CP-667 9 April 2013

DPotFit **2.0**
A Computer Program for Fitting Diatomic Molecule Spectral Data to Potential Energy Functions

Robert J. Le Roy,* Jenning Y. Seto and Yiye Huang

Guelph-Waterloo Centre for Graduate Work in Chemistry
University of Waterloo, Waterloo, Ontario N2L 3G1, Canada
**e-mail: leroy@UWaterloo.ca*

University of Waterloo

Chemical Physics Research Report

DPotFit **2.0**
A Computer Program for Fitting Diatomic Molecule Spectral Data to Potential Energy Functions

Robert J. Le Roy, Yiye Huang and Jenning Seto

Guelph-Waterloo Centre for Graduate Work in Chemistry
University of Waterloo, Waterloo, Ontario N2L 3G1, Canada
Electronic mail: leroy@UWaterloo.ca

This manual 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, to determine analytic potential energy functions defining the observed levels of each state. Four families of analytical potential functions are available in the current version of DPotFit: the Expanded Morse Oscillator (EMO) potential, the the Morse/Long-Range (MLR) potential, the Double Exponential/Long-Range (DELR) potential, and Šurkus' Generalized Potential Energy Function (GPEF), which incorporates a variety of polynomial potential forms. DPotFit also allows the fit 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 of the 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 which must be retained when rounding, 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. The program can also use a set of read-in constants to make predictions and calculate deviations $[y_i^{\rm calc} - y_i^{\rm obs}]$ for any chosen input data set.

1 General – The Radial Hamiltonian

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 some parameterized analytic potential energy function. This effective radial Hamiltonian may also include radial strength functions characterizing the atomic-mass-dependent adiabatic and non-adiabatic 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 eigenvalue differences with respect to the parameters defining the potential energy and other radial functions are then used in least-squares fits to determine an optimized radial Hamiltonian for the system. This report describes a robust and flexible computer program for performing this type of analysis that may be downloaded freely from the www site `http://leroy.uwaterloo.ca/programs/`.

As in most direct-potential-fit (DPF) data analyses reported to date, the present code is based on an effective radial Schrödinger equation derived by Watson [1, 2], in which atomic-mass-dependent nonadiabatic contributions to the kinetic energy operator are incorporated into an effective "adiabatic" contribution to the electronic potential energy function and into the non-adiabatic BOB contribution to the effective centrifugal potential of the rotating molecule. Following the conventions of Refs. [3, 4, 5], 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

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

Here, $V_{\rm ad}^{(1)}(r)$ is the total electronic internuclear potential for the chosen reference isotopologue (labeled $\alpha = 1$), $\Delta V_{\rm ad}^{(\alpha)}(r)$ is the *difference* between the effective adiabatic potentials for isotopologue α and that for the reference species ($\alpha = 1$), $g^{(\alpha)}(r)$ is the non-adiabatic centrifugal potential correction function for isotopologue α, and μ_α is Watson's "charge-modified reduced mass" [1]:

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

in which `CHARGE` is the net ±(integer) charge on the molecule, $m_{\rm e}$ is the electron mass, and $M_{\rm A}^{(\alpha)}$ and $M_{\rm B}^{(\alpha)}$ are the isotope masses of the neutral atoms A and B forming isotopologue α of species A–B$^{\texttt{CHARGE}}$. Each of $\Delta V_{\rm ad}^{(\alpha)}(r)$ and $g^{(\alpha)}(r)$ can be expressed as the sum of two terms, one for each atom, whose components have magnitudes inversely proportional to the masses of the specific atomic isotopes [1, 2, 6, 3]:

$$\Delta V_{\rm ad}^{(\alpha)}(r) \;=\; \frac{\Delta M_{\rm A}^{(\alpha)}}{M_{\rm A}^{(\alpha)}}\;\widetilde{S}_{\rm ad}^{\rm A}(r) \;+\; \frac{\Delta M_{\rm B}^{(\alpha)}}{M_{\rm B}^{(\alpha)}}\;\widetilde{S}_{\rm ad}^{\rm B}(r) \quad , \quad (3)$$

$$g^{(\alpha)}(r) \;=\; \frac{M_{\rm A}^{(1)}}{M_{\rm A}^{(\alpha)}}\;\widetilde{R}_{\rm na}^{\rm A}(r) \;+\; \frac{M_{\rm B}^{(1)}}{M_{\rm B}^{(\alpha)}}\;\widetilde{R}_{\rm na}^{\rm B}(r) \quad . \quad (4)$$

Here, $\Delta M_{\rm A}^{(\alpha)} \equiv M_{\rm A}^{(\alpha)} - M_{\rm A}^{(1)}$ is the difference between the mass of the isotope of atom A in isotopologue α and in that the reference isotopologue ($\alpha = 1$), while the expressions employed to represent the mass-independent $\widetilde{S}_{\rm ad}^{\rm A/B}(r)$ and $\widetilde{R}_{\rm na}^{\rm A/B}(r)$ radial functions can be found in §2.3. Straightforward extensions of Eq. (1) to take account of the *e/f* Λ-doubling splittings that occur for singlet states with $\Lambda \neq 0$ [5], or of the doublet splittings of the rotational levels of $^2\Sigma$ states, are presented in §§2.4 and 2.5.

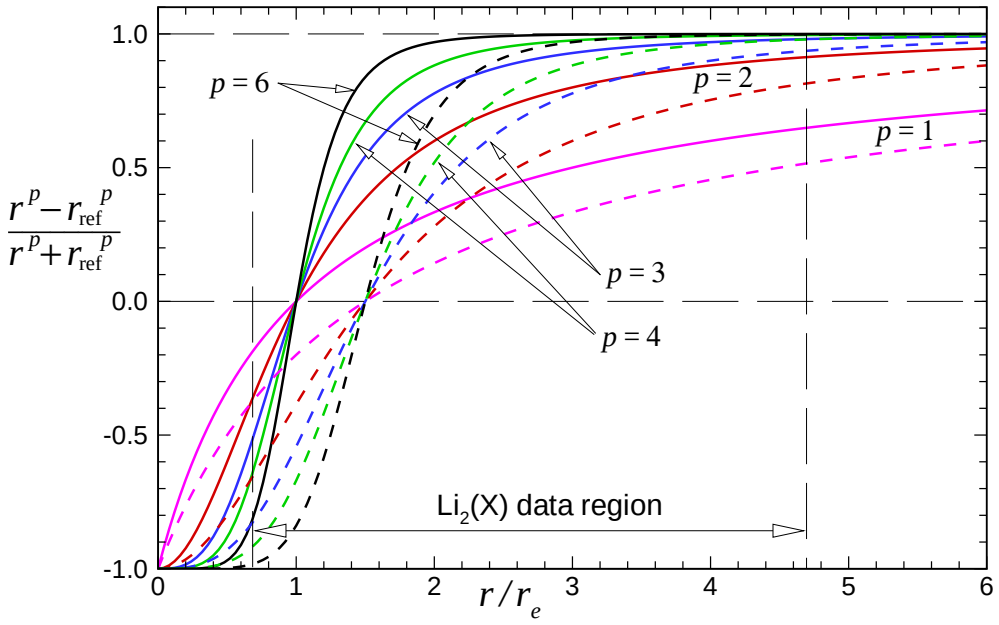

Figure 1. Illustrative plots of the expansion variables of Eqs. (5) and (6) showing the 'data range' associated with an analysis for the $X\,^1\Sigma_g^+$ state of Li_2 [10]. The solid curves are for the case $r_{\rm ref} = r_{\rm e}$ (variable $y_p^{\rm eq}(r)$), while the dashed curves correspond to $r_{\rm ref} = 1.5\,r_{\rm e}$ ($y_p^{1.5r_{\rm e}}(r)$).

