Journal of Quantitative Spectroscopy & Radiative Transfer 186 (2017) 179–196

Contents lists available at ScienceDirect

Journal of Quantitative Spectroscopy & Radiative Transfer

journal homepage: www.elsevier.com/locate/jqsrt

dPotFit: A computer program to fit diatomic molecule spectral data to potential energy functions

Robert J. Le Roy

Department of Chemistry, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1

ARTICLE INFO

Article history:
Received 17 February 2016
Received in revised form
1 June 2016
Accepted 1 June 2016
Available online 8 June 2016

Keywords:
Diatom PEF determination
Diatom data analysis
Multi-property/multi-isotopolog fits
Analytic diatom potentials

ABSTRACT

This paper describes program **dPotFit**, which performs least-squares fits of diatomic molecule spectroscopic data consisting of any combination of microwave, infrared or electronic vibrational bands, fluorescence series, and tunneling predissociation level widths, involving one or more electronic states and one or more isotopologs, and for appropriate systems, second virial coefficient data, to determine analytic potential energy functions defining the observed levels and other properties of each state. Four families of analytical potential functions are available for fitting in the current version of **dPotFit**: the Expanded Morse Oscillator (EMO) function, the Morse/Long-Range (MLR) function, the Double-Exponential/Long-Range (DELR) function, and the 'Generalized Potential Energy Function' (GPEF) of Šurkus, which incorporates a variety of polynomial functional forms. In addition, **dPotFit** allows sets of experimental data to be tested against predictions generated from three other families of analytic functions, namely, the 'Hannover Polynomial' (or "X-expansion") function, and the 'Tang–Toennies' and Scoles–Aziz 'HFD', exponential-plus-van der Waals functions, and from interpolation-smoothed pointwise potential energies, such as those obtained from *ab initio* or RKR calculations. **dPotFit** also allows the fits to determine atomic-mass-dependent Born–Oppenheimer breakdown functions, and singlet-state Λ-doubling, or $^2\Sigma$ splitting radial strength functions for one or more electronic states.

dPotFit always reports both the 95% confidence limit uncertainty and the "sensitivity" of each fitted parameter; the latter indicates the number of significant digits that must be retained when rounding fitted parameters, in order to ensure that predictions remain in full agreement with experiment. It will also, if requested, apply a "sequential rounding and refitting" procedure to yield a final parameter set defined by a minimum number of significant digits, while ensuring no significant loss of accuracy in the predictions yielded by those parameters.

1. Introduction

In recent years, it has become increasingly common to analyze diatomic molecule spectroscopic data by performing "direct potential fits", in which observed transition energies are compared with eigenvalue differences calculated from an effective radial Schrödinger equation based on a parameterized analytic potential energy function. This effective radial Hamiltonian may also include radial strength functions characterizing the atomic-mass-dependent potential energy and centrifugal Born–Oppenheimer breakdown (BOB) functions, and (if appropriate) radial strength functions that account for Λ-doubling in singlet states or doublet splittings in $^2\Sigma$ states. Partial derivatives of calculated eigenvalues with respect to the parameters defining the potential energy and other radial functions, such as those characterizing BOB corrections,

E-mail address: leroy@UWaterloo.ca

http://dx.doi.org/10.1016/j.jqsrt.2016.06.002
0022-4073/

Λ-doubling, or $^2\Sigma$ splittings, are then used in least-squares fits to determine an optimized radial Hamiltonian for the system. This paper describes a robust and flexible computer program for performing this type of analysis.

One of the most important features of a 'direct potential fit' (DPF) data analysis is that it provides a compact way of summarizing and reproducing very large bodies of very different types of experimental data in a compact manner that facilitates making predictions for unobserved data and properties. **dPotFit** is able to fit to any combination of pure rotational ('microwave'), vib-rotational ('infrared'), and electronic data in the form of fully assigned bands or fluorescence series, to the energies and/or tunneling predissociation widths of 'quasibound' levels, to binding energies from photoassociation spectroscopy (PAS), and to the 'pressure virial' and 'acoustic virial' coefficients that reflect the non-ideal-gas properties of the atomic dissociation products. It also allows one to combine data for different isotopologs of a given molecular species and different electronic states in a single unified analysis. Moreover, if the experimental information for a particular electronic state is not sufficiently extensive or systematic to define a full potential energy function (PEF) for it, **dPotFit** allows its energy levels to be represented by (often quite large) sets of independent term values $\{T_{vJ}\}$, or by a set of 'band constants' $\{G_v, B_v, -D_v, H_v, \ldots\}$ for each vibrational level v of each isotopolog. These last capabilities can be particularly important in the early stage of a multi-state analysis, as it allows one to perform a DPF analysis to determine an initial PEF for one state at a time.

DPF fits to spectroscopic data are, of course, non-linear least-squares fits, since the experimental observables are not linear functions of the parameters defining the analytic PEFs. As in any non-linear least-squares fit, it is necessary to provide the fitting procedure with some realistic preliminary estimate of the parameters of the model: this is particularly important when working with sophisticated models, such as the analytic PEFs described below. In some cases one may start with a very rudimentary version of the model, determine its parameters from fits that consider only data spanning a very limited vibrational range, and then gradually expand the vibrational range considered and the number of free parameters in the PEF model until the whole data set is included. However, it is often much more efficient to start with some realistic set of parameters for the full model.

Such realistic preliminary parameter sets can be obtained by fitting the chosen PEF model to an approximate set of PEF values obtained from some other source, such as *ab initio* calculations, or by application of the first-order semiclassical Rydberg–Klein–Rees procedure [1–4] to results obtained from a traditional 'parameter-fit' analysis [56]. A procedure for doing this is presented in a separate companion work [5]. As Ref. [5] provides detailed descriptions of the three main analytic PEFs considered here, the presentations below focus mainly on presenting the potential forms, identifying the parameters required in the input data files, and illustrating some features of their use based on consideration of published applications.

Section 2 provides a description of the Hamiltonian upon which **dPotFit** is based, while Section 3 describes the four families of PEF forms for which **dPotFit** can perform DPFs. Section 3 also describes the manner in which sets of band constants or individual term values may be utilized to represent the vib-rotational levels of a given state. Section 4 then describes how three other analytic PEFs, or a pointwise *ab initio* or RKR potential, can be imported and tested against the range of experimental data allowed by **dPotFit**. Section 5 then describes how the radial strength functions characterizing Born–Oppenheimer breakdown (BOB), Λ–doubling, and $^2\Sigma$ splittings are represented in the code. Finally, Section 6 describes the methods used for calculating predicted values of the various experimental observables, and of the partial derivatives required by the least-squares fitting procedure, from an assumed knowledge of the potential energy and BOB functions, while Section 7 describes some strategies that have proven useful in applying **dPotFit** to the analysis of experimental data, and Section 8 describes some practical details regarding the use of the code. An Appendix then provides a brief overview of the online Supplementary Material associated with this paper.

2. The radial Hamiltonian

As in most direct-potential-fit (DPF) data analyses reported to date, the present code is based upon an effective radial Schrödinger equation derived by Watson [6,7], in which atomic-mass-dependent nonadiabatic contributions to the kinetic energy operator have been incorporated into an effective "adiabatic" contribution to the electronic potential energy function and a BOB contribution to the effective centrifugal potential of the rotating molecule. Following the conventions of Refs. [8–10], the resulting effective radial Schrödinger equation for isotopolog 'α' of molecule A–B in a singlet electronic state with electronic angular momentum projection quantum number Λ, may be written as

$$\left\{-\frac{\hbar^2}{2\mu_\alpha}\frac{d^2}{dr^2}+\left[V_{\rm ad}^{(1)}(r)+\Delta V_{\rm ad}^{(\alpha)}(r)\right]+\frac{[J(J+1)-\Lambda^2]\hbar^2}{2\mu_\alpha r^2}\left[1+g^{(\alpha)}(r)\right]\right\}\psi_{vJ}(r)=E_{vJ}\psi_{vJ}(r). \quad (1)$$

Here, $V_{\rm ad}^{(1)}(r)$ is the total electronic internuclear potential for a chosen reference isotopolog (labelled $\alpha=1$), $\Delta V_{\rm ad}^{(\alpha)}(r)$ is the difference between the effective adiabatic potential for isotopolog α and that for the reference species ($\alpha=1$), and $g^{(\alpha)}(r)$ is the centrifugal BOB potential correction function for isotopolog α with effective reduced mass μ_α. The default definition of μ_α is Watson's "charge-modified reduced mass" [6],

$$\mu_\alpha=\mu_\alpha^{\rm W}\equiv M_{\rm A}^{(\alpha)}M_{\rm B}^{(\alpha)}/\left(M_{\rm A}^{(\alpha)}+M_{\rm B}^{(\alpha)}-{\rm CHARGE}\times m_{\rm e}\right), \quad (2)$$

in which 'CHARGE' is the net $\pm$(integer) charge on the molecule, $m_{\rm e}$ is the electron mass, and $M_{\rm A}^{(\alpha)}$ and $M_{\rm B}^{(\alpha)}$ are the masses of the neutral atoms A and B forming isotopolog α of species A–B $^{\rm CHARGE}$ respectively. However, other choices are also possible (see Appendix B). Of course,

for CHARGE = 0 Eq. (2) becomes the conventional two-body reduced mass for particles of mass $M_A^{(\alpha)}$ and $M_B^{(\alpha)}$.

Each of the quantities $\Delta V_{ad}^{(\alpha)}(r)$ and $g^{(\alpha)}(r)$ in Eq. (1) is expressed as a sum of two terms, one for each atom, whose components have magnitudes inversely proportional to the masses of the specific atomic isotopes [6–8,11], and given specifically by

$$\Delta V_{ad}^{(\alpha)}(r) = \frac{\Delta M_A^{(\alpha)}}{M_A^{(\alpha)}} \tilde{S}_{ad}^{A}(r) + \frac{\Delta M_B^{(\alpha)}}{M_B^{(\alpha)}} \tilde{S}_{ad}^{B}(r), \tag{3}$$

and

$$g^{(\alpha)}(r) = \frac{M_A^{(1)}}{M_A^{(\alpha)}} \tilde{R}_{na}^{A}(r) + \frac{M_B^{(1)}}{M_B^{(\alpha)}} \tilde{R}_{na}^{B}(r). \tag{4}$$

The quantity $\Delta M_A^{(\alpha)} \equiv M_A^{(\alpha)} - M_A^{(1)}$ is the difference between the isotopic masses of atom A in isotopolog α and in the reference isotopolog ($\alpha = 1$). The expressions employed to represent the mass-independent $\tilde{S}_{ad}^{A/B}(r)$ and $\tilde{R}_{na}^{A/B}(r)$ radial functions can be found in Sections 5.1 and 5.2. Straightforward extensions of Eq. (1) that account for the *e/f* Λ-doubling splittings that occur for singlet states with $\Lambda \neq 0$ [10], or for doublet splittings of the rotational levels of $^2\Sigma$ states, are presented in Sections 5.3 and 5.4, respectively.

3. Potential function forms that can be used in the fits

Program **dPotFit** can currently perform fits to PEFs that are represented by one of four families of analytic functions. These functions are all expressed in terms of dimensionless radial variables of the form

$$y_p^{r_e}(r) = \frac{r^p - r_e^p}{r^p + r_e^p}, \tag{5}$$

$$y_q^{\rm ref}(r) = \frac{r^q - r_{\rm ref}^q}{r^q + r_{\rm ref}^q} \quad \text{and/or} \quad y_p^{\rm ref}(r) = \frac{r^p - r_{\rm ref}^p}{r^p + r_{\rm ref}^p}, \tag{6}$$

in which p and q are small positive integers ($\{p/q\} = 1, 2, 3, 4, \ldots$), r_e is the equilibrium internuclear distance of the effective adiabatic potential energy function for the chosen reference isotopolog, $V_{ad}^{(1)}(r)$, and $r_{\rm ref}$ is a reference distance chosen as the expansion center for the variables of Eq. (6) (usually $r_{\rm ref} > r_e$). Most of the early work employing this type of variable fixed $r_{\rm ref} = r_e$ and $p = q = 1$ [12–14]. However, it was later found that allowing p and q to be greater than 1, and fixing $r_{\rm ref}$ at some distance between r_e and the outer end of the data-sensitive region allowed high quality fits to be achieved with a smaller number of expansion parameters [15–21].

The fact that $y_p^{r_e}(r)$ and $y_{\{q/p\}}^{\rm ref}(r)$ approach finite limits both as $r \to 0$ and as $r \to \infty$ means that functions of these variables will also approach finite values in these limits. At the same time, the fact that $y_p^{r_e}(r) \propto (r - r_e)$ and $y_q^{\rm ref}(r) \propto (r - r_{\rm ref})$ at distances near their respective expansion centers means that they will be effective expansion variables for properties that change significantly in those regions. This mapping of the infinite radial domain $r \in [0, \infty)$ onto the finite interval $y_{\{q/p\}}(r) \in [-1, +1]$ greatly facilitates the imposition of proper theoretical constraints onto the behavior of the potential function both at long range, and in the very short-range region. Moreover, it means that functions defined as finite power series in one of these variables will not have singularities at either very small or very large values of r. Furthermore, experience has shown that for larger values of q ($q = 2, 3, 4, \ldots$) the resulting potential energy functions are increasingly strongly inhibited from having implausible spurious extrema in the extrapolation regions at small or large values of r [9,10,15,22–24].

In any case, values of p, q, and $r_{\rm ref}$ must be selected by the user and specified in the main input data file. For the Morse/Long-Range (MLR) potential of Section 3.2, theory places some restrictions on the allowed value of p, while some guidance regarding how to choose appropriate values of the other parameters may be found below and in Refs. [16,25].

3.1. The Expanded Morse Oscillator (EMO) potential function

The first type of potential function form considered here is the *Expanded Morse Oscillator* or EMO function [12], which has the form of a Morse potential [26] in which the exponent coefficient varies with distance. Other functions of this type were introduced by Coxon and Hajigeorgiou (the "GMO" potential) [27] and by Dulick and co-workers (the "MMO" potential) [28], but because of its simpler form and better extrapolation behavior, only the EMO function is considered here.

