dPotFit 17: 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 N2L 3G1, Canada

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 isotopologues, and for appropriate systems, second virial coefficient data, to determine analytic potential energy functions defining the observed levels 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.

[‡] Electronic mail: leroy@UWaterloo.ca

1. Introduction

In recent years, it has become increasingly common to analyse 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, 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 isotopologues 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) fir it, **dPotFit** allows its energy levels to be represented by (often quite large) sets of independent term values $\{T_{v,J}\}$, or by a set of 'band constants' $\{G_v,\ B_v,\ -D_v,\ H_v,\ \dots\}$ for each vibrational level v of each isotopologue. 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 [5]. A procedure for doing this is presented in a separate companion work [6]. As Ref. [6] 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. Section 9 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 [7, 8], 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. [9–11], the resulting effective radial Schrödinger equation for isotopologue 'α' of molecule A–B in a singlet electronic state with electronic angular momentum projection quantum number Λ, may be written as

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

$$\left. + \frac{[J(J+1) - \Lambda^2]\hbar^2}{2\mu_\alpha r^2} \left[1 + g^{(\alpha)}(r) \right] \right\} \psi_{v,J}(r) = E_{v,J}\, \psi_{v,J}(r) \quad . \tag{1}$$

Here, $V_{\text{ad}}^{(1)}(r)$ is the total electronic internuclear potential for a chosen reference isotopologue (labelled $\alpha = 1$), $\Delta V_{\text{ad}}^{(\alpha)}(r)$ is the difference between the effective adiabatic potential for isotopologue α and that for the reference species ($\alpha = 1$), and $g^{(\alpha)}(r)$ is the centrifugal BOB potential correction function for isotopologue–α with effective reduced mass μ_α. The default definition of μ_α is Watson's "charge-modified reduced mass" [7],

$$\mu_\alpha = \mu_\alpha^{\text{W}} \equiv M_A^{(\alpha)} M_B^{(\alpha)} \Big/ \left(M_A^{(\alpha)} + M_B^{(\alpha)} - \text{CHARGE} \times m_e \right) \quad , \tag{2}$$

in which 'CHARGE' is the net $\pm$(integer) charge on the molecule, m_e is the electron mass, and $M_A^{(\alpha)}$ and $M_B^{(\alpha)}$ are the masses of the neutral atoms A and B forming isotopologue α of species A–B$^{\text{CHARGE}}$. However, other choices are also possible (see Appendix B). Of course, for $\text{CHARGE} = 0$ Eq. (2) becomes the conventional two-body reduced mass for particles of mass $M_A^{(\alpha)}$ and $M_B^{(\alpha)}$.

Each of the quantities $\Delta V_{\text{ad}}^{(\alpha)}(r)$ 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 [7–9, 12], and given specifically by

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

and

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

The quantity $\Delta M_A^{(\alpha)} \equiv M_A^{(\alpha)} - M_A^{(1)}$ is the difference between the isotopic masses of atoms A in isotopologue–α and in the reference isotopologue ($\alpha = 1$). The expressions employed to represent the mass-independent $\widetilde{S}_{\text{ad}}^{A/B}(r)$ and $\tilde{R}_{\text{na}}^{A/B}(r)$ radial functions can be found in §§ 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 $\Lambda \neq 0$ [11], or for doublet splittings of the rotational levels of $^2\Sigma$ states, are presented in §§ 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

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

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

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 isotopologue, $V_{\text{ad}}^{(1)}(r)$, and r_{ref} is a reference distance chosen as the expansion centre for the variables of Eq. (6) (usually $r_{\text{ref}} > r_e$). Most of the early work employing this type of variable fixed $r_{\text{ref}} = r_e$ and $p = q = 1$ [13–15]. However, it was later found that allowing p and q to be greater than 1, and fixing r_{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 [16–22].

The fact that $y_p^{r_e}(r)$ and $y_{\{q/p\}}^{\text{ref}}(r)$ 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 $y_q^{\text{ref}}(r) \propto (r - r_{\text{ref}})$ at distances near their respective expansion centres 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 behaviour 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 ($q = 2, 3, 4, \cdots$) 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 [10, 11, 16, 23–25].

In any case, values of p, q, and r_{ref} must be selected by the user and specified in the main input data file. For the Morse/Long-Range (MLR) potential of §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. [17] and [26].

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 [13], which has the form of a Morse potential [27] in which the exponent coefficient varies with distance. Other functions of this type had been introduced by Coxon and Hajigeorgiou (the "GMO" potential) [28] and by Dulick and co-workers (the "MMO" potential) [29], but because of its simpler form and better extrapolation behaviour, only the EMO function is considered here.

An EMO potential function has the form

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

5

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

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

As discussed in Refs. [11, 23, 24], for cases in which $r_{\mathrm{ref}} = r_e$, an appropriate choice of q (usually ≥ 2) in the definition of $y_q^{\mathrm{ref}}(r)$ 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 centre r_{ref} at some distance greater than r_e [30]. In earlier versions of this code [5, 31] r_{ref} had to be varied manually in the parameter optimization procedure, but in the current code it may also be treated as a free fitting parameter.

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 [13, 14, 25] and a state whose potential function has an additional 'ripple' or incipient shelf [24]. However, the fact that $[\mathfrak{D}_e - V_{\mathrm{EMO}}(r)]$ 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 behaviour being

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

in which the powers m_1, m_2, $\cdots$ etc., are determined by the nature of the atoms into which the given molecular state dissociates [32, 33], 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 behaviour 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 [16–18, 20, 21],

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

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

$$u_{\mathrm{LR}}(r) \; = \; D_{m_1}(r) \, \frac{C_{m_1}}{r^{m_1}} \; + \; D_{m_2}(r) \, \frac{C_{m_2}}{r^{m_2}} \; + \; \dots \; + \; D_{m_{\mathrm{Last}}}(r) \, \frac{C_{m_{\mathrm{Last}}}}{r^{m_{\mathrm{Last}}}} \quad , \tag{11}$$

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 three families of damping functions currently allowed by **dPotFit** are a generalized [26] version of the damping functions of Douketis *et al.* [34],

6

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

a simplified version of Eq. (12) that has the form

$$D_m^{\mathrm{nb}(s)}(r) \;=\; \left(1 \;-\; e^{-\rho\cdot r} \right)^{m+s} \quad, \tag{13}$$

and a generalized version [26] of the Tang-Toennies damping function [35],

$$D_m^{\mathrm{tt}(s)}(r) \;=\; 1 \;-\; e^{-b^{\mathrm{tt}}(s)\cdot(\rho\, r)} \sum_{k=0}^{m-1+s} \frac{[b^{\mathrm{tt}}(s)\cdot(\rho\, r)]^k}{k!} \quad, \tag{14}$$

Values of the system-independent parameters $b^{\mathrm{ds}}(s)$ and $c^{\mathrm{ds}}(s)$ in Eq. (12) and $b^{\mathrm{tt}}(s)$ in Eq. (14) were determined in Ref. [26], 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 behaviour of these functions, which is [26]

$$\lim_{r\to 0} \left\{ D_m^{(s)}(r)/r^m \right\} \;\propto\; r^s \tag{15}$$

for all values of m, for all three families. The original versions of these functional forms corresponded to the cases $s = 0$ for the Douketis *et al.* form [34] and $s = 1$ for the Tang-Toennies form [35]. However, the *ab initio* results of Kreek and Meath suggest that the actual very short-range behaviour of dispersion energy damping functions corresponds to $s = 0$ [36]. Moreover, as the very short-range behaviour of an MLR potential is defined by the quadratic term in Eq. (10), values $s > 0$ are physically *un*acceptable, 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. [26] recommended the use of Eq. (12) with $s = -1$ (i.e., setting input parameter `IVSR` $= -2$), but it is tantalizing to note that setting $s = -1/2$, which is possible only for Eqs. (12) or (13), will yield an MLR potential which at very short range would take on the limiting $1/r$ functional Coulomb behaviour predicted by theory. However, any of these functional forms with $s \leq 0$ will yield sensible short-range behaviour for an MLR PEF.

The constant ρ appearing in Eqs. (12) and (14) is a system-dependent range parameter devised by Douketis *et al.* [34]. For a pair of interacting atoms A and B, they defined it by the combining rule

$$\rho \;=\; \rho^{\mathrm{AB}} \;\equiv\; [2\,\rho^{\mathrm{A}}\,\rho^{\mathrm{B}}/(\rho^{\mathrm{A}} + \rho^{\mathrm{B}})] \quad, \tag{16}$$

in which $\rho^{\mathrm{A/B}}$ are defined in terms of the ratio of the ionization potential of atom A or B to that of atomic hydrogen as $\rho_{\{\mathrm{A/B}\}} = \left(I_p^{\{\mathrm{A/B}\}}/I_p^{\mathrm{H}} \right)^{2/3}$. However, other choices for this range-scaling parameter may be made by the user, and in the case of Eq. (13) it is often treated as a free fitting parameter.

As a final comment regarding the $u_{\mathrm{LR}}(r)$ function that defines the long-range behaviour 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 $\mathrm{Li}(^2P_{1/2}) + \mathrm{Li}_2(^2S_{1/2})$ asymptote, $u_{\mathrm{LR}}(r)$ has been represented by one of the roots of a diagonalization arising from two-state [20] or three-state [22] 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 $u_{\mathrm{LR}}(r)$ to be defined by one of the roots of the 2×2 or 3×3 coupling matrix that defines the long-range behaviour predicted by theory for these states [37, 38]. The ease with which it can treat such special cases further illustrates capabilities of the MLR functional form.

As was discussed in Ref.[6], a defining property of the MLR function is the fact that

$$\lim_{r \to \infty} \beta_{\mathrm{MLR}}(r) \equiv \beta_\infty = \ln\{2\,\mathfrak{D}_e / u_{\mathrm{LR}}(r_e)\} \quad . \tag{17}$$

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

$$\begin{aligned}
V_{\mathrm{MLR}}(r) &\simeq \mathfrak{D}_e - u_{\mathrm{LR}}(r) + \frac{1}{4\mathfrak{D}_e}[u_{\mathrm{LR}}(r)]^2 \\
&\simeq \mathfrak{D}_e - \sum_{i=1}^{\mathrm{Last}} \frac{C_{m_i}}{r^{m_i}} + \frac{1}{4\mathfrak{D}_e}\left(\sum_{i=1}^{\mathrm{Last}} \frac{C_{m_i}}{r^{m_i}}\right)^2 \quad .
\end{aligned} \tag{18}$$

This indicates that although the MLR function does take on the form of Eq. (9) at long range, unless $m_{\mathrm{Last}} < 2\,m_1$, the long-range behaviour of the quadratic term in Eqs. (10) and (18) would change the long-range behaviour specified by Eq. (11). This problem is most serious for cases in which $m_1 = 3$ [20], but should also be considered for systems with larger m_1 values [39]. 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 $C_6^{\mathrm{adj}} = C_6 + (C_3)^2/4\mathfrak{D}_e$ and $C_9^{\mathrm{adj}} = C_3\,C_6^{\mathrm{adj}}/2\mathfrak{D}_e$ for the $m_1 = 3$ case that was presented in Ref. 20.

In order to use the MLR potential form, it is clearly necessary to know appropriate values for the powers m_i [32, 33], 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 often has no significant advantages over the simpler function, whose simpler form would tend to make it more 'robust', especially if there are no data for levels lying near the disssociation limit. 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. [40, 41], because more reasonable long-range extrapolation behaviour would be imposed [16]. 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 [17, 20, 22].

The algebraic form of Eq. (10) means that at sufficiently long range $V_{\mathrm{MLR}}(r)$ always takes on the form of Eq. (18). To achieve this, the exponent coefficient function $\beta(r)$ is required both to approach asymptotically the value β_∞ defined by Eq. (17), 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 **Potfit**. In the first, $\beta(r)$ is written as a constrained polynomial in the variable $y_q^{\mathrm{ref}_q}(r)$, in which the separate variable $y_p^{\mathrm{ref}_p}(r)$ acts as a switching function [10, 11, 16, 20, 23, 42], that is,

$$\beta(r) \; \equiv \; \beta_{\mathrm{PE-MLR}}\left(y_{\{p,q\}}^{\mathrm{ref}_{\{p,q\}}}(r)\right) \; = \; \beta_\infty \, y_p^{\mathrm{ref}_p}(r) \; + \; \left[1 - y_p^{\mathrm{ref}_p}(r)\right] \sum_{i=0}^{N_\beta} \beta_i \, y_q^{\mathrm{ref}_q}(r)^i \quad . \tag{19}$$

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 (usually with $q < p$) in the power-series portion of Eq. (19) can lead to more compact and robust potential functions [20–22, 30]. Note too that in the present case two distinct reference distances have been introduced, ref_p and ref_q, the former being the midpoint of the switching between the polynomial expansion and the long-range tail behaviour of the exponent coefficient function, and the latter the polynomial variable expansion centre. In earlier work a preference for simplicity led these two reference distances to be constrained to always be the same, but a greater appreciation of their distinctly different roles later led them to be treated as separate parameters. Moreover, while the single reference distance 'ref' was a fixed read-in parameter in earlier versions of this code [4, 31], and had to be optimized manually, in the present version the reference distances ref_q and/or ref_q may (in principle) be optmized automatically by the fitting procedure.

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 $u_{\mathrm{LR}}(r)$. The algebraic form of the exponent coefficient function of Eq. (19) implies that at large r the exponential term in the MLR function takes the form $e^{-\beta_\infty}(1 + A/r^p + \ldots)$. This has the effect of adding a term having the form $(A\,C_{m_1})/r^{m_1+p}$ to Eq. (18) [17, 20]. As a consequence, the leading contributions to long-range behaviour of $V_{\mathrm{MLR}}(r)$ will only truly be defined by the specified version of Eq. (11) if the power p defining the exponent variables satisfies the constraint $p > (m_{\mathrm{Last}} - m_1)$, where m_{Last} is the power of the last (i.e., highest-power) term contributing to $u_{\mathrm{LR}}(r)$ [17, 20, 21]. It may also be desirable to set $p = m_{\mathrm{next}} - m_1$, in which m_{next} is the (inverse) power associated with the first long-range term predicted by theory that is *not* included in the chosen definition of $u_{\mathrm{LR}}(r)$.

9

As discussed in Ref. [20], there are no formal restrictions on the choice of the power q defining the radial variable in the power series part of Eq. (19). Experience suggests [20, 22, 43] 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 $p > (m_{\mathrm{Last}} - m_1)$) must also be guided by consideration of the manner in which the potential function approaches its limiting long-range behaviour. In particular, $V_{\mathrm{MLR}}(r)$ always takes on the limiting behaviour of Eq. (9), a rearrangement of which yields

$$C_{m_1}^{\mathrm{eff}}(r) \;\equiv\; r^{m_1}\,[\,\mathfrak{D}_e - V_{\mathrm{MLR}}(r)\,] \;\simeq\; C_{m_1} + \frac{C_{m_2}}{r^{m_2-m_1}} + \cdots \tag{20}$$

Thus, a plot of $C_{m_1}^{\mathrm{eff}}(r)$ 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 this limiting behaviour in a smooth, monotone fashion. Figure 1 shows plots of this type for a number of otherwise equivalent potentials for the ground-state of Br_2 [43], for which the leading terms in the long-range potential correspond to $m = 6, 8,$ and 10 [44]. While all of the six potentials considered here do eventually achieve the predicted linear approach to an intercept of C_6^{theory} with a slope of C_8^{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 $C_6^{\mathrm{eff}}(r)$ plots are *not* accompanied by discernably irregular behaviour in plots of the potential functions $V_{\mathrm{MLR}}(r)$ themselves. This illustrates the practical importance of using plots such as Fig. 1 to detect such implausible extrapolation behaviour! Moreover, the fact that the $C_6^{\mathrm{eff}}(r)$ 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 behaviour for this particular species.

As with the values of p and q, there are no formal constraints on the values of ref_p and ref_q, other than a nominal expectation that they should both be larger than r_e (at least for a single-minimum potential). However, some time ago Young showed [45] that for a data-sensitive region $[r_{\min}, r_{\max}]$, independent of the value of q the value of ref_q for which $|1 - y_q^{\mathrm{ref}_q}(r_{\max})| = |y_q^{\mathrm{ref}_q}(r_{\min}) + 1)|$ was simply the geometric mean distance $\mathrm{ref}_q = \sqrt{r_{\max} \times r_{\min}}$, which makes this seem a plausible initial estimate for this parameter. Unfortunately, there is no analogous plausible preliminary estimate for ref_p, other that that it would normally be larger, and sometimes *much* larger, than ref_q. Hence, when performing fits to a PE-MLR function it is necessary to consider a range values of q, p, $r_{\mathrm{ref}_{q,p}}$, and N_β in a search for an optimal model. Figure 2 illustrates how the quality of fits to data for ground-state Ca_2 depends upon three of these parameters (results obtained with $\mathrm{ref}_q \equiv \mathrm{ref}_q$). 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,
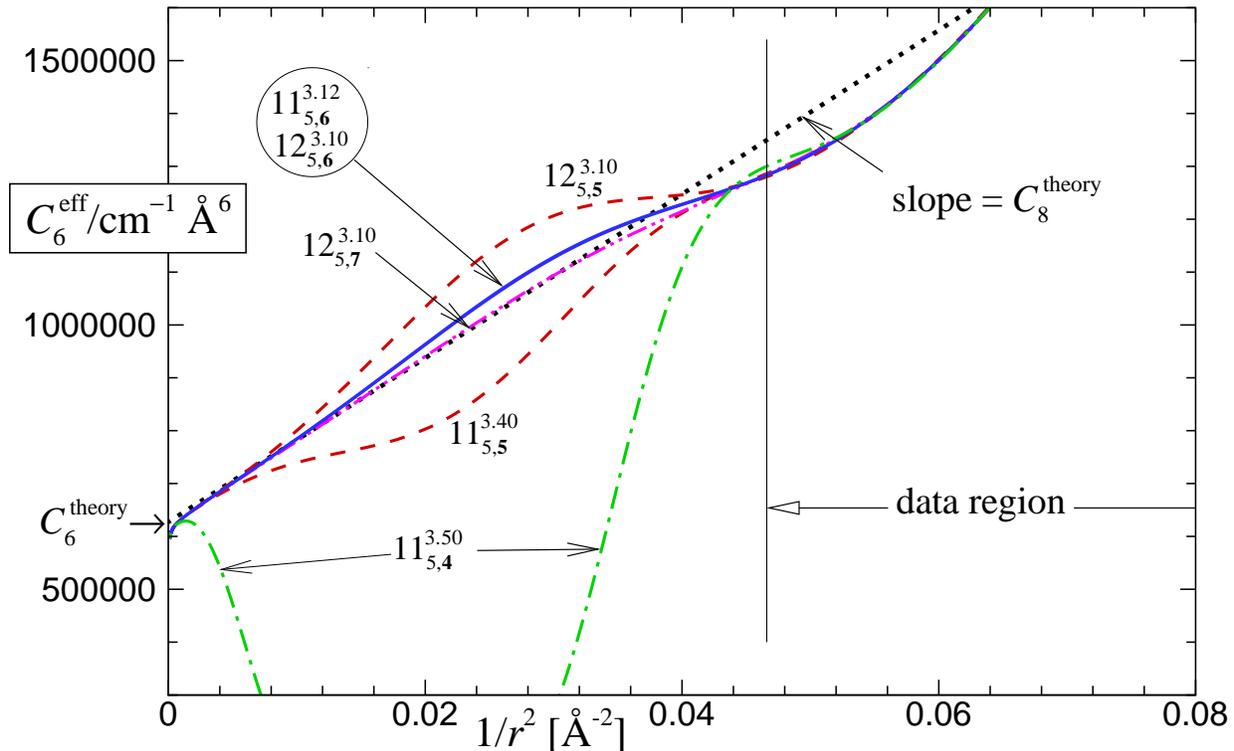
Figure 1. Tests of the long-range extrapolation behaviour of various fitted potentials for the $X\,^1\Sigma_g^+$ state of Br_2 associated with different MLR models $N_{p,q}^{r_{\mathrm{ref}}}$, where $N = N_\beta$ [Adapted from Fig. 7 of Ref. [43]].

