed in terms of effective internuclear potential

and consists of three parts [35],

$$\mathcal{E}^{(4,1)}(R) = \mathcal{E}_1^{(4,1)}(R) + \mathcal{E}_2^{(4,1)}(R) + \mathcal{E}_3^{(4,1)}(R), \quad (62)$$

where

$$\mathcal{E}_1^{(4,1)}(R) = \frac{1}{\mu_{\mathrm{n}}} \langle \vec{\nabla}_R \phi_{\mathrm{rel}} | \vec{\nabla}_R \phi_{\mathrm{el}} \rangle, \quad (63)$$

$$\mathcal{E}_2^{(4,1)}(R) = -\frac{1}{\mu_{\mathrm{n}}} \langle \phi_{\mathrm{rel}} | \vec{\nabla}_{\mathrm{el}}^2 | \phi_{\mathrm{el}} \rangle, \quad (64)$$

$$\mathcal{E}_3^{(4,1)}(R) = \langle \phi_{\mathrm{el}} | H^{(4,1)} | \phi_{\mathrm{el}} \rangle, \quad (65)$$

and $\phi_{\mathrm{rel}}$ is a relativistic correction to the BO electronic wave function:

$$|\phi_{\mathrm{rel}}\rangle = \frac{1}{(\mathcal{E}_{\mathrm{el}} - H_{\mathrm{el}})'} H^{(4,0)} |\phi_{\mathrm{el}}\rangle. \quad (66)$$

The Hamiltonian $H^{(4,1)}$ describes the electron-nucleus Breit interaction, which in the coordinate system assumed in this work takes the form

$$H^{(4,1)} = \quad (67)$$
$$-\frac{1}{4\mu_{\mathrm{n}}} \sum_{a=1,2} \nabla_a^i \left( \frac{\delta^{ij}}{r_{aA}} + \frac{r_{aA}^i r_{aA}^j}{r_{aA}^3} - \frac{\delta^{ij}}{r_{aB}} - \frac{r_{aB}^i r_{aB}^j}{r_{aB}^3} \right) \nabla_R^j$$
$$+\frac{1}{4\mu_{\mathrm{n}}} \sum_{a=1,2} \nabla_a^i \left( \frac{\delta^{ij}}{r_{aA}} + \frac{r_{aA}^i r_{aA}^j}{r_{aA}^3} + \frac{\delta^{ij}}{r_{aB}} + \frac{r_{aB}^i r_{aB}^j}{r_{aB}^3} \right) \nabla_{\mathrm{el}}^j.$$

This effective internuclear potential $\mathcal{E}^{(4,1)}$ is used to obtain the relativistic recoil correction to rovibrational levels, using

$$E^{(4,1)} = \langle \chi | \mathcal{E}^{(4,1)}(R) | \chi \rangle \quad (68)$$
$$+ 2 \langle \chi | \mathcal{E}^{(4,0)}(R) \frac{1}{(E^{(2,0)} - H_{\mathrm{n}})'} \mathcal{E}^{(2,1)}(R) | \chi \rangle,$$

where $\mathcal{E}^{(2,1)}(R) = \mathcal{E}_{\mathrm{a}}(R)$. The potential $\mathcal{E}^{(4,1)}(R)$ has been reported recently in Ref. [35]. The numerical error contributed by the potential was estimated to be $2 \times 10^{-4} \langle \chi | \mathcal{E}^{(4,1)} | \chi \rangle$. Furthermore, because currently no higher finite-nuclear-mass relativistic corrections are known, the effect of their omission is approximated by $E^{(4,1)} m_{\mathrm{e}}/\mu_{\mathrm{n}}$ and included in the total $E^{(4)}$ error estimate.

## V. QED CORRECTIONS

### A. Leading-order QED correction $E^{(5)}$

The complete formula for the leading quantum electrodynamic correction $E^{(5)}$ for $H_2$ and its isotopologues was obtained in Refs. [38, 40]. Direct (four-body) numerical calculations have been performed only for the ground molecular level, whereas for all the excited levels we use the BO approximation. The leading QED Born-Oppenheimer contribution can be expressed as

$$E^{(5,0)} = \langle \chi | \mathcal{E}^{(5,0)}(R) | \chi \rangle, \quad (69)$$

where

$$\mathcal{E}^{(5,0)}(R) = \frac{4}{3} \left[ \frac{19}{30} - 2 \ln \alpha - \ln k_0(R) \right] \sum_{a,X} \langle \delta^3(r_{aX}) \rangle_{\mathrm{el}}$$
$$+ \left[ \frac{164}{15} + \frac{14}{3} \ln \alpha \right] \langle \delta^3(r_{12}) \rangle_{\mathrm{el}} - \frac{7}{6\pi} \left\langle \frac{1}{r_{12}^3} \right\rangle_{\mathrm{el},\epsilon}. \quad (70)$$