An EMO potential function has the form

$$V_{\rm EMO}(r) = \mathfrak{D}_e \left[1 - e^{-\beta(r)\cdot(r - r_e)}\right]^2, \tag{7}$$

in which $\mathfrak{D}_e$ is the well depth, r_e is the equilibrium internuclear distance, and

$$\beta(r) = \beta_{\rm EMO}(y_q^{\rm ref}(r)) = \sum_{i=0}^{N_\beta} \beta_i y_q^{\rm ref}(r)^i. \tag{8}$$

As discussed in Refs. [10,22,23], for cases in which $r_{\rm ref} = r_e$, an appropriate choice of q (usually ≥ 2) in the definition of $y_q^{\rm ref}(r)$ can often prevent extrapolation problems at large r, but may not always resolve such problems at small r. However, most such extrapolation problems can normally be resolved by setting the expansion center $r_{\rm ref}$ at some distance greater than r_e [29].

The EMO (or EMO$_p$) potential is a very flexible form that has been used successfully in a number of demanding data analyses involving both 'normal' single well potentials [12,13,24] and a state whose potential function has an additional 'ripple' or incipient shelf [23]. However, the fact that $[\mathfrak{D}_e - V_{\rm EMO}(r)]$ dies off exponentially at large r makes it a less than ideal function for representing states for which the data extend fairly close to the dissociation limit. This problem stimulated the development of the next two potential function forms discussed below.

3.2. The Morse/Long-Range (MLR) potential

At long range, all intermolecular potential functions may be described as a sum of inverse-power terms, with

the limiting long-range behavior being

$$V(r) \simeq \mathfrak{D} - C_{m_1}/r^{m_1} - C_{m_2}/r^{m_2} - \cdots, \tag{9}$$

in which the powers m_1, $m_2, \ldots$, are determined by the nature of the atoms into which the given molecular state dissociates [30,31], and the coefficients C_{m_i} may often be calculated from theory. It is therefore desirable to use a potential form that has the limiting behavior of Eq. (9), especially if the data set includes information from vibrational levels lying fairly close to dissociation. This consideration stimulated the development of the Morse/Long-Range (MLR) potential form [15–17,19,20],

$$V_{\rm MLR}(r) = \mathfrak{D}_e \left\{ 1 - \frac{u_{\rm LR}(r)}{u_{\rm LR}(r_e)} e^{-\beta(r)\cdot y_p^{r_e}(r)} \right\}^2, \tag{10}$$

in which $\mathfrak{D}_e$ is the well depth, r_e is the equilibrium internuclear distance, and the exponent coefficient $\beta(r) = \beta_{\rm MLR}(r)$ is a (fairly) slowly varying function of r. The desired long-range behavior is defined by a generalized version of the attractive contribution to Eq. (9):

$$u_{\rm LR}(r) = D_{m_1}(r)\frac{C_{m_1}}{r^{m_1}} + D_{m_2}(r)\frac{C_{m_2}}{r^{m_2}} + \cdots + D_{m_{\rm Last}}(r)\frac{C_{m_{\rm Last}}}{r^{m_{\rm Last}}}, \tag{11}$$

in which the $D_{m_i}(r)$ are "damping functions" that have the key property that all $D_{m_i}(r) \to 1$ at large r. The two families of damping functions allowed by **dPotFit** are a generalized [25] version of the damping function of Douketis et al. [32],

$$D_m^{{\rm ds}(s)}(r) = \left(1 - e^{-\frac{b^{\rm ds}(s)\cdot(\rho r)}{m} - \frac{c^{\rm ds}(s)\cdot(\rho r)^2}{\sqrt{m}}} \right)^{m+s}, \tag{12}$$

and a generalized version [25] of the Tang–Toennies damping function [33],

$$D_m^{{\rm tt}(s)}(r) = 1 - e^{-b^{\rm tt}(s)\cdot(\rho r)} \sum_{k=0}^{m-1+s} \frac{\left[b^{\rm tt}(s)\cdot(\rho r) \right]^k}{k!}. \tag{13}$$

Values of the system-independent parameters $b^{\rm ds}(s)$ and $c^{\rm ds}(s)$ in Eq. (12) and $b^{\rm tt}(s)$ in Eq. (13) were determined in Ref. [25], and are stored in the code, so that the only parameters that must be set by the user are the system-dependent range-scaling parameter ρ, and the integer s that defines the limiting short-range behavior of all of these functions, which is [25]

$$\lim_{r\to 0} \left\{ D_m^{(s)}(r)/r^m \right\} \propto r^s \tag{14}$$

for all values of m, for both families. The original versions of these functional forms corresponded to the cases $s = 0$ for the Douketis et al. form [32] and $s = 1$ for the Tang–Toennies form [33]. However, the *ab initio* results of Kreek and Meath suggest that the actual very short-range behavior of dispersion energy damping functions corresponds to $s = 0$ [34]. Moreover, as the very short-range behavior of an MLR potential is defined by the quadratic term in Eq. (10), values $s > 0$ are physically unacceptable, since an MLR potential will turn over and approach zero when $r \to 0$ unless either there are no damping functions, or those chosen correspond to a value of s that is ≤ 0. The study of Ref. [25] recommended the use of Eq. (12) with $s = -1$ (*i.e.*, setting input parameter `IVSR` $= -2$), but use of either functional form with $s \leq 0$ will yield sensible short-range behavior for an MLR PEF.

The constant ρ appearing in Eqs. (12) and (13) is a system-dependent range parameter devised by Douketis et al. [32]. For a pair of interacting atoms A and B, it is defined by the combining rule

$$\rho = \rho^{\rm AB} \equiv [2\rho^{\rm A}\rho^{\rm B}/(\rho^{\rm A} + \rho^{\rm B})], \tag{15}$$

in which $\rho^{\rm A/B}$ are defined in terms of the ratio of the ionization potential of atom A or B to that of atomic hydrogen as $\rho^{\{\rm A/B\}} = \left(I_p^{\{\rm A/B\}}/I_p^{\rm H} \right)^{2/3}$. However, other choices for this range-scaling parameter may be made by the user.

As a final comment regarding the $u_{\rm LR}(r)$ function that defines the long-range behavior of an MLR-type potential, note that it is not necessary to restrict it to having the explicit inverse-power-sum form of Eq. (9) or (11), as it can take on any form dictated by theory. For example, in recent applications to states of Li_2 dissociating to the Li $(^2P_{1/2}) + Li_2(^2S_{1/2})$ asymptote, $u_{\rm LR}(r)$ has been represented by one of the roots of a diagonalization arising from two-state [19] or three-state [21] coupling near that asymptote. As a result, an option coded in **dPotFit** to deal with alkali homodimers in states dissociating to the lowest $nS + nP$ atomic limit allows $u_{\rm LR}(r)$ to be defined by one of the roots of the 2×2 or 3×3 coupling matrix that defines the long-range behavior predicted by theory for these states [35,36]. The ease with which it can treat such special cases further illustrates capabilities of the MLR functional form.

As was discussed in Ref. [5], a defining property of the MLR function is the fact that

$$\lim_{r\to\infty} \beta_{\rm MLR}(r) \equiv \beta_\infty = \ln\{2\mathfrak{D}_e/u_{\rm LR}(r_e)\}. \tag{16}$$

Since the pre-factor to the exponential term in Eq. (10) is equal to 1 at the equilibrium distance, r_e, and $y_{p,}^{r_e}(r) \to +1$ as $r \to \infty$, this means that at large distances Eq. (10) becomes

$$\begin{aligned} V_{\rm MLR}(r) &\simeq \mathfrak{D}_e - u_{\rm LR}(r) + \frac{1}{4\mathfrak{D}_e}[u_{\rm LR}(r)]^2 \\ &\simeq \mathfrak{D}_e - \sum_{i=1}^{\rm Last} \frac{C_{m_i}}{r^{m_i}} + \frac{1}{4\mathfrak{D}_e}\left(\sum_{i=1}^{\rm Last} \frac{C_{m_i}}{r^{m_i}} \right)^2. \end{aligned} \tag{17}$$

This indicates that although the MLR function does take on the form of Eq. (9) at long range, unless $m_{\rm Last} < 2m_1$, the long-range behavior of the quadratic term in Eqs. (10) and (17) would change the long-range behavior specified by Eq. (11). This problem is most serious for cases in which $m_1 = 3$ [19], but should also be considered for systems with larger m_1 values [37]. However, it is resolved here by the fact that, as necessary, the code internally modifies the input C_{m_i} coefficients, and/or adds additional terms, so as to precisely cancel the effect of those non-physical quadratic terms, as illustrated by the discussion of the coefficients $C_6^{\rm adj} = C_6 + (C_3)^2/4\mathfrak{D}_e$ and $C_9^{\rm adj} = C_3 C_6^{\rm adj}/2\mathfrak{D}_e$ for the $m_1 = 3$ case given in Ref. [19].

In order to use the MLR potential form, it is clearly necessary to know appropriate values for the powers m_i [30,31], and to have realistic estimates of the coefficients C_{m_i}. When no realistic estimate of the leading (smallest-power) coefficient C_{m_1} is available, the MLR form has no significant advantages over the simpler EMO function, which will likely be more 'robust'. However, if the leading inverse-power coefficient C_{m_1} is known, but no calculated value for the

second coefficient C_{m_2} is available, it may be desirable to make a plausible *ad hoc* estimate of the latter, and employ the two-term MLR form rather than the simple one-term MLJ function of Refs. [38,39], because more reasonable long-range extrapolation behavior would be imposed [15]. In cases for which vibrational data extend *very* close to the dissociation limit, it may also be possible to treat one or more of the long-range potential coefficients as free parameters to be optimized in the fit [16,19,21].

The algebraic form of Eq. (10) means that at sufficiently long range $V_{\rm MLR}(r)$ always takes on the form of Eq. (17). To achieve this, the exponent coefficient function $\beta(r)$ is required both to approach asymptotically the value β_∞ defined by Eq. (16), and to be sufficiently flexible to describe accurately the shape of the potential function well. Its functional form should also prevent or discourage the potential function from having unphysical extrema in the two extrapolation intervals, namely, for $r\to 0$, and between the data region and the asymptotic limit. Two approaches to this problem are offered in **dPotFit** . In the first, $\beta(r)$ is written as a constrained polynomial in the variable $y_q^{\rm ref}(r)$, in which the separate variable $y_p^{\rm ref}(r)$ acts as a switching function [9,10,15,19,22,40], that is

$$\beta(r) \equiv \beta_{\rm PE-MLR}\left(y_{p,q}^{\rm ref}(r)\right) = \beta_\infty y_p^{\rm ref}(r) + \left[1 - y_p^{\rm ref}(r)\right] \sum_{i=0}^{N_\beta} \beta_i y_q^{\rm ref}(r)^i. \tag{18}$$

Note that the power p appearing here is the same as that used to define the distance parameter $y_p^{r_e}(r)$ in the exponential term in Eq. (10). While most of the early work with this model was performed with $q=p$, it has since been shown that use of a separate power $q<p$ in the power-series portion of Eq. (18) can lead to more compact and robust potential functions [19–21,29].

One restriction associated with this 'polynomial exponent' MLR (PE-MLR) form is a limitation on the allowed value of p, which depends on the particular set of powers m_i that define the terms contributing $u_{\rm LR}(r)$. The algebraic form of the exponent coefficient function of Eq. (18) implies that at large r the exponential term in the MLR function takes the form $e^{-\beta_\infty}\left(1+A/r^p+\cdots\right)$. This has the effect of adding a term having the form $\left(AC_{m_1}\right)/r^{m_1+p}$ to Eq. (17) [16,19]. As a consequence, the leading contributions to long-range behavior of $V_{\rm MLR}(r)$ will only truly be defined by the specified version of Eq. (11) if the power p defining the exponent variables satisfies the constraint $p>(m_{\rm Last}-m_1)$, where $m_{\rm Last}$ is the power of the last (*i.e.*, highest-power) term contributing to $u_{\rm LR}(r)$ [16,19,20]. It also may be desirable to set $p=m_{\rm next}-m_1$, in which $m_{\rm next}$ is the (inverse) power associated with the first long-range term predicted by theory that is *not* included in the chosen definition of $u_{\rm LR}(r)$.

As discussed in Ref. [19], there are no formal restrictions on the choice of the power q defining the radial variable in the power series part of Eq. (18). Experience suggests [19,21,41] that when q is very small (say, 1 or 2), the potential is more likely to be unstable in the extrapolation region(s). However, the optimum choices for q (and p, subject to the constraint $p>(m_{\rm Last}-m_1)$) must also be guided by consideration of the manner in which the potential function approaches its limiting long-range behavior. In particular, $V_{\rm MLR}(r)$ always takes on the limiting behavior of Eq. (9), a rearrangement of which yields

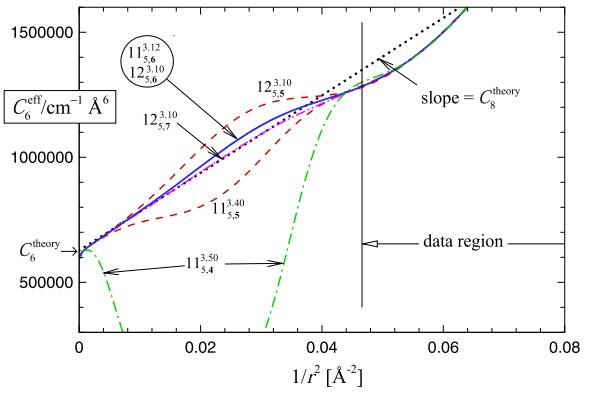

Fig. 1. Tests of the long-range extrapolation behavior of various fitted potentials for the $X^1\Sigma_g^+$ state of Br_2 associated with different MLR models $N_{p,q}^{r_{\rm ref}}$, where $N=N_\beta$ [adapted from Fig. 7 of Ref. [41]].

$$C_{m_1}^{\rm eff}(r) \equiv r^{m_1}[\mathfrak{D}_e - V_{\rm MLR}(r)] \simeq C_{m_1} + \frac{C_{m_2}}{r^{m_2-m_1}} + \cdots \tag{19}$$

Thus, a plot of $C_{m_1}^{\rm eff}(r)$ *vs.* $1/r^{m_2-m_1}$ must approach the intercept C_{m_1} with slope C_{m_2}, and it should bridge the extrapolation interval between the "data region" and the limiting behavior in a smooth, monotone fashion. Fig. 1 shows plots of this type for a number of otherwise equivalent potentials for the ground-state of Br_2 [41], for which the leading terms in the long-range potential correspond to m=6, 8, and 10 [42]. While all of the six potentials considered here do eventually achieve the predicted linear approach to an intercept of $C_6^{\rm theory}$ with a slope of $C_8^{\rm theory}$, it is clear that the plots for model potentials with $q=4$ or 5 have physically implausible extrema in the extrapolation interval between the 'data-sensitive region' and that intercept. Note, however, that these implausible extrema in $C_6^{\rm eff}(r)$ plots are *not* accompanied by discernably irregular behavior in plots of the potential functions $V_{\rm MLR}(r)$ themselves. This illustrates the practical importance of using plots such as Fig. 1 to detect such implausible extrapolation behavior! Moreover, the fact that the $C_6^{\rm eff}(r)$ plot for the $12_{5,7}^{3.10}$ potential *fails* to display the positive curvature away from this limiting slope implied by the existence of an attractive C_{10}/r^{10} term probably reflects the excessive 'stiffness' of the q=7 radial variable. Thus, it appears that the potentials for which q=6 have the optimum extrapolation behavior for this particular species.