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_{ref} over which values of $\overline{dd}$ lie near their minimum is narrower for $\{q = 4,\, N_\beta = 7\}$ (solid square points) than for $\{q = 3,\, N_\beta = 7\}$ (open square 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 $\{q = 4,\, N_\beta = 7,\, r_{\mathrm{ref}} = 6.85\,\text{Å}\}$ 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. [6]. All existing applications of the MLR potential have employed the constrained polynomial expansion of Eq. (19) to represent the exponent coefficient function $\beta(r)$. However, while preliminary work showed some promise [49, 50], 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 co-workers [51–53], 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
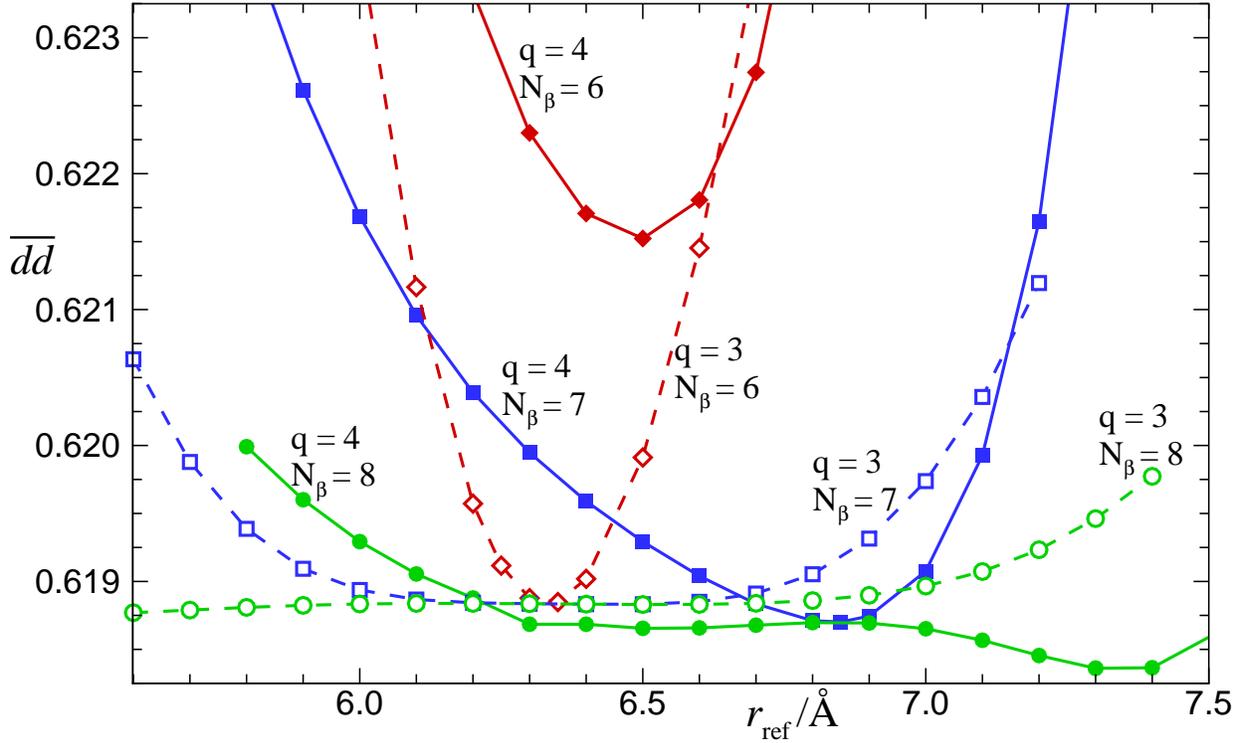
Figure 2. Dependence of the dimensionless root-mean-square deviation $\overline{dd}$ obtained from fits to 3553 fluorescence series data for the $X\,^1\Sigma_g^+$ state of Ca_2 [46, 47] on the parameters r_{ref}, q, and N_β. Displayed here are the results of fits to MLR potentials with $p = 5$ whose long-range tails are defined by Douketis-type damping with $s = -1$ and fixed theoretical values of the C_8, and C_{10} [48] coefficients, but with C_6 allowed to be a fitted parameter [21].

(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 [6, 54] 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 $\beta(r_j) = \beta(y_q^{\mathrm{ref}_q}(r_j))$ values at a specified set of $y_q^{\mathrm{ref}_q}(r_j)$ 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

$$\beta(r) \;=\; \sum_{k=1}^{N_\beta} S_k(y_q^{\mathrm{ref}_q}(r))\,\beta_k \;\;, \tag{21}$$

in which the spline 'basis functions' $S_k(y_p^{\mathrm{ref}_q}(r))$ are completely defined by the chosen mesh of values of $y_q^{\mathrm{ref}_q}(r_j)$. A straightforward application of the chain rule of calculus then yields the partial

derivatives required for the least-squares fit procedure:

$$\frac{\partial V(r)}{\partial \beta_k} = 2\,\mathfrak{D}_e \left\{ 1 - \frac{u_{\mathrm{LR}}(r)}{u_{\mathrm{LR}}(r_e)}\, e^{-\beta(r)\cdot y_p^{r_e}(r)} \right\} \left(\frac{u_{\mathrm{LR}}(r)}{u_{\mathrm{LR}}(r_e)}\, e^{-\beta(r)\cdot y_p^{\mathrm{eq}}(r)} \right) y_p^{r_e}(r)\ S_k(y_q^{\mathrm{ref}_q}(r)) \ . \tag{22}$$

Although the least-squares problem is non-linear, the fact that the $S_k(y_q^{\mathrm{ref})q}(r))$ 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 $y_q^{\mathrm{ref}_q}(r)$ used to define $\beta(r)$ should normally be defined by a value of ref_q that is greater than r_e, and the value of q defining the radial variable in Eq. (21) should be greater than 1, in order to assure that the chosen mesh of points sample the full range of $\beta(y_q^{\mathrm{ref}_q}(r))$ values appropriately. However, in this case *both* q and p (rather than just p) must be set at values greater than $(m_{\mathrm{Last}} - m_1)$ in order to prevent the long-range behaviour of the exponential term from disrupting the long-range behaviour specified by Eq. (11). Note that the switching function variable $y_{[}^{\mathrm{ref}_p}(r))$ does not appear in the SE-MLR form, so the second reference distance ref_p is a redundant parameter.

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 [11, 23],

$$V_{\mathrm{DELR}}(r) \;=\; \left\{ A\,e^{-2\beta(r)\cdot(r-r_e)} \;-\; B\,e^{-\beta(r)\cdot(r-r_e)} \;+\; \mathfrak{D}_e \right\} \;-\; u_{\mathrm{LR}}(r) \ , \tag{23}$$

in which the exponent coefficient $\beta(r)$ is defined as the same type of simple power series in $y_p^{\mathrm{ref}}(r)$ appearing in the EMO potential (see Eq. (8)) [30]. In this case, the presence of a repulsive leading term in the additive long-range function $u_{\mathrm{LR}}(r)$ (which could, in principle, be attractive or repulsive), served to introduce the potential function barrier that was being modelled [11]. The pre-exponential coefficients A and B in Eq. (23) 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

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

$$B \;=\; 2\,\mathfrak{D}_e \;-\; 2\,u_{\mathrm{LR}}(r_e) \;-\; u'_{\mathrm{LR}}(r_e)/\beta_0 \ , \tag{25}$$

in which $u'_{\mathrm{LR}}(r_e) \equiv [du_{\mathrm{LR}}(r)/dr]_{r=r_e}$. If $u_{\mathrm{LR}}(r) = 0$ the DELR potential becomes the EMO function of Eq. (7). However, other choices of $u_{\mathrm{LR}}(r)$ allow it to represent the outer wall of a potential function with a barrier [11], 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 $u_{\mathrm{LR}}(r)$ in Eq. (23) is assumed to be defined either by a sum of damped or undamped inverse-power terms, as in Eq. (11) or, for alkali

13

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 $u_{\mathrm{LR}}(r)$ function (positive coefficients being *attractive*) is opposite to that used in Ref.[11]. As in §3.2, damping functions $D_m(r)$ may be defined by either Eq. (12) or (14), but since u_{LR} is an additive, rather than a multiplicative contribution to the potential, the parameter s defining the very short-range behaviour of these damping functions *is only* allowed to have *positive* values, as negative values will cause the short-range potential wall to turn over. Of course, other damping function expressions [55, 56] or entirely different types of expressions for $V_{\mathrm{LR}}(r)$ 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 [57] that Šurkus *et al.* [58] introduced and called the 'Generalized Potential Energy Function' (GPEF). A modified, but exactly equivalent expression for his expansion variable (devised by Seto [15]) leads to the GPEF potential form

$$V_{\mathrm{GPEF}}(r) \;=\; c_0\, z_q^2 \left[1 \;+\; \sum_{i=1}^{N_\beta} c_i\, z_q^{\,i} \right] \;, \qquad \text{with} \qquad z_q \;\equiv\; \frac{(r^q - r_e^{\,q})}{(a_S\, r^q + b_S\, r_e^{\,q})} \;. \qquad (26)$$

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 [57];
- setting $q = 1$, $a_S = 1$ and $b_S = 0$ yields Simons-Parr-Finlan (SPF) expansions [59];
- setting $q = 1$, $a_S = b_S = 0.5$ yields the Ogilvie-Tipping (OT) potential expansion [60];
- setting $a_S = b_S = 1$ yields the expansion variable $y_p^{\mathrm{eq}}(r)$ 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 behaviour. 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 behaviour of Eq. (9) [58, 61]. 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) [58], 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:

$$
\begin{aligned}
E(v, J) &= G_v + B_v[(J(J+1)] - D_v[(J(J+1)]^2 + H_v[(J(J+1)]^3 + \cdots \\
&= \sum_{m=0} K_m[J(J+1)]^m \quad .
\end{aligned}
\tag{27}
$$

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

$$
\delta E_\Lambda^{(\alpha)}(v, J; \{e/f\}) = \text{sg}(\{e/f\})[J(J+1)]^\Lambda \sum_{m=\Lambda} q_m^{(\alpha)}(v) \, [J(J+1) - \Lambda^2]^{m-\Lambda} \quad ,
\tag{28}
$$

in which the $\text{sg}(\{e/f\})$ 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 §5.3): the choice of which pair of splitting factors to use for the term $\text{sg}(\{e/f\})$ must be specified by the user [5].

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:

$$
\begin{aligned}
\delta E_\Lambda^{(\alpha)}(v, N; \{e\}) &= +\frac{1}{2} N \sum_{m=1} \gamma_m^{(\alpha)}(v) \, [N(N+1)]^{m-1} \quad , \\
\delta E_\Lambda^{(\alpha)}(v, N; \{f\}) &= -\frac{1}{2} (N+1) \sum_{m=1} \gamma_m^{(\alpha)}(v) \, [N(N+1)]^{m-1} \quad .
\end{aligned}
\tag{29}
$$

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 [62, 63]:

$$V_{\mathrm{HPP}}(r) = A_{\mathrm{I}}\, e^{-B_{\mathrm{I}}(r-R_{\mathrm{I}})} \qquad\qquad \text{for} \qquad r < R_{\mathrm{I}} \qquad (30)$$

$$= \{\mathtt{VLIM} - \mathfrak{D}_e\} + V_X(r) \qquad\qquad \text{for} \qquad R_{\mathrm{I}} \le r \le R_{\mathrm{O}}$$

$$= \{\mathtt{VLIM} - \mathfrak{D}_e + V_X(R_{\mathrm{O}}))\} - \sum_{m_1}^{m_{\mathrm{last}}} \frac{C_m}{r^m} + \frac{C_{m_{\mathrm{last}}+2}}{r^{m_{\mathrm{last}}+2}} \qquad \text{for} \qquad r > R_{\mathrm{O}} \qquad (31)$$

in which

$$V_X(r) = \sum_{i=2}^{N_\beta} \beta_i\, X(r)^i \qquad\qquad \text{and} \qquad\qquad X = \frac{r - R_m}{r + b\,R_m} \quad . \qquad (32)$$

Here, R_m is an "arbitrarily chosen" expansion centre 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_{I} and B_{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 $r = R_{\mathrm{I}}$. Similarly, with the other long-range coefficients C_m defined by theory, the value of the coefficient $C_{m_{\mathrm{last}}+2}$ in Eq. (31) is defined by requiring the potential to be continuous at the outer switching point $r = R_{\mathrm{O}}$. 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 behaviour of Eq. (30) and/or the long-range tail of Eq. (31): 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. (14). This structure has been used in a number of published potentials [64–66], and the extended version of this form available in **dPotFit** incorporates most variants found in the recent literature. This generalized form is

$$V_{\mathrm{TT}}(r) \equiv \left(\beta_5 + \beta_6\, r + \frac{\beta_7}{r} + \beta_8\, r^2 + \beta_9\, r^3 \right) \times \qquad\qquad (33)$$

$$e^{-\{\beta_1 r + \beta_2 r^2 + \beta_3/r + \beta_4/r^2\}} - u_{\mathrm{LR}}^{\mathrm{TT}}(r) \quad ,$$

in which the attractive outer-wall function $u_{\mathrm{LR}}^{\mathrm{TT}}(r)$ has the form of Eq. (11) with the damping functions defined by $s = +1$ version of Eq. (14) with its shape parameter fixed at $b^{\mathrm{tt}}(s) = 1$.

16

However, while the 'basic' Tang-Toennies function [35, 67] 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. (33), 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 $m_i \geq 6$, **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. (14) and $\beta_1 - \beta_9$ of Eq. (33), 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 1970's 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.*,

$$V_{\mathrm{HFD}}(r) \equiv A_{\mathrm{HFD}} \, x^\gamma \, e^{-\beta_1 r - \beta_2 r^2} - u_{\mathrm{LR}}^{\mathrm{HFD}}(R) \ . \tag{34}$$

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

$$u_{\mathrm{LR}}^{\mathrm{HFD}}(r) = f_1^{\mathrm{HFD}}(r) \sum_{m=m_1}^{m_{\mathrm{Last}}} \frac{C_m}{r^m} \ , \tag{35}$$

with

$$f_1^{\mathrm{HFD}}(r) = \exp \left\{ -\alpha_1 \left(\frac{\alpha_2 \, r_e}{r} \right)^{\alpha_3} \right\} \qquad \text{for} \qquad r < \alpha_2 \ , \tag{36}$$
$$= 1 \qquad \text{for} \qquad r \geq \alpha_2 \ .$$

Since all of the 'A', 'B' and 'C' variations are incorporated in Eqs. (34)–(36), 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 [68], *viz.*,

$$u_{\mathrm{LR}}^{\mathrm{HFD}}(r) = f_2^{\mathrm{HFD}}(\rho r) \sum_{m=m_1}^{m_{\mathrm{Last}}} D_m(r) \frac{C_m}{r^m} \ , \tag{37}$$

in which

$$f_2^{\mathrm{HFD}}(\rho \, r) = \left\{ 1 - \left(\frac{\rho \, r}{b_{\mathrm{r}}} \right)^{1.68} e^{-0.78 \, (\rho \, r)/b_{\mathrm{r}}} \right\} \ , \tag{38}$$

17

where 'b$_\mathrm{r}$' 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 §3.2. Both this "HFD-D" form of $u_{\mathrm{LR}}(r)$ and the "HFD-ABC" form of Eq. (35) are implemented in **dPotFit**.

For all HFD cases, the present code defines the values of the pre-exponential constant A_{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. (36).

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 Isotopologue' Reference Species

Following the approach of Ref. [10], 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:

$$\widetilde{S}_{\mathrm{ad}}^{\mathrm{A}}(r) \;=\; y_{p_{\mathrm{ad}}}^{r_e}(r)\, u_\infty^{\mathrm{A}} \;+\; [1 - y_{p_{\mathrm{ad}}}^{r_e}(r)] \sum_{i=0}^{N_{\mathrm{ad}}^{\mathrm{A}}} u_i^{\mathrm{A}}\, y_{q_{\mathrm{ad}}}^{r_e}(r)^i \quad, \tag{39}$$

$$\widetilde{R}_{\mathrm{na}}^{\mathrm{A}}(r) \;=\; y_{p_{\mathrm{na}}}^{r_e}(r)\, t_\infty^{\mathrm{A}} \;+\; [1 - y_{p_{\mathrm{na}}}^{r_e}(r)] \sum_{i=0}^{N_{\mathrm{na}}^{A}} t_i^{\mathrm{A}}\, y_{q_{\mathrm{na}}}^{r_e}(r)^i \quad. \tag{40}$$

The structure of Eqs. (39) and (40) allows the asymptotic behaviours 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 $\widetilde{S}_{\mathrm{ad}}^{\mathrm{A}}(r)$ is u_∞^{A}. 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 $u_\infty^{\mathrm{A}} \equiv 0$ for all electronic states that dissociate to yield ground-state atoms, while for a state that dissociates to

18

yield atom A in an excited electronic state, u_∞^A is determined by the associated atomic isotope shift [10, 20]. For all states that dissociate into ground-state atoms, u_0^A and u_0^B define the difference between the well depths of those states for the various atom–A and atom–B isotopologues, such that

$$\delta\mathfrak{D}_e^{(\alpha)}(X) \;=\; \frac{\Delta M_A^{(\alpha)}}{M_A^{(\alpha)}}\, u_0^A \;+\; \frac{\Delta M_B^{(\alpha)}}{M_B^{(\alpha)}}\, u_0^B \quad . \tag{41}$$

Alternatively, if one wishes to define the potential minimum of the ground state as the absolute zero of energy, the values of $u_0^A(X)$ and $u_0^B(X)$ would be fixed at zero, and the values of u_∞^A and u_∞^B would then define the isotopologue dependence of the well depth *via* an expression analogous to Eq. (41). A user of **dPotFit** may select this (not recommended) alternate convention by choosing, in the input data file, to fix $u_0^A(X) = u_0^B(X) = 0$ while allowing u_∞^A and u_∞^B to be varied freely. In either convention, u_0^A and u_0^B would determine the electronic isotope shift for excited states.

The algebraic form of Eq. (39) and its expansion variables means that the limiting long-range behaviour 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 $1/r^{p_{ad}}$ with a coefficient defined by the vagaries of the $u_i^{A/B}$ values obtained from the fit. Thus, for cases in which the reference-isotopologue potential has an inverse-power tail with the form of Eq. (11), one must set $p_{ad} > m_{Last}$ in order to ensure that the limiting asymptotic behaviour of the effective adiabatic potentials $V_{ad}^{\{\alpha\}}(r)$ are the same for all isotopologues [69].

For the few special cases in which differences between the long-range tail coefficients C_m for different isotopologues are known, that knowledge may be incorporated by replacing the constant $u_\infty^{A/B}$ in Eq. (39) by the expansion

$$u_\infty \;-\; \frac{\delta C_{m_1}}{r^{m_1}} \;-\; \frac{\delta C_{m_2}}{r^{m_2}} \;-\; \ldots \;-\; \frac{\delta C_{m_{Last}}}{r^{m_{Last}}} \quad . \tag{42}$$

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

$$\delta C_{m_i}^{A/B} \;\equiv\; \left(\frac{M_{A/B}^{(2)}}{M_{A/B}^{(2)} - M_{A/B}^{(1)}} \right) \left(C_{m_i}^{(2)} - C_{m_i}^{(1)} \right) \quad , \tag{43}$$

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

By analogy with the case for Eq. (39), the algebraic form of Eq. (40) and its expansion variables means its limiting long-range behaviour will be to approach the limiting value of $t_\infty^{A/B}$ with a term proportional to $1/r^{p_{na}}$ whose coefficient is defined the difference between $t_\infty^{A/B}$ and the sum of the

19

$\{t_i^{A/B}\}$ coefficients. However, the absence of any general theoretical knowledge about the long-range behaviour of these functions leads to the recommendation that one should simply fix $p_{na} = q_{na}$.

There are on physical constraints on the values of the integers q_{ad} and q_{na} that define the expansion variables in Eqs. (39) and (40), so they must be selected by the user by trial-and-error ($q = 3, 4, \cdots$ 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. [10]). It may be convenient to set $q_{na} = p_{na}$, and sometimes also to set $q_{ad} = p_{ad}$, in order to obtain expressions for $\widetilde{R}_{ad}^{A/B}(r)$ and $\widetilde{S}_{na}^{A/B}(r)$ that involve only a single type of radial variable [16]. However, too-large values of q_{ad} or q_{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 [10]: thus, case-by-case experimentation is required. Note that since EMO potential energy functions do not have *any* inverse-power limiting long-range behaviour, **dPotFit** internally sets $p_{ad} = q_{ad}$ in Eq. (39) when an EMO function is used for the potential energy.

As was discussed in Ref. [10], the limiting asymptotic value of the *centrifugal* BOB correction function $\widetilde{R}_{na}^A(r)$ should always be $t_\infty^A = 0$, unless the species in question is a molecular ion that yields A^+ or A^- upon dissociation, in which case q_∞^A would have a small non-zero value [10]. For example, for a molecular ion AB^{CHARGE} that dissociates to yield a neutral atom B plus the atomic ion A^{CHARGE}, this limit is

$$t_\infty^A = \frac{\mu_W}{\mu(A^{CHARGE}, B)} - 1 \approx \frac{Q\, m_e}{M_A} + \left(\frac{Q\, m_e}{M_A}\right)^2 + \left(\frac{Q\, m_e}{M_A}\right)^3 + \cdots \,, \qquad (44)$$

in which μ^W is the charge-modified reduced mass of Eq. (2), and $\mu(A^{CHARGE}, B)$ is the usual two-body reduced mass of the ion A^{CHARGE} with the neutral atom B. In practise however, the magnitude of this calculated value of t_∞^A is often negligible relative to the variation of the function $\widetilde{R}_{na}^A(r)$ across the data-sensitive region, so no significant loss of accuracy is introduced by simply fixing $t_\infty^A = 0$, as one does for neutral species [39].