2 The Potential Energy Function

2.1 The Radial Expansion Variable

Program DPotFit currently allows the potential energy function for a given electronic state to be represented by one of four families of analytic model potentials. Most of these functions are expressed in terms of radial variables of the form

$$y_p^{\rm eq}(r) \;=\; \frac{r^{\,p} - {r_{\rm e}}^{\,p}}{r^{\,p} + {r_{\rm e}}^{\,p}} \quad , \tag{5}$$

$$y_p^{\rm ref}(r) \;=\; \frac{r^{\,p} - {r_{\rm ref}}^{\,p}}{r^{\,p} + {r_{\rm ref}}^{\,p}} \quad , \tag{6}$$

in which p is a small positive integer ($p = 1,\ 2,\ 3,\ 4,\ \ldots$), $r_{\rm e}$ is the equilibrium internuclear distance of $V_{\rm ad}^{(1)}(r)$, and $r_{\rm ref}$ is a reference distance chosen as the expansion centre for this variable (usually $r_{\rm ref} > r_{\rm e}$). Most of the early work employing this type of variable fixed $r_{\rm ref} = r_{\rm e}$ [7, 8, 9]. However, that is not an essential constraint, and it has been shown that fixing $r_{\rm ref}$ at some distance between $r_{\rm e}$ and the outer end of the data-sensitive region tends to allow accurate fits to be achieved with a smaller number of expansion parameters [10, 11, 12]. Note that one of the models described below uses two different expansion variables of this type, defined by different values of the (integer) power. In that case, a second (integer) label q ($q \neq p$) is introduced to identify that second radial variable.

The nature of these variables is illustrated by Fig. 1 for a range of values of p [10] and two values of $r_{\rm ref}$. The fact that $y_p^{\rm eq}(r)$ and $y_p^{\rm ref}(r)$ approach finite limits both as $r \to 0$ and as $r \to \infty$ means that functions of these variables will also approach finite values in these limits. At the same time, the fact that $y_p^{\rm eq}(r) \propto (r - r_{\rm e})$ and $y_p^{r_{\rm ref}}(r) \propto (r - r_{\rm ref})$ 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 $y_p(r) \in [-1, +1]$ 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. Furthernore, for larger values of p the resulting potential ebergy functions will be increasingly strongly inhibited from having implausible spurious extrema in the extrapolation regions at small or very large values of r [7, 8, 9].

Early applications [13, 14, 15, 16] of this type of radial variable were based on the single variable $y_p^{\rm eq}(r)$ of Eq. (5). However, in later applications [10, 12, 17] it was found that use of an expansion variable $y_p^{r_{\rm ref}}(r)$ centred at a distance $r_{r_{\rm ref}} > r_{\rm e}$ can, with no loss of accuracy, lead to much more compact and robust potential function expressions than could otherwise be obtained. Such results are readily explained by examining the differences the difference between the solid and dashed curves in Fig. 1, for a given value of p. For example, for $p=4$, the solid curve for $y_4^{\rm eq}(r)$ is flat and lies very close to its upper-limit value of +1 over a substantial fraction of the 'data region'. As a consequence, expansions in that variable would be unable to represent accurately any properties that vary significantly across this domain. In contrast, the corresponding variable $r_4^{r_{\rm ref}}(r) = y_4^{1.5r_{\rm e}}(r)$ changes significantly across the entire data region, and hence may be expected to provide a robust and compact description of properties that vary across this whole domain.

The radial variable of Eq. (5) is a special case of a more general expansion variable introduced by Šurkus *et al.* [18] (see §2.7). They chose to represent the overall potential function as a power series in their version of $y_p^{\rm eq}(r)$, and specified that the value of p should be based on the asymptotically dominant inverse-power term in the intermolecular potential for the state of interest. However, our present use of this variable is based upon our finding [19, 4, 5, 20, 21, 13] that an expansion in powers of $y_p^{\rm eq}(r)$ or $y_q^{r_{\rm ref}}(r)$ for some small integer power of p or q that is greater than 1 (say, $\{p,q\} = 2, \cdots, 6$) greatly reduces any propensity for the resulting expansion to "extrapolate badly" outside the radial interval to which the data are most sensitive.

In any case, values of p, q, and $r_{\rm ref}$ must be selected by the user and specified in the main input data file. Some guidance regarding how to choose appropriate values of these parameters may be found below and in Refs. [14] and [17].

2.2 The Expanded Morse Oscillator (EMO) Potential Function

The first type of potential function form considered here is the *Expanded Morse Oscillator* or EMO function [7], which has the form of a Morse potential [22] in which the exponent coefficient varies with distance. Other functions of this type have been introduced by Coxon and Hajigeorgiou (the "GMO" potential) [23] and by Dulick and co-workers (the "MMO" potential) [24], but because of its simpler form and better extrapolation behaviour, only the EMO function is considered here. An EMO (or EMO$_p$) potential has the form

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

in which $\mathfrak{D}_{\rm e}$ is the well depth, $r_{\rm e}$ the equilibrium internuclear distance, and

$$\beta(r) \;=\; \beta_{\rm EMO}(y_p^{r_{\rm ref}}(r)) \;=\; \sum_{i=0}^{N_\beta} \beta_i \; y_q^{r_{\rm ref}}(r)^i \quad , \tag{8}$$

As discussed in Refs. [19, 5, 20], for cases in which $r_{\rm ref} = r_{\rm e}$, an appropriate choice of q (usually >1) in the definition of $y_q^{r_{\rm ref}}(r)$ can prevent extrapolation problems at large r, but does not always resolve such problems at small r. However, such residual extrapolation problems can usually be solved by setting the expansion centre $r_{\rm ref}$ at some distance greater than $r_{\rm e}$.[1]

[1] An early attempt to address problems associated with extrapolation to short distances involved allowing the exponent polynomial of Eq. (8) to have a lower order for $r < r_{\rm e}$ than for $r > r_{\rm e}$ [5, 20, 21, 13, 16]. However, the introduction of a variable $r_{\rm ref}$ parameter obviated this option. See the example of AgH illustrated in §9.4.2

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 [7, 8, 21] and a state whose potential function has an additional 'ripple' [20]. However, the fact that $[\mathfrak{D}_e - V_{\rm EMO}(r)]$ dies off exponentially at large r makes it a less than ideal function for representing states for which the data extend fairly close to the dissociation limit. This problem stimulated the development of two of the other potential function forms discussed below.

2.3 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

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

in which the powers m_1, m_2, ... etc., are determined by the nature of the atoms to which the given molecular state dissociates [25, 26], 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 vibrational levels lying fairly close to dissociation. This consideration stimulated the development of the Morse/Long-Range (MLR) potential form [13, 14, 15, 10, 11]:

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

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

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

while $u_{\rm LR}(r_{\rm e})$ is the value of this function at $r_{\rm e}$. If the long-range function of Eq. (11) has only a single term, then the MLR potential reduces to the *Morse/Lennard-Jones* (or MLJ) potential,[2] which was introduced almost two decades ago [27, 28], and has been used in a number of detailed data analyses [29-37]. However, the MLJ potential is merely a simpler version of the general MLR potential form.