In the above formula the expectation values are evaluated with the nonrelativistic wave function $\phi_{\mathrm{el}}$, and the notation $\langle 1/r_{ij}^3 \rangle_\epsilon$ means the following:

$$\left\langle \frac{1}{r_{ij}^3} \right\rangle_\epsilon = \lim_{\epsilon \to 0} \left[ \left\langle \frac{\theta(r_{ij} - \epsilon)}{r_{ij}^3} \right\rangle + 4\pi(\gamma + \ln \epsilon) \langle \delta^3(r_{ij}) \rangle \right], \quad (71)$$

where the symbol $\gamma$ denotes the Euler-Mascheroni constant, and $\theta$ is the Heaviside function. The Bethe logarithm $\ln k_0(R)$ is

$$\ln k_0(R) = \frac{\langle \phi_{\mathrm{el}} | \vec{j} (H_{\mathrm{el}} - \mathcal{E}_{\mathrm{el}}) \ln[2(H_{\mathrm{el}} - \mathcal{E}_{\mathrm{el}})] \vec{j} | \phi_{\mathrm{el}} \rangle}{\langle \phi_{\mathrm{el}} | \vec{j} (H_{\mathrm{el}} - \mathcal{E}_{\mathrm{el}}) \vec{j} | \phi_{\mathrm{el}} \rangle}, \quad (72)$$

with $\vec{j} = -\vec{p}_1/m_{\mathrm{e}} - \vec{p}_2/m_{\mathrm{e}}$. It has been calculated in Ref. [41], whereas the results for the Araki-Sucher term and Dirac $\delta$ are taken from newer calculations reported in Refs. [39, 42]. The numerical uncertainty is estimated to be ca. $5 \times 10^{-4} \langle \chi | \mathcal{E}^{(5,0)}(R) | \chi \rangle$. The greatest source of error in this term comes from the uncalculated finite-nuclear-mass contribution, estimated as $E^{(5,0)} m_{\mathrm{e}}/\mu_{\mathrm{n}}$. For the levels and transitions where the nonrelativistic contribution is calculated directly (so the NAPT error is removed), it dominates the total theoretical uncertainty.

### B. Higher-order QED correction $E^{(6)}$

The higher-order QED contribution is calculated within the BO approximation and is given by

$$\mathcal{E}^{(6,0)}(R) = \langle \phi_{\mathrm{el}} | H^{(6,0)} | \phi_{\mathrm{el}} \rangle$$
$$+ \langle \phi_{\mathrm{el}} | H^{(4,0)} \frac{1}{(\mathcal{E}_{\mathrm{el}} - H_{\mathrm{el}})'} H^{(4,0)} | \phi_{\mathrm{el}} \rangle, \quad (73)$$

where $H^{(4,0)}$ is the Breit Hamiltonian in the nonrecoil limit, and $H^{(6,0)}$ is the $O(\alpha^2)$ correction to this Hamiltonian. The explicit formulas for $\mathcal{E}^{(6,0)}(R)$ are far too extensive to be presented here. They can be found in Ref. [42]. The total energy contribution in this order is

$$E^{(6,0)} = \langle \chi | \mathcal{E}^{(6,0)}(R) | \chi \rangle$$
$$+ \langle \chi | \mathcal{E}^{(4,0)}(R) \frac{1}{(E^{(2,0)} - H_{\mathrm{n}})'} \mathcal{E}^{(4,0)}(R) | \chi \rangle. \quad (74)$$

The second term in the above equation is again the second-order relativistic correction with respect to the relativistic

BO potential (in our former works presented separately as $E_{\text{sec}}^{(6)}$ correction). The $\mathcal{E}^{(6,0)}(R)$ potential was calculated in Ref. [42] and $\mathcal{E}^{(4,0)}(R)$ in Ref. [39]. The numerical error was estimated as $3 \times 10^{-3} E^{(6,0)}$, whereas the missing finite-nuclear-mass correction was estimated as $E^{(6,0)} m_{\text{e}}/\mu_{\text{n}}$.