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. (40) at $t_0^A = 0$, as theory shows that this parameter is 100% correlated with the leading adiabatic-correction shape parameter u_1^A, and that it represents an indeterminate integration constant in the theory [7, 8, 10]. However, the value of $g^{(\alpha)}(r \approx r_e)$ is related to observable magnetic properties of the molecule [39, 70–73], so that when measurements of those properties are available, it may be appropriate to fix t_0^A at some specific non-zero value, or even to allow it to be varied in the fit [39]. **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 [9], the formally equivalent alternate parameterization of Watson's original paper [7] has also sometimes been employed, so **dPotFit** allows the user to employ either formulation. In the Watson approach, the effective adiabatic potential $V_{\text{ad}}^{(1)}(r)$ for reference isotopologue $\alpha = 1$ is replaced by the "clamped nuclei" potential $V_{\text{CN}}(r)$ to give the radial equation:

$$\left\{ -\frac{\hbar^2}{2\mu_\alpha}\frac{d^2}{dr^2} \;+\; \left[V_{\text{CN}}(r) + \Delta V_{\text{ad,W}}^{(\alpha)}(r) \right] \right. \tag{45}$$

$$\left. + \frac{[J(J+1) - \Lambda^2]\hbar^2}{2\mu_\alpha\, r^2} \left[1 + g_{\text{W}}^{(\alpha)}(r) \right] \right\} \psi_{v,J}(r) \;=\; E_{v,J}\,\psi_{v,J}(r) \quad,$$

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

$$\Delta V_{\text{ad,W}}^{(\alpha)}(r) \;=\; \frac{m_{\text{e}}}{M_{\text{A}}^{(\alpha)}}\, \widetilde{S}_{\text{ad,W}}^{\text{A}}(r) \;+\; \frac{m_{\text{e}}}{M_{\text{B}}^{(\alpha)}}\, \widetilde{S}_{\text{ad,W}}^{\text{B}}(r) \quad, \tag{46}$$

$$g_{\text{W}}^{(\alpha)}(r) \;=\; \frac{m_{\text{e}}}{M_{\text{A}}^{(\alpha)}}\, \widetilde{R}_{\text{na,W}}^{\text{A}}(r) \;+\; \frac{m_{\text{e}}}{M_{\text{B}}^{(\alpha)}}\, \widetilde{R}_{\text{na,W}}^{\text{B}}(r) \quad. \tag{47}$$

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 $\widetilde{S}_{\text{ad}}^{\text{A/B}}(r)$ and $\widetilde{R}_{\text{na}}^{\text{A/B}}(r)$ are expanded as in Eqs. (39) and (40), but the magnitudes of the expansion parameters will differ by the factors $\{\Delta M_{\text{A/B}}^{(\alpha)}/m_{\text{e}}\}$ for $S_{\text{ad,W}}^{\text{A/B}}(r)$ and $\{M_{\text{A/B}}^{(1)}/m_{\text{e}}\}$ for $R_{\text{ad,W}}^{\text{A/B}}(r)$.

5.3 Λ-Doubling Splittings for Singlet States

It was shown in Ref. [11] 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

$$\left\{ -\frac{\hbar^2}{2\mu_\alpha}\frac{d^2}{dr^2} \;+\; \left[V_{\text{ad}}^{(1)}(r) + \Delta V_{\text{ad}}^{(\alpha)}(r) \right] \right. \tag{48}$$

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

$$\left. + \text{sg}_\Lambda(e/f)\, \Delta V_\Lambda^{(\alpha)}(r)\, [J(J+1)]^\Lambda \right\} \psi_{v,J}(r) \;=\; E_{v,J}\,\psi_{v,J}(r) \quad,$$

in which $\text{sg}_\Lambda(e/f)$ 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 [74],

$$\Delta V_\Lambda^{(\alpha)}(r) \;=\; \left(\frac{\hbar^2}{2\mu_\alpha\, r^2} \right)^{2\Lambda} f_\Lambda(r) \quad. \tag{49}$$

No general theoretical predictions regarding the long-range behaviour 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. (49). The mass-independent radial function $f_\Lambda(r)$ is therefore written as a simple polynomial expansion in the reduced variable $y_{p_\Sigma}^{re}(r)$ of Eq. (5), that is,

$$f_\Lambda(r) \;=\; \sum_{i=0}^{N_\Lambda} w_i^\Lambda \, y_{q_\Lambda}^{re}(r)^i \;\;, \tag{50}$$

in which the expansion coefficients w_i^Λ have units $1/(\mathrm{cm}^{-1})^{2\Lambda-1}$ [75]. The fact that $y_q^{re}(r) \to 1$ 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. **dPot-Fit** 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 [11] yields the following effective radial Hamiltonian for an electronic state with $^2\Sigma$ symmetry:

$$\left\{ -\frac{\hbar^2}{2\mu_\alpha}\frac{d^2}{dr^2} \;+\; \left[V_{\mathrm{ad}}^{(1)}(r) + \Delta V_{\mathrm{ad}}^{(\alpha)}(r) \right] \right. \tag{51}$$

$$+ \frac{N(N+1)\,\hbar^2}{2\mu_\alpha\,r^2}\left[1 + g^{(\alpha)}(r) \right] \tag{52}$$

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

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

$$\Delta V_\Sigma^{(\alpha)} \;=\; \left(\frac{\hbar^2}{2\mu_\alpha\,r^2} \right)\, f_\Sigma(r) \;\;, \tag{53}$$

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

$$f_\Sigma(r) \;=\; \sum_{i=0} w_i^\Sigma \, y_{q_\Sigma}^{re}(r)^i \;\;, \tag{54}$$

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 are 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), (45), (48) or (51) many hundreds or thousands or tens of thousands of times. In particular, in each cycle of the iterative non-linear fit, 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

$$\frac{\partial E_{v,J}}{\partial p_j} \;=\; \left\langle \psi_{v,J}(r) \left| \frac{\partial \widehat{\mathcal{H}}}{\partial p_j} \right| \psi_{v,J}(r) \right\rangle \;. \tag{55}$$

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 [76–80]. The present version of that routine incorporates several unique features, such as the ability to locate quasibound (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, 11, 81, 82]. This last capability is required for cases in which measured tunneling predissociation level widths are included in the experimental data set being analyzed [11]. Most details and features of the Schrödinger-solver routine SCHRQ used by **dPotFit** are described in Ref. [4] and the the online manual for earlier versions of that program [83], 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 Eq. (1), (45), (48) or (51). 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 ~ 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

$$\mathtt{RH} \;=\; \pi / \left(\mathtt{NPN} \times [\mu \times \max\{E - V(r)\}/\mathbb{C}_{\mathrm{u}}]^{1/2} \right) \tag{56}$$

in which $\mathtt{NPN}$ 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 $\mathbb{C}_{\mathrm{u}} \equiv \hbar^2/2 = 16.857\,629\,206\;[\mathrm{amu} \cdot \text{Å}^2 \cdot \mathrm{cm}^{-1}]$. A value of $\mathtt{NPN}$ 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. (56) may be a useful starting point, a careful user *should always* examine the effect of different $\mathtt{RH}$ values on the calculated band constants written to output 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 $\mathtt{RMIN}$ to $\mathtt{RMAX}$ (see READ statement #11) using the Numerov algorithm [76, 84]. These bounds must lie sufficiently far into the inner and outer classically-forbidden regions (for which $V_{\mathrm{eff}}(r) > E_{v,J}$) 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 $\mathtt{RMIN}$ or larger $\mathtt{RMAX}$ value should be used to remove them and to ensure that the desired accuracy is achieved. However, if $\mathtt{RMIN}$ 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 $\mathtt{RMIN}$ would cause such warning messages to disappear and (marginally) reduce the computational effort. For most diatomic molecules, a reasonable value of $\mathtt{RMIN}$ 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 $\mathtt{RMAX}$ or the largest distance consistent with the specified mesh and the internally-defined (see §8.2) potential energy array dimension. As with $\mathtt{RMIN}$, the choice of $\mathtt{RMAX}$ 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 $\mathtt{RMAX}$ are much larger for highly excited vibrational levels than for those lying near the potential minimum. Moreover, for quasibound levels, $\mathtt{RMAX}$ 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 [85]. "Pressure" second virial coefficients are computed semi-classically as the sum of the classical term and the first and second (translational) quantum corrections [86–88], that is,

$$B_2(T) \;=\; B_{\rm cl}(T) \;+\; \left(\frac{\hbar^2}{2\mu}\right) B_{\rm Q}^{\rm I}(T) \;+\; \left(\frac{\hbar^2}{2\mu}\right)^2 B_{\rm Q}^{\rm II}(T) \;+\; \cdots \;. \tag{57}$$

The classical term may be written as

$$B_{\rm cl}(T) \;=\; -\,2\pi\,N_{\rm A}\,\int_0^\infty \left[e^{-V(r)/k_{\rm B}T} - 1\right] r^2\,dr \;, \tag{58}$$

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

$$B_{\rm Q}^{\rm I}(T) \;=\; \pi N_{\rm A}\left(\frac{1}{6\,(k_{\rm B}T)^3}\right)\int_0^\infty e^{-V(r)/k_{\rm B}T}\left(\frac{dV(r)}{dr}\right)^2 r^2\,dr \;, \tag{59}$$

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

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

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 behaviour of a gas from the 'ideal gas Law' $PV = nRT$, so the "acoustic" second virial coefficients provide analogous explanations for non-ideal-gas behaviour 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 [89, 90], 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. [88]), which means that it may be expressed in terms of a sum of integrals analogous to those of Eqs. (58)–(60).

All of Eqs. (58)-(60) 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 $y_{p_{\mathrm{vir}}}(r; r_e) \in [-1, +1]$ where $y_{p_{\mathrm{vir}}}^{r_e}(r)$ 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 $p_{\mathrm{vir}} = 1$. In any case, standard formulae suggest that the errors in the the resulting virial coefficients should decrease by a factor of $1/2^{16}$ with each stage of bisection (see § 25.4.30 of Ref. [91].).

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 isotopologue $V_{\mathrm{ad}}^{(1)}(r)$ has been obtained.

For the potential energy function $V_{\mathrm{ad}}^{(1)}(r)$ (or $V_{\mathrm{CN}}(r)$) 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 [5], 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 [5] to determine analytic level energy expressions, and then use the resulting functions to generate a pointwise RKR potential for that state [92, 93]. 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 [6, 94].

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 isotopologue. 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 [31]. 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 [cm3/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), (45), (48) or (51). This numerical constant is based on the 2010 CODATA recommended physical constant values [95], while the atomic isotope masses stored in subroutine MASSES were taken from the 2012 compilation of Ref. [96].

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 $\{y_i^{\text{calc}}\}$ is characterized by the value of the dimensionless root mean square deviation $\overline{dd}$, defined as

$$\text{DRMSD} \equiv \overline{dd} = \left\{ \frac{1}{N} \sum_{i=1}^{N} \left[\frac{y_i^{\text{calc}} - y_i^{\text{obs}}}{u_i} \right]^2 \right\}^{1/2}, \tag{61}$$

or by the related dimensionless standard error σ_f, defined as $\text{DSE} \equiv \bar{\sigma}_f = \overline{dd} \sqrt{N/(N-M)}$. 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. [97] by also always listing the associated "parameter sensitivity" (identified as PS in the output). This quantity is defined (see Eq. (4) of Ref. [97]) 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 $\bar{\sigma}_f$ by a maximum of $(0.1/M)\,\bar{\sigma}_f$. 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. [97], to three significant digits, rounding off all parameters at the first significant digit of their sensitivity had no meaningful effect on the values of $\bar{\sigma}_f$ 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. [97], 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_{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. [97], 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 $|\mathtt{IROUND}| > 1$ 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. [98] 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. [98] and replaces the normal least-squares data weights $w_i = 1/(u_i)^2$ by the '*robust*' weights $w_i^{\mathrm{rob}} = 1/\left[(u_i)^2 + (y_i^{\mathrm{calc}} - y_i^{\mathrm{obs}})^2/3\right]$. 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 $[y_i^{\mathrm{calc}} - y_i^{\mathrm{obs}}]$ 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 $[y_i^{\text{calc}} - y_i^{\text{obs}}]$ 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 $[y_i^{\text{calc}} - y_i^{\text{obs}}]/u_i$ lies between 2 and 4 is followed 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 isotopologues 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 $[y_i^{\text{calc}} - y_i^{\text{obs}}]$ results.