When performing fits to a PE-MLR function it is necessary to consider ranges of values of $r_{\rm ref}$, N_β and q in a search for an optimal model. Fig. 2 illustrates how the quality of fits to data for ground-state Ca_2 depends upon these three parameters. As expected, for either q=3 or q=4, the quality of fit improves, and the breadth of the region over which the dimensionless root-mean-square deviation $\overline{dd}$ has a minimum increases with N_β. However, for larger q values the expansion variable is "stiffer", and larger N_β values tend to be required to attain the same quality of fit. This is the reason that the range of $r_{\rm ref}$ over

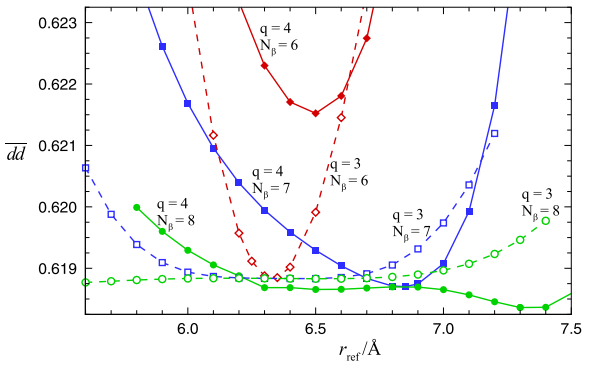

Fig. 2. Dependence of the dimensionless root-mean-square deviation $\overline{dd}$ obtained from fits to 3553 fluorescence series data for the $X^1\Sigma_g^+$ state of Ca_2 [100,101] on the parameters r_{ref}, q, and N_β. Displayed here are the results of fits to MLR potentials with p=5 whose long-range tails are defined by Douketis-type damping with $s=-1$ and fixed theoretical values of the C_8, and C_{10} [102] coefficients, but with C_6 allowed to be a fitted parameter [20].

which values of $\overline{dd}$ lie near their minimum is narrower for $\{q=4, N_\beta=7\}$ (solid square blue points) than for $\{q=3, N_\beta=7\}$ (open square blue points), and the $\overline{dd}$ minimum for $\{q=4, N_\beta=6\}$ lies well above those for the other cases. Moreover, for most of the $N_\beta=8$ cases associated with either value of q, between one and three of the fitted β_i values have uncertainties greater that 100%. Thus, the results shown in this figure lead to a recommendation of $\{q=4, N_\beta=7, r_{ref}=6.85$ Å$\}$ as the optimum MLR model for this Ca_2 system.

A second way of representing the exponent coefficient function in Eq. (10) is described now; more details regarding its properties may be found in Ref. [5]. All existing applications of the MLR potential have employed the constrained polynomial expansion of Eq. (18) to represent the exponent coefficient function $\beta(r)$. However, while preliminary work showed some promise [43,44], it has not yet been demonstrated that that form can provide a practical, compact, and accurate representation of double-minimum or shelf-state potential energy functions. A successful model for representing and fitting to such potential functions was introduced by Pashov and coworkers [45–47], in which the potential function defined as a cubic spline through a set of points whose ordinate values serve as the parameters of the model. However, a relatively large number of points/parameters (typically $\gtrsim$50) seem to be required to define a potential function accurately in this way, and such functions can only be extrapolated sensibly outside the data region if an analytic repulsive wall and the theoretically predicted inverse-power long-range tail are attached in some *ad hoc* manner, a procedure that always introduces discontinuities in the higher-order PEF derivative at the chosen attachment points.

An alternative approach now being developed [5,48] is the 'Spline-Exponent-MLR' (SE-MLR) function which uses Pashov's 'spline-pointwise' approach to define the exponent coefficient $\beta(r)$ in the MLR potential function form of Eq. (10). In particular, $\beta(r)$ is defined as a 'natural' cubic spline function passing through $\beta(r_j)=\beta(y_q^{ref}(r_j))$ values at a specified set of $y_q^{ref}(r_j)$ values, and those $\beta(r_j)$ values become the parameters defining the shape of the potential. Following this approach, the exponent coefficient function is written as

$$\beta(r)=\sum_{k=1}^{N_\beta} S_k(y_q^{ref}(r))\beta_k, \tag{20}$$

in which the spline 'basis functions' $S_k(y_p^{ref}(r))$ are completely defined by the chosen mesh of values of $y_q^{ref}(r_j)$. A straightforward application of the chain rule of calculus then yields the partial derivatives required for the least-squares fit procedure:

$$\frac{\partial V(r)}{\partial\beta_k}=2\mathfrak{D}_e\left\{1-\frac{u_{LR}(r)}{u_{LR}(r_e)}e^{-\beta(r)\cdot y_p^{re}(r)}\right\}\left(\frac{u_{LR}(r)}{u_{LR}(r_e)}e^{-\beta(r)\cdot y_p^{eq}(r)}\right)y_p^{r_e}(r)S_k\left(y_q^{ref}(r)\right). \tag{21}$$

Although the least-squares problem is non-linear, the fact that the $S_k(y_p^{ref}(r))$ functions are independent of the parameter values $\{\beta_k\}$ yields some computational simplifications.

In using the SE-MLR form, it is important to realize that the radial variable $y_q^{ref}(r)$ used to define $\beta(r)$ should normally be defined by a value of r_{ref} that is greater than r_e, and the value of q defining the radial variable in Eq. (20) should be greater than 1, in order to assure that the chosen mesh of points sample the full range of $\beta(y_q^{ref}(r))$ values appropriately. However, in this case *both* q and p (rather than just p) must be set at values greater than $(m_{Last}-m_1)$ in order to prevent the long-range behavior of the exponential term from disrupting the long-range behavior specified by Eq. (11).

3.3. The double-exponential/long-range (DELR) model potential

The need for a flexible analytic potential function with a barrier that protrudes above the potential asymptote for distances $r>r_e$ stimulated the development of the 'double-exponential/long-range' (DELR) potential function form [10,22],

$$V_{DELR}(r)=\left\{Ae^{-2\beta(r)\cdot(r-r_e)}-Be^{-\beta(r)\cdot(r-r_e)}+\mathfrak{D}_e\right\}-u_{LR}(r), \tag{22}$$

in which the exponent coefficient $\beta(r)$ is defined as the same type of simple power series in $y_p^{ref}(r)$ appearing in the EMO potential (see Eq. (8)) [29]. In this case, the presence of a repulsive leading term in the additive long-range function $u_{LR}(r)$ (which could, in principle, be attractive or repulsive), served to introduce the potential function barrier that was being modelled [10]. The pre-exponential coefficients A and B in Eq. (22) are defined in terms of the well depth $\mathfrak{D}_e$ (relative to the potential asymptote) and the equilibrium distance r_e by the expressions

$$A=\mathfrak{D}_e-u_{LR}(r_e)-u'_{LR}(r_e)/\beta_0, \tag{23}$$

$$B=2\mathfrak{D}_e-2u_{LR}(r_e)-u'_{LR}(r_e)/\beta_0, \tag{24}$$

in which $u'_{LR}(r_e)\equiv[du_{LR}(r)/dr]_{r=r_e}$. If $u_{LR}(r)=0$ the DELR potential becomes the EMO function of Eq. (7). However, other choices of $u_{LR}(r)$ allow it to represent the outer wall

of a potential function with a barrier [10], the multi-term attractive inverse-power long-range tail of a normal single-well potential energy function, or even (in principle) the outer wall of a double-minimum or shelf-state potential function.

In the present version of **dPotFit**, the long-range function $u_{\rm LR}(r)$ in Eq. (22) is assumed to be defined either by a sum of damped or undamped inverse-power terms, as in Eq. (11) or, for alkali homodimers, by one of the roots of the 2×2 or 3×3 long-range coupling matrices described in the previous subsection. Note, however, that the present sign convention for this $u_{\rm LR}(r)$ function (positive coefficients being *attractive*) is opposite to that used in Ref. [10]. As in Section 3.2, damping functions $D_m(r)$ may be defined by either Eq. (12) or (13), but since $u_{\rm LR}$ is an additive, rather than a multiplicative contribution to the potential, the parameter s defining the very short-range behavior of these damping functions *is* allowed to have positive values, while negative values will cause the short-range potential wall to turn over. Of course, other damping function expressions [49,50] or entirely different types of expressions for $V_{\rm LR}(r)$ could equally well be used in the DELR type of potential form.

3.4. The Generalized Potential Energy Function (GPEF) of Šurkus

The fourth family of potential functions is a generalization of the familiar Dunham polynomial potential [51] that Šurkus et al. [52] introduced and called the 'Generalized Potential Energy Function' (GPEF). A modified, but exactly equivalent expression for his expansion variable (devised by Seto [14]) leads to the GPEF potential form

$$V_{\rm GPEF}(r) = c_0 z_q^2\left[1+\sum_{i=1}^{N_\beta} c_i z_q^i\right], \quad \text{with } z_q \equiv \frac{(r^q - r_e^q)}{(a_S r^q + b_S r_e^q)}. \tag{25}$$

For appropriate choices of the (fixed) parameters a_S and b_S, this expansion takes on a number of familiar forms:

- setting $q=1$, $a_S=0$ and $b_S=1$ yields Dunham expansions [51];
- setting $q=1$, $a_S=1$ and $b_S=0$ yields Simons–Parr–Finlan (SPF) expansions [53];
- setting $q=1$, $a_S=b_S=0.5$ yields the Ogilvie–Tipping (OT) potential expansion [54];
- setting $a_S=b_S=1$ yields the expansion variable $y_p^{\rm eq}(r)$ of Eq. (5).

For $a_S\neq 0$, PEFs with the GPEF form always asymptotically approach a finite asymptote with a $1/r^q$ functional behavior. Thus, if appropriate constraints are applied to the coefficients, the GPEF form can, in principle, be required to have the theoretically predicted limiting long-range behavior of Eq. (9) [52,55]. However, apart from the relatively simple case in which q is set equal to the power of the leading long-range term in Eq. (9) [52], such constraints have proven to be too unwieldy for practical use.

3.5. Term-value and band-constant representations

One often encounters cases in which there are too few data to allow an analytic potential function to be determined for a state that still must be taken into account, because it is at one end (usually as the upper state) of a set of transitions to a state for which a potential function is being determined, and those transitions contain information about the state of interest. In some cases the levels of that upper state may be accounted for as the origins of fluorescence series, but it is often more convenient to treat all of the assigned levels of that state as independent term values $T_{v,J,p}$ in the fit. Alternatively, it may be convenient to represent all of those observed (upper-state) levels with sets of band constants:

$$\begin{aligned} E(v,J) &= G_v + B_v[J(J+1)] - D_v[J(J+1)]^2 + H_v[J(J+1)]^3 + \cdots \\ &= \sum_{m=0} K_m[J(J+1)]^m. \end{aligned} \tag{26}$$

For states for which $\Lambda>0$ and the associated Λ-doubling is observed, the '$q_m^{(\alpha)}(v)$'-band constants characterizing the splitting of rotational sublevels J of vibrational levels v may also be determined in the fit by inclusion of the additional term

$$\delta E_\Lambda^{(\alpha)}(v,J;\{e/f\}) = {\rm sg}(\{e/f\})[J(J+1)]^\Lambda \sum_{m=\Lambda} q_m^{(\alpha)}(v)[J(J+1)-\Lambda^2]^{m-\Lambda}, \tag{27}$$

in which the sg($\{e/f\}$) are standard Λ-doubling splitting parameters, with values of $+1$ and 0, 0 and -1, or $\pm 1/2$, for e-parity and f-parity levels, respectively (see Section 5.3): the choice of which pair of splitting factors to use for the term sg($\{e/f\}$) must be specified by the user [56].

Similarly, use of a band-constant representation for a $^2\Sigma$ state will invoke the analogous representations for the splitting of the e and f components of its rotational N–levels:

$$\begin{aligned} \delta E_\Lambda^{(\alpha)}(v,N;\{e\}) &= +\frac{1}{2}N\sum_{m=1}\gamma_m^{(\alpha)}(v)[N(N+1)]^{m-1}, \\ \delta E_\Lambda^{(\alpha)}(v,N;\{f\}) &= -\frac{1}{2}(N+1)\sum_{m=1}\gamma_m^{(\alpha)}(v)[N(N+1)]^{m-1}. \end{aligned} \tag{28}$$

All of these cases are allowed by **dPotFit**. While their use will tend to increase greatly the number of independent parameters being determined (especially if term values are used), often by hundreds or even thousands(!), this usually presents little difficulty, since the fit will be linear with respect to these parameters. Use of one of these representations for all-states-but one can also be a convenient way of reducing or removing interparameter correlations involving those states in order to facilitate the determination of a good preliminary potential function model for a given state in the early stages of an analysis.