### C.  Estimation of $E^{(7)}$

The $E^{(7)}$ correction is of the highest order considered so far for the hydrogen molecule. Currently, its complete form is unknown. Here, we follow Ref. [40] and include the leading one- and two-loop radiative corrections known from the hydrogen atom (see [43]) in the BO approximation

$$E^{(7)} = \langle \chi | \mathcal{E}^{(7)}(R) | \chi \rangle \,, \tag{75}$$

where

$$\mathcal{E}^{(7)}(R) \approx \pi \, \langle \phi_{\text{el}} | \sum_{a,X} \delta^3(r_{aX}) | \phi_{\text{el}} \rangle_{\text{el}} \left\{ \frac{1}{\pi} \big[ A_{60} \right. \tag{76}$$

$$\left. + A_{61} \, \ln \alpha^{-2} + A_{62} \, \ln^2 \alpha^{-2} \big] + \frac{1}{\pi^2} \, B_{50} + \frac{1}{\pi^3} \, C_{40} \right\}.$$

As an uncertainty of the $E^{(7)}$ correction, following Ref. [40], we assume 25% of its value.

### VI.  FINITE NUCLEAR SIZE EFFECT $E_{\text{FS}}$

At the achieved accuracy level, the nuclear finite size effect cannot be neglected anymore. This correction, when evaluated in the BO approximation,

$$E_{\text{FS}}^{(4)} = \langle \chi | \mathcal{E}_{\text{FS}}^{(4)}(R) | \chi \rangle \,, \tag{77}$$

is accounted for by the following formula:

$$\mathcal{E}_{\text{FS}}^{(4)}(R) = \frac{2\pi}{3} \, \langle \phi_{\text{el}} | \sum_{a,X} \delta^3(r_{aX}) | \phi_{\text{el}} \rangle_{\text{el}} \, \frac{(r_{C,A}^2 + r_{C,B}^2)}{2 \, \lambdabar^2}, \tag{78}$$

where $\lambdabar$ is the reduced electron Compton wavelength; $r_{C,X}^2$ is the mean square charge radius of the nucleus $X = A, B$, with $r_p = 0.8414(19)$ fm [44], $r_d = 2.12799(74)$ fm [44], and $r_t = 1.7591(363)$ fm [45]. Any higher-order effects due to the nuclear size or nuclear polarizability are neglected. The accuracy of $E_{\text{FS}}^{(4)}$ is limited by the accuracy of the charge radii. The connection between the dissociation energy and the charge radius of a nucleus can potentially be utilized to determine the latter one provided that both theoretical and experimental dissociation energy are known to a sufficient accuracy, which is about $10^{-7}$ cm$^{-1}$.

### VII.  UNCERTAINTY ESTIMATION

Previous sections devoted to individual components $E^{(i)}$ of the $\alpha$ expansion (1) contain a short description of the uncertainty estimates $\delta E^{(i)}$, which are assumed to be uncorrelated.

Therefore, the total uncertainty is the square-root of the sum of squares of all the partial uncertainties. Depending on the availability of the direct nonadiabatic results for a given rovibrational level we can distinguish three different cases. In the first case, the direct nonadiabatic results are available for the nonrelativistic energy as well as for relativistic and QED corrections. In this case, currently represented by the ground levels of all the isotopologues, the dominating uncertainty comes from the incomplete knowledge of the $E^{(7)}$ term. In the second case, only the nonrelativistic energy $E^{(2)}$ is known with high accuracy from the direct nonadiabatic calculations. In such a case, the overall accuracy is limited by the lack of the recoil correction to the leading QED term $E^{(5)}$. Finally, in the third and the most common case, all the energy components are evaluated from the NAPT. Then, the limitations in accuracy originate either from the nonrelativistic or QED component of the energy.

The estimation of the uncertainty assigned to a transition energy is more complicated. Depending on the pair of the states involved in a given transition, we observe smaller or larger cancellation of different energy components. A systematic description of this cancellation is difficult, and we assumed in general that the uncertainty assigned the a transition energy is equal to the larger uncertainty out of these two states. However, in particular cases, like the fundamental $\nu = 0 \rightarrow 1$ transitions, the cancellation of uncertainties is significant, and we associate relative $m_{\text{e}}/\mu_{\text{n}}$ uncertainty to the energy difference, as demonstrated in Table I.

We also note that, because at the long-distance points the accuracy of the potentials usually deteriorates, our error estimates for highly excited levels can be inaccurate.