Since $y_p^{\rm eq}(r) \to +1$ as $r \to \infty$, at long range the MLR function of Eq. (10) becomes

$$V_{\rm MLR}(r) \simeq \mathfrak{D}_{\rm e} - \left\{\frac{2\,\mathfrak{D}_{\rm e}\, e^{-\beta_\infty}}{u_{\rm LR}(r_{\rm e})}\right\} u_{\rm LR}(r) = \mathfrak{D}_{\rm e} - \frac{C_{m_1}}{r^{m_1}} - \frac{C_{m_2}}{r^{m_2}} - \ldots \tag{12}$$

in which

$$\beta_\infty \equiv \lim_{r\to\infty} \beta_{\rm MLR}(r) = \ln\left\{\frac{2\,\mathfrak{D}_{\rm e}}{u_{\rm LR}(r_{\rm e})}\right\} . \tag{13}$$

Thus, the limiting asymptotic value of the exponent coefficient function $\beta_{\rm MLR}(r)$ is defined by the values of $\mathfrak{D}_{\rm e}$, $r_{\rm e}$, and the coefficients C_{m_i} of the inverse-power terms included in $u_{\rm LR}(r)$. In order to use the MLR potential form, it is clearly necessary to know appropriate values for the powers m_i [25, 26], and to have realistic estimates of the coefficients C_{m_i}. When no realistic estimate of the leading (smallest-power) coefficient C_{m_1} is available, the MLR form has no significant advantages over the simpler EMO function, which will probably be more 'robust'. However, if the leading inverse-power coefficient C_{m_1} is known, but no proper estimate of the value of the second coefficient C_{m_2} is available, it may be desirable to make a

[2] In early work [27, 28] this was called a *Modified Lennard-Jones* (MLJ) oscillator. However, since elimination of the exponential term yields a Lennard-Jones$(2n, n)$ potential, while elimination of the $(r_{\rm e}/r)^n$ pre-factor yields a Morse-like potential function, its algebraic form suggests that *Morse/Lennard-Jones* is probably a more appropriate label.

plausible *ad hoc* estimate of the latter, and employ the two-term MLR form rather than the simpler MLJ function, because more reasonable long-range extrapolation behaviour is thereby imposed [13]. In the rare cases in 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 [14, 10, 12].

The algebraic form of Eq. (10) means that at sufficiently long range $V_{\rm MLR}(r)$ always takes on the form of Eq. (12). To achieve this, the exponent coefficient function $\beta(r)$ is required both to approach the value β_∞ defined by Eq. (13) asymptotically, and to be sufficiently flexible to describe the shape of the potential function well accurately. Its functional form should also prevent or discourage the potential from having unphysical extrema in the two extrapolation intervals: where $r \to 0$, and between the data region and the asymptotic limit. To this end, $\beta(r)$ is written as a constrained polynomial in the variable $y_q^{r_{\rm ref}}(r)$, in which the separate variable $y_p^{r_{\rm ref}}(r)$ acts as a switching function [19, 4, 5, 13, 10]:[1]

$$\beta(r) \;=\; \beta_{\rm MLR}(y_{p,q}^{r_{\rm ref}}(r) \;=\; y_p^{r_{\rm ref}}(r)\;\beta_\infty \;+\; \left[1 - y_p^{r_{\rm ref}}(r)\right]\;\sum_{i=0}^{N_\beta}\beta_i\;y_q^{r_{\rm ref}}(r)^i \quad . \tag{14}$$

Note that the power p appearing here is the same as that used to define the distance parameter $y_p^{\rm eq}(r)$ in the exponential term in Eq. (10). While most of the early work with this model was performed with $q = p$, it has since been shown that use of a separate power $q < p$ in the power-series portion of Eq. (14) can lead more compact and robust potential functions [10, 11, 12].[1]

One restriction associated with the MLR form is a limitation on the allowed value of p, depending on the particular set of powers m_i that define the terms contributing $u_{\rm LR}(r)$. The algebraic form of the exponent coefficient function of Eq. (14) 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. (12) [14, 10]. As a consequence, the leading contributions to long-range behaviour of $V_{\rm MLR}(r)$ will only truly be defined by the specified version of Eq. (11) if the power p defining the exponent variables satisfies the constraint $p > (m_{\rm Last} - m_1)$, where $m_{\rm Last}$ is the power of the last (i.e., highest-power) term contributing to $u_{\rm LR}(r)$ [14, 10, 11]. It may also be desirable to set $p = m_{\rm next} - m_1$, in which $m_{\rm next}$ is the (inverse) power associated with the first long-range term predicted by theory that is not included in the chosen definition of $u_{\rm LR}(r)$.

As discussed in Ref. [10], there are no formal restrictions on the choice of the power q defining the the radial variable in the power series part of Eq. (14). Practical experience suggests [10, 12, 29] that when values $q = 1$ or 2, the potential is more likely to be unstable in the extrapolation region(s). However, the optimum choices for q (and p, subject to the constraint $p > (m_{\rm Last} - m_1)$) may be guided by consideration of the manner in which the potential function approaches its limiting long-range behaviour. In particular, $V_{\rm MLR}(r)$ always takes on the limiting behaviour of Eq. (12), a rearrangement of which yields

$$C_{m_1}^{\rm eff}(r) \;\equiv\; r^{m_1}\,[\,\mathfrak{D}_{\rm e} - V_{\rm MLR}(r)] \;\simeq\; C_{m_1} + \frac{C_{m_2}}{r^{m_2-m_1}} \;+\; \ldots \tag{15}$$

Thus, a plot of $C_{m_1}^{\rm eff}(r)$ *vs.* $1/r^{m_2-m_1}$ must approach the intercept C_{m_1} with slope C_{m_2}, and it should bridge the extrapolation interval between the "data region" and this limiting behaviour in a smooth, monotone fashion. Figure 2 shows plots of this type for a number of otherwise equivalent potentials for the ground-state of Br_2 [29], for which the leading terms in the long-range potential correspond to $m = 6$, 8, and 10.[3] While all of the six potentials considered do eventually achieve the predicted linear approach to the intercept $C_6^{\rm theory}$ with slope $C_8^{\rm theory}$, it is clear that for model potentials with $=4$ or 5, there are physically

[3] There is also a very weak repulsive $m = 5$ term [30], but it has no discernable effect outside the immediate neighbourhood of the intercept [29].