4. Other potential function forms that can be used in dPotFit

The next four subsections describe potential function forms that have been introduced into the **dPotFit** code in order to allow users to make comparisons with the

predictions of other potential functions that have been proposed for a state of interest. However, use of **dPotFit** to perform fits using these additional potential function forms has not (yet) been implemented

4.1. Hannover polynomial potentials (HPP)

Perhaps the most widely used alternate potential function form is the "Hannover polynomial potential" (HPP) or "X-representation" function, which has the form [57,58]:

$$V_{\rm HPP}(r) = A_{\rm I} e^{-B_{\rm I}(r-R_{\rm I})} \quad \text{for } r < R_{\rm I} \tag{29}$$

$$= \{\texttt{VLIM} - \mathfrak{D}_e\} + V_X(r) \quad \text{for } R_{\rm I} \le r \le R_{\rm O}$$

$$V_{\rm HPP}(r) = \left\{\texttt{VLIM} - \mathfrak{D}_e + V_X(R_{\rm O})\right\} - \sum_{m_1}^{m_{\rm last}} \frac{C_m}{r^m} + \frac{C_{m_{\rm last}+2}}{r^{m_{\rm last}+2}}$$

$$\text{for } r > R_{\rm O} \tag{30}$$

in which

$$V_X(r) = \sum_{i=2}^{N_\beta} \beta_i X(r)^i \quad \text{and} \quad X = \frac{r - R_m}{r + bR_m}. \tag{31}$$

Here, R_m is an "arbitrarily chosen" expansion center that lies close to the position of the potential minimum, and the parameter b has been manually adjusted to optimize the potential function slope at short distances, and the expansion coefficients β_i have been determined from a fit to experimental data. The parameters $A_{\rm I}$ and $B_{\rm I}$ defining the exponential function used to extrapolate to short distances are defined by the requirement that there be a smooth connection to the polynomial function at $r = R_{\rm I}$. Similarly, with the other long-range coefficients C_m defined by theory, the value of the coefficient $C_{m_{\rm last}+2}$ in Eq. (30) is defined by requiring the potential to be continuous at the outer switching point $r = R_{\rm O}$. Note that while the HPP potential form has been implemented in **dPotFit** as described above, some of its published applications use other definitions for the short-range behavior of Eq. (29) and/or the long-range tail of Eq. (30): however, those differences have little effect on the data-sensitive region.

4.2. 'Tang–Toennies'-type (TT) exponential/Van der Waals potentials

Generalized Tang–Toennies-type potentials consist of a repulsive exponential term, whose exponent and pre-factor may consist of multiple terms, minus a sum of attractive inverse-power terms that are damped by the $s = +1$ version of the generalized TT-type damping function of Eq. (13). This structure has been used in a number of published potentials [59–61], and the extended version of this form available in **dPotFit** incorporates most variants found in the recent literature. This generalized form is

$$V_{\rm TT}(r) \equiv \left(\beta_5 + \beta_6 r + \frac{\beta_7}{r} + \beta_8 r^2 + \beta_9 r^3\right) \times e^{-\{\beta_1 r + \beta_2 r^2 + \beta_3/r + \beta_4/r^2\}} - u_{\rm LR}^{\rm TT}(r), \tag{32}$$

in which the attractive outer-wall function $u_{\rm LR}^{\rm TT}(r)$ has the form of Eq. (11) with the damping functions defined by $s = +1$ version of Eq. (13) with its shape parameter fixed at $b^{\rm tt}(s) = 1$. However, while the 'basic' Tang–Toennies function [33,62] constrained the damping-function range-scaling parameter ρ to be identical to β_1, the leading coefficient in the exponent of the exponential term in of Eq. (32), in the present implementation they are independent parameters. Moreover,while the original Tang–Toennies model only allowed for even powers m_i in Eq. (11), with $m_i \ge 6$, **dPotFit** will accept any user-supplied set of powers $\{m_i\}$.

The input parameters for this model are then the set of powers $\{m_i\}$ and coefficients $\{C_{m_i}\}$ for the attractive outer branch of the potential, and the values of the parameters ρ of Eq. (13) and $\beta_1 - \beta_9$ of Eq. (32), while the well depth $\mathfrak{D}_e$ and equilibrium distance r_e are *properties*, rather than defining parameters of the resulting potential.

4.3. The Scoles–Aziz 'Hartree–Fock dispersion' (HFD) potential form

The period from the mid 1970s until the turn of the century saw a large body of published work, mostly concerned with rare gas dimers, that used what came to be known as "Hartree–Fock Dispersion" or HFD-type potentials. Like the TT form, it consists of a repulsive exponential term minus a sum of damped dispersion terms, and it too allowed for a variety of parameterizations for the repulsive term, *viz.*,

$$V_{\rm HFD}(r) \equiv A_{\rm HFD} x^\gamma e^{-\beta_1 r - \beta_2 r^2} - u_{\rm LR}^{\rm HFD}(R). \tag{33}$$

In particular, in some cases the exponent included both a linear and a quadratic term, and in other cases the exponential term had a pre-factor of a power of $x \equiv \left(\frac{r}{r_e}\right)$: all of these variations are allowed in **dPotFit**.

In the earlier work, with what were then called the HFD-A, HFD-B and HFD-C forms, the attractive 'dispersion' term was written using a single common damping function, to give

$$u_{\rm LR}^{\rm HFD}(r) = f_1^{\rm HFD}(r) \sum_{m=m_1}^{m_{\rm Last}} \frac{C_m}{r^m}, \tag{34}$$

with

$$f_1^{\rm HFD}(r) = \exp\left\{-\alpha_1 \left(\frac{\alpha_2 r_e}{r}\right)^{\alpha_3}\right\} \quad \text{for } r < \alpha_2,$$

$$= 1 \quad \text{for } r \ge \alpha_2. \tag{35}$$

Since all of the 'A', 'B' and 'C' variations are incorporated in Eqs. (33)–(35), we call this the "HFD-ABC" function.

Some later work with this form involved use of both a common overall damping function and individual damping functions for each inverse-power term [63], *viz.*,

$$u_{\rm LR}^{\rm HFD}(r) = f_2^{\rm HFD}(\rho r) \sum_{m=m_1}^{m_{\rm Last}} D_m(r) \frac{C_m}{r^m}, \tag{36}$$

in which

$$f_2^{\rm HFD}(\rho r) = \left\{1 - \left(\frac{\rho r}{\mathrm{b_r}}\right)^{1.68} e^{-0.78(\rho r)/\mathrm{b_r}}\right\}, \tag{37}$$

where '$\mathrm{b_r}$' is the Bohr radius, distances have units Å, and the $D_m(\rho r)$ functions are the original ($s = 0$) versions of the

generalized Douketis damping functions of Eq. (12), while the system-dependent range-scaling parameter ρ is defined as described in Section 3.2. Both this "HFD-D" form of $u_{\rm LR}(r)$ and the "HFD-ABC" form of Eq. (34) are implemented in **dPotFit**.

For all HFD cases, the present code defines the values of the pre-exponential constant $A_{\rm HFD}$ and the linear exponent coefficient β_1 by applying the constraint that the potential function has a minimum of depth $\mathfrak{D}_e$ at the equilibrium distance, r_e. The input parameters defining such potentials are therefore: $\mathfrak{D}_e$, r_e, γ and β_2, together with the $\{C_m\}$ dispersion coefficients, the range-scaling parameter ρ and, for HFD-ABC functions, the parameters $\{\alpha_i\}$ required to define the damping function of Eq. (35).

4.4. Pointwise potential

The final type of potential function that may be used by **dPotFit** to define the vibration–rotation levels of a given electronic state is one defined by a fixed set of read-in turning points. The dense grid of potential function values required for solving the radial Schrödinger equation for that state is then generated by interpolating over and extrapolating beyond the read-in points using user-specified procedures (see Appendix B). Potentials of this type are fixed, having no free parameters that can be varied in a fit to experimental data. Inclusion of this type of function allows fits to make use of and/or test previously reported pointwise potentials for one (or more) of the states of interest.

5. Born–Oppenheimer breakdown (BOB) radial strength functions

5.1. BOB functions defined with a 'selected isotopolog' reference species

Following the approach of Ref. [9], the radial strength functions characterizing the atom-dependent potential-energy and centrifugal BOB corrections of Eqs. (3) and (4) are expanded in the same form utilized for the exponent coefficient-function of the PE-MLR potential:

$$\tilde{S}^{\rm A}_{\rm ad}(r) = y^{r_e}_{p_{\rm ad}}(r)u^{\rm A}_\infty + \left[1 - y^{r_e}_{p_{\rm ad}}(r)\right]\sum_{i=0}^{N^{\rm A}_{\rm ad}} u^{\rm A}_i y^{r_e}_{q_{\rm ad}}(r)^i, \tag{38}$$

$$\tilde{R}^{\rm A}_{\rm na}(r) = y^{r_e}_{p_{\rm na}}(r)t^{\rm A}_\infty + \left[1 - y^{r_e}_{p_{\rm na}}(r)\right]\sum_{i=0}^{N^{\rm A}_{\rm na}} t^{\rm A}_i y^{r_e}_{q_{\rm na}}(r)^i. \tag{39}$$

The structure of Eqs. (38) and (39) allows the asymptotic behaviors and equilibrium properties of these functions either to be specified by the user, or to be determined in a fit. In particular, the limiting asymptotic value of $\tilde{S}^{\rm A}_{\rm ad}(r)$ is $u^{\rm A}_\infty$. Hence, if the zero of energy is defined as the energy of ground-state atoms separated at $r\to\infty$ (the convention recommended here), then $u^{\rm A}_\infty \equiv 0$ for all electronic states that dissociate to yield ground-state atoms, while for a state that dissociates to yield atom A in an excited electronic state, $u^{\rm A}_\infty$ is determined by the associated atomic isotope shift [9,19]. For all states that dissociate into ground-state atoms, $u^{\rm A}_0$ and $u^{\rm B}_0$ define the difference between the well depths of those states for the various atom-A and atom-B isotopologs, such that

$$\delta\mathfrak{D}^{(\alpha)}_e(X) = \frac{\Delta M^{(\alpha)}_{\rm A}}{M^{(\alpha)}_{\rm A}}u^{\rm A}_0 + \frac{\Delta M^{(\alpha)}_{\rm B}}{M^{(\alpha)}_{\rm B}}u^{\rm B}_0. \tag{40}$$

Alternatively, if one wishes to define the potential minimum of the ground state as the absolute zero of energy, the values of $u^{\rm A}_0(X)$ and $u^{\rm B}_0(X)$ would be fixed at zero, and the values of $u^{\rm A}_\infty$ and $u^{\rm B}_\infty$ would then define the isotopolog dependence of the well depth *via* an expression analogous to Eq. (40). A user of **dPotFit** may select this (not recommended) alternate convention by choosing, in the input data file, to fix $u^{\rm A}_0(X) = u^{\rm B}_0(X) = 0$ while allowing $u^{\rm A}_\infty$ and $u^{\rm B}_\infty$ to be varied freely. In either convention, $u^{\rm A}_0$ and $u^{\rm B}_0$ would determine the electronic isotope shift for excited states.

The algebraic form of Eq. (38) and its expansion variables means that the limiting long-range behavior of these functions, and hence also of the adiabatic correction difference potential of Eqs. (1) and (3), will be a term that is proportional to $1/r^{p_{\rm ad}}$ with a coefficient defined by the vagaries of the $u^{\rm A/B}_i$ values obtained from the fit. Thus, for cases in which the reference-isotopolog potential has an inverse-power tail with the form of Eq. (11), one must set $p_{\rm ad} > m_{\rm Last}$ in order to ensure that the limiting asymptotic behavior of the effective adiabatic potentials $V^{\{\alpha\}}_{\rm ad}(r)$ is the same for all isotopologs [64].

For the few special cases in which differences between the long-range tail coefficients C_m for different isotopologs are known, that knowledge may be incorporated by replacing the constant $u^{\rm A/B}_\infty$ in Eq. (38) by the expansion

$$u_\infty - \frac{\delta C_{m_1}}{r^{m_1}} - \frac{\delta C_{m_2}}{r^{m_2}} - \cdots - \frac{\delta C_{m_{\rm Last}}}{r^{m_{\rm Last}}}. \tag{41}$$

For cases in which the isotope-dependence of (some of) the C_{m_i} is not known, these terms either may be ignored, or the input values of the δC_{m_i} coefficients may be fixed at zero. However, in the special cases in which such isotope-dependence can actually be calculated, one should define the fixed input value of this coefficient for atom A or B as

$$\delta C^{\rm A/B}_{m_i} \equiv \left(\frac{M_{\rm A/B^{(2)}}}{M^{(2)}_{\rm A/B} - M^{(1)}_{\rm A/B}}\right)\left(C^{(2)}_{m_i} - C^{(1)}_{m_i}\right), \tag{42}$$

in which $\alpha = 1$ for the reference isotopolog and $\alpha = 2$ for the selected second isotopolog for which the isotopically different $C^{(\alpha)}_{m_i}$ coefficient has been determined.

By analogy with the case for Eq. (38), the algebraic form of Eq. (39) and its expansion variables means that its limiting long-range behavior will be to approach the limiting value of $t^{\rm A/B}_\infty$ with a term proportional to $1/r^{p_{\rm na}}$ whose coefficient is defined as the difference between $t^{\rm A/B}_\infty$ and the sum of the $\{t^{\rm A/B}_i\}$ coefficients. However, the absence of any general theoretical knowledge about the long-range behavior of these functions leads to the recommendation that one should simply fix $p_{\rm na} = q_{\rm na}$.

There are no physical constraints on the values of the integers $q_{\rm ad}$ and $q_{\rm na}$ that define the expansion variables in Eqs. (38) and (39), so they must be selected by the user by trial-and-error ($q = 3, 4, \ldots$ are reasonable trial values)

subject to the twin objectives that the resulting functions provide a compact and accurate representation of the data, and that they approach their asymptotic values without having spurious extrema in the intervals outside the data-sensitive region (see Fig. 3 of Ref. [9]). It may be convenient to set $q_{\rm na}=p_{\rm na}$, and sometimes also to set $q_{\rm ad}=p_{\rm ad}$, in order to obtain expressions for $\tilde{R}_{\rm ad}^{\rm A/B}(r)$ and $\tilde{S}_{\rm na}^{\rm A/B}(r)$ that involve only a single type of radial variable [15]. However, too-large values of $q_{\rm ad}$ or $q_{\rm na}$ tend to require the expansions to have an excessive number of terms, while too-small values may lead to functions with implausible extrema in the extrapolation regions [9]: thus, case-by-case experimentation is required. Note that since EMO potential energy functions do not have *any* inverse-power limiting long-range behavior, **dPotFit** internally sets $p_{\rm ad}=q_{\rm ad}$ in Eq. (38) when an EMO function is used for the potential energy.