Channel–7 output contains values of the band constants for all levels of all isotopologues 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 $\{[y_i^{\text{calc}} - y_i^{\text{obs}}]\}$ and $\{[y_i^{\text{calc}} - y_i^{\text{obs}}]/u_i\}$ 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 isotopologue $V_{\text{ad}}^{(1)}(r)$ (or $V_{\text{CN}}(r)$),
>
> **Channels–12 & 13** contain the adiabatic BOB radial strength functions $\widetilde{S}_{\text{ad}}^{\text{A}}(r)$ and $\widetilde{S}_{\text{ad}}^{\text{B}}(r)$, respectively,
>
> **Channels–14 & 15** contain the centrifugal BOB radial strength functions $\widetilde{R}_{\text{na}}^{\text{A}}(r)$ and $\widetilde{R}_{\text{na}}^{\text{B}}(r)$, 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:

```sh
#!/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 $\{y_i^{\text{calc}}\}$ 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`.

9. Supplementary Material

Supplementary material associated with this paper consist 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 [99–108]

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".

References

[1] R. Rydberg, Z. Physik **73**, 376 (1931).

[2] O. Klein, Z. Physik **76**, 226 (1932).

[3] A. L. G. Rees, Proc. Phys. Soc. (London) **59**, 998 (1947).

[4] R. J. Le Roy, J. Quant. Spectrosc. Radiat. Transfer **186**, 179 (2016), **dPotFit :** A Computer Program to Fit Diatomic Molecule Spectral Data to Potential Energy Functions: [http//dx.doi.org.10.1016/j.jqsrt.2016.06.002].

[5] R. J. Le Roy and A. Pashov, J. Quant. Spectrosc. Radiat. Transfer **186**, 210 (2016), **betaFIT:** A Computer Program to Fit Pointwise Potentials to Selected Analytic Functions: [http//dx.doi.org.10.1016/j.jqsrt.2016.03.036].

[6] R. J. Le Roy, J. Quant. Spectrosc. Radiat. Transfer **186**, 167 (2016), **LEVEL :** A Computer Program for Solving the Radial Schrödinger Equation for Bound and Quasibound Levels.

[7] J. K. G. Watson, J. Mol. Spectrosc. **80**, 411 (1980).

[8] J. K. G. Watson, J. Mol. Spectrosc. **223**, 39 (2004).

[9] R. J. Le Roy, J. Mol. Spectrosc. **194**, 189 (1999).

[10] R. J. Le Roy and Y. Huang, J. Mol. Struct. (Theochem) **591**, 175 (2002).

[11] Y. Huang and R. J. Le Roy, J. Chem. Phys. **119**, 7398 (2003), erratum: *ibid* **126** 169904 (2007).

[12] J. F. Ogilvie, J. Phys. B: At. Mol. Opt. Phys. **27**, 47 (1994).

[13] E. G. Lee, J. Y. Seto, T. Hirao, P. F. Bernath, and R. J. Le Roy, J. Mol. Spectrosc. **194**, 197 (1999).

[14] J. Y. Seto, Z. Morbi, F. Charron, S. K. Lee, P. F. Bernath, and R. J. Le Roy, J. Chem. Phys. **110**, 11756 (1999).

[15] J. Y. Seto, *Direct Fitting of Analytic Potential Functions to Diatomic Molecule Spectroscopic Data*, M.Sc. Thesis, Department of Chemistry, University of Waterloo (2000).

[16] R. J. Le Roy, Y. Huang, and C. Jary, J. Chem. Phys. **125**, 164310 (2006).

[17] R. J. Le Roy and R. D. E. Henderson, Mol. Phys. **105**, 663 (2007).

[18] H. Salami, A. J. Ross, P. Crozet, W. Jastrzebski, P. Kowalczyk, and R. J. Le Roy, J. Chem. Phys. **126**, 194313 (2007).

[19] A. Shayesteh, R. D. E. Henderson, R. J. Le Roy, and P. F. Bernath, J. Phys. Chem. A **111**, 12495 (2007).

[20] R. J. Le Roy, N. Dattani, J. A. Coxon, A. J. Ross, P. Crozet, and C. Linton, J. Chem. Phys. **131**, 204309 (2009).

[21] R. J. Le Roy, *Determining Equilibrium Structures and Potential Energy Functions for Diatomic Molecules*, Chapter 6, pp. 159-203, of *Equilibrium Structures of Molecules*, J. Demaison and A. G. Csaszar editors, Taylor & Francis, London (2011).

[22] N. Dattani and R. J. Le Roy, J. Mol. Spectrosc. **268**, 199 (2011).

[23] Y. Huang, *Determining Analytical Potential Energy Functions of Diatomic Molecules by Direct Fitting*, M.Sc. Thesis, Department of Chemistry, University of Waterloo (2001).

[24] R. J. Le Roy, D. R. T. Appadoo, K. Anderson, A. Shayesteh, I. E. Gordon, and P. F. Bernath, J. Chem. Phys. **123**, 204304 (2005).

[25] R. J. Le Roy, D. R. T. Appadoo, R. Colin, and P. F. Bernath, J. Mol. Spectrosc. **236**, 178 (2006).

[26] R. J. Le Roy, C. C. Haugen, J. Tao, and H. Li, Mol. Phys. **109**, 435 (2011).

[27] P. M. Morse, Phys. Rev. **34**, 57 (1929).

[28] J. A. Coxon and P. G. Hajigeorgiou, J. Mol. Spectrosc. **139**, 84 (1990).

[29] H. G. Hedderich, M. Dulick, and P. F. Bernath, J. Chem. Phys. **99**, 8363 (1993).

[30] In early work [11, 31], 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 $y_q^{\text{ref}}(r)$ as the expansion variable instead of $y_q^{r_e}(r)$ has made this complication unnecessary, and has enabled the discontinuities in its higher-derivative at $r = r_e$ to be eliminated.

[31] R. J. Le Roy, J. Seto, and Y. Huang, (a) `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); *ibid* `DPotFit 1.2` CP-664 (2007); (c) *ibid*, `DPotFit 1.1`, CP-662R (2006); see also `http://leroy.uwaterloo.ca/programs/`.

[32] R. J. Le Roy and R. B. Bernstein, J. Chem. Phys. **52**, 3869 (1970).

[33] R. J. Le Roy, in *Molecular Spectroscopy*, edited by R. N. Barrow, D. A. Long, and D. J. Millen (Chemical Society of London, London, 1973), Vol. 1, Specialist Periodical Report 3, pp. 113–176.

[34] C. Douketis, G. Scoles, S. Marchetti, M. Zen, and A. J. Thakkar, J. Chem. Phys. **76**, 3057 (1982).

[35] K. T. Tang and J. P. Toennies, J. Chem. Phys. **80**, 3726 (1984).

[36] H. Kreek and W. J. Meath, J. Chem. Phys. **50**, 2289 (1969).

[37] F. Martin, M. Aubert-Frécon, R. Bacis, P. Crozet, C. Linton, S. Magnier, A. Ross, and I. Russier, Phys. Rev. A **55**, 3458 (1997).

[38] M. Aubert-Frécon, G. Hadinger, S. Magnier, and S. Rousseau, J. Mol. Spectrosc. **188**, 182 (1998).

[39] Y.-S. Cho and R. J. Le Roy, J. Chem. Phys. **144**, 024311 (2016).

[40] P. G. Hajigeorgiou and R. J. Le Roy, in *49th Ohio State University International Symposium on Molecular Spectroscopy* (Columbus, Ohio, 1994), paper WE04.

[41] P. G. Hajigeorgiou and R. J. Le Roy, J. Chem. Phys. **112**, 3949 (2000).

[42] In earlier versions of the model an exponential switching function was used [40, 41, 109–112], but that form involves greater complexity and has no significant advantages.

[43] T. Yukiya, N. Nishimiya, Y. Samajima, K. Yamaguchi, M. Suzuki, C. D. Boone, I. Ozier, and R. J. Le Roy, J. Mol. Spectrosc. **283**, 32 (2013).

[44] There is also a very weak repulsive $m = 5$ term for this state [113], but it has no discernable effect outside the immediate neighbourhood of the intercept [43].

[45] M. Young, unpublished notes (2005).

[46] O. Allard, A. Pashov, H. Knöckel, and E. Tiemann, Phys. Rev. A **66**, 042503:1 (2002).

[47] O. Allard, C. Samuelis, A. Pashov, H. Knöckel, and E. Tiemann, Eur. Phys. J. D **26**, 155 (2003).

[48] J. Mitroy and J.-Y. Zhang, J. Chem. Phys. **128**, 134305 (2008).

[49] Y. Huang and R. J. Le Roy, in *55th Ohio State University International Symposium on Molecular Spectroscopy* (Columbus, Ohio, 2000), paper RA01.

[50] J. Y. Seto and R. J. Le Roy, in *55th Ohio State University International Symposium on Molecular Spectroscopy* (Columbus, Ohio, 2000), paper RA02.

[51] A. Pashov, W. Jastrzębski, and P. Kowalczyk, Comp. Phys. Comm. **128**, 622 (2000).

[52] A. Pashov, W. Jastrzębski, and P. Kowalczyk, J. Chem. Phys. **113**, 6624 (2000).

[53] A. Pashov, W. Jastrzębski, W. Jaśniecki, V. Bednarska, and P. Kowalczyk, J. Mol. Spectrosc. **203**, 264 (2000).

[54] J. Tao, R. J. Le Roy, and A. Pashov, in *65th Ohio State University International Symposium on*

Molecular Spectroscopy (Columbus, Ohio, 2010), paper MG07.

[55] A. Koide, W. J. Meath, and A. R. Allnatt, Chem. Phys. **58**, 105 (1981).

[56] C. Douketis, J. M. Hutson, B. J. Orr, and G. Scoles, Mol. Phys. **52**, 763 (1984).

[57] J. L. Dunham, Phys. Rev. **41**, 721 (1932), Applies 3'rd order WKB quantization condition to vibrating rotor to get energy expansion.

[58] A. A. Šurkus, R. J. Rakauskas, and A. B. Bolotin, Chem. Phys. Lett. **105**, 291 (1984).

[59] G. Simons, R. G. Parr, and J. M. Finlan, J. Chem. Phys. **59**, 3229 (1973).

[60] J. F. Ogilvie, Proc. Roy. Soc. (London) A **378**, 287 (1981).

[61] J. F. Ogilvie, J. Chem. Phys. **88**, 2804 (1988).

[62] H. Knöckel, B. Bodermann, and E. Tiemann, Eur. Phys. J. D **28**, 199 (2004).

[63] E. J. Salumbides, K. S. E. Eikema, W. Ubachs, U. Hollenstein, H. Knöckel, and E. Tiemann, Eur. Phys. J. D **47**, 171 (2008).

[64] R. Hellmann, E. Bich, and E. Vogel, Mol. Phys. **106**, 133 (2008).

[65] K. Patkowski and K. Szalewicz, J. Chem. Phys. **133**, 094304 (2010).

[66] J. M. Waldrop, B. Song, K. Patkowski, and X. Wang, J. Chem. Phys. **142**, 204307 (2015).

[67] K. T. Tang and J. P. Toennies, J. Chem. Phys. **118**, 4976 (2003).

[68] R. A. Aziz, J. Chem. Phys. **99**, 4518 (1993).

[69] On reconsideration, our previous recommendation [31] simply to set $p_{\mathrm{ad}} \geq m_1$, the power of the leading inverse-power term in Eqs. (11) is not sufficiently restrictive.

[70] J. F. Ogilvie, *The Vibrational and Rotational Spectrometry of Diatomic Molecules* (Academic Press, Toronto, 1998).

[71] R. M. Herman and J. F. Ogilvie, Adv. Chem. Phys. **103**, 187 (1998).

[72] S. P. A. Sauer, Chem. Phys. Lett. **297**, 475 (1998).

[73] S. P. A. Sauer and V. Špirko, J. Chem. Phys. **138**, 024315 (2013).

[74] Note that while the derivation of Ref. [11] only addressed the case of Λ-doubling in Π states, the present implementation has been extended to handle states for which $\Lambda > 1$.

[75] This assumes, of course, that the factor $\hbar^2/(2\mu_\alpha r^2)$ has units cm^{-1}.

[76] J. W. Cooley, Math. Computations **15**, 363 (1961).

[77] J. K. Cashion, J. Mol. Spectrosc. **10**, 182 (1963).

[78] R. N. Zare and J. K. Cashion, *The IBM Share Program D2 NU SCHR 1072 for Solution of the Schrödinger Radial Equation, by J.W. Cooley: Necessary and Useful Modifications for its Use on an IBM 7090*, University of California Lawrence Radiation Laboratory Report UCRL-10881 (1963).

[79] R. N. Zare, University of California Lawrence Radiation Laboratory Report UCRL-10925 (1963).

[80] R. N. Zare, J. Chem. Phys. **40**, 1934 (1964).

[81] R. J. Le Roy and R. B. Bernstein, J. Chem. Phys. **54**, 5114 (1971).

[82] R. J. Le Roy and W.-K. Liu, J. Chem. Phys. **69**, 3622 (1978).

[83] a) R. J. Le Roy, LEVEL 1.0, University of Wisconsin Theoretical Chemistry Institute Report WIS-TCI-429G (1971). b) R. J. Le Roy, LEVEL 2.0 University of Waterloo Chemical Physics Research Report CP-58 (1976). c) *ibid*, LEVEL 3.0, CP-110 (1979). d) *ibid*, LEVEL 4.0, CP-230 (1983). e) *ibid*, LEVEL 5.0, CP-330 (1991). f) *ibid, LEVEL 5.1*, CP-330R (1992). g) *ibid, LEVEL 5.2* CP-330R^2 (1993). h) *ibid*, LEVEL 5.3, CP-330R^3 (1994). i) *ibid*, LEVEL 6.0, CP-555 (1995). j) *ibid*, LEVEL 6.1, CP-555R (1996) k) *ibid*, LEVEL $7.0 - 7.7$, CP-$642 - 642$R^3, CP-655 & CP-661 $(2000 - 2005)$; *ibid* LEVEL 8.0, CP-663 $(2007 - 2014)$; ℓ) *ibid* LEVEL 8.2, CP-668 (May, 2014); see `http://leroy.uwaterloo.ca/programs/`.

[84] K. Smith, *The Calculation of Atomic Collision Processes* (Wiley-Interscience, New York, 1971), Chapter 4.

[85] L. Piticco, F. Merkt, A. A. Cholewinski, F. R. W. McCourt, and R. J. Le Roy, J. Mol. Spectrosc. **264**, 83 (2010).

[86] J. O. Hirschfelder, C. F. Curtiss, and R. B. Bird, *Molecular Theory of Gases and Liquids* (Wiley, New York, 1964).

[87] E. Mason and T. Spurling, in *International Encyclopedia of Physical Chemistry and Chemical Physics, Topic 10, "The Fluid State"*, edited by J. Rowlinson (Pergamon Press, Oxford, 1969), Vol. 2, Chap. 2, pp. 17–66.

[88] F. R. W. McCourt, *Virial Coefficients*, Chapter 27, pp. 673-711 of "Handbook of Molecular Physics and Quantum Chemistry", S. Wilson Editor, Volume 3: *Molecular Physics in the Physicochemical Environment: Spectroscopy, Dynamics and Bulk Properties*, P. F. Bernath, Volume Editor, Wiley (2003).

[89] J. P. M. Trusler, *Physical Acoustics and Metrology of Fluids* (Adam Hilger, Bristol, 1991).

[90] S. J. Boyes, Chem. Phys. Lett. **221**, 467 (1994).

[91] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions* (Dover, New York, 1970), ninth Printing.

[92] R. J. Le Roy, **186**, 158 (2016), **RKR1**: A Computer Program Implementing the First-Order RKR Method for Determining Diatomic Molecule Potential Energy Functions: [http//dx.doi.org.10.1016/j.jqsrt.2016.03.030].

[93] R. J. Le Roy, 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: *ibid*, CP-657 (2003).

[94] R. J. Le Roy, (a) R. J. Le Roy. betaFIT 2.1, University of Waterloo Chemical Physics Research Report CP-666 (2013); (b) *ibid*, betaFIT 2.0, CP-665 (2009); (b) *ibid*, phiFIT 1.1, CP-663R (2006); (c) *ibid*, phiFIT 1.0, CP-663 (2006); see also http://leroy.uwaterloo.ca/programs/.

[95] P. Mohr, B. Taylor, and D. Newell, Rev. Mod. Phys. **84**, 1527 (2012).

[96] M. Wang, G. Audi, A. Wapstra, F. Kondev, M. MacCormick, X. Xu, and B. Pfeiffer, Chin. Phys. C **36**, 1603 (2012).

[97] R. J. Le Roy, J. Mol. Spectrosc. **191**, 223 (1998).

[98] J. K. G. Watson, J. Mol. Spectrosc. **219**, 326 (2003).

[99] A. Carrington and P. J. Sarre, J. Phys. (Paris) **40**, C1 (1979).

[100] P. C. Cosby, H. Helm, and J. T. Moseley, ApJ **235**, 52 (1980).

[101] H. Helm, P. C. Cosby, M. M. Graff, and J. T. Moseley, Phys. Rev. A **25**, 304 (1982).

[102] W. Jäger, Y. Xu, and M. Gerry, J. Chem. Phys. **99**, 919 (1993).

[103] B. Schramm, H. Schmiedel, R. Gehrmann, and R. Bartl, Ber. Bunsenges. Phys. Chem. **81**, 316 (1977).

[104] H. P. Rentschler and B. Schramm, Ber. Bunsenges. Phys. Chem. **81**, 319 (1977).

[105] J. H. Dymond, K.N. Marsh, R.C. Wilhoit, and K.C. Wong, in *Landolt-Börstein: Numerical Data and Functional Relationships in Science and Technology*, New Series, Group IV: *Physical Chemistry*, Vol. 21, M. Frenkel and K.N. Marsh, Eds. (2003).

[106] W. Gunton, M. Semczuk, N. S. Dattani, and K. W. Madison, Phys. Rev. A **88**, 062510 (2013).

[107] M. Semczuk, X. Li, W. Gunton, M. Haw, N. S. Dattani, J. Witz, A. K. Mills, D. J. Jones, and K. W. Madison, Phys. Rev. A **87**, 052505 (2013).

[108] R. D. E. Henderson, A. Shayesteh, J. Tao, C. C. Haugen, P. F. Bernath, and R. J. Le Roy, J. Phys. Chem. A **117**, 13373 (2013).

[109] J. A. Coxon and R. Colin, J. Mol. Spectrosc. **181**, 215 (1997).

[110] J. A. Coxon and P. G. Hajigeorgiou, J. Mol. Spectrosc. **193**, 306 (1999).

[111] J. A. Coxon and P. G. Hajigeorgiou, J. Mol. Spectrosc. **203**, 49 (2000).

[112] J. Y. Seto, R. J. Le Roy, J. Vergès, and C. Amiot, J. Chem. Phys. **113**, 3067 (2000).

[113] M. Saute, M. Bussery, and M. Aubert-Frécon, Mol. Phys. **51**, 1459 (1984).

Supplementary Material

This document contains Supplementary Material associated with the paper "**dPotFit**: A Computer Program to Fit Diatomic Molecule Spectral Data to Potential Energy Functions", submitted to the *Journal of Quantitative Spectroscopy and Radiative Transfer* in February 2016. It consists of the five Appendices enumerated below. Note that Equation and Reference numbering appearing herein refer to the equation and reference numbering in the Journal Article.

Appendix A: Experimental Data File Structure and Control Parameters

The experimental data are read from a file whose name is specified in the regular Channel–5 input data file *via* READ #2. The data must be arranged as separate vibrational bands (or fluorescence series, or sets of photo-association spectroscopy (PAS) binding energies, tunneling predissociation level widths, or potential function values), each characterized by the upper- and lower-state vibrational quantum numbers $v' = $ VP and $v'' = $ VPP, respectively, by the three-alphanumeric-character labels LABLP and LABLPP (enclosed between single quotes; e.g. 'X1S') that have been chosen to label the particular upper and lower electronic states, by the (integer) mass numbers MN1 and MN2 of the atoms forming that particular isotopologue, and by a label indicating the source of the data in this band/group, consisting of up to 30 alphanumeric characters (e.g,. 'Linton, JMS 151, 159 (1992)'). Note that the electronic state labels must correspond to names used to identify the different electronic states in READ #6 of the Channel–5 input (see below).

For each such band, the data are read, one per line, with each datum consisting of the upper and lower rotational quantum numbers, $J' = $ JP and $J'' = $ JPP, respectively, the integer $+1$ (for e–parity levels) or -1 (for f–parity levels) defining the e/f parity of the upper ($p' = $ EFP) and lower ($p'' = $ EFPP) levels, the experimental datum $y_i^{\text{obs}} = $ FREQ(i) and its uncertainty $u_i = $ UFREQ(i). A datum line with JP < 0 signals the end of the data set for this band or group, and asks the program to start the input for a new band. The overall data input is assumed to be complete either at the end of the data file, or when a negative value of the band-label quantum number $v' = $ VP is encountered.

```
       IBAND= 0
       COUNT= 1
    10 IBAND= IBAND+ 1
 #1    READ(4,*,END=20)  VP(IBAND), VPP(IBAND), LABLP, LABLPP, MN1, MN2, BANDNAME(IBAND)
       IF(VP(IBAND.LT.0) GOTO 20
 #2  5 READ(4,*)  JP(COUNT), EFP(COUNT), JPP(COUNT), EFPP(COUNT),FREQ(COUNT), UFREQ(COUNT)
       IF(JP(COUNT).GE.0) THEN
           COUNT= COUNT+1
```

```
         GOTO 5
       ELSE
          COUNT= COUNT-1
          GOTO 10
       ENDIF
   20 CONTINUE
```

For a fluorescence series, this band-type data structure is retained, but the definitions of the 'band' and individual datum labels differ. In particular, setting the upper-state input parameter LABLP = 'FLS' identifies the data group as a fluorescence series and causes the 'band' parameters VP and VPP (see READ $^{\#}1$ above) to be defined as the vibration-rotation quantum numbers v' and J', respectively, for the emitting level, and the transition labels JP and JPP (READ $^{\#}2$ above) as the corresponding final (lower) state vibration-rotation quantum numbers v'' and J'', respectively. The parity label parameters EFP and EFPP retain their usual meaning, so the former has the same value for all lines in a given fluorescence series 'band' or data-group. Note that for a fluorescence series, the quantities v', J' and EFP are merely used to label the various series, and they have no physical significance as far as the analysis is concerned. Indeed, program operation is not affected if the same values of these quantities are used for several different series; it just means that the distinct fitted origins of those several fluorescence series will have the same label in the Channel-6 output file.

A third distinct data type that can be input using this same band-type data structure is a set of individual-level binding energies, such as those yielded by photoassociation spectroscopy (PAS). In this case the nature of the data-group is identified by setting the input upper-electronic-state band label as LABLP = 'PAS'. As with fluorescence series, the data parameters JP, JPP and EFPP are defined as the final-state level parameters v'', J'', and parity p'', respectively, while the 'band parameters' VP and VPP, and the data parameter EFP are all dummy variables that are ignored by the analysis. This same input data file structure may also be used to input the energies (relative to the associated asymptote) of tunneling-predissociation levels, as measured by 'translational spectroscopy' [99–101], except that in this case, the "binding energies" are input as negative numbers [39].

PAS-type data are expressed as (positive) binding energies, relative to the dissociation asymptote of the given electronic state. They have special significance in the data analysis only if the vibrational energies are represented by an expression in which the dissociation limit is an explicit parameter, i.e., only for PSEL = $1 - 5$ potential energy functions. If this is not the case, **dPotFit** simply treats a PAS-type data-group as an ordinary fluorescence series whose upper level is a free fitting parameter.

A fourth type of experimental data that uses this same band-type input structure is a set of of tunneling predissociation level widths. In this case the input parameter value LABLP = 'WID' identifies this data-group as a set of FWHM level widths $\Gamma(v, J)$ for levels of electronic state LABLPP, and signals that the band parameters VP & VPP and the parity label EFP are dummy variables. The input parameter JP is then the vibrational quantum number v, JPP is the rotational quantum number J, and EFPP is the parity label of the predissociating level, while FREQ is its full width at half maximum $\Gamma(v, J)$ and UFREQ the estimated uncertainty in that value (both in mcm^{-1}).

A different type of 'experimental' data that can be used in a fit is a set of assumed-known potential function values. This allows one to incorporate into an analysis, for example, *ab initio* potential energy values in the very short-range region that is inaccessible to normal spectroscopic data. This type of input data is identified by setting the input parameter value LABLP = 'VVV'.

The other band labels `VP`, `VPP`, `MN1` and `MN2` are then all dummy parameters, and the 'band' input commands shown above are replaced by

```
        COUNT= 1
     10 IBAND= IBAND+ 1
#1      READ(4,*,END=20)  VP(IBAND), VPP(IBAND), LABLP, LABLPP, MN1, MN2, BANDNAME(IBAND)
        IF(VP(IBAND.LT.0) GOTO 20
        IF(LABLP.EQ.'VV') THEN
#2   12    READ(4,*) R(COUNT), VV(COUNT), uVV(COUNT)
           IF(R(COUNT).GE.0) THEN
              COUNT= COUNT+1
              GOTO 12
            ELSE
              COUNT= COUNT-1
              GOTO 10
            ENDIF
        ENDIF
     20 CONTINUE
```

in which `VV(COUNT)` and `uVV(COUNT)` are the value and estimated uncertainty (in units cm^{-1}) of the potential energy function at the distance `R(COUNT)` Å. This type of data should normally be associated with the reference isotopologue, or with `MN1 = MN2 = 0`, if it is an allowed 'species'.

The final types of experimental data that can be used in a fit are sets of measured second 'pressure virial' coefficients [85] or 'acoustic virial' coefficients. In this case, one of the isotopologues specified in READ $^{\#}4$ of Appendices B and C should correspond to mass numbers `MN(1,IISTP)` = `MN(2,IISTP)` = 0 in order to specify that the abundance-averaged isotopic atomic masses to be used to define the reduced mass in the virial coefficient calculation. This type of input data is identified by setting the input parameter value `LABLP = 'VIR'` for ordinary 'pressure virial' coefficients, of `LABLP = 'VAC'` for 'acoustic' virial coefficients. The vibrational band labels `VP` and `VPP` are then dummy parameters, while the isotope labels should be set as `MN1 = MN2 = 0`. The 'band' input instructions for this case then have the same structure as those for the preceding case of potential-function-value data, except that the variables have different meanings:

```
        COUNT= 1
     10 IBAND= IBAND+ 1
#1      READ(4,*,END=20)  VP(IBAND), VPP(IBAND), LABLP, LABLPP, MN1, MN2
      1                   BANDNAME(IBAND)
        IF(VP(IBAND.LT.0) GOTO 20
        IF(LABLP.EQ.'VV') THEN
#2   14    READ(4,*) TEMP(COUNT), BVIR(COUNT), uBVIR(COUNT)
           IF(TEMP(COUNT).GE.0) THEN
              COUNT= COUNT+1
              GOTO 14
            ELSE
              COUNT= COUNT-1
              GOTO 10
            ENDIF
        ENDIF
     20 CONTINUE
```

Here, `Bvir(COUNT)` and `uBvir(COUNT)` are the value and estimated uncertainty (in units $cm^3 mol^{-1}$) of the second virial coefficient of the atomic species to which the molecule dissociates, at temperature $T = $ `TEMP(COUNT)` K.

As an illustration of this data file structure, the listing below presents portions of an experimental data file used in an analysis to determine the potential energy function for the ground $^1\Sigma_g^+$ state of ArXe [85] from a combination of microwave data [102], high-resolution vacuum laser spectroscopy [85], and virial coefficient data [103–105]. Note that text beginning at the "%" sign on a line of the input data file are comments that are ignored by the program.