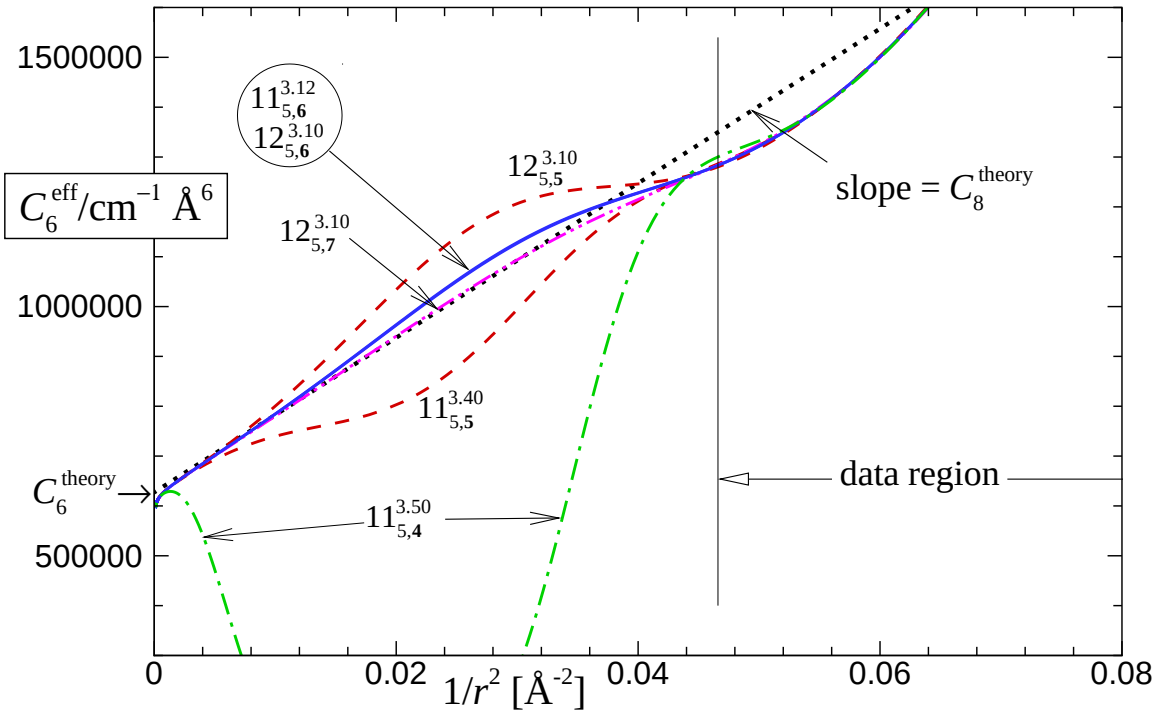

Figure 2. Tests of the long-range extrapolation behaviour of various fitted potentials for the $X\,^1\Sigma_g^+$ state of Br_2 associated with different MLR models $N_{p,q}^{r_{\rm ref}}$, where $N = N_\beta$. [Adapted from Fig. 7 of Ref. [29].]

implausible extrema in the extrapolation interval between the 'data-sensitive region' and that intercept.[4] Moreover, the $C_6^{\rm eff}(r)$ plot for the $11^{3.10}_{5,7}$ potential fails to display the positive curvature away from this limiting slope implied by the existence of the C_{10}/r^{10} term, probably because of 'stiffness' of the $q = 7$ radial variable. Thus, only potentials for which $q = 6$ have physically reasonable extrapolation behaviours for this species.

When performing fits to an MLR form it is necessary to consider a range values of $r_{\rm ref}$, N_β and q in order to determine an optimal model. Figure 3 illustrates how the quality of fits to data for ground-state Ca_2 depends upon these three parameters. As expected, for either $q = 3$ or $q = 4$, the quality of fit improves and the breadth of the region over which $\overline{dd}$ has a minimum increases with N_β. However, for larger q values the expansion variable is "stiffer", and higher N_β values tend to be required to give the same quality of fit. This is the reason that the range of $r_{\rm ref}$ 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 on this figure lead to a recommendation of $\{q = 4,\ N_\beta = 7,\ r_{\rm ref} = 6.85\,$Å$\}$ as the optimum MLR model for this Ca_2 system.

2.3.1 A More General Definition of $u_{\rm LR}(r)$: Inclusion of Damping Functions

While all potential energy functions take on the limiting behaviour of Eq. (9) at very large r, at shorter distances, overlap of the electron distributions of the interacting atoms reduces the strength of the interaction energies associated with the even-inverse-power 'dispersion' terms contributing to this expression. This consideration led a number of groups to propose models for representing this 'damping' behaviour

[4] Note, however, that these implausible extrema in $C_6^{\rm eff}(r)$ plots are *not* accompanied by discernably irregular behaviour in plots of the potential functions $V_{\rm MLR}(r)$ themselves.

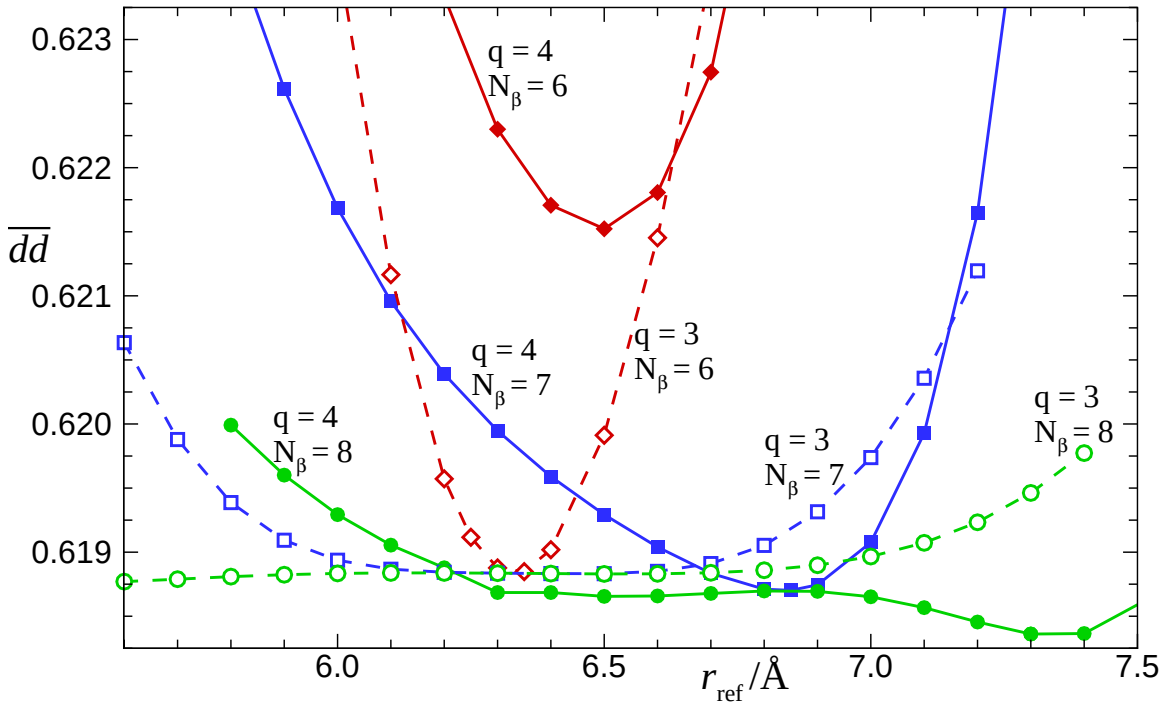

Figure 3. Dependence of dimensionless root-mean-square deviation $\overline{dd}$ of fits to 3553 fluorescence series data for the $X\,^1\Sigma_g^+$ state of Ca_2 [31, 32] on the parameters $r_{\rm ref}$, q, and N_β. These are results of fits to MLR potentials whose long-range tails are defined by Douketis-type damping with $s = -1$ and fixed theoretical values of the C_8, and C_{10} [33] coefficients, but with C_6 allowed to be a fitted parameter [11].

[38-42]. Another consideration here is the fact that the quadratic term in Eq. (10), which defines the short range repulsive wall of an MLR potential, contains the factor $[u_{\rm LR}(r)]^2$. As a result, if there is no damping, at very short range the strength of this term will grow as $1/r^{2m_{\rm last}}$, which is much steeper than the exponential-type behaviour expected for the repulsive wall of a normal potential energy function. These concerns led to the introduction of an (optional) alternate definition for the long-range potential incorporated in Eq. (10), namely,

$$u_{\rm LR}(r) \;=\; D_{m_1}(r)\,\frac{C_{m_1}}{r^{m_1}} \;+\; D_{m_1}(r)\,\frac{C_{m_2}}{r^{m_2}} \;+\; \ldots \;+\; D_{m_{\rm Last}}(r)\,\frac{C_{m_{\rm Last}}}{r^{m_{\rm Last}}} \quad . \tag{16}$$