As was discussed in Ref. [9], the limiting asymptotic value of the *centrifugal* BOB correction function $\tilde{R}_{\rm na}^{\rm A}(r)$ should always be $t_\infty^{\rm A}=0$, unless the species in question is a molecular ion that yields A^+ or A^- upon dissociation, in which case $q_\infty^{\rm A}$ would have a small non-zero value [9]. For example, for a molecular ion $AB^{\rm CHARGE}$ that dissociates to yield a neutral atom B plus the atomic ion $A^{\rm CHARGE}$, this limit is

$$t_\infty^{\rm A}=\frac{\mu_{\rm W}}{\mu(A^{\rm CHARGE},B)}-1\approx\frac{Qm_e}{M_A}+\left(\frac{Qm_e}{M_A}\right)^2+\left(\frac{Qm_e}{M_A}\right)^3+\cdots, \tag{43}$$

in which $\mu^{\rm W}$ is the charge-modified reduced mass of Eq. (2), and $\mu(A^{\rm CHARGE},B)$ is the usual two-body reduced mass of the ion $A^{\rm CHARGE}$ with the neutral atom B. In practice however, the magnitude of this calculated value of $t_\infty^{\rm A}$ is often negligible relative to the variation of the function $\tilde{R}_{\rm na}^{\rm A}(r)$ across the data-sensitive region, so no significant loss of accuracy is introduced by simply fixing $t_\infty^{\rm A}=0$, as one does for neutral species [37].

At the other limit of $r=r_e$, a convention commonly associated with use of the Watson radial Hamiltonian of Eq. (1) is to fix the leading power-series coefficient of Eq. (39) at $t_0^{\rm A}=0$, as theory shows that this parameter is 100% correlated with the leading adiabatic-correction shape parameter $u_1^{\rm A}$, and that it represents an indeterminate integration constant in the theory [6,7,9]. However, the value of $g^{(\alpha)}(r\approx r_e)$ is related to observable magnetic properties of the molecule [37,65–68], so that when measurements of those properties are available, it may be appropriate to fix $t_0^{\rm A}$ at some specific non-zero value, or even to allow it to be varied in the fit [37]. **dPotFit** allows a user to select any of these options.

5.2. BOB functions defined with a 'clamped nuclei' reference species

Although the BOB parameterization of Eqs. (1), (3) and (4) is preferred for a number of reasons [8], the formally equivalent alternate parameterization of Watson's original paper [6] has also sometimes been employed, so **dPotFit** allows the user to employ either formulation. In the Watson approach, the effective adiabatic potential $V_{\rm ad}^{(1)}(r)$ for reference isotopolog $\alpha=1$ is replaced by the "clamped nuclei" potential $V_{\rm CN}(r)$ to give the radial equation:

$$\left\{-\frac{\hbar^2}{2\mu_\alpha}\frac{d^2}{dr^2}+\left[V_{\rm CN}(r)+\Delta V_{\rm ad,W}^{(\alpha)}(r)\right]+\frac{[J(J+1)-\Lambda^2]\hbar^2}{2\mu_\alpha r^2}\left[1+g_{\rm W}^{(\alpha)}(r)\right]\right\}\psi_{vJ}(r)=E_{vJ}\psi_{vJ}(r), \tag{44}$$

while the mass-independent radial functions in the potential energy and centrifugal BOB terms are scaled by the factors $m_e/M_{\rm A}^{(\alpha)}$ and $m_e/M_{\rm B}^{(\alpha)}$, in which m_e is the electron mass,

$$\Delta V_{\rm ad,W}^{(\alpha)}(r)=\frac{m_e}{M_{\rm A}^{(\alpha)}}\tilde{S}_{\rm ad,W}^{\rm A}(r)+\frac{m_e}{M_{\rm B}^{(\alpha)}}\tilde{S}_{\rm ad,W}^{\rm B}(r), \tag{45}$$

$$g_{\rm W}^{(\alpha)}(r)=\frac{m_e}{M_{\rm A}^{(\alpha)}}\tilde{R}_{\rm na,W}^{\rm A}(r)+\frac{m_e}{M_{\rm B}^{(\alpha)}}\tilde{R}_{\rm na,W}^{\rm B}(r). \tag{46}$$

Program **dPotFit** requires the user to select either this parameterization or that of Eqs. (3) and (4) by specifying an appropriate value of the parameter BOBCN (see READ #9 in the input data description of Appendix C). In either case, the radial strength functions $\tilde{S}_{\rm ad}^{\rm A/B}(r)$ and $\tilde{R}_{\rm na}^{\rm A/B}(r)$ are expanded as in Eqs. (38) and (39), but the magnitudes of the expansion parameters will differ by the factors $\left\{\Delta M_{\rm A/B}^{(\alpha)}/m_e\right\}$ for $S_{\rm ad,W}^{\rm A/B}(r)$ and $\left\{M_{\rm A/B}^{(1)}/m_e\right\}$ for $R_{\rm ad,W}^{\rm A/B}(r)$.

5.3. Λ-Doubling splittings for singlet states

It was shown in Ref. [10] that for singlet states with $\Lambda>0$, the effect of Λ-doubling splittings may be taken into account by inclusion of an additional term in the effective radial Schrödinger equation, to yield

$$\left\{-\frac{\hbar^2}{2\mu_\alpha}\frac{d^2}{dr^2}+\left[V_{\rm ad}^{(1)}(r)+\Delta V_{\rm ad}^{(\alpha)}(r)\right]+\frac{[J(J+1)-\Lambda^2]\hbar^2}{2\mu_\alpha r^2}[1+g^{(\alpha)}(r)]+{\rm sg}_\Lambda(e/f)\Delta V_\Lambda^{(\alpha)}(r)[J(J+1)]^\Lambda\right\}\psi_{vJ}(r)=E_{vJ}\psi_{vJ}(r), \tag{47}$$

in which ${\rm sg}_\Lambda(e/f)$ is a dimensionless numerical factor defined by the e/f parity of the level of interest (see below), and the overall Λ-doubling function is defined as [69],

$$\Delta V_\Lambda^{(\alpha)}(r)=\left(\frac{\hbar^2}{2\mu_\alpha r^2}\right)^{2\Lambda}f_\Lambda(r). \tag{48}$$

No general theoretical predictions regarding the long-range behavior expected for the Λ-doubling function $\Delta V_\Lambda^{(\alpha)}(r)$ are known to the author, other than that imposed by the first factor on the right-hand side of Eq. (48). The mass-independent radial function $f_\Lambda(r)$ is therefore written as a simple polynomial expansion in the reduced variable $y_{p_\Sigma}^{r_e}(r)$ of Eq. (5), that is,

$$f_\Lambda(r)=\sum_{i=0}^{N_\Lambda}w_i^\Lambda y_{q_\Lambda}^{r_e}(r)^i, \tag{49}$$

in which the expansion coefficients w_i^{Λ} have units $1/(\mathrm{cm}^{-1})^{2\Lambda-1}$ [70]. The fact that $y_q^{r_e}(r)\to 1$ as $r\to\infty$ means that $f_{\Lambda}(r)$ will necessarily approach a finite value in this limit, and hence that at long range $\Delta V_{\Lambda}^{(\alpha)}(r)\propto 1/r^{2\Lambda}$.

If the dominant perturbing state giving rise to the Λ-doubling has $^1\Sigma^+$ symmetry, then $\mathrm{sg}_{\Lambda}(e/f)=+1$ for e-parity levels, and equals 0 for f-parity levels. Similarly, if the perturbing state has $^1\Sigma^-$ symmetry, then $\mathrm{sg}_{\Lambda}(e)=0$ and $\mathrm{sg}_{\Lambda}(f)=-1$. Should if the identity of the dominant perturbing state be unknown, or if one does not wish to make any *a priori* assumption about its symmetry, $\mathrm{sg}_{\Lambda}(e/f)$ is normally set to $+\frac{1}{2}$ for e-parity levels and $-\frac{1}{2}$ for f levels. **dPotFit** requires the user to select one of these conventions for $\mathrm{sg}_{\Lambda}(e/f)$ when fitting to or predicting Λ-doubling splittings for a given electronic state.

5.4. Doublet splittings for $^2\Sigma$ states

In $^2\Sigma$ state molecules, the quantum-number label J is normally assigned to the total angular momentum, which is the vector sum of the spin ($\vec{S}$) and nuclear rotational ($\vec{N}$) angular momenta, *i.e.*, $\vec{J}=\vec{N}+\vec{S}$. The interaction of $\vec{N}$ with the total electron spin angular momentum $\vec{S}$ gives rise to a term in the Hamiltonian with the form $\gamma\vec{N}\cdot\vec{S}$, which causes shifts of the e and f parity components of a given rotational level that increase linearly with N. A derivation analogous to that used for Λ-doubling [10] yields the following effective radial Hamiltonian for an electronic state with $^2\Sigma$ symmetry:

$$\left\{-\frac{\hbar^2}{2\mu_\alpha}\frac{d^2}{dr^2}+\left[V_{\mathrm{ad}}^{(1)}(r)+\Delta V_{\mathrm{ad}}^{(\alpha)}(r)\right]\right. \tag{50}$$

$$+\frac{N(N+1)\hbar^2}{2\mu_\alpha r^2}\left[1+g^{(\alpha)}(r)\right]$$

$$\left.+\mathrm{sg}_{\Sigma}(e/f;N)\Delta V_{\Sigma}^{(\alpha)}(r)\right\}\psi_{v,J}(r)=E_{v,J}\psi_{v,J}(r), \tag{51}$$

in which $\mathrm{sg}_{\Sigma}(e;N)=+N/2$, $\mathrm{sg}_{\Sigma}(f;N)=-(N+1)/2$ and

$$\Delta V_{\Sigma}^{(\alpha)}=\left(\frac{\hbar^2}{2\mu_\alpha r^2}\right)f_{\Sigma}(r), \tag{52}$$

and as for the case of Λ doubling, the radial function is expanded as

$$f_{\Sigma}(r)=\sum_{i=0} w_i^{\Sigma} y_{q_{\Sigma}}^{r_e}(r)^i, \tag{53}$$

in which the expansion coefficients w_i^{Σ} are dimensionless.

In spite of their different mass and quantum-number dependence, the formal structure of the treatments of Λ-doubling $\{e/f\}$ splittings and $^2\Sigma$ state $\{e/f\}$ splittings is quite similar. As a result, the control parameters governing this treatment are input to **dPotFit** through the same set of READ statements, and the integer input parameter IOMEG formally associated with the definition of the value of Λ (see READ #7 in Appendix B) is used to distinguish between the two cases.

6. Computational methods

6.1. Solving the radial Schrödinger equation

The central computational activity of program **dPotFit** is solving the effective radial Schrödinger equation of (1), (44), (47) or (50) many hundreds or thousands or tens of thousands of times. In particular, in each cycle of the iterative non-linear fits, it must solve one of these equations in order to determine the upper- and/or lower-state eigenvalue(s) of every transition in the data set with a numerical accuracy at least an order of magnitude better than the experimental uncertainty for that datum. It also must determine the associated radial eigenfunctions in order to generate the partial derivatives of every eigenvalue with respect to each of the parameters in the Hamiltonian for that electronic state using the Hellmann–Feynman theorem expression

$$\frac{\partial E_{v,J}}{\partial p_j}=\left\langle \psi_{v,J}(r)\left|\frac{\partial\widehat{\mathcal{H}}}{\partial p_j}\right|\psi_{v,J}(r)\right\rangle. \tag{54}$$

These quantities are required in order to provide the partial derivatives of each datum with respect to all parameters of the model that are required by the least-squares fitting procedure.

dPotFit performs these eigenvalue/eigenfunction calculations using a numerical propagation algorithm based on the famous Cooley–Cashion–Zare subroutine SCHR [71–75]. The present version of that routine incorporates several unique features, such as the ability to locate quasi-bound (or tunneling-predissociation) levels automatically, and to calculate both their widths and the partial derivatives of those widths with respect to the potential function parameters [4,10,76,77]. This last capability is required for cases in which measured tunneling predissociation level widths are included in the experimental data set being analyzed [10]. Most details and features of the Schrödinger-solver routine SCHRQ used by **dPotFit** are described in Ref. [4] and the online manual for earlier versions of that program [78], and hence need not be discussed here. However, it is important to point out the role and significance of three parameters that control the numerical propagation procedure, and must be specified in the input data file.

The accuracy of the eigenfunctions and eigenvalues obtained using subroutine SCHRQ is largely determined by the size of the fixed radial mesh RH (read on line #11 of the data file) used in the numerical integration of Eqs. (1), (44), (47) or (50). For potentials that are not too steep or too sharply curved, sufficient accuracy is normally obtained by using an RH value that yields a minimum of $\sim$50 mesh points between adjacent wavefunction nodes in the classically allowed region. An appropriate mesh size (in units Å) may be estimated using the 'particle-in-a-box' expression

$$\mathrm{RH}=\pi/\left(\mathrm{NPN}\times\left[\mu\times\max\{E-V(r)\}/\mathbb{C}_{\mathrm{u}}\right]^{1/2}\right) \tag{55}$$

in which NPN is the selected minimum number of mesh points per wavefunction node (say 50), '$\max\{E-V(r)\}$' is the maximum of the local kinetic energy (in cm^{-1}) for the levels

under consideration (an upper bound to which is the potential well depth $\mathfrak{D}_e$), the reduced mass μ is in amu, and the numerical factor $\mathbb{C}_u \equiv \hbar^2/2 = 16.857629206$ [amu Å^{2}cm^{-1}]. A value of NPN that is too small yields unreliable results, while a value that is too large may consume excessive computational effort and require that array dimensions be made inconveniently large. Note that while Eq. (55) may be a useful starting point, a careful user *should always* examine the effect of different RH values on the calculated band constants written to Channel 7 in order to ensure that the computation yields results that have an accuracy that is adequate for their particular application.