```
0  0     'X1S'  'X1S'      40    128        'MW data of Jaeger et al. (1993)'
   3 +1  2 +1  0.19375109  3.3D-08          % J' p' J" p" FREQ UFREQ
   4 +1  3 +1  0.25830956  3.3D-08
   5 +1  4 +1  0.32284639  3.3D-08
   6 +1  5 +1  0.38735613  3.3D-08
  -1 -1 -1 -1       -1.d0    -1.d0

0  0     'X1S'  'X1S'      40    129        'MW data of Jaeger et al. (1993)'
   3 +1  2 +1  0.19339688  3.3D-08          % J' p' J" p" FREQ UFREQ8
   4 +1  3 +1  0.25783738  3.3D-08
   5 +1  4 +1  0.32225632  3.3D-08
   6 +1  5 +1  0.38664857  3.3D-08
   7 +1  6 +1  0.45100778  3.3D-08
   8 +1  7 +1  0.51532937  3.3D-08
   9 +1  8 +1  0.57960754  3.3D-08
  -1 -1 -1 -1       -1.d0    -1.d0

0  0     'X1S'  'X1S'      40    130        'MW data of Jaeger et al. (1993)'
   3 +1  2 +1  0.19304902  3.3D-08          % J' p' J" p" FREQ UFREQ8
   4 +1  3 +1  0.25737363  3.3D-08          % MW data of Jaeger et al. (1993)
   5 +1  4 +1  0.32167676  3.3D-08
   6 +1  5 +1  0.38595299  3.3D-08
   7 +1  6 +1  0.45019693  3.3D-08
  -1 -1 -1 -1       -1.d0    -1.d0

2  0     'D0'   'X1S'      40    132        'Piticco [JMS 264, 83 (2010)]'
   1 +1  0 +1  77239.1655  0.0012           % J' p' J" p" FREQ UFREQ
   2 +1  1 +1  77239.2240  0.0012
   3 +1  2 +1  77239.2817  0.0010
   4 +1  3 +1  77239.3348  0.0011
   5 +1  4 +1  77239.3867  0.0008
   6 +1  5 +1  77239.4358  0.0011
...... omit 18 intermediate data to save space ...............................
  13 +1 14 +1  77237.9631  0.0020
  16 +1 17 +1  77237.6269  0.0010
  -1 -1 -1 -1       -1.d0    -1.d0

0  0     'VIR'  'X1S'       0     0        'Press virials, various sources'
  173.2   -141.2   2.d0          % Virial coefft. from Landbolt & Boernstein (2003)
  198.2   -108.0   2.d0          % Virial coefft. from Landbolt & Boernstein (2003)
  223.2    -86.5   2.d0          % Virial coefft. from Landbolt & Boernstein (2003)
  273.2    -56.2   2.d0          % Virial coefft. from Landbolt & Boernstein (2003)
  323.2    -36.7   2.d0          % Virial coefft. from Landbolt & Boernstein (2003)
  203.0    -95.0   6.d0          % Virial coefft. from Schramm et al. (1977)
  213.0    -84.5   6.d0          % Virial coefft. from Schramm et al. (1977)
...... omit 15 intermediate data to save space ...............................
  482.0     -5.1   6.d0          % Virial coefft. from Rentschler & Schramm (1977)
  555.0      2.4   6.d0          % Virial coefft. from Rentschler & Schramm (1977)
  626.0      4.9   6.d0          % Virial coefft. from Rentschler & Schramm (1977)
  695.0      9.5   6.d0          % Virial coefft. from Rentschler & Schramm (1977)
   -1      -1.0  -1.d0
```

Appendix B. Structure of The Channel–5 'Instruction' Data File

The logical structure and read statements of the Channel–5 data input that describes the system to be treated, specifies the type of fit to be carried out, and inputs any necessary system parameters, is shown below. The following subsection then provides a detailed description of the nature and options associated with each of the input variables.

```
#1    READ(5,*)  AN(1), AN(2), CHARGE, NISTP, NSTATES, LPRINT, PRINP
#2    READ(5,*)  DATAFILE
#3    READ(5,*)  WRITFILE
#4    IF(CHARGE.ne.0) READ(5,*)  CHARGE1,CHARGE2
*******************************************************************************
*     Loop over isotopologues to read isotope masses
*******************************************************************************
      DO  IISTP= 1,NISTP
#5      READ(5,*) MN(1,IISTP), MN(2,IISTP)
      ENDDO
*******************************************************************************
*     End loop over isotopologues
*******************************************************************************
 #6   READ(5,*)  UCUTOFF, NOWIDTHS, IROUND, ROBUST, CYCMAX, uBv
*******************************************************************************
*     Begin loop over electronic states
*******************************************************************************
      DO 20 ISTATE=1,NSTATES
#7        READ(5,*) SLABL(ISTATE), IOMEG(ISTATE), VMIN(ISTATE), VMAX(ISTATE),
     1                                      JTRUNC(ISTATE), EFSEL(ISTATE)
#8        IF(VMAX(ISTATE.1).LT.0) READ(5,*) (VMAX(ISTATE,ISOT), ISOT= 1, NISTP)

#9        READ(5,*) PSEL(ISTATE), VLIM(ISTATE), MAXMIN(ISTATE). BOBCN(ISTATE),
     1                                                  OSEL(ISTATE)
          IF(PSEL(ISTATE).EQ.-2) GOTO 20
          IF(PSEL(ISTATE).EQ.-1) THEN
*******************************************************************************
*  If fitting to band constants, must specify no. for each v and isotop.
*******************************************************************************
            DO  I= VMIN(ISTATE,1),VMAX(ISTATE,1)
#10            READ(5,*) VTST,(NBC(I,IISTP,ISTATE),IISTP= 1,NISTP)
              ENDDO
            ENDIF
#11       READ(5,*) RMIN(ISTATE), RMAX(ISTATE), RH(ISTATE)
          IF(PSEL(ISTATE).EQ.0) THEN
#12         READ(5,*) NPT, NUSE, IR2, ILR, NCN, CNN
#13         READ(5,*) RFACT, EFACT, VSHIFT
#14         READ(5,*) (XI(I), YI(I), I= 1,NTP)
            EXIT
            ENDIF
          IF((PSEL(ISTATE).GE.2).AND.(PSEL(ISTATE).LE.6)) THEN
    c** For an MLR or DELR or HPP or TT-type potential ...
#15         READ(5,*) NCMM(ISTATE), rhoAB(ISTATE), IVSR(ISTATE, IDSTT(ISTATE)
            DO  m= 1,NCMM(ISTATE)
#16           READ(5,*) MMLR(m,ISTATE), CmVAL(m,ISTATE), IFXCm(m,ISTATE)
#17           IF(IFXCm(m,ISTATE).LT.0) READ(5,*) SCCm(m.ISTATE)
              ENDDO
```

```
             ENDIF
     c** For a GPEF potential ...
#18          IF(PSEL(ISTATE).EQ.6) READ(5,*) AGPEF(ISTATE), BGPEF(ISTATE)
     c** For all  PSEL > 0  cases ...
#19          READ(5,*) DE(ISTATE), IFXDE(ISTATE)
#20          READ(5,*) RE(ISTATE), IFXRE(ISTATE)
#21          READ(5,*) qPOT(ISTATE),RREFq(ISTATE),IFXrefq(ISTATE)
#22          IF(PSEL(ISTATE).EQ.2) READ(5,*) pPOT(ISTATE),RREFp(ISTATE),IFXrefp(ISTATE)
#23          READ(5,*) Nbeta(ISTATE), APSE(ISTATE), nPB(ISTATE), qPOT(ISTATE), RREF(ISTATE)
             IF((PSEL(ISTATE).EQ.2).AND.(APSE(ISTATE).GE.0) THEN
                 DO  I= 1,Nbeta(ISTATE)
#24                  READ(5,*) ypBETA(I,ISTATE), BETA(I,ISTATE), IFXBETA(I,ISTATE)
                     ENDDO
               ELSE
                 I1=0
                 IF(PSEL(ISTATE).GE.6) I1=1
                 DO   I= I1, Nbeta(ISTATE))
#25                  READ(5,*) BETA(I,ISTATE), IFXBETA(I,ISTATE)
                     ENDDO

#26          READ(5,*) NUA(ISTATE), NUB(ISTATE), qAD(ISTATE), pAD(ISTATE), LRad(ISTATE)
     c** if appropriate, read isotope shift coeffts as part of BOB correction
#27          IF(LRad(ISTATE)).GT.0) READ(5,*) (dCMVALA(m,ISTATE),m=1, NCMM(ISTATE))
#28          IF(LRad(ISTATE)).GT.0) READ(5,*) (dCMVALB(m,ISTATE),m=1, NCMM(ISTATE))
     c** Adiabatic potential energy BOB function parameters for atom-A
             IF(NUA(ISTATE).GE.0) THEN
                 DO  I= 0, NUA(ISTATE)
#29                  READ(5,*) UA(I,ISTATE), IFXUA(I,ISTATE)
                     ENDDO
#30              READ(5,*) UAinf, IFXUAinf
                 ENDIF
     c** Adiabatic potential energy BOB function parameters for atom-B
             IF(NUB(ISTATE).GE.0) THEN
                 DO  I= 0, NUB(ISTATE)
#31                  READ(5,*) UB(I,ISTATE), IFXUB(I,ISTATE)
                     ENDDO
#32              READ(5,*) UBinf, IFXUBinf
                 ENDIF
#33          READ(5,*) NTA(ISTATE), NTB(ISTATE), qNA(ISTATE) pNA(ISTATE)
     c** Centrifugal BOB function parameters for atom-A
             IF(NTA(ISTATE).GE.0) THEN
                 DO  I= 0, NTA(ISTATE)
#34                  READ(5,*) TA(I,ISTATE), IFXTA(I,ISTATE)
                     ENDDO
#35              READ(5,*) TAinf, IFXTAinf
                 ENDIF
     c** Centrifugal BOB function parameters for atom-B
             IF(NTB(ISTATE).GE.0) THEN
                 DO  I= 0, NTB(ISTATE)
#36                  READ(5,*) TB(I,ISTATE), IFXTB(I,ISTATE)
                     ENDDO
#37              READ(5,*) TBinf, IFXTBinf
                 ENDIF
             IF(IOMEG(ISTATE).NE.0) THEN
```

```
      c** Lambda-doubling or 2\Sigma splitting expansion parameters
#38            READ(5,*) NwCFT(ISTATE), Pqw(ISTATE), efREF(ISTATE)
               IF(NwCFT(ISTATE).GE.0) THEN
                   DO  I= 0, NwCFT(ISTATE)
#39                    READ(5,*) wCft(I,ISTATE), IFXwCft(I,ISTATE)
                     ENDDO
                 ENDIF
             ENDIF
         ENDDO
    C************************************************************************
    c          End of loop over electronic states
    C************************************************************************
         IF(DATAFILE.EQ.MAKEPRED) THEN
    c** To generate a predicted spectrum, loop over bands & read specifications
    70       CONTINUE
#40          READ(5,*,end=99) VP(IBAND), VPP(IBAND), LABLP, LABLPP, MN1, MN2,
      1                                 PP, PPP, JMAXX, J2DL, J2DU, J2DD
             IBAND= IBAND+ 1
             IF(VP(IBAND).GE.0) GO TO 70
             ENDIF
    C************************************************************************
    c          End of Channel-5 input data file
    C************************************************************************
```

Appendix C: Definitions of the Parameters in the 'Instruction' Data File

Read integers identifying the molecule and system.

$^{\#}$1. READ(5,*) AN(1), AN(2), CHARGE, NISTP, NSTATES, LPRINT, PRINP

> AN(1) & AN(2) are integer atomic numbers of the atoms 1 & 2 forming the molecule.
>
> CHARGE is a ($\pm$) integer for the total charge on the molecule. It is used to generate Watson's charge-modified reduced mass for molecular ions; see Eq. (2).
>
> NISTP is the number of isotopologues to be considered in the particular analysis (their identifying mass numbers are read below).
>
> NSTATES is the number of different electronic states associated with the data to be input and analyzed.
>
> LPRINT is an integer specifying the level of output to be generated by the least-squares subroutine package NLLSSRR.
> - If LPRINT $= 0$ no internal printout except for convergence-failure warning messages;
> - If LPRINT $\neq 0$, also print convergence test on each fitting cycle and indicate nature of convergence; LPRINT ≤ 0 is recommended when there are no problems with the fits.
> - If LPRINT ≥ 1, also print the change and uncertainty for each parameter in each fitting cycle;
> - If LPRINT ≥ 2, also print parameter changes on each rounding step.
>
> PRINP is an integer parameter that controls whether (PRINP > 0) or not (PRINP ≤ 0) a summary of the experimental data is printed immediately following its input. This option is useful only for helping to identify problems in the input data; in most cases one should set PRINP ≤ 0, since an analogous summary is always printed at the end of every (successful) run.

Read the name of the file containing the experimental data to be fitted in the analysis.

$^{\#}$2. READ(5,*) DATAFILE

> DATAFILE is the name for the file containing the experimental data to be fitted; it may consist of up to 40 alphanumeric characters and must be enclosed in single quotes with no leading blanks (e.g., 'Li2B(A-X).4'). If the file does not reside in the current directory, this name must include the absolute or relative path. Note that if this file contains data for more states and/or isotopologues than one wishes to consider in a particular analysis, they are simply ignored, so it is not necessary to construct a separate data file if one wishes to consider only a subset of the data (e.g., only the microwave transitions, or only the data for one particular isotopologue).
> - If the program is being asked to predict data from a set of known input parameters (see § 5.3), the filename read here *must* be MAKEPRED (i.e., the data file entry must be 'MAKEPRED').

$^{\#}$3. READ(5,*) WRITFILE

> WRITFILE is the root of the names used for the output files written to Channels 6–16 & 20, that will have the names WRITFILE.6, WRITFILE.7, WRITFILE.8, ... , WRITFILE.20, respectively. WRITFILE may consist of up to 20 alphanumeric characters, and should be enclosed in single quotes with no leading blanks (e.g., 'EMO4(6,8)u3t1'). A distinct name should be chosen to identify the results of each particular run; if the same name is used for subsequent runs, previous files with those names will be over-written.

If CHARGE $\neq 0$, read integers specifying how the mass(es) of the addd/subtracted e^- are to be taken into account in the definition of the molecular ion reduced masses.

$^{\#}$4 If(CHARGE $\neq 0$) READ(5,*) CHARGE1, CHARGE2

CHARGE1 & CHARGE2: the integer number of $m_e/2$ masses to be added to or subtracted from the normal masses of atoms $^{\#}1$ and $^{\#}2$ before calculating a conventional 2-body reduced mass for each isotopologue. If CHARGE1 = CHARGE2 = 0 use Watson's charged-modified reduced mass for all species. Otherwise, necessarily, $1/2$(CHARGE1 + CHARGE2) = CHARGE .

Loop over the NISTP isotopologues, and for each, read the integer mass numbers of the two atoms. The first isotopologue is defined as the reference species (ISOT$=\alpha=1$) for the mass scaling of BOB radial strength functions. The mass numbers for each isotopologue must be on a separate line.

$^{\#}5$ READ(5,*) MN(1,ISOT), MN(2,ISOT)

MN(1,ISOT) & MN(2,ISOT): integer mass numbers of the atoms/particles 1 & 2 forming the isotopologue $\alpha =$ ISOT . The mass of a normal stable atomic isotope is taken from the tabulation in subroutine MASSES; if MN is outside the range for the normal stable isotopes for the atom, then the abundance-averaged atomic mass is used.

Read parameters allowing one to set global restrictions on the data set to be used without editing the actual input DATAFILE, and to specify general features of the fit and output.

$^{\#}6$ READ(5,*) UCUTOFF, NOWIDTHS, IROUND, ROBUST, CYCMAX, uBv

UCUTOFF: causes any input spectroscopic data with uncertainties $u_i >$ UCUTOFF (a real number) to be neglected. It does not apply to virial coefficient data or read-in potential function values and their uncertainties.

NOWIDTHS: if NOWIDTHS > 0, omit calculation of the partial derivatives of tunneling-predissociation level widths of quasibound levels, and ignore any tunneling-level-width data (defined by SLABLPP= 'WID') in the data file; otherwise, set NOWIDTHS $= 0$.

IROUND: Setting (integer) IROUND $\neq 0$ causes the "sequential rounding and refitting" procedure of Ref. [97] to be implemented, with each parameter being rounded at the |IROUND|'th significant digit of its uncertainty. If IROUND > 0 the rounding is applied sequentially to the remaining free parameter having the largest relative uncertainty; if IROUND < 0 the rounding proceeds systematically from the last free parameter to the first. If IROUND$=0$ the fit simply stops after full convergence and performs no parameter rounding; this last option saves considerable computation time, and would normally be chosen except for a "final" fit to obtain parameters for publication and/or distribution.

ROBUST is an integer that is set greater than 0 to cause the *robust* fitting procedure described at the end of § 8.1 to be applied; otherwise (the normal case), it should be set equal to 0 (or negative).

CYCMAX is an integer that sets an upper bound on the number of least-squares cycles allowed in NLLSSRR. Normally MAXCYC ~ 30 suffices. If larger values are required, there may be something wrong either with the model or with the precision of the calculations.

uBv Setting the integer uBv > 0 causes the program to calculate and write to Channel–7 the uncertainties in the final B_v values due to the correlated-parameter uncertainties determined in the fit. Otherwise, set uBv ≤ 0.

Now begin the loop over the NSTATES electronic states to be considered in the analysis. This loop from $s \equiv$ ISTATE $=1$ to NSTATES includes almost all of the rest of the Channel–5 input data.

First, read integer parameters characterizing the state, and (if desired) set limits on the ranges of the rotational and vibrational levels to be considered for this fit.

$^{\#}7$. READ(5,*) SLABL(s), IOMEG(s), VMIN(s), VMAX(s), JTRUNC(s), EFSEL(s)

SLABL   is a three-alphanumeric-character label for this electronic state, enclosed in single quotes (e.g., 'X1S', 'a1P', ...), to identify it in the output and in the experimental data input file.

IOMEG : If IOMEG ≥ 0 the electronic state is treated as a spin singlet with integer electronic angular momentum projection quantum number IOMEG. For a state with $^2\Sigma$ symmetry, set IOMEG $= -1$.

VMIN & VMAX   are data-range selection parameters. All data for this state associated with vibrational levels outside the range VMIN to VMAX are ignored and are omitted from the analysis.

JTRUNC   is an integer data-selection parameter. If JTRUNC > 0, all data for this electronic state for which $J >$ JTRUNC are omitted from the analysis; if JTRUNC < 0 all data with $J < |$JTRUNC$|$ are omitted.

EFSEL   is an integer that allows a user to consider:
   a) only transitions involving e–parity levels of this state, if EFSEL > 0,
   b) only transitions involving f–parity levels of this state if EFSEL < 0,
   c) all transitions if EFSEL $= 0$ (the normal case).

The code normally sets VMAX for minor isotopologues using first-order semiclassical scaling from the value for isotopologue-1, but if (VMAX(ISTATE,1).LT.0), we override that approach and read in a 'specified VMAX value for each isotopologue. This is sometimes needed when a level search for a small-μ minor isotopologue causes the code to stop because the highest level(s) cannot be found.

$^\#$8. IF(VMAX(ISTATE,1).LT.0) READ(5,*) (VMAX(ISTATE,ISOT), ISOT= 1, NISTP)

Read parameters to select the analytical potential function model, to specify the potential asymptote, to select the representation to be used for BOB corrections, and to control the size of output print arrays.

$^\#$9. READ(5,*) PSEL(s), VLIM(s), MAXMIN(s), BOBCN(s), OSEL(s)

PSEL   is an integer which specifies the type of analytic function used for the potential.

   If PSEL $= 0$ use a fixed potential defined by interpolating over and extrapolating beyond a set of input turning points using the program **LEVEL** subroutine package PREPOT [4, 83].

   If PSEL $= 1$ use the Expanded Morse Oscillator (EMO) form of §3.1.

   If PSEL $= 2$ use the Morse/Long-Range (MLR) potential form of §3.2.

   If PSEL $= 3$ use the Double-Exponential/Long-Range (DELR) form of §3.3.

   If PSEL $= 4$ use Seto's modification [15] of Šurkus' Generalized Potential Energy Function (GPEF) [58], as described in §3.4. Its parameters a_S, b_S and p, are input via READs $^\#$17 and 20.

   • Dunham expansions are generated by setting $p = 1$, $a_S = 0$ & $b_S = 1$.
   • SPF expansions are generated by setting $p = 1$, $a_S = 1$ & $b_S = 0$.
   • Ogilvie–Tipping expansions are generated by setting $p = 1$, $a_S = b_S = 0.5$.
   • These polynomial-type potentials usually have an undefined (or at best, indirectly-defined) asymptote, so in these cases parameter VLIM defines the energy at the potential minimum.

   If PSEL $= 5$ use the Tiemann/Hannover polynomial of §4.1.

   If PSEL $= 6$ use the Tang-Toennies type potential of Eq. (33) in §4.2. In this case Nbeta $= 9$ and the the values of the well depth $\mathfrak{D}_e$ and equilibrium distance r_e are determined internally as properties of the potential.

If $\mathtt{PSEL}=7$ use the Scoles-Aziz HFD-type potential of Eq. (34) in §4.3. In this case $\mathtt{Nbeta} = N_\beta = 5$ or 2 (see below), and the the values of β_1 and A_{HFD} are determined internally from the other parameters.

If $\mathtt{PSEL} = -2$ the energy levels of the state are represented by independent terms values which are labelled by vibrational quantum number v, rotational quantum number J and e/f parity quantum number p. In this case, VLIM, BOBCN and OSEL are dummy parameters, and the remainder of the READs for this state, $^\#10$–36, are ignored.

If $\mathtt{PSEL} = -1$ the energy levels of the state are represented by a set of band constants $\{G_v,\ B_v,\ -D_v,\ H_v,\ \ldots\}$ for each vibrational level of each isotopologue. In this case, VLIM, BOBCN and OSEL are dummy parameters, READ $^\#10$ is used to input the number of such constants for each state, and the remainder of the READs for this state, $^\#11$–36, are ignored.

VLIM : Parameter VLIM specifies the fixed absolute energy (in cm^{-1}) at the potential asymptote for this state. The set of values of VLIM(ISTATE) define the absolute energy scale for this system.

MAXMIN : specifies the maximum number of potential function minima allowed for a given state. If this number is exceeded (e.g., in the midst of a multi-cycle many-parameter fit), the code prints a warning message and STOPS. Normally set $\mathtt{MAXMIN} = 1$.

BOBCN specifies the manner in which the BOB mass-scaling is to be done.
If $\mathtt{BOBCN}=0$, use the mass scaling of Eqs. (3) and (4) with isotopologue $^\#1$ as the reference isotopologue.
If $\mathtt{BOBCN}=1$, combine use of the clamped-nuclei reference potential with mass scaling factors $m_\mathrm{e}/M_A^{(\alpha)}$, of Eqs. (46) and (47).

OSEL : if (integer) $\mathtt{OSEL} \neq 0$, write every |OSEL|'th point of the final potential energy array to channel-10. If $\mathtt{OSEL} < 0$ also write any BOB radial function arrays to Channels 12-16 on that same mesh; if $\mathtt{OSEL} = 0$, omit all such output. Note that smaller values of |OSEL| yield larger output files written on a finer radial mesh.

If $\mathtt{PSEL} = -1$, loop over vibrational levels VMIN to VMAX, reading in the vibrational quantum number $\mathtt{VTST} = v$ and the number of Band Constants $\{G_v,\ B_v,\ -D_v,\ H_v,\ \ldots\}$ to be used for that level, for each isotopologue, NBC(I,IISTP,ISTATE). If $\mathtt{PSEL} \neq -1$, ignore READ $^\#10$.