Two types of damping function are currently allowed by DPotFit. One is a generalized [17] version of the damping functions of Douketis *et al.* [34]:

$$D_m^{{\rm ds}(s)}(r) \;=\; \left(1 \;-\; e^{-\frac{b^{\rm ds}(s)\cdot(\rho\, r)}{m} - \frac{c^{\rm ds}(s)\cdot(\rho\, r)^2}{\sqrt{m}}}\right)^{m+s} \quad , \tag{17}$$

and the other is a generalized version [17] of the Tang-Toennies damping function [35]:

$$D_m^{{\rm tt}(s)}(r) \;=\; 1 \;-\; e^{-b^{\rm tt}(s)\cdot(\rho\, r)} \sum_{k=0}^{m-1+s} \frac{[b^{\rm tt}(s)\cdot(\rho\, r)]^k}{k!} \quad . \tag{18}$$

The parameters $b^{\rm ds}(s)$ and $c^{\rm ds}(s)$ in Eq. (17) and $b^{\rm tt}(s)$ in Eq. (18) are *system-independent* constants determined in Ref. [17] from fits to *ab initio* $m = 6$, 8 and 10 damping functions for two ground-state hydrogen atoms (for which $\rho \equiv 1$, see below) obtained by Kreek and Meath [36]. The original versions of these functional forms corresponded to the cases $s = 0$ for the Douketis *et al.* [34] form and $s = 1$ for

the Tang-Toennies form [35]. The constant ρ appearing in Eqs. (17) and (18) is a *system-dependent* range parameter introduced by Douketis *et al.* [34]: for a pair of interacting atoms A and B, it is defined by the combining rule [34]

$$\rho \;\equiv\; \rho_{\rm AB} \;=\; 2\,\rho_{\rm A}\cdot\rho_{\rm B}/(\rho_{\rm A}+\rho_{\rm B}) \;\;, \tag{19}$$

in which $\rho_{\rm A} = (I_p^{\rm A}/I_p^{\rm H})^{2/3}$ is defined in terms of the ratio of the ionization potential of the atom in question ($I_p^{\rm A}$) to that of an H atom ($I_p^{\rm H}$). This is a physically plausible parameterization, since the diffuseness of an atom's electron distribution tends to vary inversely with its ionization potential, and it is very easy to apply, since accurate $I_p^{\rm A}$ values are readily available for all atomic species.

The nature of these generalized definitions of the two families of damping functions means that at very small r,

$$\lim_{r\to 0}\left\{D_m^{(s)}(r)/r^m\right\} \;\propto\; r^s \tag{20}$$

for all values of m and s. The results of Kreek and Meath [36] show that the actual very short-range behaviour of dispersion energy damping functions correspond to $s=0$. If the overall potential function model consists simply of a sum of damped dispersion terms plus a short-range repulsion term, the $s=1$ limiting behaviour of the original Tang-Toennies function [35] presents no practical problem, since the essential requirement of preventing the attractive inverse-power terms from becoming infinite as $r\to 0$ is still achieved. However when the attractive long-range tail of the potential appears as a multiplicative factor, rather than an additive term, as in the case of the MLR potential, damping functions for which $s>0$ are physically unacceptable, as they would cause the repulsive wall of the potential to turn over and go to zero as $r\to 0$. As discussed in Ref. [17], these considerations led to our recommendation that the damping functions used in the MLR potential be $s=-1$ (`IDF` $=-2$) versions of the Douketis-type damping function of Eq. (17). However, the code currently allows a user to make other choices for s or to use the generalized Tang-Toennies function of Eq. (18).

2.3.2 Treatment of Interstate Coupling near the Asymptote: States of Li_2 dissociating to Li($2S$) + Li($2P$)

The above discussion of the MLR form focused upon use of the inverse-power sums (11) or (16) to represent the long-range potential. However, it is not necessary to restrict $u_{\rm LR}(r)$ to these forms, and they can take on any form dictated by theory. For example, in recent applications to states of Li_2 dissociating to the $Li(^2P_{1/2}) + Li_2(^2S_{1/2})$ asymptote, $u_{\rm LR}(r)$ has been represented by one of the roots of a diagonalization arising from two-state [10] or three-state [12] coupling near that asymptote. In particular, the theory of this interstate mixing presented in Refs. [37, 38] shows that the expression for the 2×2 roots may be written in closed form as

$$\begin{aligned} u_{\rm LR}^{\rm A\text{-}F}(r) \;=\; & -\frac{A_{\rm so}}{2} + \frac{C_3^\Sigma + C_3^\Pi}{2\,r^3} + \frac{C_6^\Sigma + C_6^\Pi}{2\,r^6} + \frac{C_8^\Sigma + C_8^\Pi}{2\,r^8} \\ & \pm \frac{1}{2}\left\{\left(\frac{C_3^\Sigma - C_3^\Pi}{3\,r^3} + \frac{C_6^\Sigma - C_6^\Pi}{3\,r^6} + \frac{C_8^\Sigma - C_8^\Pi}{3\,r^8} - A_{\rm so}\right)^2\right. \\ & \left. + 8\left(\frac{C_3^\Sigma - C_3^\Pi}{3\,r^3} + \frac{C_6^\Sigma - C_6^\Pi}{3\,r^6} + \frac{C_8^\Sigma - C_8^\Pi}{3\,r^8}\right)^2\right\}^{1/2} \end{aligned} \tag{21}$$

in which $A_{\rm so}$ is the (positive) difference between the $^2P_{1/2}$ and $^2P_{3/2}$ atomic spin-orbit level energies, and C_m^Σ / C_m^Π are the long-range coefficients associated with the relevant coupled states [10]. Use of the '+' sign in front of the square root term yields the correct long-range tail for the $A\,^1\Sigma^+$ state, while use of the '−'

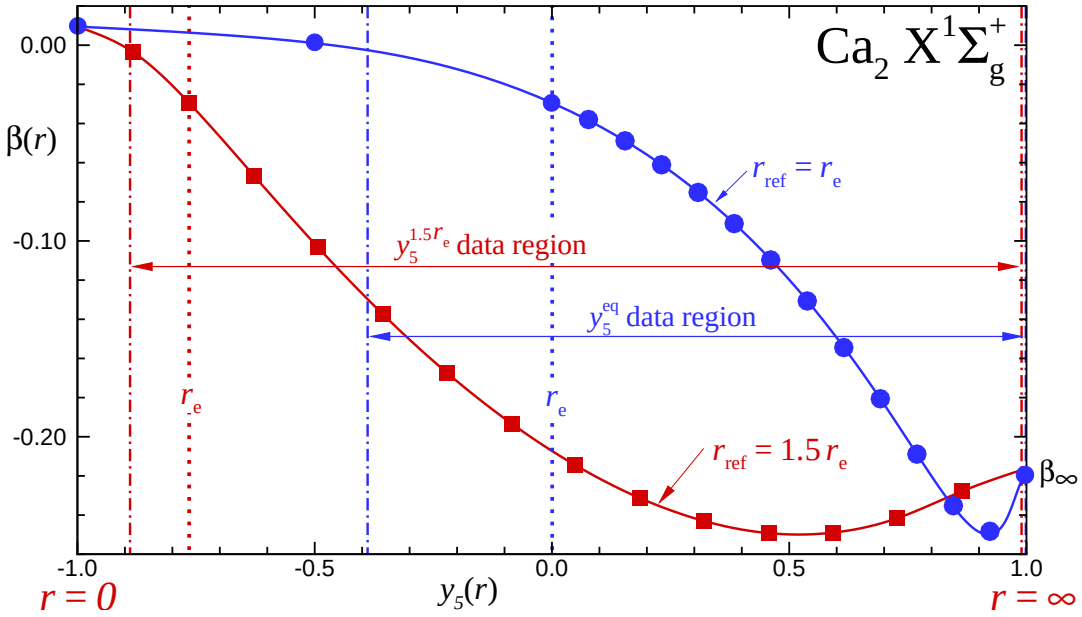

Figure 4. Comparison of $\beta(y_q^{r_{\rm ref}}(r))$ functions for ground-state Ca_2 determined from DPFs of spectroscopic data to SE-MLR models based on $r_{\rm ref} = r_{\rm e}$ (round points, blue curve and lines), and $r_{\rm ref} = 1.5\,r_{\rm e}$ (square points, red curve and lines) [Figure taken from Ref. [44], with permission].

sign yields that for the 0_u^+ component of the $b\,^3\Pi_u$ state [37]. The analogous theory for the long-range tail of the $1\,^3\Sigma_g^+$ state potential, which involves 3-state coupling near the asymptote is described in Ref. [12].