The numerical integration is performed on the interval from RMIN to RMAX (see READ statement #11) using the Numerov algorithm [71,79]. These bounds must lie sufficiently far into the inner and outer classically-forbidden regions (for which $V_{\rm eff}(r) > E_{v,J}$) that the wavefunction has decayed by several orders of magnitude relative to its amplitude in the classically-allowed region. The present version of the code prints warning messages if this decay is not smaller by a factor of at least 10^{-9}; if such warnings are printed, a smaller RMIN or larger RMAX value should be used to remove them and to ensure that the desired accuracy is achieved. However, if RMIN lies too far into the classically-forbidden regions and $[V_J(r) - E]$ becomes extremely large, then the integration algorithm may become numerically unstable for the specified mesh size. If this occurs, a warning message is printed, and the beginning of the integration range is automatically shifted outward until the problem disappears. Use of a slightly larger value of RMIN would cause such warning messages to disappear and (marginally) reduce the computational effort. For most diatomic molecules, a reasonable value of RMIN is ca. $0.5-0.8$ times the smallest inner turning point for the levels involved in the data set, but for hydrides or other species of low reduced mass, even smaller values may be required.

The program internally defines the upper bound on the range of numerical integration as the smaller of the read-in value of RMAX or the largest distance consistent with the specified mesh and the internally-defined (see Section 8.2) potential energy array dimension. As with RMIN, the choice of RMAX is not critical so long as (for truly bound states) the wave function has decayed to an amplitude much smaller than that in the classically allowed region, and the same relative amplitude decay test (of 10^{-9}) is used for it. However, due to the anharmonicity of typical molecular potential curves, the requisite values of RMAX are much larger for highly excited vibrational levels than for those lying near the potential minimum. Moreover, for quasibound levels, RMAX should lie in the classically-allowed region beyond the outermost potential function turning point for the level in question.

6.2. Calculating 'pressure' and 'acoustic' second virial coefficients

For molecules formed from ground-state rare-gas atoms, measured "pressure" and "acoustic" second virial coefficients, $B_2(T)$ and $\beta_{2a}(T)$, respectively, may be employed together with spectroscopic data in a combined analysis to determine a potential energy function for the species in question. **dPotFit** is able to fit to a combination of virial coefficients with the various types of spectroscopic data [80]. "Pressure" second virial coefficients are computed semi-classically as the sum of the classical term and the first and second (translational) quantum corrections [81–83], that is,

$$B_2(T) = B_{\rm cl}(T) + \left(\frac{\hbar^2}{2\mu}\right) B_{\rm Q}^{\rm I}(T) + \left(\frac{\hbar^2}{2\mu}\right)^2 B_{\rm Q}^{\rm II}(T) + \cdots. \tag{56}$$

The classical term may be written as

$$B_{\rm cl}(T) = -2\pi N_{\rm A} \int_0^\infty \left[e^{-V(r)/k_{\rm B}T} - 1\right] r^2 dr, \tag{57}$$

in which $N_{\rm A}$ is the Avogadro number, $k_{\rm B}$ is the Boltzmann constant, and m is the atomic mass, while the first quantum correction, $B_{\rm Q}^{\rm I}(T)$, is given by

$$B_{\rm Q}^{\rm I}(T) = \pi N_{\rm A} \left(\frac{1}{6(k_{\rm B}T)^3}\right) \int_0^\infty e^{-V(r)/k_{\rm B}T} \left(\frac{dV(r)}{dr}\right)^2 r^2 dr, \tag{58}$$

and the second quantum correction, $B_{\rm Q}^{\rm II}(T)$, is given by

$$B_{\rm Q}^{\rm II}(T) = -\pi N_{\rm A} \left(\frac{1}{60(k_{\rm B}T)^4}\right) \int_0^\infty e^{-V(r)/k_{\rm B}T} \left[\left(\frac{d^2V(r)}{dr^2}\right)^2 + \frac{2}{r^2}\left(\frac{dV(r)}{dr}\right)^2 + \frac{10}{9k_{\rm B}T}\frac{1}{r}\left(\frac{dV(r)}{dr}\right)^3 - \frac{5}{36(k_{\rm B}T)^2}\left(\frac{dV(r)}{dr}\right)^4\right] r^2 dr. \tag{59}$$

Since all three terms are explicit functions of the potential energy function and its radial derivatives, it is a straightforward matter to determine expressions for the partial derivatives of $B_2(T)$ with respect to the potential function parameters that are required for the least-squares fit procedure.

Just as "pressure" second virial coefficients explain the leading deviation of the *P vs. V* behavior of a gas from the 'ideal gas law' $PV = nRT$, so the "acoustic" second virial coefficients provide analogous explanations for non-ideal-gas behavior of the speed of sound. As it has been argued that the temperature dependence of the acoustic second virial coefficient is an even more sensitive test of a proposed empirical or theoretical PEF [84,85], acoustic second virial coefficients have been included as an additional type of experimental data to which **dPotFit** can fit. The "acoustic" second virial coefficient $\beta_{2a}(T)$ may be written as a linear combination of the usual "pressure virial" coefficient $B_2(T)$ with its first and second derivatives with respect to temperature (see Eq. (19) of Ref. [83]), which means that it may be expressed in terms of a sum of integrals analogous to those of Eqs. (57)–(59).

All of Eqs. (57)–(59) involve numerical integration over the range $r \in [0, \infty)$ of integrands which, for physically realistic potential functions, are everywhere 'well behaved'. Program **dPotFit** performs all of these integrations at the same time, after first mapping this semi-infinite domain onto the finite interval $y_{p_{\rm vir}}(r; r_e) \in [-1, +1]$ where $y_{p_{\rm vir}}^{r_e}(r)$ is defined by Eq. (5), and applying a standard eight-point Gauss–Legendre quadrature procedure, with a twist. Following an initial eight-point integration across the interval $[-1, +1]$, that interval is divided in two, analogous quadratures are performed over both halves, and their sum is compared to the result from the

preceding step. This bisection procedure is then applied to each subinterval, and the procedure repeated until the desired degree of convergence (currently 0.0001%) is achieved. The present code uses a mapping coordinate defined by $p_{\rm vir} = 1$. In any case, standard formulae suggest that the errors in the resulting virial coefficients should decrease by a factor of $1/2^{16}$ with each stage of bisection (see Section 25.4.30 of Ref. [86]).

7. Fitting strategies

7.1. Initial trial parameters

In a DPF treatment of experimental data, the observables – the transition energies and/or tunneling level widths – are not linear functions of the parameters of the radial functions characterizing the effective Hamiltonian. As in any non-linear least-squares fitting procedure, it is therefore essential to have a set of realistic initial trial values for all fit-parameters. For BOB radial functions and Λ-doubling or $^2\Sigma$-splitting radial functions, this presents little practical difficulty. All of those functions are relatively weak, and practical experience indicates that if their parameters are initially all set to zero or (for the $w_0^{\Lambda/\Sigma}$ coefficient for Λ-doubling or $^2\Sigma$ splitting) given some plausible small initial trial value, and then let go free, the fits are stable and well-behaved. However, one would not normally try to obtain an accurate final determination of those supplementary radial strength functions until a realistic description of the potential energy function for the reference isotopolog $V_{\rm ad}^{(1)}(r)$ has been obtained.

For the potential energy function $V_{\rm ad}^{(1)}(r)$ (or $V_{\rm CN}(r)$) itself, the problem of determining initial trial parameters is somewhat more challenging, and a number of strategies have been utilized. From a conventional preliminary parameter-fit analysis of the data [56], it is normally relatively straightforward to obtain a good estimate of the potential minimum position r_e, which is a central parameter for all model potentials. It is normally also not difficult to obtain a plausible initial estimate for the well depth $\mathfrak{D}_e$, which is a central parameter in the EMO, MLR and DELR models. Moreover, experience suggests that this initial trial value of $\mathfrak{D}_e$ should often be held fixed until a good fit to some associated set of exponent expansion coefficients $\{\beta_i\}$ is obtained.

The most generally useful method for generating a realistic set of initial trial β_i values is to fit a set of approximate potential function points generated in some other manner to the chosen potential form. For example, one might use a conventional "parameter-fit" analysis of the data in terms of Dunham expansions or near-dissociation expansions [56] to determine analytic level energy expressions, and then use the resulting functions to generate a pointwise RKR potential for that state [87,88]. Alternatively, one may use *ab initio* predictions to define such a preliminary potential, or (say, for a double minimum potential) a combination of RKR and *ab initio* points. A companion program named **betaFIT** has been developed for fitting such input potential arrays to determine realistic estimates of the exponent expansion coefficients $\{\beta_i\}$ of an EMO, MLR or DELR potential, or of the power-series coefficients $\{c_i\}$ of a GPEF potential [5,89].

It is very important to realize that direct potential fits of the type performed by **dPotFit** are *highly* non-linear, and care must be taken to prevent them from diverging. To that end, when initial trial values of each of the fitting parameters are read in, **dPotFit** requires the user to specify, one-by-one, whether each parameter is to be held fixed or varied in that particular fit. This allows a user to release a small number of parameters (say, r_e, β_0 and β_1) initially, while holding all others fixed, and once preliminary optimized values of those parameters have been determined, use then to update the original trial values in the input data file and a new fit be performed that allows additional parameters (say β_2, β_3 and β_4) to be also free. This sort of stepwise procedure is often necessary to ensure that the fit to determine the main supporting potential remains stable. However, once converged values for r_e and set of $\{\beta_i\}$ parameters have been determined, the basic description of the system is normally sufficiently well defined that all BOB and/or Λ-doubling or $^2\Sigma$ splitting parameters may be released at the same time.

As an alternative to the use of a code such as BETAFIT to determine a complete set of trial $\{\beta_i\}$ parameters in a single step, one may also proceed in a stepwise manner by initially considering only a fraction of the data and a small number of parameters, and then progressively extending the range until the whole data set has been included. By specifying parameters VMIN, VMAX and JTRUNC in READ #7, **dPotFit** allows the user to limit the range of data to be utilized in a given fit without having to edit the data file. This makes it quite straightforward to restrict the vibrational range of the data to be considered to (say) $v = 0-3$ and fit to a potential model that has only (say) r_e, β_0 and β_1 as free parameters, with all higher-order β_i (for $i \geq 2$) fixed at zero. Once a converged fit to that restricted data set has been obtained, the vibrational range of data and number of β_i parameters may be extended, using initial trial values of zero for the added higher-order coefficients β_i, and the process repeated until the entire data set has been included.

7.2. Multi-state fits

Should the fit involve more than one electronic state, it is often necessary to utilize a stepwise procedure – initially determining estimates of the parameters for one state at a time – prior to a final step in which all parameters are freed simultaneously. This tends to be necessary because a relatively poor initial representation of one state can inhibit one's ability to determine an optimum representation for another. In some cases, this might be a simple matter of first performing a one-state fit to the pure rotational and vibration–rotation data for a selected state, and then holding its parameters fixed while performing a two-state fit that includes the electronic transition data and varies only the parameters of the second state. When a good model is determined for the second state, the two-state fit would then be repeated while allowing the parameters for both states to be fitted simultaneously. However, when electronic transition data are available, they are

often the only source of information about the upper vibrational levels of the state of interest, so they cannot be ignored when one is attempting to obtain a good description of the lower state.

The best way to treat this problem then is firstly, to fit to *all* of the data (electronic and other) involving the first (usually ground) state, while representing the levels of all other electronic states either as individual term values, or by sets of band constants for each isotopolog. While a relatively large number of parameters (sometimes thousands, when term values are fitted!) tend to be required for such fits, use of this approach means that the determination of parameters for the first state is not affected by the model(s) chosen to represent the other state(s). This ability to represent the levels of a given state by individual term values was first introduced in Version 1.2 of this code, and the extension to allow fits to band constants was introduced in Version 2.0 [90]. These options are invoked by giving the 'potential-type' parameter PSEL that is input via READ #9 the value '−2' for a state whose levels are to be represented by term values or '−1' for a state whose levels are to be represented by band constants (see Appendix C). Once a 'good' fit is obtained to a model for the potential energy and any other radial functions required to describe the first electronic state fully, the parameters describing that state may be held fixed in a two-state fit to determine an optimum model for a second electronic state. After that has been done, the parameters for both states should be fitted simultaneously. This stepwise procedure may then continue until all of the data have been fitted simultaneously, preferably to potential function models for all of the electronic states involved in the data set.

Of course, the PSEL = −1 or −2 options may also prove useful for cases in which only fragments of information are available for a given state, or when level energy irregularities due to perturbations make a potential function treatment impractical.

Note too that because of the vagaries on the manner in which parameters are enumerated and stored in **dPotFit**, for a multi-state analysis in which the level energies of one or more states are represented by band constants or term value, those states should be listed last in the main (Channel 5) input data file.

8. Practical details ··· using dPotFit

8.1. Units, uncertainties and parameter rounding

The units of mass, length and energy used throughout this program, and assumed for all input data, are u (amu), Å and cm^{-1}, respectively, except for virial coefficients, whose units are [cm^3/mole]. The values of the relevant physical constants occur in the program as the single factor $\mathbb{C}_u \equiv \hbar^2/2 = 16.857629206$ [amu cm^{-1} $Å^2$] appearing in the effective radial Schrödinger Eqs. (1), (44), (47) or (50). This numerical constant is based on the 2010 CODATA recommended physical constant values [91], while the atomic isotope masses stored in subroutine MASSES were taken from the 2012 compilation of Ref. [92].

Because **dPotFit** performs weighted least-squares fits, each input datum must be accompanied by an estimated uncertainty u_i in the same units (usually cm^{-1}) as the observable. The quality of a fit of an M–parameter model to N input data which yields the predicted quantities $\{y_i^{calc}\}$ is characterized by the value of the dimensionless root mean square deviation $\overline{dd}$, defined as

$$\text{DRMSD} \equiv \overline{dd} = \left\{ \frac{1}{N} \sum_{i=1}^{N} \left[\frac{y_i^{calc} - y_i^{obs}}{u_i} \right]^2 \right\}^{1/2}, \tag{60}$$

or by the related dimensionless standard error σ_f, defined as $\text{DSE} \equiv \overline{\sigma}_f = \overline{dd}\sqrt{N/(N-M)}$. This data weighting allows observables with very different magnitudes and very different absolute uncertainties (*e.g.*, microwave *vs*. electronic band head data), and even different dimensions (*e.g.*, virial coefficients), to be treated concurrently in an appropriately balanced manner. A "good" fit is one that yields DSE and $\overline{dd}$ values close to unity: a $\overline{dd}$ value of (say) 3.7 would mean that, on average, the predictions of the model disagree with the input data by 3.7 times the estimated experimental uncertainties. Convergence to a value significantly greater than unity will reflect the fact that either the experimental uncertainties assigned to the data were overly optimistic, that the model(s) used for the PEF and/or BOB corrections are inadequate, or that the data are affected by non-mechanical perturbations that are not allowed for in the Hamiltonian of Eq. (1). The non-linear least-squares procedure implemented in **dPotFit** is simply an iterative series of steepest-descent steps which proceed until $\overline{dd}$ converges to 7 significant digits or it appears that cumulative numerical round-off errors are inhibiting full convergence.