```
        DO i= VMIN(ISTATE,1),VMAX(ISTATE,1)
#10.        READ(5,*) VTST,(NBC(I,IISTP,ISTATE),IISTP= 1, NISTP)
        ENDDO
```

For all cases for which $\mathtt{PSEL} \geq 0$, read READ $^\#11$.

$^\#11$. READ(5,*) RMIN(s), RMAX(s), RH(s)

RMIN & RMAX : are the inner and outer limits, respectively, of the range of numerical integration (in Å). As zeroth order estimates, one may set $\mathtt{RMIN} \approx 0.6\times$(potential inner wall position) and RMAX very large (say $= 99$ Å); see §6.1.

RH : the numerical integration mesh size; see Eq. (56) in §6.1.

If $\mathtt{PSEL} = 0$, define a fixed potential for this state by reading, interpolating over and extrapolating beyond a given set of turning points, in the manner specified here. For $\mathtt{PSEL} \neq 0$, skip READs $^\#12 - 14$.

$^\#12$. READ(5,*) NTP, NUSE, IR2, ILR, NCN, CNN

NTP : gives the number of turning point pairs to be input via READ $^\#14$ to define the potential.

NUSE : Specifies how the interpolation is to be done. If $\texttt{NUSE} > 0$, use $\texttt{NUSE}$–point piecewise polynomials; if $\texttt{NUSE} \leq 0$, perform cubic spline interpolation. For highly precise and smooth input points, such as those generated from an RKR calculation, $\texttt{NUSE} = 8$, 10 or 12 is usually most appropriate; for less precise or less dense points, such as those from *ab initio* calculations, low-order piecewise polynomials ($\texttt{NUSE} = 4$) or splines ($\texttt{NUSE} \leq 0$) are usually best [4, 83].

IR2 : For very steep repulsive potential walls, better interpolation is often attained by interpolating over $r^2 \times V(r)$ rather than over $V(r)$ itself; setting the integer $\texttt{IR2} > 0$ causes this to be done (normally recommended). A comparison between results obtained with this option turned on *vs.* off (setting $\texttt{IR2} \leq 0$ causes interpolation to be performed over $V(r)$ itself) provides an indication of the magnitude of "interpolation noise" uncertainties in the final results.

ILR : Specifies how to extrapolate from the outermost read-in turning points to the asymptote. For a long energy extrapolation, a choice of $\texttt{ILR} = -1$, 0, or 1 is often most appropriate; however, if the outer turning points extend moderately close to the dissociation limit (at $\texttt{VLIM}$), one should set $\texttt{ILR} \geq 2$ and specify the theoretically appropriate value of $\texttt{NCN}$ (≥ 1), and if it is available, also input an estimate of $\texttt{CNN}$ (see below).

For $\texttt{ILR} < 0$, fit the last 3 points to: $V(r) = \texttt{VLIM} - A \times \exp[-b(r - r_o)^2]$

For $\texttt{ILR} = 0$, fit the last 3 points to: $V(r) = \texttt{VLIM} - A \times r^p \times \exp[-b\,r]$.

For $\texttt{ILR} = 1$, fit the last 2 points to: $V(r) = \texttt{VLIM} - A/r^B$.

For $\texttt{ILR} = 2$ or 3, respectively, fit the outermost 2 or 3 points to a sum of 2 or 3 inverse-power terms, with powers differing by 2: $V(r) = \texttt{VLIM} - \sum_{m=0}^{\texttt{ILR}-1} C_{\texttt{NCN}+2m}/r^{\texttt{NCN}+2m}$.

For $\texttt{ILR} \geq 4$, fit outermost $\texttt{ILR}$ turning points to a sum of $\texttt{ILR}$ inverse-power terms, with powers differing by 1: $V(r) = \texttt{VLIM} - \sum_{m=0}^{\texttt{ILR}-1} C_{\texttt{NCN}+m}/r^{\texttt{NCN}+m}$.

NCN : For inverse-power potential extrapolation with $\texttt{ILR} \geq 2$, $\texttt{NCN}$ (> 0) specifies the (inverse) power of the asymptotically dominant long-range term: $V(r) \propto \texttt{VLIM} - \texttt{CNN}/r^{\texttt{NCN}}$. Otherwise (for $\texttt{ILR} \leq 1$) it is a dummy parameter.

CNN : For inverse-power potential extrapolation with $\texttt{ILR} \geq 2$, setting $\texttt{CNN} \neq 0$ causes the leading inverse-power coefficient to be fixed at the read-in value $\texttt{CNN} = C_{\texttt{NCN}}$ [cm^{-1} Å$^{\texttt{NCN}}$] rather than to be determined from a fit to the outermost turning points.

$^{\#}$13. `READ(5,*) RFACT, EFACT, VSHIFT`
$^{\#}$14. `READ(5,*) (XI(I), YI(I), I= 1,NTP)`

RFACT & EFACT : are multiplicative factors required to convert units of the $\texttt{NTP}$ input turning point distances $\texttt{XI}(i)$ and energies $\texttt{YI}(i)$ to Å and cm^{-1}, respectively. If no conversion is required, read in factors of 1.0d+00.

VSHIFT : An energy shift (in cm^{-1}) to be added to the input potential point energies to make them consistent with $\texttt{VLIM}$. It addresses the fact that *ab initio* or RKR turning points may be expressed relative to an energy zero that is inconsistent with the user-specified asymptote energy $\texttt{VLIM}$.

XI(i) & YI(i) : are the (distance, energy) input turning points defining the potential function.

If $\texttt{PSEL} = 2$, 3, or $5-7$ (for an MLR, DELR, HPP, TT or HFD-type potential), specify the long-range potential energy tail function $u_{\text{LR}}(r)$. For other cases, skip READs $^{\#}$15 & 16.

$^{\#}$15. `READ(5,*) NCMM(s), rhoAB(s), IDF(s), IDSTT(s)`
   `DO i= 1, NCMM(s)`

[#]16.   `READ(5,*) MMLR(i,s), CmVAL(i,s), IFXCm(i,s)`
[#]17.   `IF(IFXCm(i,s) < 0,  READ(5,*) SCCm(i,s)`
         `ENDDO`

`NCMM(s)` is the number of inverse-power terms included in the definition of $u_{\mathrm{LR}}(r)$ in Eq. (11).

If `rhoAB` ≤ 0.0, the long-range potential is the pure inverse-power sum of Eq. (9);i.e., all $D_{m_i}(r) = 1.0$.

If `rhoAB` > 0.0, include damping functions in the definition of $u_{\mathrm{LR}}(r)$ by Eq. (11), in which and `rhoAB` $= \rho \equiv \rho_{\mathrm{AB}}$ is the system-dependent scaling factor of Eq. (16). In this case:

- `IVSR` $\equiv 2s$ selects the damping function from the family specified by parameter `IDSTT` that has the limiting short-range behaviour $D_m(r)/r^m \propto r^s = r^{\mathrm{IVSR}/2}$.
- If `IDSTT` > 0, use the generalized Douketis-type damping functions of Eq. (12) (recommended). Allows any integer $-4 \leq$ `IVSR` ≤ 0, $+2$ or $+4$ (`IVSR` $= -2, -1$, or 0 recommended for an MLR).
- If `IDSTT` $= 0$, use the simple damping functions of of Eq. (13). Allows any even or odd integers `IVSR`.
- If `IDSTT` < 0, use the generalized Tang-Toennies-type damping functions of Eq. (14). Allows even integers $-4 \leq$ `IVSR` ≤ 4.

`MMLR(i,s)` is the power and `CmVAL(i,s)` the coefficient of the i'th contribution to $u_{\mathrm{LR}}(r)$ in Eq. (11), C_{m_i}/r^{m_i}. Positive `CmVAL` values yield attractive potential energy terms.

`IFXCm(i,s)` controls whether $C_{m_i} =$ `CmVAL(i,s)` is to be fitted (`IFXCm` $= 0$) or constrained (`IFXCm` $\neq 0$) in the fit. Fitting to one or more C_{m_i} values should only be considered if a substantial amount of data is available for levels lying very near dissociation [20].

- Setting `IFXCm` > 0 causes this C_m coefficient to be held fixed at the read-in value.

- Setting `IFXCm` < 0 causes this C_m coefficient to be constrained to equal the value of parameter [#]|`IFXCm`|, which has a smaller index in the parameter count than its own, multiplied by the fixed scaling parameter `SCCm`. E.g. to constrain C_m for state `ISTATE` to equal precisely half the value of the leading C_m value for the first electronic state (`ISTATE` $= 1$), set `IFXCm(s)` $= -3$ (since r_e and $\mathfrak{D}_e$ are parameters [#]1 and 2, respectively), and read in the constant `SCCm(ISTATE,1)=0.5d0` . This option allows one to fit simultaneously to potentials for two states with this leading C_m coefficient being allowed to vary, subject to the constraint that the value for the second state will always be precisely half that for the first state.

For the special case of states of a homonuclear alkali dimer dissociating to the $(nS) + (nP)$ asymptote, to use the Aubert-Frécon 2×2 diagonalization long-range form [20, 38, 106], set `NCMM` $= 7$ with `MMLR(i)` $= x$, 3, 3, 6, 6, 8, & 8, for which the input values of `CmVAL(i)` are A_{so}, C_3^{Σ}, C_3^{Π}, C_6^{Σ}, C_6^{Π}, C_8^{Σ} and C_8^{Π}, for $i = 1 \cdots 7$, respectively, and

- For the 2×2 tail of the $A\,^1\Sigma_u^+$ state, set $x =$ `MMLR(1)` $= 0$.
- For the 2×2 tail of the $b\,^3\Pi_u$ state, set $x =$ `MMLR(1)` $= -1$.

Similarly, to use the Aubert-Frécon 3×3 diagonalization long-range form [22, 38, 107], set `NCMM` $= 10$ with `MMLR(i)` $= x$, 3, 3, 3, 6, 6, 6, 8, 8, & 8, for which the input values of `CmVAL(i)` are A_{so}, C_3^{Σ}, $C_3^{1\Pi}$, $C_3^{3\Pi}$, C_6^{Σ}, $C_6^{1\Pi}$, $C_6^{3\Pi}$, C_8^{Σ} $C_8^{1\Pi}$, and $C_8^{3\Pi}$, for $i = 1 \cdots 10$, respectively, and

- For the 3×3 tail of the $1\,^3\Sigma_g$ state, set $x =$ `MMLR(1)` $= -2$.
- For the 3×3 (mainly repulsive) tail of the $B\,^1\Pi_u$ state, set $x =$ `MMLR(1)` $= -3$.

If `PSEL` $= 7$, read in parameters defining the expansion variable in the GPEF potential of Eq. (26); otherwise, skip READ $^\#18$.

$^\#18.$ `READ(5,*) AGPEF`(s), `BGPEF`(s)

> In the GPEF radial expansion variable of Eq. (26): $a_S = $ `AGPEF` and $b_S = $ `BGPEF`, while $p = $ `nPB` is input via READ $^\#21$.

Read the dissociation energy and equilibrium distance of the potential function for this state.

$^\#19.$ `READ(5,*) DE`(s), `IFXDE`(s)

$^\#20.$ `READ(5,*) RE`(s), `IFXRE`(s)

> `DE` $\equiv \mathfrak{D}_e$ is the well depth. It is a dummy variable for the case of a GPEF potential. For a TT potential (`PSEL` $= 5$) the reported values of $\mathfrak{D}_e$ and r_e should be read in, but the code will report more accurate values. variable for the ca

> `IFXDE` controls whether $\mathfrak{D}_e$ is to be fitted (`IFXDE` ≤ 0) or held fixed (`IFXDE` > 0) in the fit.

> `RE` is the equilibrium radial distance r_e for this state, and `IFXRE` controls whether it is to be fitted (`IFXRE` ≤ 0) or held fixed (`IFXRE` > 0) in the fit.

Read the power and expansion centre for the dimensionless PEF expansion variable.

$^\#21.$ `READ(5,*) qPOT`(s)`,RREFq`(s)`,IFXrefq`(s)

> Integer `qPOT` is the power and `RREFq` radial position of the expansion centre for the (exponent) expansion variable $y_q^{\text{ref}_q}(r)$ of Eqs. (6), (8), (19), (21) and (26), while `IFXrefq` controls whether `RREFq` is to be held fixed (if > 0) or fitted (if ≤ 0). For GPEF potentials (`PSEL` $= 4$) `RREFq` and `IFXrefq` are dummy variables.

For MLR potentials, specify the form of the exponent coefficient function $\beta(r)$ and the power and form of the switching function variable.

$^\#22.$ `IF(PSEL.EQ.2) READ(5,*) APSE`(s), `pPOT`(s), `RREFp`(s), `IFXrefp`(s)

> If (integer) `APSE`$(s) > 0$ use the 'Pashov-Spline' expression for the MLR exponent coefficient $\beta_{\text{SE}-\text{MLR}}(r)$, Eq. (21) [6, 54].
> If `APSE`$(s) \leq 0$ use the constrained-polynomial expansion of Eq. (19) for the exponent-coefficient function $\beta(r) = \beta_{\text{PE}-\text{MLR}}(r)$ of the MLR potential.

> Integer `pPOT` is the power for the dimensionless radial variable $y_p^{r_e}(r)$ of Eqs. (5), (10) and for the switching function variable $y_p^{\text{ref}_p}(r)$ of Eq. (19), while `RREFp` the location of the centre for that switching function variable, and `IFXrefp` controls whether `RREFp` is to be held fixed (if > 0) or fitted (if ≤ 0). Note that for a SE-MLR function (`APSE`$(s) > 0$), `RREFp` and `IFXrefp` are dummy variables.

Read integer specifying the order of the (exponent) polynomial defining the potential function $V_{\text{ad}}^{(1)}(r)$ (or $V_{\text{CN}}(r)$).

$^\#23.$ `READ(5,*) Nbeta`(s)

> For `PSEL` $= 1$ (an EMO potential), `Nbeta` is the order of the potential function exponent-coefficient polynomial expansion of Eq. (8).

> For `PSEL` $= 2$ (an MLR potential), `Nbeta` specifies the order of the exponent polynomial of Eq. (19), when `APSE` ≤ 0, or the number of $\beta(r_j)$ values to be input, when `APSE` > 0.

> For `PSEL` $= 3$ (a DELR potential), `Nbeta(ISTATE)` is the order of the DELR exponent polynomial of Eq. (8).

> For `PSEL` $= 4$, `Nbeta` is the order of the GPEF polynomial expansion of Eq. (26).

> For all `PSEL` ≥ 5 cases, `qPOT` and `RREFq` are dummy variables, and we omit READs $^\#25$-38.

For $\texttt{PSEL} = 5$, (an HPP potential)

$\quad$ $\texttt{Nbeta}$ is the order of the "X–representation" polynomial expansion of Eq. (31).

For $\texttt{PSEL} = 6$ (a Tang-Toennies potential), set $\texttt{Nbeta} = 9$.

For $\texttt{PSEL} = 7$ with $\texttt{Nbeta} = 5$ generate the HFD-ABC potentials of Eqs. (34)–(36), where
$\quad$ $\beta(i, s)$ for i=1-3 are the α_i parameters of Eq.(36), $\beta(4, s) = \beta_2$ and $\beta(5, s) = \gamma$.

For $\texttt{PSEL} = 7$ with $\texttt{Nbeta} = 2$ generate the HFD-D potential of Eqs. (34), (37) and (38).
$\quad$ with $\beta(1, s) = \beta_2$ and $\beta(2, s) = \gamma$.

For $\texttt{PSEL} = 2$ and $\texttt{APSE(ISTATE)} > 0$, loop over the range $i = 1$ to $N_\beta = \texttt{Nbeta}(s)$, and for the fixed values of the reduced distance $\texttt{yqBETA}$(I,s), read initial trial values of the exponent function values $\beta_i = \texttt{BETA(i)}$ of Eqs. (21), while $\texttt{IFXBETA(I,}s\texttt{)}$ controls whether β_i is to be fitted ($\texttt{IFXBETA} \leq 0$) or held fixed ($\texttt{IFXBETA} > 0$).

$^{\#}24.$ $\texttt{READ(5,*) ypBETA(I,}s\texttt{), BETA(I,}s\texttt{), IFXBETA(I,}s\texttt{)}$

For other $\texttt{PSEL} > 0$ cases, loop over the range $i = 0$ to $N_\beta = \texttt{Nbeta}(s)$ (but $i = 1$ to $N_\beta = \texttt{Nbeta}(s)$ $\texttt{PSEL} = 6$ or 7) and read initial trial values of the expansion parameters $\beta_i = \texttt{BETA(i)}$ of Eqs. (8), (19), (26), (32) or (33).

$\quad$ Skip these loops-over-i READ $^{\#}21$ or $^{\#}21$ statements if $\texttt{PSEL} \leq 0$.

$^{\#}25.$ $\texttt{READ(5,*) (BETA(I,}s\texttt{), IFXBETA(I,}s\texttt{), I=0, Nbeta(}s\texttt{))}$

$\quad$ $\texttt{BETA(I,}s\texttt{)}$ is the initial trial value of the potential exponent expansion coefficients β_i or (if $\texttt{PSEL=6}$) of the GPEF expansion coefficients $c_i = \beta_i$.

$\quad$ $\texttt{IFXBETA(I,}s\texttt{)}$ controls whether β_i is to be fitted ($\texttt{IFXBETA} \leq 0$) or held fixed ($\texttt{IFXBETA} > 0$).

Now, read in parameters specifying the Born-Oppenheimer breakdown radial functions. First, for the effective adiabatic functions $\widetilde{S}_{\mathrm{ad}}^A$ and $\widetilde{S}_{\mathrm{ad}}^B$ of Eqs. (3) and (39).

$^{\#}26.$ $\texttt{READ(5,*) NUA(ISTATE), NUB(ISTATE), qAD(ISTATE), pAD(ISTATE), LRad(ISTATE)}$

$\quad$ $\texttt{NUA} = N_{\mathrm{ad}}^A$ and $\texttt{NUB} = N_{\mathrm{ad}}^B$ are the (integer) orders of the polynomial in $y_{q_{\mathrm{ad}}}(r)$ used to define the effective adiabatic BOB radial function of Eq. (39) for atoms A and B, respectively. If $\texttt{NUA} < 0$ or $\texttt{NUB} < 0$, omit reading the adiabatic BOB terms for that atom; i.e., skip READs $^{\#}25$, $^{\#}27$ and $^{\#}28$, or READs $^{\#}26$, $^{\#}29$ and $^{\#}30$ respectively.
$\quad$ • *Note*: for a chemically homonuclear molecule ($\texttt{AN(1)} = \texttt{AN(2)}$), set $\texttt{NUB} = -1$, as the parameters for atom A are used for both atoms.

$\quad$ $\texttt{QAD}(s) = q_{\mathrm{ad}}$ and $\texttt{PAD}(s) = p_{\mathrm{ad}}$ define the nature of the radial variables $y_{q_{\mathrm{ad}}}(r)$ and $y_{p_{\mathrm{ad}}}(r)$ appearing in Eq. (39). In cases for which $\texttt{NCMM}(s) > 0$ normally set $\texttt{pAD}(s) \geq \texttt{MMLR}(\texttt{NCMM}(s), s)$.

$\quad$ $\texttt{LRad}(s) = 0$ for cases in which the function $\widetilde{S}_{\mathrm{ad}}$ is defined by the simple constrained polynomial of Eq. (39) with u_∞ being a constant.

$\quad$ $\texttt{LRad}(s) = 1$ for cases in which the factor u_∞ in Eq. (39) is to be represented by the expansion of Eq. (42). In this case, values of $\overline{\delta C}_{m_i}^{\mathrm{A/B}}$ must be read in for $i = 1$, $\texttt{NCMM}(s)$, for both atoms.

If $\texttt{NLad}(s) \geq 0$ read in, one per line, values of the isotopologue-*in*dependent parameters $\overline{\delta C}_m(\mathrm{A/B})$ that define how the long-range coefficients $C_m^{\mathrm{A/B},\alpha}$ for atoms 'A' and 'B' differ from one isotopologue to another. For cases in which no values are known, read in values of 0.0.

$^{\#}27.$ $\texttt{IF(NLad(ISTATE} > 0) \texttt{ READ(5,*) dCmVALA((m), m=1,NCMM(ISTATE)}$

$^{\#}28.$ $\texttt{IF(NLad(ISTATE} > 0) \texttt{ READ(5,*) dCmVALB((m), m=1,NCMM(ISTATE)}$

If $\mathtt{NUA}(s) \geq 0$, loop over $i = 0$ to $\mathtt{NUA}(s)$ while reading in initial trial values of parameters $\mathtt{UA}(i, s) = u_i^{\mathrm{A}}$ of Eq. (39) and integer parameter $\mathtt{IFXUA}(i, s)$ to control whether u_i^{A} is to be fitted ($\mathtt{IFXUA}(i, s) \leq 0$) or held fixed ($\mathtt{IFXUA}(i, s) > 0$). Also read $\mathtt{UAinf} = u_\infty^{\mathrm{A}}$, the limiting asymptotic value of the atom-A radial BOB function $\widetilde{S}_{\mathrm{ad}}^{A}(R)$, and integer parameter $\mathtt{IFXUAinf}$ that controls whether $\mathtt{UAinf}$ is to be fitted ($\mathtt{IFXUAinf} \leq 0$) or held fixed ($\mathtt{IFXUAinf} > 0$).

$^{\#}$29. `READ(5,*) UA(I,`s`), IFXUA(I,`s`), I=0, NUA(`s`)`
$^{\#}$30. `READ(5,*) UAinf, IFXUAinf`

If $\mathtt{NUB}(s) \geq 0$, loop over $i = 0$ to $\mathtt{NUB}(s)$ while reading in initial trial values of parameters $\mathtt{UB}(i, s) = u_i^{\mathrm{B}}$ of Eq. (39) and integer parameter $\mathtt{IFXUB}(i, s)$ to control whether u_i^{B} is to be fitted ($\mathtt{IFXUB}(i, s) \leq 0$) or held fixed ($\mathtt{IFXUB}(i, s) > 0$). Also, read $\mathtt{UBinf} = u_\infty^{\mathrm{B}}$, the limiting asymptotic value of the atom-B radial BOB function $\widetilde{S}_{\mathrm{ad}}^{B}(R)$, and integer parameter $\mathtt{IFXUBinf}$ that controls whether $\mathtt{UBinf}$ is to be fitted ($\mathtt{IFXUBinf} \leq 0$) or held fixed ($\mathtt{IFXUBinf} > 0$).

$^{\#}$31. `READ(5,*) (UB(I,`s`), IFXUB(I,`s`), I=0, NUB(`s`))`
$^{\#}$32. `READ(5,*) UBinf, IFXUBinf`

Now, read in parameters defining the centrifugal BOB radial functions $\widetilde{R}_{\mathrm{na}}^{\mathrm{A}}$ and $\widetilde{R}_{\mathrm{na}}^{\mathrm{B}}$ of Eqs. (4) and (40).

$^{\#}$33. `READ(5,*) NTA(ISTATE), NTB(ISTATE), qNA(ISTATE), pNA(ISTATE)`

> $\mathtt{NTA} = N_{\mathrm{na}}^{\mathrm{A}}$ and $\mathtt{NTB} = N_{\mathrm{na}}^{\mathrm{B}}$ are the (integer) orders of the polynomial in $y_{q_{\mathrm{na}}}(r)$ used to define the centrifugal BOB radial function of Eq. (40) for atoms A and B, respectively. If $\mathtt{NTA} < 0$ or $\mathtt{NTB} < 0$, omit the centrifugal BOB terms for that atom; i.e., skip READS $^{\#}$32 and $^{\#}$33, or READS $^{\#}$36 and $^{\#}$37, respectively.
>
> • *Note*: for a chemically homonuclear molecule ($\mathtt{AN}(1) = \mathtt{AN}(2)$), set $\mathtt{NTB} = -1$, as the parameters for atom A are automatically used for both atoms.
>
> $\mathtt{pNA}(s) = p_{\mathrm{na}}$ and $\mathtt{qNA}(s) = q_{\mathrm{na}}$ define the nature of the radial variables $y_{p_{\mathrm{na}}}(r)$ and $y_{q_{\mathrm{na}}}(r)$ appearing in Eq. (40).

If $\mathtt{NTA}(s) \geq 0$, loop over $i = 0$ to $\mathtt{NTA}(s)$ while reading in initial trial values of parameters $\mathtt{TA}(i, s) = t_i^{\mathrm{A}}$ of Eq. (39), and integer parameter $\mathtt{IFXTA}(i, s)$ to control whether t_i^{A} is to be fitted ($\mathtt{IFXTA}(i, s) \leq 0$) or held fixed ($\mathtt{IFXTA}(i, s) > 0$). Also, read $\mathtt{TAinf} = t_\infty^{\mathrm{A}}$, the limiting asymptotic value of the atom-A radial BOB function $\widetilde{R}_{\mathrm{na}}^{\mathrm{A}}(R)$, and the integer parameter $\mathtt{IFXTAinf}$ that controls whether $\mathtt{TAinf}$ is to be fitted ($\mathtt{IFXTAinf} \leq 0$) or held fixed ($\mathtt{IFXTAinf} > 0$).

$^{\#}$34. `READ(5,*) TA(I,`s`), IFXTA(I,`s`)a`

$^{\#}$35. `READ(5,*) TAinf, IFXTAinf`

If $\mathtt{NTB}(s) \geq 0$, loop over $i = 0$ to $\mathtt{NTB}(s)$ while reading in initial trial values of parameters $\mathtt{TB}(i, s) = t_i^{\mathrm{B}}$ of Eq. (40), and integer parameter $\mathtt{IFXTB}(i, s)$ to control whether t_i^{B} is to be fitted ($\mathtt{IFXTB}(i, s) \leq 0$) or held fixed ($\mathtt{IFXTB}(i, s) > 0$). Also read $\mathtt{TBinf} = t_\infty^{\mathrm{B}}$, the limiting asymptotic value of the atom-B radial BOB function $\widetilde{R}_{\mathrm{na}}^{\mathrm{B}}(R)$, and the integer parameter $\mathtt{IFXTBinf}$ which controls whether $\mathtt{TBinf}$ is to be fitted ($\mathtt{IFXTBinf} \leq 0$) or held fixed ($\mathtt{IFXTBinf} > 0$).

$^{\#}$36. `READ(5,*) (TB(I,`s`), IFXTB(I,`s`), I=0, NTB(`s`))`

$^{\#}$37. `READ(5,*) TBinf, IFXTBinf`

Finally, if the electronic angular momentum quantum number $\mathtt{IOMEG}(s)$ is not zero, read in parameters defining the radial Λ-doubling or $^2\Sigma$ spin-splitting radial strength function of Eqs. (50) or (54), respectively.

$^{\#}$38. `READ(5,*) NwCFT(`s`), Pqw(`s`), efREF(`s`)`

NwCFT is the order of the polynomial defining the radial strength function $f_\Lambda(r)$ or $f_\Sigma(r)$.
If NwCFT$(s) < 0$, omit READ $^{\#}37$.

Pqw is the integer q_Λ or q_Σ defining the radial expansion variable $y_{q_\Lambda}^{re}(r)$ in Eq. (50) or $y_{q_\Sigma}^{re}(r)$ in Eq. (54).

efREF : specifies the choice of reference (or zero-shift) parity levels for Λ–doubling splittings. Select them as the f–parity sublevels when efREF $= -1$, the e–parity sublevels when efREF $= +1$, or their mid-point when efREF $= 0$; e.g., for a $^1\Pi$ state, if efREF $= -1$ the f–levels are treated as unperturbed and the e–levels shifted by $+ q_B(v)[J(J + 1)]$, ... etc. For $^2\Sigma$ splitting, efREF is a dummy parameter.

If NwCFT$(s) \geq 0$, loop over $i = 1$ to NwCFT(s) while reading in initial trial values of parameters wCFT$(i, s) = w_i$ of Eq. (50) or (54), and integer parameter IFXwCFT(i, s) to control whether thia s w_i is to be fitted (IFXwCFT$(i, s) \leq 0$) or held fixed (IFXwCFT$(i, s) > 0$).

$^{\#}$39. READ(5,*) wCFT(I,s), IFXwCFT(I,s)

To Generate predictions from a set of fixed system parameters.

Loop over READ $^{\#}40$ for all bands for which predictions are desired, and for each one read the following. Input stops at the end of the data file or if the read-in value of VP for a band is negative.

$^{\#}$40. READ(5,*) VP(IBAND),VPP(IBAND), LABLP,LABLP, MN1, MN2, PP, PPP,
   1   JMAXX,J2DL,J2DU,J2DD

VP(IBAND) & VPP(IBAND) : are the vibrational indices v' and v'', respectively, labeling that band. Set $v' =$ VP(IBAND) < 0 to indicate the end of the prediction data set,

LABLP & LABLPP : are the three-alphanumeric-character names (enclosed in single quotes, as in 'X1S') identifying the upper and lower electronic state for which the predicted band is to be generated. Set LABLP $=$ LABLPP to generate IR or MW transitions for that state.

MN1 & MN2 : are the integer mass numbers identifying the isotopologue for which the predictions are to be generated.

PP & PPP : are integers with values $+1$, 0 or -1 to indicate the e/f parity of the upper ($p' =$ PP) and lower ($p'' =$ PPP) levels in the band.

JMAXX, J2DL, J2DU & J2DD : generate predicted transition energies for $J'' = 0$ to JMAXX , subject to the selection rule that $\Delta J = J' - J''$ runs from J2DL to J2DU in steps of J2DD .

Appendix D: Illustrative Application: Determine an MLR Potential for $X\,^1\Sigma_g^+$-State Ca_2

This program has been successfully applied to a wide variety of problems [10, 11, 13, 14, 16–22, 24, 25, 41, 43, 85, 108, 112]. The present subsection briefly describes one of these cases and presents the associated Channel–5 'instruction' input data file and portions of the Channel-6 main output file.

Appendix D.1 Input Instruction Data File for the Ground $X\,^1\Sigma_g^+$ State of Ca_2

This is the case described in Refs. [17] and [21], in which 3553 fluorescence series transitions from 180 levels of the $B\,^1\Sigma_u^+$ state of Ca_2 into 924 levels of the $X\,^1\Sigma_g^+$ state spanning 99.97% of its well [46, 47] are fitted to an MLR potential whose long-range tail is defined by accurate fixed theoretical C_8, and C_{10} values, but with the C_6 coefficient treated as a free parameter. The following input 'instruction file' is for the fit to the updated version of the model described in Ref. [21]. The Channel-4 file containing the 3553 experimental data used in this analysis may be obtained from the author on request.