The ease with which DPotFit can treat these special cases further illustrates capabilities of the MLR functional form.

2.3.3 A Spline-Pointwise Representation of the MLR Exponent Coefficient $\beta(r)$

All existing applications of the MLR potential have been based upon employing a polynomial function such, as Eq. (14) to represent the exponent coefficient $\beta(r)$. However, while preliminary work showed some promise [39, 40], it is not yet clear whether that form can provide a practical, compact, and accurate representation of double-minimum potentials or of shelf-state potentials. A novel approach introduced by Pashov and co-workers [41, 42, 43], in which the potential is defined as a cubic spline through a set of points whose energies are the parameters of the model, has proved remarkably successful for treating such systems. However, a relatively large number of points (typically $\gtrsim 50$) are required to define a potential 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.

An alternative approach now under investigation [44] 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 $\beta(r_j) = \beta(y_q^{r_{\rm ref}}(r_j))$ values at a specified set of $y_q^{r_{\rm ref}}(r_j)$ values, and those $\beta(r_j)$ values become the parameters defining the shape of the potential. Following the Pashov approach [41, 45], the exponent coefficient function is written as

$$\beta(r) \;=\; \sum_{k=1}^{N_\beta} S_k(y_q^{\rm ref}(r))\;\beta_k \quad , \tag{22}$$

in which the spline 'basis functions' $S_k(y_p^{\rm ref}(r))$ are completely defined by the chosen mesh of values of

$y_q^{\rm ref}(r_j)$. A straightforward application of the chain rule of calculus then yields the partial derivatives required for the least-squares fit procedure:

$$\frac{\partial V(r)}{\partial \beta_k} = 2\,\mathfrak{D}_{\rm e}\left\{1 - \frac{u_{\rm LR}(r)}{u_{\rm LR}(r_{\rm e})}\; e^{-\beta(r)\cdot y_p^{\rm eq}(r)}\right\}\left(\frac{u_{\rm LR}(r)}{u_{\rm LR}(r_{\rm e})}\; e^{-\beta(r)\cdot y_p^{\rm eq}(r)}\right) y_p^{\rm eq}(r)\; S_k(y_q^{\rm ref}(r)) \quad . \tag{23}$$

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

In using the SE-MLR form, it is important to realize that the radial variable $y_q^{\rm ref}(r)$ used to define $\beta(r)$ should normally be defined by a value of $r_{\rm ref}$ that is significantly greater than $r_{\rm e}$. This is necessary in order to assure that the chosen mesh of points samples the full range of $\beta(y_q^{\rm ref}(r))$ values appropriately. This consideration is illustrated by Fig. 4, which compares the SE-MLR exponent coefficient functions $\beta(r)$ determined from of fits to an extensive data set for the $X\,^1\Sigma_g^+$ state of Ca_2 [32, 14, 44] that were performed using, in turn $r_{\rm ref} = r_{\rm e}$ (blue points, curve and lines), and $r_{\rm ref} = 1.5\,r_{\rm e}$ (red points, curve and lines). Both of these cases placed two points, equally-spaced in $y_q^{\rm ref}(r)$, at r less than $r_{\rm e}$ and 14 points, equally-spaced in $y_q^{\rm ref}(r)$, at r greater than $r_{\rm e}$, together with one point at $r = r_{\rm e}$ and one at $y_q^{\rm ref} = 1$, which corresponds to the limit $r \to \infty$. Consideration of this figure makes it quite clear that when using the same number of spline points, it will be easier to obtain an accurate description of $\beta(r)$ using points based (in this case) on $r_{\rm ref} = 1.5\,r_{\rm e}$.[5]

2.4 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 at distances $r > r_{\rm e}$ stimulated the development of the *double-exponential/long-range* (DELR) potential function form [19, 5],

$$V_{\rm DELR}(r) \;=\; \left\{A\,e^{-2\,\beta(r)\cdot(r-r_{\rm e})} \;-\; B\,e^{-\beta(r)\cdot(r-r_{\rm e})} \;+\; \mathfrak{D}_{\rm e}\right\} \;-\; u_{\rm LR}(r) \quad , \tag{24}$$

in which the exponent coefficient $\beta(r)$ is defined as the same type of simple power series in $y_p^{r_{\rm ref}}(r)$ used for the EMO potential (see Eq. (8)). The only published application of this form to date [5] used different power-series orders for $r \le r_{\rm e}$ and $r > r_{\rm e}$ in order to avoid unphysical extrapolation behaviour at small r. However, subsequent experience with the MLR form has led us to treat $\beta(r)$ as a single simple polynomial, with the introduction of an expansion centre parameter $r_{\rm ref}$ located at some distance greater than $r_{\rm e}$ serving to ensure stable short-range behaviour. In this case, inclusion of a repulsive term in the additive long-range function $u_{\rm LR}(r)$ (which in principle may be attractive or repulsive) served to introduce the potential function barrier that was being modeled.

The pre-exponential coefficients A and B in Eq. (24) are defined in terms of the well depth $\mathfrak{D}_{\rm e}$ (relative to the potential asymptote) and the equilibrium distance $r_{\rm e}$ by the expressions

$$A \;=\; \mathfrak{D}_{\rm e} \;-\; u_{\rm LR}(r_{\rm e}) \;-\; u'_{\rm LR}(r_{\rm e})/\beta_0 \quad , \tag{25}$$

$$B \;=\; 2\,\mathfrak{D}_{\rm e} \;-\; 2\,u_{\rm LR}(r_{\rm e}) \;-\; u'_{\rm LR}(r_{\rm e})/\beta_0 \quad , \tag{26}$$

in which $u'_{\rm LR}(r_{\rm e}) \equiv [du_{\rm LR}(r)/dr]_{r=r_{\rm e}}$. If $u_{\rm LR}(r) = 0$ the DELR potential becomes the EMO function of Eq. (7). However, other choices of $u_{\rm LR}(r)$ allow it to represent the outer wall of a potential function with a barrier [5], a multi-term attractive inverse-power long-range potential, or even the outer well of a double-minimum or shelf-state potential.

[5] Note that the SE-MLR option is not yet fully implemented.

In the present version of DPotFit, the long-range function $u_{\rm LR}(r)$ in Eq. (24) is assumed to be defined by the same type of sum of damped or undamped inverse-power terms represented by Eqs. (11) or (16). Note, however, that the present sign convention for this $u_{\rm LR}(r)$ function is the opposite of that used in Ref.[5]. The damping functions $D_m(r)$ may be defined by either of by Eqs. (17) or (18), but since $u_{\rm LR}$ is an additive, rather than multiplicative contribution to the potential, the damping function parameter s is allowed to have positive values, while negative values should be used only with care. Of course, other damping function expressions [46, 47] or entirely different types of expressions for $V_{\rm LR}(r)$ could equally well be used in the DELR type of potential form.