In addition to reporting the 95% confidence limit (approximately 'two-sigma') uncertainty in each fitted parameter, **dPotFit** follows the approach of Ref. [93] by also always listing the associated "parameter sensitivity" (identified as PS in the output). This quantity is defined (see Eq. (4) of Ref. [93]) as the magnitude of the largest change in the given parameter whose effect on the predictions of the model, if all other parameters are held fixed, could increase $\overline{\sigma}_f$ by a maximum of $(0.1/M)\overline{\sigma}_f$. The parameter sensitivity indicates the degree to which any particular fitted parameter value may be rounded off while having no significant effect (within the data uncertainties) on the ability of the resulting parameter set to predict the input data accurately. For the illustrative cases considered in Ref. [93], to three significant digits, rounding off all parameters at the first significant digit of their sensitivity had no meaningful effect on the values of $\overline{\sigma}_f$ or $\overline{dd}$. In general, these 'parameter sensitivities' are orders-of-magnitude smaller than the overall 'parameter uncertainties', which also take account of interparameter correlation.

Note that as a flag for the user, in the main (Channel 6) output file the listed name of any fitted parameter for which the (95% confidence limit) parameter uncertainty is greater than 100% is preceded by a '∗'. If this occurs for the last term in a potential function exponent polynomial or BOB radial strength function expansion, it usually indicates that the user has more free parameters in the model than the data can support.

Another feature of **dPotFit** is its implementation (within subroutine NLLSSRR) of the automated "sequential rounding and refitting" (SRR) procedure of Ref. [93], which minimizes the total number of significant digits required to represent the overall parameter set with no (significant) loss of accuracy. Application of this procedure involves a substantial increase (by up to a factor of $M/2$) in computational effort relative to that required for an ordinary fit, so it is usually desirable to omit it in the many trial fits involved in any global data analysis, such as the sets of trial fits required to determine manually the optimum values of q, $r_{\rm ref}$ and N_β. Application of this SRR procedure is turned on or off by the value of the flag IROUND that is set by the user in READ #6 of the input data file. One would normally turn this flag off (set IROUND=0) for preliminary analyses, and only turn it on when one wishes to generate a final parameter set to report and distribute. As discussed in Ref. [93], in most cases setting IROUND $= \pm 1$ yields an excellent degree of rounding without significant loss of precision, but in some cases it may be necessary to set |IROUND| > 1 in order to reduce the growth of $\overline{dd}$ during the sequential rounding process.

One final choice regarding the manner in which the least-squares fits are performed is whether or not to perform "robust" fits. As described in Ref. [94] and references therein, *robust* least-squares fits attempt to minimize the effect of data "outliers", which are defined as observations that yield anomalously large discrepancies with the model. When this choice is invoked, **dPotFit** adopts the approach of Ref. [94] and replaces the normal least-squares data weights $w_i = 1/(u_i)^2$ by the '*robust*' weights $w_i^{\rm rob} = 1/\left[(u_i)^2 + (y_i^{\rm calc} - y_i^{\rm obs})^2/3\right]$. Because the latter depend on the then-current degree of agreement of the data with the model, fits of this type are repeated iteratively, with the parameter values and the weights being updated in each cycle until self-consistency is achieved. As a result, *robust* fits require substantially more computer time than normal fits do. Moreover, the fact that *robust* weighting reduces the effect of large $[y_i^{\rm calc} - y_i^{\rm obs}]$ values on DSE and $\overline{dd}$ makes it more difficult to interpret differences in those quantities obtained from fits to different versions of a model, and it may tend to obscure the presence of systematic discrepancies that indicate shortcomings of the model, rather than of the data. However, our (limited) experience with this option indicates that it can facilitate identifying 'bad' data as well as local and/or systematic discrepancies from a model, when one examines the $[y_i^{\rm calc} - y_i^{\rm obs}]$ results in the Channel 8 output file.

Independent of whether or not a 'robust' fit has been performed, flags for data outliers are provided in the Channel 8 output file by the fact that the entry for a datum for which the absolute value of $[y_i^{\rm calc} - y_i^{\rm obs}]/u_i$ lies between 2 and 4 is followed by the symbol '$*$', that for a datum for which this quantity lies between 4 and 8 is followed by '$**$', while if this quantity is larger than 8, the entry for the datum is followed by the symbol '$***$'. What should be done with this information is up to the user.

8.2. Array dimensions, input/output conventions, and program execution

The operation of program **dPotFit** involves the use of a number of moderately large multi-dimensional integer and double-precision real-number arrays whose size is specified at the time the program is compiled. If those arrays are unnecessarily large, it could slow or hinder computations on some computers. The current version of **dPotFit** assumes (but does not require, as it is also compatible with most modern Fortran compilers) the use of a Fortran-77-type compiler that does not allow run-time array dimensioning. Thus, since one does not wish to recompile the code case-by-case, setting those array dimensions at modest (but adequate) values should facilitate computations by minimizing computer memory requirements. The parameters that set the upper bounds on the sizes of the large arrays are set by PARAMETER statements contained in the block data subroutine arrsizes.h, which is embedded at a numerous of locations in the code. A user who wishes to change any of these dimensions must be sure to make that change at *all* occurrences of this bock.

Parameters defined in this way include the maximum number of isotopologs being considered, NISTPMX, the maximum number of electronic states, NSTATEMX, the maximum number of fitting parameters, NPARMX, the maximum number of data, NDATAMX, the maximum number of observed vibrational levels in any of the electronic states considered, NVIBMX, and the maximum dimension for the radial arrays used to store the potential energy and related functions, NPNTMX. A user should examine the first occurrence of this block in the source code before compiling it, and ensure that these parameters have values adequate for the types of systems that they will be considering.

dPotFit reads two separate input data files. The first one contains the experimental data being fitted; its name is read in line #2 of the second data file. The second data file is the 'instruction' data file that contains the initial trial parameters defining the model, plus the control variables that characterize the problem and specify which type of fit is to be performed. It is the FORTRAN 'standard input' file read on Channel 5. The structure of these data files and the definitions of and options for the various input quantities are presented in Appendices A–C.

The program writes standard output to Channel 6 and supplementary output files to a selection of Channels 7–20. The output to Channel X is written to the file WRITFILE.X, where WRITFILE is a user-specified output filename that is input *via* line #3 of the (Channel 5) 'instruction' data file.

Channel 6 *output* summarizes the input data, describes the nature of the fit being performed, reports the results of the fit, lists fitted parameters, their sensitivities, and their 95% confidence limit uncertainties, and presents a band-by-band summary of the $[y_i^{\rm calc} - y_i^{\rm obs}]$ results.

Channel 7 *output* contains values of the band constants for all specified levels of all isotopologs in all states involved in the analysis, as generated from the final results of the fit.

Channel 8 *output* consists of a full listing of the $\{[y_i^{\rm calc}-y_i^{\rm obs}]\}$ and $\{[y_i^{\rm calc}-y_i^{\rm obs}]/u_i\}$ values for all data utilized in the fit.

Channels 10–16 *output* files contain arrays of values of the various radial functions associated with the model Hamiltonian, and the associated 95% confidence limit uncertainties.

Channel 10 contains the effective radial potential for the reference isotopolog $V_{\rm ad}^{(1)}(r)$ (or $V_{\rm CN}(r)$),

Channels 12 *and* 13 contain the adiabatic BOB radial strength functions $\tilde{S}_{\rm ad}^{\rm A}(r)$ and $\tilde{S}_{\rm ad}^{\rm B}(r)$, respectively,

Channels 14 *and* 15 contain the centrifugal BOB radial strength functions $\tilde{R}_{\rm na}^{\rm A}(r)$ and $\tilde{R}_{\rm na}^{\rm B}(r)$, respectively, and

Channel 16 contains the radial strength function associated with Λ-doubling or $^2\Sigma$ level splittings, $f_\Lambda(r)$ or $f_\Sigma(r)$, as appropriate.

Channel 20 contains a listing of the potential function parameters determined by the fit, formatted so as to facilitate their inclusion in the 'instruction' data file for a subsequent fit.

Those executing **dPotFit** in a UNIX or LINUX operating system environment may find it convenient to do so using a shell named (say) `rdpot`, such as that shown below, which may be stored in the system 'bin' directory or the user's 'bin' directory:

```
#!/bin/sh
# UNIX shell 'rdpot' to execute the compiled version of program dPotFit named
#   dpot.x, which is stored in the user directory /userpath/. The Channel-5
#   input data file  $1.5  and the output files WRITFILE.6, WRITFILE.7, ...
#   etc., will be in the same directory.
#
time /userpath/dpot.x  < $1.5

if [ -e MAKEPRED ]; then
   rm MAKEPRED
   fi
```

Note that `userpath` is a path specifying the location of the executable file `dpot.x` on the user's computer, and `MAKEPRED` is defined below. This shell allows the program to be executed with the simple command:

```
rdpot filename
```

in which `filename.5` is the data file containing the instructions regarding the type of fit to be performed. `filename` may be any name chosen by the user, but it is usually convenient if it has a name which identifies the particular case. If this file does not reside in the current directory, this name must also include the relative path.

8.3. Generating sets of predicted transition energies

`dPotFit` may also be employed to generate a set of predicted data $\{y_i^{\rm calc}\}$ from a given (fixed) set of input parameters. This option is invoked by setting the value of the input variable in READ #2, which is normally the name of the file containing the experimental data, to be `MAKEPRED`. In this case the program will use READ #37 in the Channel 5 instruction data file to read specifications and selection rules for bands for which the user wishes to generate predictions. The resulting predictions are written in the normal Channel 8 output format to file `filename.8`, and in 'data input' format to file `filename.4`.

Anyone who wishes to be Registered with the author as a user of this code, eligible to be sent any future bug fixes or updates, should fill in the online form at the www address "http://scienide2.uwaterloo.ca/rleroy/dPotFit16".

Acknowledgments

I am pleased to acknowledge the contributions of Dr. Dominique Appadoo, who designed and 'built' the initial version of the procedures for handling the many different kinds of experimental data allowed by the code, and Mr. Jenning Seto and Mr. Yiye Huang who helped develop the early versions of the overall program. I am also grateful to Professor George McBane of Grand Valley State University for helpful comments and suggestions, to Mr. Philip Myatt and especially Dr. James Hornkohl, for assistance with debugging the final version of the code, and to Professor F.R.W. McCourt of the University of Waterloo for a thorough, critical reading of the manuscript. Finally, I am pleased to thank the Natural Sciences and Engineering Research Council of Canada which has generously supported my work in this area for the past four decades.

Appendix. Supplementary data

Supplementary material associated with this paper consists of: a text file containing the full standalone FORTRAN source code for program **dPotFit**, a text file containing the sample data files that are presented in Appendix D, and a PDF document containing: (i) Appendix A, a detailed illustrative description of the manner in which the various types of experimental data are to be arranged in the input data file whose name is input through READ #2 of the Channel 5 'instruction' input file, (ii) Appendix B, a structured listing of the READ statements used in the program, (iii) Appendix C, full definitions of all of the parameters and quantities read in the input 'instruction' data file, and descriptions of the associated program options, (iv) Appendix D, commented listings of illustrative sample input data files, (iv) Appendix E,

listings of the main (Channel 6) output files for the illustrative cases of Appendix D, and (v) Appendix F – an overview of the program structure.

Note that the Equation and reference numbering appearing in this Supplemental Material document refer to those in this Journal article, so for completeness we hereby cite ten papers referenced there that are associated with particular applications of this code, but had not been referred to above [103–113].

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.jqsrt.2016.06.002.