```
 20 20  0  1  1  0  0            % AN(1) AN(2) CHARGE NISTP NSTATES LPRINT PRINP
 '../Ca2_X_FS.4'                 % Name of dPotFit input data file
 'N7q4C6rr5_55'                  % Writefile
  40   40                        % MN1(1) MN2(1)
 1.d0 1 -0   0   30   0          % UCUTOFF NOWIDTHS IROUND ROBUST CYCMAX uBv
 'X1S'  0   0 38 300 0           %(1) SLABL IOMEG VMIN VMAX JTRUNC EFSEL
 2    0.d0    1   0   20         % PSEL VLIM MAXMIN BOBCN OSEL
 3.00 99.0   0.0005d0            % RMIN RMAX RH

  3   0.5870d0  -2   1           % NCMM rhoAB IVSR IDSTT
  6   1.04600000D+07     0       % MMLR(1) CMVAL(1) IFXCm(1)
  8   3.06080000D+08     1       % MMLR(2) CMVAL(2) IFXCm(2)
 10   8.34400000D+09     1       % MMLR(3) CMVAL(3) IFXCm(3)

   1.102080000000D+03   0
   4.277820000000D+00   0

   7 0    3  5  5.55D+00         % NSpow NLpow pPOT qPOT RREF
  -1.4672501   0
  -0.20012     0
  -0.72633     0
   0.1229      0
  -0.2763      0
   0.506       0
   0.357       0
   1.022       0

  -6 -6   3  3    0              % NUA NUB  qAD pAD  LRad
  -6 -6   3  3                   % NTA NTB  qNA pNA
```

Appendix D.2 Main Output File for an MLR Fit to Ground($X\,{}^1\Sigma_g^+$)-State Ca_2

Note that the value of DRMSD $= \overline{dd}$ reported at the end of the first cycle of fit is very large ($\sim 10^6$), while those for the next two cycles are small and identical to one another. This behaviour arises because the reported value for each cycle is calculated using the trial parameters available at the *beginning* of the cycle, and at the beginning of the very first cycle the trial values of all fluorescence series origins were zero. In this initial phase of a fit, the code fits only those parameters w.r.t. which the data are are strictly linearly dependent, while holding all PEF parameters fixed. As for any truly linear least-squares fit, one cycle gives the final result, but in our case an additional (third) cycle is then run because the convergence criterion requires the $\overline{dd}$ values in two consecutive cycles to agree to 1 part in 10^7,

The above behaviour of occurs for all fits involving fluorescence series data, or in which the energy levels for one or more states are represented by term values or band constants. Following that initial 3-cycle convergence, the full fit in which *all* parameters are free commences, and it then continues until it too converges fully.

A prominent feature of the output shown below is the nest of 48 lines of warning messages immediately preceding the " After Cycle # 1 DRMSD=" message. These indicate that 18 of the reported transitions involve five lower-state levels that the code was unable to locate, and hence that those 48 lines were omitted from the analysis. In the present case those five lower levels in are very high-J quasibound levels lying near the centrifugal barrier maximum for which the procedure for generating initial trial values seems to have failed. Such problems are not encountered often, but when they do, the omission of the associated handful of lines (here 18 out of 3553) has little effect on the fit. While these warnings occur on every cycle of the current fit, they are only show once below in order to make the structure of the rest of the results easier to see.

```
 Input data for  1  isotopologues(s)
 *******************************
     Isotopologues    Mass of atom-1   Mass of atom-2    Reduced mass
     -----------------   --------------   --------------   --------------
  Ca( 40) - Ca( 40)      39.962590864      39.962590864     19.981295432

 Use experimental data input file:  Ca2_Robbie.4
 Uncertainties for transitions involving quasibound levels modified to:
                  SQRT{(u(i;exp)**2 + ( 0.20*width)**2}
 For electrically homonuclear molecules, BO correction functions are the same
    for both atoms, so only the first sets of correction coefficients
    UA(s) and TA(s) are used, and the mass scaling factors are sums over
    the two individual atoms.
 Fit uses standard  1/[uncertainty(i)]**2  data weighting
 Non-linear fits are allowed a maximum of  CYCMAX=  25 cycles

 For state   X0
    integrate from   RMIN= 3.00   to   RMAX= 48.00   with mesh   RH= 0.00050
 ================================================================================

 State X0  represented by an MLR(q= 3, p= 5) potential defined in terms of
 ========   exponent coefficient:  beta(R)= betaINF*y5 +(1-y5)*Sum{beta_i*y3^i}
                                                      for  i= 0 to  7
        with radial variable:   y_{p,q} = (R^q - 5.550000^q)/(R^q + 5.550000^q)
    uLR inverse-power terms incorporate NO damping functions
                                         C6= 1.0460000D+07[cm-1 Ang^6]
                                         C8= 3.0608000D+08[cm-1 Ang^8]
                                         C10= 8.3440000D+09[cm-1 Ang^{10}]
               These constants yield:   betaINF= -1.3499558834
   Parameter    Initial Value   Uncertainty   Sensitivity
     VLIM       0.000000000000D+00      --          --
      De        1.102080000000D+03   0.0D+00      0.0D+00
```

```
      Re        4.277820000000D+00      0.0D+00          0.0D+00
     C 6        1.046000000000D+07      0.0D+00          0.0D+00
     C 8        3.060800000000D+08        --               --
     C10        8.344000000000D+09        --               --
  beta( 0)     -1.467250100000D+00      0.0D+00          0.0D+00
  beta( 1)     -2.001200000000D-01      0.0D+00          0.0D+00
  beta( 2)     -7.263300000000D-01      0.0D+00          0.0D+00
  beta( 3)      1.229000000000D-01      0.0D+00          0.0D+00
  beta( 4)     -2.763000000000D-01      0.0D+00          0.0D+00
  beta( 5)      5.060000000000D-01      0.0D+00          0.0D+00
  beta( 6)      3.570000000000D-01      0.0D+00          0.0D+00
  beta( 7)      1.022000000000D+00      0.0D+00          0.0D+00
  beta_INF     -1.349955883434D+00        --               --
  C11{exp}     -6.107145366294D+10        --               --
==============================================================================

Neglect data with:  Uncertainties > UCUTOFF=  1.00D+00 (cm-1)
      and State X0  data with  J > JTRUNC= 300
                          or  v  outside range  0  to  38    for   ISOT= 1
State X0   Energy Convergence criterion EPS is 1.0D-06 cm-1

Search for v=  8   J=162  starting from  E=    299.89 finds  E(v= -1)=    308.57
*** SCECOR failed   1 times,   Currently Seeking   v=  8, J=162;  Found   v= -1
*** FAIL to find level(X0 )    v=  8   J=162  so ignore  YOBS(   85)=  19370.3320
Search for v=  9   J=160  starting from  E=    298.10 finds  E(v= -1)=    295.27
*** SCECOR failed   2 times,   Currently Seeking   v=  9, J=160;  Found   v= -1
*** FAIL to find level(X0 )    v=  9   J=160  so ignore  YOBS(   86)=  19373.9930
Search for v=  8   J=162  starting from  E=    299.89 finds  E(v= -1)=    308.57
*** SCECOR failed   3 times,   Currently Seeking   v=  8, J=162;  Found   v= -1
*** FAIL to find level(X0 )    v=  8   J=162  so ignore  YOBS(  154)=  19406.1530
Search for v=  9   J=158  starting from  E=    280.50 finds  E(v= -1)=    288.24
*** SCECOR failed   4 times,   Currently Seeking   v=  9, J=158;  Found   v= -1
*** FAIL to find level(X0 )    v=  9   J=158  so ignore  YOBS(  165)=  19354.4750
Search for v= 17   J=120  starting from  E=    124.90 finds  E(v= -1)=    131.05
*** SCECOR failed   5 times,   Currently Seeking   v= 17, J=120;  Found   v= -1
*** FAIL to find level(X0 )    v= 17   J=120  so ignore  YOBS(  474)=  20531.1610
Search for v= 21   J=102  starting from  E=     81.01 finds  E(v= -1)=     80.83
*** SCECOR failed   6 times,   Currently Seeking   v= 21, J=102;  Found   v= -1
*** FAIL to find level(X0 )    v= 21   J=102  so ignore  YOBS(  904)=  18444.0030
Search for v= 21   J=102  starting from  E=     81.01 finds  E(v= -1)=     80.83
*** SCECOR failed   7 times,   Currently Seeking   v= 21, J=102;  Found   v= -1
*** FAIL to find level(X0 )    v= 21   J=102  so ignore  YOBS( 1253)=  18444.0000
Search for v=  8   J=162  starting from  E=    299.89 finds  E(v= -1)=    308.57
*** SCECOR failed   8 times,   Currently Seeking   v=  8, J=162;  Found   v= -1
*** FAIL to find level(X0 )    v=  8   J=162  so ignore  YOBS( 1962)=  19370.3340
Search for v=  9   J=160  starting from  E=    298.10 finds  E(v= -1)=    295.27
*** SCECOR failed   9 times,   Currently Seeking   v=  9, J=160;  Found   v= -1
*** FAIL to find level(X0 )    v=  9   J=160  so ignore  YOBS( 1963)=  19373.9660
Search for v=  8   J=162  starting from  E=    299.89 finds  E(v= -1)=    308.57
*** SCECOR failed  10 times,   Currently Seeking   v=  8, J=162;  Found   v= -1
*** FAIL to find level(X0 )    v=  8   J=162  so ignore  YOBS( 2030)=  19406.1440
Search for v= 21   J=102  starting from  E=     81.01 finds  E(v= -1)=     80.83
*** SCECOR failed  11 times,   Currently Seeking   v= 21, J=102;  Found   v= -1
*** FAIL to find level(X0 )    v= 21   J=102  so ignore  YOBS( 2142)=  18444.0430
Search for v= 21   J=102  starting from  E=     81.01 finds  E(v= -1)=     80.83
*** SCECOR failed  12 times,   Currently Seeking   v= 21, J=102;  Found   v= -1
*** FAIL to find level(X0 )    v= 21   J=102  so ignore  YOBS( 2438)=  18444.0500
Search for v= 18   J=116  starting from  E=    115.09 finds  E(v= -1)=    119.39
*** SCECOR failed  13 times,   Currently Seeking   v= 18, J=116;  Found   v= -1
*** FAIL to find level(X0 )    v= 18   J=116  so ignore  YOBS( 2939)=  19345.3040
Search for v= 18   J=116  starting from  E=    115.09 finds  E(v= -1)=    119.39
*** SCECOR failed  14 times,   Currently Seeking   v= 18, J=116;  Found   v= -1
*** FAIL to find level(X0 )    v= 18   J=116  so ignore  YOBS( 3138)=  19345.3200
Search for v= 18   J=116  starting from  E=    115.09 finds  E(v= -1)=    119.39
*** SCECOR failed  15 times,   Currently Seeking   v= 18, J=116;  Found   v= -1
*** FAIL to find level(X0 )    v= 18   J=116  so ignore  YOBS( 3249)=  19345.3200
Search for v= 18   J=116  starting from  E=    115.09 finds  E(v= -1)=    119.39
*** SCECOR failed  16 times,   Currently Seeking   v= 18, J=116;  Found   v= -1
*** FAIL to find level(X0 )    v= 18   J=116  so ignore  YOBS( 3383)=  19345.3030
```