### VIII.  RESULTS AND SUMMARY

One of the most pronounced features of NAPT combined with $\alpha$ expansion in Eq. (1) is that it gives access to an arbitrary bound rovibrational energy level within the electronic ground state. Consequently, it enables all transitions to be obtained within this manifold of levels. Another merit of NAPT with $\alpha$ expansion is the possibility of full control of the accuracy of the results, as well as the potential of gradually increasing this accuracy by improving the existing, and adding new, terms to the expansion series.

The theoretical underpinning presented in this work as well as numerical calculations performed over the past years enabled construction of a computer program [15] serving the numerical values of the rovibrational energy levels and splittings between them for all the isotopologues of the hydrogen molecule. There are several thousands of such levels and many more transitions available from this program. It would be impractical to present such a large amount of data in printed form. Therefore, this program has been made publicly available to the scientific community so that numerical results for levels or transitions of interest (in particular also for all of them) can be easily generated by the reader. This form of the presentation of the results has also another important advantage—it is our intention to support the program in the

TABLE I. Selected transition energies (in cm$^{-1}$) obtained for H$_2$ from the NAPT and direct nonadiabatic calculations with breakdown into components. CODATA 2018 [44] values of physical constants are used. For a shorthand notation, we shall identify the $\alpha^n m E^{(n)}$ terms with the bare coefficients $E^{(n)}$.

|  | (0,1) — (0,0) | (1,0) — (0,0) | (1,1) — (0,1) | (2,1) — (0,3) | (3,5) — (0,3) |
|---|---|---|---|---|---|
| $E^{(2)}$(NAPT) | 118.485 262(7) | 4 161.164 2(9) | 4 155.252 0(9) | 7 488.283 3(17) | 12 559.750 0(25) |
| $E^{(2)}$(direct) | 118.485 260 5(1) | 4 161.164 070 0(1) | 4 155.251 869 3(1) | 7 488.283 212 0(1) | 12 559.749 918 5(1) |
| $E^{(4)}$ | 0.002 583 6 | 0.023 553 9(2) | 0.023 333 3(2) | 0.028 570 7(3) | 0.065 877 6(6) |
| $E^{(5)}$ | −0.001 022 7(12) | −0.021 318(26) | −0.021 257(25) | −0.036 018(43) | −0.065 815(79) |
| $E^{(6)}$ | −0.000 008 9 | −0.000 191 3(6) | −0.000 190 8(6) | −0.000 326 3(10) | −0.000 594 9(19) |
| $E^{(7)}$ | 0.000 000 5(1) | 0.000 010 3(26) | 0.000 010 3(26) | 0.000 017 4(44) | 0.000 031 9(80) |
| $E^{(4)}_{\mathrm{FS}}$ | −0.000 000 2 | −0.000 003 2 | −0.000 003 2 | −0.000 005 4 | −0.000 009 8 |
| $E$ | 118.486 812 8(12) | 4 161.166 122(26) | 4 155.253 762(26) | 7 488.275 451(43) | 12 559.749 408(79) |
| Exp. | 118.486 8(1) [46] | 4 161.166 36(15) [2] | 4 155.254 00(21) [2] | 7 488.275 3(10) [47] | 12 559.749 39(22) [48] |
| Diff. | 0.000 0(1) | −0.000 24(15) | −0.000 24(21) | +0.000 2(10) | +0.000 02(23) |

future by updating the input potentials and physical constants and possibly by adding new functionalities—a guarantee that it has the best currently available data.

Here, we present only a small selection of the numerical results to illustrate the most important features of NAPT combined with $\alpha$ expansion. In Table II we show the total dissociation energy $E$ for the ground level of three lightest isotopologues. These energies are compared with the reference theoretical results obtained from direct nonadiabatic calculations and with the best available experimental data. This comparison shows the current accuracy limitations of NAPT but simultaneously demonstrates that this method performs very well because its results agree within uncertainties with direct variational calculations.

The finite nuclear mass effects are the most significant in H$_2$ because it is the lightest isotopologue. Therefore, Table I contains a few examples of transitions between rovibrational levels of H$_2$ with growing energies. In most cases, theoretical energies are an order of magnitude more accurate than experimental values and in agreement with them. However, we observe a significant $3\sigma$ discrepancy between the measured dissociation energy and our calculations for HD, in spite of a good agreement for H$_2$ and D$_2$; see Table II. Before drawing any conclusions, this experimental value should be verified.