2.5 Hannover Polynomial Potentials (HPP)

The three following subsections describe potential forms that have been introduced into the DPotFit code to allow us to reproduce and compare with the results of published fits that were performed using them. Perforance of fits using these optional potential forms has not been rigorously tested, and the use of the MLR or EMO forms described above is generally recommended.

One of the most widely used of these alternate potential function forms is the 'Hannover polynomial potential' (HPP) or "X-representation" function, which has the form [48, 49]:

$$V_{\rm HPP}(r) = A_{\rm I}\, e^{-B_{\rm I}(r-R_{\rm I})} \qquad \text{for} \quad r < R_{\rm I} \tag{27}$$

$$= \sum_{i=0}^{N_\beta} \beta_i\, {X_p}^i \qquad \text{for} \quad R_{\rm I} \le r \le R_{\rm O} \tag{28}$$

$$= \mathfrak{D}_{\rm e} - \sum_m \frac{C_m}{r^m} + A_{\rm O}\, e^{-B_{\rm O}(r-R_{\rm O})} \qquad \text{for} \quad r > R_{\rm O} \tag{29}$$

in which

$$X = \frac{r - R_m}{r + b\, R_m} \tag{30}$$

R_m is an 'arbitrarily chosen' expansion centre that is close to the position of the potential minimum, and the parameter b is manually chosen to optimize the potential slope at short distances. The expansion coefficients β_i are determined from the fit to the data. The parameters $A_{\rm I}$ and $B_{\rm I}$ defining the exponential function used to extrapolate to short distances are defined by the requirement that there be a smooth connection to the polynomial function at $r = R_{\rm I}$. Similarly, with the long-range coefficients C_m defined by theory, the parameters $A_{\rm O}$ and $B_{\rm O}$ defining the exponential term parameters in Eq. (29) must be chosen so as to provide a smooth connection to the polynomial expansion function at the outer switching point $r = R_{\rm O}$.

2.6 Tang-Toennies Exponential/Van der Waals Potentials (TTP)

The Tang-Toennies potential (TTP) consists of a simple exponential repulsive term and a sum of damped attractive inverse-power terms [35, 50]:

$$V_{\rm TTP}(r) = A_{\rm TT}\, e^{-br} - \sum_m D_m(r)\, \frac{C_m}{r^m} \quad . \tag{31}$$

Only even values of $m \ge 6$ were considered by Tang and Toennies [35, 50, 51], and their damping functions were always the $s = +1$ version of Eq. (18). However DPotFit allows the use of any specified powers m_i, allows the damping functions to be defined by either of Eqs. (17) or (18), and allows the user to choose the parameter 's' which determines the limiting short-range behaviour of the damped inverse-power terms (see Eq. (20)). Note, however, that for positive (attractive) C_m values, necessarily $s \ge 0$. In the DPotFit

implementation of the TT potential, the input values of the exponent coefficient b, the equilibrium distance r_e, and the Van der Waals coefficients $\{C_m\}$ are employed to define $\mathfrak{D}_e$ and the pre-exponential factor $A = A_{TT}$.

2.7 Šurkus' Generalized Potential Energy Function (GPEF)

The sixth family of potential functions is a generalization of the familiar Dunham polynomial potential [52] which Šurkus *et al.* [18] introduced and called the *Generalized Potential Energy Function* (GPEF). Using a modified but exactly equivalent expression for this expansion variable (devised by Seto [9]), the GPEF potential form is

$$V_{GPEF}(r) \;=\; c_0\, z_q^2 \left[1 \;+\; \sum_{i=1}^{N_\beta} c_i\, {z_q}^i \right] \quad , \qquad \text{with} \qquad z_q \;=\; \frac{(r^q - {r_e}^q)}{(a_S\, r^q + b_S\, {r_e}^q)} \quad . \tag{32}$$

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 [52].
- Setting $q=1$, $a_S=1$ and $b_S=0$ yields Simons-Parr-Finlan (SPF) expansions [53].
- Setting $q=1$, $a_S=b_S=0.5$ yields the Ogilvie-Tipping (OT) potential expansion [54].
- Setting $a_S=b_S=1$ yields the expansion variable $y_p^{eq}(r)$ of Eq. (5).

When $a_S \neq 0$ this function always asymptotically approaches a finite limit with a $1/r^q$ functional behaviour. Thus, if appropriate constraints are applied to the coefficients, it can in principle be required to have the theoretically predicted limiting long-range behaviour of Eq. (9) [18, 55]. However, except for the relatively simple case in which q is set equal to the power of the leading long-range term in Eq. (9) [18], such constraints have proven to be too unwieldy for practical use.

2.8 Fixed Pointwise Potential

The final type of potential function which can be used by DPotFit to define the vibration-rotation levels of a given electronic state is one that is 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. 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 data involving multiple electronic states to make use of previously reported pointwise potentials for one (or more) of the states of interest.

2.9 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 given state, but it must still be taken into account because it is at one end of a set of transitions to some other state for which a potential function is being determined. In some cases the levels of that state may be accounted for as the origins of fluorescence series, but it is often more convenient to treat all of the observed 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 levels by a set of band constants:

$$\begin{aligned} E(v,J) \;&=\; G_v + B_v[(J(J+1)] \;-\; D_v[(J(J+1)]^2 \;+\; H_v[(J(J+1)]^3 \;+\; \ldots \\ &=\; \sum_{m=0} K_m [J(J+1)]^m \quad . \end{aligned} \tag{33}$$

Both choices are allowed by DPotFit. While their use will tend to increase greatly the number of independent parameters being determined, often by hundreds or 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 removing interparameter correlations involving those states in order to facilitate the determination of a good preliminary potential function model for a each state in the early stages of an analysis.

3 Born-Oppenheimer Breakdown Radial Functions

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

$$\widetilde{S}^{\rm A}_{\rm ad}(r) \;=\; y^{\rm eq}_{p_{\rm ad}}(r)\; u^{\rm A}_{\infty} \;+\; [1 - y^{\rm eq}_{p_{\rm ad}}(r)] \sum_{i=0}^{N^{\rm A}_{\rm ad}} u^{\rm A}_i \; y^{\rm eq}_{q_{\rm ad}}(r)^i \quad , \tag{34}$$

$$\widetilde{R}^{\rm A}_{\rm na}(r) \;=\; y^{\rm eq}_{p_{\rm na}}(r)\; t^{\rm A}_{\infty} \;+\; [1 - y^{\rm eq}_{p_{\rm na}}(r)] \sum_{i=0}^{N^{A}_{\rm na}} t^{\rm A}_i \; y^{\rm eq}_{q_{\rm na}}(r)^i \quad . \tag{35}$$

The structure of Eqs. (34) and (35) allows the asymptotic behaviours and equilibrium properties of these functions to be explicitly specified by the user or determined in a fit. In particular, the limiting asymptotic value of $\widetilde{S}^{\rm A}_{\rm ad}(r)$ is $u^{\rm A}_{\infty}$. Hence, if the zero of energy is defined as ground-state atoms separated at $r \sim \infty$ (our recommended convention), then $u^{\rm A}_{\infty} \equiv 0$ for all electronic states that dissociate to yield ground-state atoms, while for a state that dissociates to yield atom A in an excited electronic state, $u^{\rm A}_{\infty}$ is determined by the associated atomic isotope shift [4]. For all states dissociating to ground-state atoms, $u^{\rm A}_0$ and $u^{\rm B}_0$ then define the difference between the well depths of those states for different atom–A and atom–B isotopologues,

$$\delta\mathfrak{D}^{(\alpha)}_{\rm e}(X) \;=\; \frac{\Delta M^{(\alpha)}_{\rm A}}{M^{(\alpha)}_{\rm A}}\; u^{\rm A}_0 \;+\; \frac{\Delta M^{(\alpha)}_{\rm B}}{M^{(\alpha)}_{\rm B}}\; u^{\rm B}_0 \quad , \tag{36}$$