```
After Cycle # 1:  DRMSD= 2.6142116D+06    test(PS)= 2.6D+02    test(PU)= 4.3D+00
After Cycle # 2:  DRMSD= 6.1891004D-01    test(PS)= 5.0D-07    test(PU)= 8.5D-09
After Cycle # 3:  DRMSD= 6.1891004D-01    test(PS)= 9.3D-08    test(PU)= 1.6D-09

Full  3-cycle convergence:  {ABS(RMSR/RMSRB)-1}= 5.73D-13  TSTPS= 9.3D-08

After Cycle # 1:  DRMSD= 6.1891004D-01    test(PS)= 1.1D+04    test(PU)= 2.7D-01
After Cycle # 2:  DRMSD= 6.2197176D-01    test(PS)= 3.0D+02    test(PU)= 4.1D-02
After Cycle # 3:  DRMSD= 6.1900333D-01    test(PS)= 5.9D+01    test(PU)= 9.1D-03
After Cycle # 4:  DRMSD= 6.1886942D-01    test(PS)= 1.3D+01    test(PU)= 2.0D-03
After Cycle # 5:  DRMSD= 6.1886276D-01    test(PS)= 2.9D+00    test(PU)= 4.5D-04
After Cycle # 6:  DRMSD= 6.1886243D-01    test(PS)= 6.5D-01    test(PU)= 1.0D-04
After Cycle # 7:  DRMSD= 6.1886241D-01    test(PS)= 1.4D-01    test(PU)= 2.3D-05

Full  7-cycle convergence:  {ABS(RMSR/RMSRB)-1}= 2.65D-08  TSTPS= 1.4D-01

 3553 data fit to  191 param. yields  DRMS(devn)= 6.1886241D-01  tst(PS)= 1.4D-01
  PV(   1) =  1.10207930826358D+03 (+/- 4.0D-03)    PS= 3.6D-06   PC=  1.3D-10
  PV(   2) =  4.27781557316645D+00 (+/- 1.9D-05)    PS= 1.4D-08   PC= -4.2D-10

.......... omit 189 lines to save space ...................

  PV( 192) =  2.01725642628215D+04 (+/- 8.7D-03)    PS= 7.1D-05   PC= -2.2D-10
  PV( 193) =  1.96650081930497D+04 (+/- 4.4D-03)    PS= 2.1D-05   PC= -1.8D-10

=========================================================================
Fitting  191  free parameters to  3553  transitions yields  DSE= 0.63619882
=========================================================================


=========================================================
 The following  180 Fluorescence Series Origins were determined
 ----------------------------------------------------------------
 ( v', J', p'; ISTP)    T(value)    Uncertainty  Sensitivity
 ----------------------------------------------------------------
 (   5, 67, +1; 1)    18847.271302    4.6D-03      2.6D-05
 (   6, 97, +1; 1)    19247.654565    5.7D-03      2.2D-05
 (   2,161, +1; 1)    19666.758095    1.0D-02      3.2D-05

.......... omit 174 lines to save space ...................

 (  12, 91, +1; 1)    19895.675096    6.1D-03      5.0D-05
 (  11,125, +1; 1)    20172.564263    8.7D-03      7.1D-05
 (  12, 63, +1; 1)    19665.008193    4.4D-03      2.1D-05
=========================================================================

State X0  represented by an MLR(q= 3, p= 5) potential defined in terms of
========  exponent coefficient:  beta(R)= betaINF*y5 +(1-y5)*Sum{beta_i*y3^i}
                                              for  i= 0 to  7
        with radial variable:   y_{p,q} = (R^q - 5.550000^q)/(R^q + 5.550000^q)
   uLR inverse-power terms incorporate NO damping functions
                                      C6= 1.0454962D+07[cm-1 Ang^6]
                                      C8= 3.0608000D+08[cm-1 Ang^8]
                                      C10= 8.3440000D+09[cm-1 Ang^{10}]
                 These constants yield:   betaINF= -1.3498686606
   Parameter      Final Value    Uncertainty   Sensitivity
     VLIM       0.000000000000D+00      --           --
      De        1.102079308264D+03    4.0D-03      3.6D-06
      Re        4.277815573166D+00    1.9D-05      1.4D-08
     C 6        1.045496176200D+07    3.2D+04      6.6D-01
     C 8        3.060800000000D+08      --           --
     C10        8.344000000000D+09      --           --
    beta( 0)   -1.467382773954D+00    8.4D-04      1.2D-08
    beta( 1)   -2.005944207706D-01    3.2D-03      5.3D-08
    beta( 2)   -7.272203056978D-01    6.0D-03      1.1D-07
    beta( 3)    1.217583441954D-01    8.9D-03      2.4D-07
    beta( 4)   -2.771089629303D-01    7.0D-03      4.5D-07
    beta( 5)    5.036641109431D-01    1.5D-02      9.3D-07
    beta( 6)    3.515368692379D-01    3.6D-02      1.7D-06
```

```
  beta( 7)   1.019354103000D+00    5.2D-02         3.4D-06
  beta_INF  -1.349868660575D+00      --              --
  C11{exp}  -6.062033969002D+10      --              --
================================================================================


   ==============================================================
   *** Discrepancies for  180 bands/series of Ca( 40)-Ca( 40) ***
   ==============================================================


   ================================================================================
   3553 Fluorescence transitions into State X0   Ca( 40)-Ca( 40) in  180 series
   ================================================== Avge. ===============
    v'  j' p' #data  v"min  v"max  AvgeUnc  Max.Unc.   Err/Unc   DRMSD
   -----------------------------------------------------------------------
     5  67  e    33      0     22  8.3D-03  1.2D-02   -0.02059   0.650  as per Mol.P. 105,663(2007)
     6  97  e    37      0     21  7.6D-03  1.5D-02    0.03076   0.492  as per Mol.P. 105,663(2007)
     2 161  e    16      1      9  8.8D-03  5.0D-02   -0.00001   0.338  as per Mol.P. 105,663(2007)

   .......... omit 174 lines to save space ....................

    12  91  e    12      0      9  1.1D-02  1.5D-02   -0.17667   0.753  as per Mol.P. 105,663(2007)
    11 125  e     9      0      8  1.2D-02  1.5D-02    0.19259   0.978  as per Mol.P. 105,663(2007)
    12  63  e    40      0     21  8.3D-03  1.8D-02    0.00945   0.600  as per Mol.P. 105,663(2007)
   ----------------- For these  3553 lines, overall:   -0.00371   0.619

   ================================================================================
   Fit of   191 total param to  3553 data yields   DRMS(devn.)= 0.61886241
   ================================================================================
```

Appendix E: Application to Determine an MLR PEF and BOB functions for $X\,^2\Sigma^+$-State MgH

This case involves the simultaneous analysis of data for six isotopologues, formed by ^{1}H and ^{2}H with ^{24}Mg, ^{25}Mg, and ^{26}Mg, in a $^2\Sigma^+$ electronic state. Thus, as well as determining the main 'mechanical' potential energy function itself, it was possible to determine potential-energy ('adiabatic') and centrifugal BOB radial strength functions, as well as the Λ-doubling radial strength function whose radial expectation values are the conventional v-dependent $^2\Sigma$splitting parameters.

This is the case described in Ref. [108], in which a fit to microwave, infrared and electronic data for the $X\,^2\Sigma^+$ and $A\,^1\Sigma^+$ states of the ^{24}Mg, ^{25}Mg and ^{26}Mg isotopologues of MgH and MgD was used to determine an MLR(12) (i.e., $N_\beta = 12$) potential energy function, as well as adiabatic (potential energy) and centrifugal BOB functions, and a radial strength function to characterize the $^2\Sigma$ doublet splittings. The fact that the data span almost the entire potential well while the long-range part of the potential is important only at energies very close to dissociation is the reason for the relatively high orders of the MLR exponent polynomial and the expansions for the BOB functions. The Channel-4 input file of the 20 103 experimental data used in this analysis may be obtained in a link from the HTML version of Ref. [108] on the Journal's www site, or from the author. In this case, the energy levels of all three excited states are represented by independent term values.

```
   12   1  0  6  4   4 0            % AN1 AN2 CHARGE NISTP NSTATES LPRINT PRINP
   '../data_MgHD_2012-01-27.4'      % Name of input data file
   'tst_5412rref2_74'              % WRITEFILE
   24   1                          %  MN1 MN2
   25   1
   26   1
   24   2                          %  MN1 MN2
   25   2
   26   2
   0.5  1  0  0  30  0             % UCUTOFF NOWIDTHS IROUND ROBUST CYCMAX uBv
   'X1S'  -1  0 -11  99  0         % SLABL IOMEG VMAX JTRUNC   EFSEL
    11 10 10 15 11 11              % VMAX(istp), ISTP= 1,6
    2 0.000  1  0    20            % PSEL VLIM MAXMIN BOBCN OSEL
   0.60  99.50  0.0025             % RMIN RMAX RH
   3 0.81 -2  1                    % NCMM rhoAB IVSR IDSTT
   6 2.77550000D+05     1          % MMLR, CmVAL, IFXCm
   8 3.45490000D+06     1          % MMLR, CmVAL, IFXCm
  10 4.61400000D+07     1          % MMLR, CmVAL, IFXCm

   1.110424552400D+04  0           % De   IFXDe
   1.729685380255D+00  0           % Re   IFXRe

   12 -1  4  5    2.74D+00         % Nbeta APSE qPOT pPOT RREF
    1.170476601132D+00  0           % BETA 0   IFXBETA 0
    1.080170651332D+00  0           % BETA 1   IFXBETA 1
    2.673340658164D+00  0           % BETA 2   IFXBETA 2
    2.483625825973D+00  0           % BETA 3   IFXBETA 3
    7.400354782870D-01  0           % BETA 4   IFXBETA 4
    1.921267534841D-01  0           % BETA 5   IFXBETA 5
    6.071160601044D-01  0           % BETA 6   IFXBETA 6
   -2.486057829458D+00  0           % BETA 7   IFXBETA 7
   -7.676067408675D+00  0           % BETA 8   IFXBETA 8
```

```
   -5.742858200526D+00  0              % BETA 9   IFXBETA 9
    2.850001493731D+00  0              % BETA10   IFXBETA10
    6.076990710549D+00  0              % BETA11   IFXBETA11
    2.409626277526D+00  0              % BETA12   IFXBETA12

    3 13   4  6    0                   % NUA NUB  qAD pAD   LRad
    1.300728351881D+00  0              % UA 0   IFXUA 0
    4.148196460337D+00  0              % UA 1   IFXUA 1
    2.612257669548D+00  0              % UA 2   IFXUA 2
    4.341067654483D+00  0              % UA 3   IFXUA 3
    0.000000000000D+00  1              % uAinf   IFXuAinf

   -1.517859106465D+01  0              % UB 0   IFXUB 0
    3.753517600616D+01  0              % UB 1   IFXUB 1
   -1.153600397884D+01  0              % UB 2   IFXUB 2
    3.213195588821D+01  0              % UB 3   IFXUB 3
    5.866458845349D+01  0              % UB 4   IFXUB 4
    2.347347913310D+02  0              % UB 5   IFXUB 5
   -1.543302677803D+03  0              % UB 6   IFXUB 6
   -2.076874526486D+03  0              % UB 7   IFXUB 7
    1.249385587088D+04  0              % UB 8   IFXUB 8
    5.664228044773D+03  0              % UB 9   IFXUB 9
   -4.993237165924D+04  0              % UB10   IFXUB10
    1.007715179054D+04  0              % UB11   IFXUB11
    7.172546964411D+04  0              % UB12   IFXUB12
   -4.723470281067D+04  0              % UB13   IFXUB13
    0.000000000000D+00  1              % uBinf   IFXUBinf

   -3  7   4  4    0                   % NTA NTB  qNA pNA
    0.000000000000D+00  1              % TB 0   IFXTB 0
    7.250023689953D-04  0              % TB 1   IFXTB 1
    2.862180763657D-04  0              % TB 2   IFXTB 2
    9.395028124006D-04  0              % TB 3   IFXTB 3
    2.804097571256D-03  0              % TB 4   IFXTB 4
   -1.917529464250D-03  0              % TB 5   IFXTB 5
   -6.715122481617D-03  0              % TB 6   IFXTB 6
    2.263936298448D-02  0              % TB 7   IFXTB 7
    0.000000000000D+00  1              % tBinf   IFXTBinf

    3  4  0                            % NwCFT Pqw efREF
    4.619303623412D-03  0              % wCFT 0   IFXwCFT 0
   -3.135099829025D-03  0              % wCFT 1   IFXwCFT 1
    2.688086629368D-04  0              % wCFT 2   IFXwCFT 2
   -4.537525776361D-03  0              % wCFT 3   IFXwCFT 3

   'A1' 0 0 99 99 0                    % SLABL IOMEG VMAX JRUNC EFSEL
    -2  0.1  1  0 20                   % PSEL VLIM MAXMIN BOBCN OSEL
   'A2' 0 0 99 99 0                    % SLABL IOMEG VMAX JRUNC EFSEL
    -2  0.1  1  0 20                   % PSEL VLIM MAXMIN BOBCN OSEL
   'B0' 0 0 99 99 0                    % SLABL IOMEG VMAX JRUNC EFSEL
    -2  0.1  1  0 20                   % PSEL VLIM MAXMIN BOBCN OSEL
```

Appendix F. Schematic Description of Program Structure

The present section lists the names and outlines the functions of the various subroutines used by **dPotFit**, and indicates their hierarchy. In particular, the level of indentation in this list indicates which subroutines call which others.

DPOTFIT : The main program that reads the input data characterizing the molecular system and the type of fit to be performed, calls the actual fitting routines, and prints descriptions of the system and the results.

  MASSES : A data subroutine containing the best current [96] values of the atomic masses (and other properties) of all stable atomic isotopes. Its presence obviates the need for a user to look up and type precise particle masses into the input data file.

  READPOT : The subroutine that reads in and organizes the parameters and properties of the potential energy, BOB, and (as appropriate) Λ-doubling or $^2\Sigma$ splitting radial strength functions to be used for a particular electronic state.

  WRITEPOT : Writes out a complete listing of the parameters defining the potential energy and other radial strength functions in the Hamiltonian. When called at the end of a run, it also lists the uncertainties and sensitivities of the fitted parameters.

  READATA : The data input subroutine for reading in, arranging, and characterizing the experimental data to be used in the fit.

  TVSORT : If the vibration-rotation levels of one or more states are to be represented by independent term values, sorts through global data file, and for each isotopomer in state ISTATE: (i) finds the number of transitions coupled to each level (v, J, p), (ii) for levels ordered by increasing v, J, and p, add a free parameter for each level involved in one or more transitions, and (iii) label each transition involving one of these levels by the index/counter of the parameter associated with that term value.

  PREPOT : A subroutine package, described in detail in the manual for program **LEVEL** [4, 83], that can generate a potential function by interpolating over a set of input points. This allows one to run a 'forward' calculation (without performing a fit) to test the ability of a pointwise potential function obtained from some other source to describe a given data set.

  MAPPAR : A subroutine to convert the 'external' parameters characterizing the potential energy and other radial functions into the 'internal' parameter array $\{$PV(i)$\}$ used by the fitting subroutine package NLLSSRR, and vise versa.

  NLLSSRR : A general non-linear (or linear) least-squares fitting subroutine package that can also implement (when input parameter IROUND$\neq 0$) the "sequential rounding and refitting" procedure of Ref. [97].

    DYIDPJ : The key user-supplied subroutine required by NLLSSRR, that for each datum returns the calculated value and partial derivatives with respect to all parameters of the model for the current set of trial parameters.

      VGEN : Prior to each cycle of fit, this subroutine updates the arrays of values of the potential energy functions and of their partial derivatives w.r.t. their defining parameters.

Scalc when using the Pashov cubic spline approach to represent the MLR exponent-coefficient function $\beta(r)$, this routine returns values of the $S_n(x)$ partial derivative functions.

Lkoef When using the Pashov cubic spline approach to represent the MLR exponent-coefficient function, on calling this function subroutine with a list of 'n' spline ordinate values, it returns the array of 'rKL' coefficients used for generating the 'n' $S_n(x)$ spline coefficient functions.

INITDD: At the beginning of each cycle of fit, this subroutine updates the sets of band constants employed for generating the trial eigenvalues required by the Schrödinger solver subroutine SCHRQ.

ALF: This (**A**utomatic **L**evel **F**inder) subroutine automatically determines the eigenvalues of all specified vibrational levels supported by a given potential energy curve.

SCHRQ: This is the core Schrödinger solver subroutine that provides the eigenvalue and radial wavefunctions required for the data simulation and partial derivative calculations. A number of features of this routine are described in the manual for program **LEVEL** [4, 83].

SCECOR: If the normal level-finding method fails, SCECOR uses semiclassical dE/dv estimates to attempt to hone in on the correct level energy .

CDJOEL: For a given vibrational level, this subroutine calculates the inertial rotational constant and the first six centrifugal distortion constants. These band constants are used to generate the trial eigenvalues required by SCHRQ for each datum.

DEDP: For each level energy associated with each datum, this routine generates and returns the eigenvalue and its partial derivatives with respect to the various parameters required for the least-squares fit.

DWDP: For each tunneling predissociation level width in the data set, this subroutine returns values of its partial derivatives with respect to the parameters of the Hamiltonian.

locateTP: For a predissociation level-width calculation, this routine determines the classical turning points required for the phase-integral calculations (to machine precision!).

PhaseIntegral: Evaluates the phase integrals required for calculating the tunneling predissociation level widths and their partial derivatives with respect to parameters of the Hamiltonian. '

DVIRDP: For each virial coefficient value in the data set, this subroutine generates the corresponding value implied by the current potential function, and its partial derivatives w.r.t. the various potential function parameters.

VPENp for virial coefficient calculations, generates and returns values of the potential function and its first two radial derivatives, and of all of their partial derivatives w.r.t. the potential parameters, at 8 quadrature points on each call. Currently only coded for the case of an MLR potential.

FUNUNC: On completion of a fit, this subroutine generates and writes to Channels 10–16

arrays of the values *and* the fully correlated 95% confidence uncertainties in the various radial functions determined from the fit.

DIFFSTATS : On completion of a fit, summarizes dimensionless RMS deviations for the entire data set on a band-by-band basis.

MKPREDICT : If the input parameter (see READ #2) DATAFILE is equal to 'MAKEPRED', this subroutine reads input to cause the program to prediction transitions for specified bands which are generated from by a complete fixed set of molecular-state parameters.