[1] G. D. Dickenson, M. L. Niu, E. J. Salumbides, J. Komasa, K. S. E. Eikema, K. Pachucki, and W. Ubachs, Phys. Rev. Lett. **110**, 193601 (2013).

[2] M. Niu, E. Salumbides, G. Dickenson, K. Eikema, and W. Ubachs, J. Mol. Spectrosc. **300**, 44 (2014).

[3] M. L. Niu, E. J. Salumbides, and W. Ubachs, J. Chem. Phys. **143**, 081102 (2015).

[4] D. Mondelain, S. Kassi, T. Sala, D. Romanini, D. Gatti, and A. Campargue, J. Mol. Spectrosc. **326**, 5 (2016).

[5] M. Schlösser, X. Zhao, M. Trivikram, W. Ubachs, and E. J. Salumbides, J. Phys. B **50**, 214004 (2017).

[6] R. K. Altmann, L. S. Dreissen, E. J. Salumbides, W. Ubachs, and K. S. E. Eikema, Phys. Rev. Lett. **120**, 043204 (2018).

[7] C.-F. Cheng, J. Hussels, M. Niu, H. L. Bethlem, K. S. E. Eikema, E. J. Salumbides, W. Ubachs, M. Beyer, N. J. Hölsch, J. A. Agner, et al., Phys. Rev. Lett. **121**, 013001 (2018).

[8] F. M. J. Cozijn, P. Dupré, E. J. Salumbides, K. S. E. Eikema, and W. Ubachs, Phys. Rev. Lett. **120**, 153002 (2018).

[9] E. Fasci, A. Castrillo, H. Dinesan, S. Gravina, L. Moretti, and L. Gianfrani, Phys. Rev. A **98**, 022516 (2018).

[10] R. Z. Martinez, D. Bermejo, P. Wcisło, and F. Thibault, J. Raman Spectrosc. **50**, 127 (2018).

[11] L.-G. Tao, A.-W. Liu, K. Pachucki, J. Komasa, Y. R. Sun, J. Wang, and S.-M. Hu, Phys. Rev. Lett. **120**, 153001 (2018).

[12] T. M. Trivikram, M. Schlösser, W. Ubachs, and E. J. Salumbides, Phys. Rev. Lett. **120**, 163002 (2018).

[13] P. Wcisło, F. Thibault, M. Zaborowski, S. Wójtewicz, A. Cygan, G. Kowzan, P. Masłowski, J. Komasa, M. Puchalski, K. Pachucki, et al., J. Quant. Spectrosc. Radiat. Transfer **213**, 41 (2018).

[14] N. Hölsch, M. Beyer, E. J. Salumbides, K. S. E. Eikema, W. Ubachs, C. Jungen, and F. Merkt, Phys. Rev. Lett. **122**, 103002 (2019).

[15] H2SPECTRE ver. 7.0 Fortran source code (2019), URL https://www.fuw.edu.pl/~krp; http://qcg.home.amu.edu.pl/qcg/public_html/H2Spectre.html; P. Czachorowski, Ph.D. thesis, University of Warsaw, Poland (2019).

[16] E. Roueff, H. Abgrall, P. Czachorowski, K. Pachucki, M. Puchalski, and J. Komasa, Astron. Astrophys. **630**, A58 (2019).

TABLE II. Comparison of dissociation energies (in cm$^{-1}$) of the ground levels of H$_2$, D$_2$, and HD obtained in the framework of NAPT with the results of direct nonadiabatic calculations and with experimental data. CODATA 2018 [44] values of physical constants are used.

|  | H$_2$ | D$_2$ | HD |
|---|---|---|---|
| $E$ (NAPT) | 36 118.069 45(53) | 36 748.362 27(17) | 36 405.782 37(33) |
| $E$ (direct) | 36 118.069 632(26) [40] | 36 748.362 342(26) [38] | 36 405.782 478(26) [38] |
| Experiment | 36 118.069 45(31) [6] | 36 748.362 86(68) [49] | 36 405.783 66(36) [50] |

[17] K.-F. Lai, P. Czachorowski, M. Schlosser, M. Puchalski, J. Komasa, K. Pachucki, W. Ubachs, and E. J. Salumbides (2019), submitted.

[18] M. Beyer, N. Hölsch, J. Hussels, C. -F. Cheng, E. Salumbides, K. S. E. Eikema, W. Ubachs, Ch. Jungen, and F. Merkt, Phys. Rev. Lett. (unpublished).