References

[1] Rydberg R. Z Phys 1931;73:376.
[2] Klein O. Z Phys 1932;76:226.
[3] Rees ALG. Proc Phys Soc (Lond) 1947;59:998.
[4] Le Roy RJ. **LEVEL**: a computer program for solving the radial Schrödinger equation for bound and quasibound levels. J Quant Spectrosc Radiat Transf; 2016, this issue, http://dx.doi.org/10.1016/j.jqsrt.2016.05.028.
[5] Le Roy RJ, Pashov A. **betaFIT**: a computer program to fit pointwise potentials to selected analytic functions. J Quant Spectrosc Radiat Transf; 2016, this issue, http//dx.doi.org.10.1016/j.jqsrt.2016.03.036.
[6] Watson JKG. J Mol Spectrosc 1980;80:411.
[7] Watson JKG. J Mol Spectrosc 2004;223:39.
[8] Le Roy RJ. J Mol Spectrosc 1999;194:189.
[9] Le Roy RJ, Huang Y. J Mol Struct (Theochem) 2002;591:175.
[10] Huang Y, Le Roy RJ. J Chem Phys 2003;119:7398, erratum: J Chem Phys 2007;126:169904.
[11] Ogilvie JF. J Phys B: At Mol Opt Phys 1994;27:47.
[12] Lee EG, Seto JY, Hirao T, Bernath PF, Le Roy RJ. J Mol Spectrosc 1999;194:197.
[13] Seto JY, Morbi Z, Charron F, Lee SK, Bernath PF, Le Roy RJ. J Chem Phys 1999;110:11756.
[14] Seto JY. Direct fitting of analytic potential functions to diatomic molecule spectroscopic data [M.Sc. thesis]. Department of Chemistry, University of Waterloo; 2000.
[15] Le Roy RJ, Huang Y, Jary C. J Chem Phys 2006;125:164310.
[16] Le Roy RJ, Henderson RDE. Mol Phys 2007;105:663.
[17] Salami H, Ross AJ, Crozet P, Jastrzebski W, Kowalczyk P, Le Roy RJ. J Chem Phys 2007;126:194313.
[18] Shayesteh A, Henderson RDE, Le Roy RJ, Bernath PF. J Phys Chem A 2007;111:12495.
[19] Le Roy RJ, Dattani N, Coxon JA, Ross AJ, Crozet P, Linton C. J Chem Phys 2009;131:204309.
[20] Le Roy RJ. In: Demaison J, Csaszar AG, editors. Determining equilibrium structures and potential energy functions for diatomic molecules of equilibrium structures of molecules. London: Taylor & Francis; 2011. p. 159–203 [Chapter 6].
[21] Dattani N, Le Roy RJ. J Mol Spectrosc 2011;268:199.
[22] Huang Y. Determining analytical potential energy functions of diatomic molecules by direct fitting [M.Sc. thesis]. Department of Chemistry, University of Waterloo; 2001.
[23] Le Roy RJ, Appadoo DRT, Anderson K, Shayesteh A, Gordon IE, Bernath PF. J Chem Phys 2005;123:204304.
[24] Le Roy RJ, Appadoo DRT, Colin R, Bernath PF. J Mol Spectrosc 2006;236:178.
[25] Le Roy RJ, Haugen CC, Tao J, Li H. Mol Phys 2011;109:435.
[26] Morse PM. Phys Rev 1929;34:57.
[27] Coxon JA, Hajigeorgiou PG. J Mol Spectrosc 1990;139:84.
[28] Hedderich HG, Dulick M, Bernath PF. J Chem Phys 1993;99:8363.
[29] In early work [10,90], the exponent polynomial was split into two separate polynomials with different orders, one being employed for $r > r_e$ and the other for $r < r_e$, but sharing common low-order coefficients. However, the flexibility provided by using $y_q^{\rm ref}(r)$ as the expansion variable instead of $y_q^{r_e}(r)$ has made this complication unnecessary, and has enabled its high-order-derivative discontinuities at $r = r_e$ to be eliminated.
[30] Le Roy RJ, Bernstein RB. J Chem Phys 1970;52:3869.
[31] Le Roy RJ. Molecular spectroscopy. In: Barrow RN, Long DA, Millen DJ, editors. Specialist periodical report 3, vol. 1. London: Chemical Society of London; 1973. p. 113–76.
[32] Douketis C, Scoles G, Marchetti S, Zen M, Thakkar AJ. J Chem Phys 1982;76:3057.
[33] Tang KT, Toennies JP. J Chem Phys 1984;80:3726.
[34] Kreek H, Meath WJ. J Chem Phys 1969;50:2289.
[35] Martin F, Aubert-Frécon M, Bacis R, Crozet P, Linton C, Magnier S, et al. Phys Rev A 1997;55:3458.
[36] Aubert-Frécon M, Hadinger G, Magnier S, Rousseau S. J Mol Spectrosc 1998;188:182.
[37] Cho Y-S, Le Roy RJ. J Chem Phys 2016;144:024311.
[38] Hajigeorgiou PG, Le Roy RJ. In: 49th Ohio State University international symposium on molecular spectroscopy, Columbus, Ohio, paper WE04; 1994.
[39] Hajigeorgiou PG, Le Roy RJ. J Chem Phys 2000;112:3949.
[40] In earlier versions of the model an exponential switching function was used [38,39,95–98] but that form involves greater complexity and has no significant advantages.
[41] Yukiya T, Nishimiya N, Samajima Y, Yamaguchi K, Suzuki M, Boone CD, et al. J Mol Spectrosc 2013;283:32.
[42] There is also a very weak repulsive $m=5$ term for this state [99], but it has no discernable effect outside the immediate neighbourhood of the intercept [41].
[43] Huang Y, Le Roy RJ. In: 55th Ohio State University international symposium on molecular spectroscopy, Columbus, Ohio, paper RA01; 2000.
[44] Seto JY, Le Roy RJ. In: 55th Ohio State University international symposium on molecular spectroscopy, Columbus, Ohio, paper RA02; 2000.
[45] Pashov A, Jastrzębski W, Kowalczyk P. Comput Phys Commun 2000;128:622.
[46] Pashov A, Jastrzębski W, Kowalczyk P. J Chem Phys 2000;113:6624.
[47] Pashov A, Jastrzębski W, Jaś niecki W, Bednarska V, Kowalczyk P. J Mol Spectrosc 2000;203:264.
[48] Tao J, Le Roy RJ, Pashov A. In: 65th Ohio State University international symposium on molecular spectroscopy, Columbus, Ohio, paper MG07; 2010.
[49] Koide A, Meath WJ, Allnatt AR. Chem Phys 1981;58:105.
[50] Douketis C, Hutson JM, Orr BJ, Scoles G. Mol Phys 1984;52:763.
[51] Dunham JL. Phys Rev 1932;41:721.
[52] Šurkus AA, Rakauskas RJ, Bolotin AB. Chem Phys Lett 1984;105:291.
[53] Simons G, Parr RG, Finlan JM. J Chem Phys 1973;59:3229.
[54] Ogilvie JF. Proc R Soc (Lond) A 1981;378:287.
[55] Ogilvie JF. J Chem Phys 1988;88:2804.
[56] Le Roy RJ. **dParFit**: a computer program for fitting diatomic molecule spectral data to parameterized level energy expressions. J Quant Spectrosc Radiat Transf; 2016, this issue, http//dx.doi.org.10.1016/j.jqsrt.2016.04.004.
[57] Knöckel H, Bodermann B, Tiemann E. J Eur Phys: D 2004;28:199.
[58] Salumbides EJ, Eikema KSE, Ubachs W, Hollenstein U, Knöckel H, Tiemann E. Eur Phys J D 2008;47:171.
[59] Hellmann R, Bich E, Vogel E. Mol Phys 2008;106:133.
[60] Patkowski K, Szalewicz K. J Chem Phys 2010;133:094304.
[61] Waldrop JM, Song B, Patkowski K, Wang X. J Chem Phys 2015;142:204307.
[62] Tang KT, Toennies JP. J Chem Phys 2003;118:4976.
[63] Aziz R. J Chem Phys 1993;99:4518.
[64] On reconsideration, our previous recommendation [90] simply to set $p_{\rm ad} > m_1$, the power of the leading inverse-power term in Eq. (11) is not sufficiently restrictive.
[65] Ogilvie JF. The vibrational and rotational spectrometry of diatomic molecules. Toronto: Academic Press; 1998.
[66] Herman RM, Ogilvie JF. Adv Chem Phys 1998;103:187.
[67] Sauer SPA. Chem Phys Lett 1998;297:475.
[68] Sauer SPA, Špirko V. J Chem Phys 2013;138:024315.
[69] Note that while the derivation of Ref. [10] only addressed the case of Λ-doubling in Π states, the present implementation has been extended to handle states for which $\Lambda > 1$.
[70] This assumes, of course, that the factor $\hbar^2/(2\mu_\alpha r^2)$ has units $\rm cm^{-1}$.
[71] Cooley JW. Math Comput 1961;15:363.
[72] Cashion J. J Chem Phys 1963;39:1872.
[73] Zare RN, Cashion JK. The IBM share program D2 NU SCHR 1072 for solution of the Schrödinger radial equation. In: Cooley JW, editor Necessary and useful modifications for its use on an IBM 7090. University of California Lawrence Radiation Laboratory Report UCRL-10881; 1963.
[74] Zare RN. University of California Lawrence Radiation Laboratory Report, UCRL-10925; 1963.
[75] Zare RN. J Chem Phys 1964;40:1934.

[76] Le Roy RJ, Bernstein RB. J Chem Phys 1971;54:5114.
[77] Le Roy RJ, Liu W-K. J Chem Phys 1978;69:3622.
[78] (a) Le Roy RJ. LEVEL 1.0, University of Wisconsin Theoretical Chemistry Institute Report WIS-TCI-429G; 1971. (b) Le Roy RJ. LEVEL 2.0, University of Waterloo Chemical Physics Research Report CP-58; 1976. (c) Le Roy RJ. LEVEL 3.0, CP-110; 1979. (d) Le Roy RJ. LEVEL 4.0, CP-230; 1983. (e) Le Roy RJ. LEVEL 5.0, CP-330; 1991. (f) Le Roy RJ. LEVEL 5.1, CP-330R; 1992. (g) Le Roy RJ. LEVEL 5.2 CP-330R^2; 1993. (h) Le Roy RJ. LEVEL 5.3, CP-330R^3; 1994. (i) Le Roy RJ. LEVEL 6.0, CP-555; 1995. (j) Le Roy RJ. LEVEL 6.1, CP-555R; 1996. (k) Le Roy RJ. LEVEL 7.0–7.7, CP-642–642R^3, CP-655 & CP-661; 2000–2005, Le Roy RJ. LEVEL 8.0, CP-663; 2007–2014. (ℓ) Le Roy RJ. LEVEL 8.2, CP-668; May 2014; see ⟨http://leroy.uwaterloo.ca/programs/⟩.
[79] Smith K. The calculation of atomic collision processes. New York: Wiley-Interscience; 1971 [Chapter 4].
[80] Piticco L, Merkt F, Cholewinski AA, McCourt FRW, Le Roy RJ. J Mol Spectrosc 2010;264:83.
[81] Hirschfelder JO, Curtiss CF, Bird RB. Molecular theory of gases and liquids. New York: Wiley; 1964.
[82] Mason E, Spurling T. The final state, Topic 10. In: Rowlinson J, editor. International encyclopedia of physical chemistry and chemical physics, vol. 2. Oxford: Pergamon Press; 1969. p. 17–66 [Chapter 2].
[83] McCourt FRW. Virial coefficients. In: Wilson S, editor. Molecular physics in the physicochemical environment: spectroscopy, dynamics and bulk properties. Handbook of molecular physics and quantum chemistry, vol. 3. Wiley; Chichester 2003. p. 673–711 [Chapter 27].
[84] Trusler JPM. Physical acoustics and metrology of fluids. Bristol: Adam Hilger; 1991.
[85] Boyes SJ. Chem Phys Lett 1994;221:467.
[86] Abramowitz M, Stegun IA. Handbook of mathematical functions. New York: Dover; 1970 [ninth printing].
[87] Le Roy RJ. **RKR1**: a computer program implementing the first-order RKR method for determining diatomic molecule potential energy functions, this issue, http://dx.doi.org/10.1016/j.jqsrt.2016.03.030; 2016.
[88] Le Roy RJ. `RKR1 1.0`: a computer program implementing the first-order RKR method for determining diatomic molecule potential energy functions. University of Waterloo Chemical Physics Research Report CP-425; 1992; `RKR1 2.0`. Le Roy RJ, CP-657; 2003.
[89] (a) Le Roy RJ. `betaFIT 2.1`. University of Waterloo Chemical Physics Research Report CP-666; 2013. (b) Le Roy RJ. `betaFIT 2.0`. CP-665; 2009. (c) Le Roy RJ. `phiFIT 1.1`. CP-663R; 2006. (d) Le Roy RJ. `phiFIT 1.0`. CP-663; 2006, see also ⟨http://leroy.uwaterloo.ca/programs/⟩.
[90] (a) Le Roy RJ, Seto J, Huang Y. `DPotFit 2.0`: a computer program for fitting diatomic molecule spectra to potential energy functions. University of Waterloo Chemical Physics Research Report CP-667; 2013. (b) Le Roy RJ, Seto J, Huang Y. `DPotFit 1.2`. CP-664; 2007. (c) Le Roy RJ, Seto J, Huang Y. `DPotFit 1.1`. CP-662R; 2006. See also ⟨http://leroy.uwaterloo.ca/programs/⟩.
[91] Mohr P, Taylor B, Newell D. Rev Mod Phys 2012;84:1527.
[92] Wang M, Audi G, Wapstra A, Kondev F, MacCormick M, Xu X, et al. Chin Phys C 2012;36:1603.
[93] Le Roy RJ. J Mol Spectrosc 1998;191:223.
[94] Watson JKG. J Mol Spectrosc 2003;219:326.
[95] Coxon JA, Colin R. J Mol Spectrosc 1997;181:215.
[96] Coxon JA, Hajigeorgiou PG. J Mol Spectrosc 1999;193:306.
[97] Seto JY, Le Roy RJ, Vergès J, Amiot C. J Chem Phys 2000;113:3067.
[98] Coxon JA, Hajigeorgiou PG. J Mol Spectrosc 2000;203:49.
[99] Saute M, Bussery M, Aubert-Frécon M. Mol Phys 1984;51:1459.
[100] Allard O, Pashov A, Knöckel H, Tiemann E. Phys Rev A 2002;66:042503:1.
[101] Allard O, Samuelis C, Pashov A, Knöckel H, Tiemann E. Eur Phys J D 2003;26:155.
[102] Mitroy J, Zhang J-Y. J Chem Phys 2008;128:134305.
[103] Carrington A, Sarre PJ. J Phys (Paris) 1979;40:C1.
[104] Cosby PC, Helm H, Moseley JT. Astrophys J 1980;235:52.
[105] Helm H, Cosby PC, Graff MM, Moseley JT. Phys Rev A 1982;25:304.
[106] Jäger W, Xu Y, Gerry M. J Chem Phys 1993;99:919.
[107] Schramm B, Schmiedel H, Gehrmann R, Bartl R. Ber Bunsenges Phys Chem 1977;81:316.
[108] Rentschler HP, Schramm B. Ber Bunsenges Phys Chem 1977;81:319.
[109] Dymond JH, Marsh KN, Wilhoit RC, Wong KC. Landolt–Börstein: numerical data and functional relationships in science and technology, New series. In: Frenkel M, Marsh KN, editors. Group IV: Physical chemistry, vol. 21; 2003.
[110] Gunton W, Semczuk M, Dattani NS, Madison KW. Phys Rev A 2013;88:062510.
[111] Semczuk M, Li X, Gunton W, Haw M, Dattani NS, Witz J, et al. Phys Rev A 2013;87:052505.
[112] Henderson RDE, Shayesteh A, Tao J, Haugen CC, Bernath PF, Le Roy RJ. J Phys Chem A 2013;117:13373.
[113] Meshkov V, Stolyarov A, Le Roy RJ. Phys Rev A 2008;78:052510.