[19] L. I. Bodine, D. S. Parno, and R. G. H. Robertson, Phys. Rev. C **91**, 035505 (2015).

[20] W. E. Caswell and G. P. Lepage, Phys. Lett. B **167**, 437 (1986).

[21] K. Pachucki and J. Komasa, J. Chem. Phys. **144**, 164306 (2016).

[22] H. M. James and A. S. Coolidge, J. Chem. Phys. **1**, 825 (1933).

[23] K. Pachucki and J. Komasa, Phys. Chem. Chem. Phys. **20**, 247 (2018).

[24] K. Pachucki and J. Komasa, Phys. Chem. Chem. Phys. **20**, 26297 (2018).

[25] K. Pachucki and J. Komasa, Phys. Chem. Chem. Phys. **21**, 10272 (2019).

[26] K. Pachucki and J. Komasa, J. Chem. Phys. **129**, 034102 (2008).

[27] K. Pachucki and J. Komasa, J. Chem. Phys. **130**, 164113 (2009).

[28] K. Pachucki and J. Komasa, Phys. Chem. Chem. Phys. **12**, 9188 (2010).

[29] K. Pachucki and J. Komasa, J. Chem. Phys. **143**, 034111 (2015).

[30] K. Pachucki, Phys. Rev. A **82**, 032509 (2010).

[31] K. Pachucki and J. Komasa, J. Chem. Phys. **141**, 224103 (2014).

[32] W. R. Johnson, *Atomic Structure Theory*, Lectures on Atomic Physics (Springer-Verlag, Berlin and New York, 2007).

[33] D. T. Colbert and W. H. Miller, J. Chem. Phys. **96**, 1982 (1992).

[34] G. C. Groenenboom and D. T. Colbert, J. Chem. Phys. **99**, 9681 (1993).

[35] P. Czachorowski, M. Puchalski, J. Komasa, and K. Pachucki, Phys. Rev. A **98**, 052506 (2018).

[36] M. Puchalski, A. Spyszkiewicz, J. Komasa, and K. Pachucki, Phys. Rev. Lett. **121**, 073001 (2018).

[37] H. A. Bethe and E. E. Salpeter, *Quantum Mechanics of One- and Two-Electron Systems* (Springer-Verlag, Berlin and New York, 1957).

[38] M. Puchalski, J. Komasa, A. Spyszkiewicz, and K. Pachucki, Phys. Rev. A **100**, 020503(R) (2019).

[39] M. Puchalski, J. Komasa, and K. Pachucki, Phys. Rev. A **95**, 052506 (2017).

[40] M. Puchalski, J. Komasa, P. Czachorowski, and K. Pachucki, Phys. Rev. Lett. **122**, 103003 (2019).

[41] K. Piszczatowski, G. Łach, M. Przybytek, J. Komasa, K. Pachucki, and B. Jeziorski, J. Chem. Theory Comput. **5**, 3039 (2009).

[42] M. Puchalski, J. Komasa, P. Czachorowski, and K. Pachucki, Phys. Rev. Lett. **117**, 263002 (2016).

[43] M. I. Eides, H. Grotch, and V. A. Shelyuto, Phys. Rep. **342**, 63 (2001).

[44] 2018 CODATA recommended values (2018), URL https://physics.nist.gov/cuu/Constants.

[45] I. Angeli and K. P. Marinova, Atomic Data and Nuclear Data Tables **99**, 69 (2013).

[46] D. E. Jennings, S. L. Bragg, and J. W. Brault, Astrophys. J. **282**, L85 (1984).

[47] A. Campargue, S. Kassi, K. Pachucki, and J. Komasa, Phys. Chem. Chem. Phys. **14**, 802 (2012).

[48] C.-F. Cheng, Y. R. Sun, H. Pan, J. Wang, A.-W. Liu, A. Campargue, and S.-M. Hu, Phys. Rev. A **85**, 024501 (2012); P. Wcisło, I. E. Gordon, C.-F. Cheng, S.-M. Hu, and R. Ciuryło, Phys. Rev. A **93**, 022501 (2016).

[49] J. Liu, D. Sprecher, C. Jungen, W. Ubachs, and F. Merkt, J. Chem. Phys. **132**, 154301 (2010).

[50] D. Sprecher, J. Liu, C. Jungen, W. Ubachs, and F. Merkt, J. Chem. Phys. **133**, 111102 (2010).