However, for excited electronic states the $u^{\rm A}_0$ and $u^{\rm B}_0$ parameters define the electronic isotope shift. Alternatively, if one wished to define the potential minimum of the ground state as the absolute zero of energy, one would fix $u^{\rm A}_0(X) = u^{\rm B}_0(X) = 0$ and the values of $u^{\rm A}_{\infty}$ and $u^{\rm B}_{\infty}$ would then define the isotopologue dependence of the ground-state dissociation energy *via* an experession analogous to Eq. (36). A user of DPotFit may select this (not recommended) alternate convention by choosing, in the input data file, to fix $u^{\rm A}_0(X) = u^{\rm B}_0(X) = 0$ while allowing $u^{\rm A}_{\infty}$ and $u^{\rm B}_{\infty}$ to be varied freely. In either case, $u^{\rm A}_0$ and $u^{\rm B}_0$ would determine the electronic isotope shift for excited states.

As discussed in Ref. [4], the limiting asymptotic value of the centrifugal BOB correction function $\widetilde{R}^{\rm A}_{\rm na}(r)$ should always be $t^{\rm A}_{\infty} = 0$ unless the species in question is a molecular ion that yields A^+ or A^- upon dissociation, in which case $q^{\rm A}_{\infty}$ would have a small non-zero value [4]. For example, for a molecular ion $AB^{\tt CHARGE}$ that dissociates to yield a neutral atom B plus the atomic ion $A^{\tt CHARGE}$, this limit is

$$t^{\rm A}_{\infty} \;=\; \frac{\mu^{\rm W}}{\mu(A^{\tt CHARGE}, B)}) \;-\; 1 \;\approx\; \frac{Q\, m_{\rm e}}{M_{\rm A}} \;+\; \left(\frac{Q\, m_{\rm e}}{M_{\rm A}}\right)^2 \;+\; \left(\frac{Q\, m_{\rm e}}{M_{\rm A}}\right)^3 \;+\; \ldots \quad , \tag{37}$$

in which $\mu = \mu^{\rm W}$ is the charge-modified reduced mass of Eq. (2), and $\mu(A^{\tt CHARGE}, B)$ is the usual two-particle reduced mass of the ion $A^{\tt CHARGE}$ with the neutral atom B. At the other limit, a convention commonly associated with use of the Watson radial Hamiltonian of Eq. (1) is to fix the leading power-series coefficient

of Eq. (35) as $t_0^{\rm A} = 0$, as it is 100% correlated with $u_1^{\rm A}$ and it represents an indeterinate integration constant in the theory [1, 2, 4]. However, the value of $g^{(\alpha)}(r = r_{\rm e})$ is related to observable electronic properties of the molecule [56, 57, 58], so that when measurements of those properties are available, it may be appropriate to fix $t_0^{\rm A}$ at some specific non-zero value, or even to allow it to be varied in the fit. DPotFit allows a user to select any of these options.

The integer $p_{\rm ad}$ should be set equal to the leading (smallest-power) term in the long-range potential of Eq. (11) or (16), $p_{\rm ad} = m_1$, if the effective adiabatic potentials for 'minor isotopologues' are to have the same limiting functional behaviour as that for the reference isotopologue.[6] However, there are no such physical constraints on the integers $q_{\rm ad}$, $p_{\rm na}$, and $q_{\rm na}$ that define the other radial variables appearing in Eqs. (34) and (35), so they must be selected manually by the user ($q = 3, \cdots, 6$ are reasonable trial values) subject to the twin objectives that the resulting functions provide a compact and accurate representation of the data, and that that they approach their asymptotic values without having spurious extrema in the intervals outside the data-sensitive region (see Fig. 3 of Ref. [4]). It is often convenient to set the powers $q_{\rm ad}$ equal to $p_{\rm ad}$ and $q_{\rm na}$ equal to $p_{\rm na}$ in order to yield expressions for $\widetilde{S}_{\rm ad}^{\rm A/B}(r)$ and $\widetilde{R}_{\rm na}^{\rm A/B}(r)$ that involves only a single type of radial variable [13]. Moreover, since EMO functions do not have *any* inverse-power limiting long-range behaviour, DPotFit sets $p_{\rm ad} = q_{\rm ad}$ in Eq. (34) when an EMO function is used for the potential energy.

Although the BOB parameterization of Eqs. (1), (3) and (4) is preferred for a number of reasons [3], the formally equivalent alternate parameterization of Watson's original paper [1] has sometimes been used by some other research groups, and DPotFit allows a user to employ either formulation. In the Watson approach, the effective adiabatic potential for the reference isotopologue $V_{\rm ad}^{(1)}(r)$ is replaced by the "clamped nuclei" potential $V_{\rm CN}(r)$ to give the radial equation:

$$\left\{ -\frac{\hbar^2}{2\mu_\alpha}\,\frac{d^2}{dr^2} \,+\, \left[V_{\rm CN}(r) + \Delta V_{\rm ad,W}^{(\alpha)}(r) \right] \,+\, \frac{[J(J+1) - \Lambda^2]\hbar^2}{2\mu_\alpha\, r^2} \left[1 + g_{\rm W}^{(\alpha)}(r) \right] \right\} \psi_{v,J}(r) \;=\; E_{v,J}\; \psi_{v,J}(r) \quad , \quad (38)$$

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

$$\Delta V_{\rm ad,W}^{(\alpha)}(r) \;=\; \frac{m_{\rm e}}{M_{\rm A}^{(\alpha)}}\; \widetilde{S}_{\rm ad,W}^{\rm A}(r) \,+\, \frac{m_{\rm e}}{M_{\rm B}^{(\alpha)}}\; \widetilde{S}_{\rm ad,W}^{\rm B}(r) \quad (39)$$

$$g_{\rm W}^{(\alpha)}(r) \;=\; \frac{m_{\rm e}}{M_{\rm A}^{(\alpha)}}\; \widetilde{R}_{\rm na,W}^{\rm A}(r) \,+\, \frac{m_{\rm e}}{M_{\rm B}^{(\alpha)}}\; \widetilde{R}_{\rm na,W}^{\rm B}(r) \quad (40)$$

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` in the input data file (see §9.3). In either case, the radial strength functions $\widetilde{S}_{\rm ad}^{\rm A/B}(r)$ and $\widetilde{R}_{\rm na}^{\rm A/B}(r)$ are expanded as in Eqs. (34) and (35), but the magnitudes of the expansion parameters will differ by the factors $\Delta M_{\rm A/B}^{(\alpha)}/m_{\rm e}$ for $S_{\rm ad,W}^{\rm A/B}(r)$ and $M_{\rm A/B}^{(1)}/m_{\rm e}$ for $R_{\rm ad,W}^{\rm A/B}(r)$.

4 Λ-Doubling Splittings for Singlet States

It was shown in Ref. [5] 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 Hamiltonian, to yield

[6] This assumes that the effective adiabatic potential $V_{\rm ad}^{(1)}(r)$ for the reference isotopologue has a long-range tail defined by Eq. (11) or (16).

$$\left\{ -\frac{\hbar^2}{2\mu_\alpha}\,\frac{d^2}{dr^2} \;+\; \left[V_{\rm ad}^{(1)}(r) + \Delta V_{\rm ad}^{(\alpha)}(r)\right] \;+\; \frac{[J(J+1)-\Lambda^2]\hbar^2}{2\mu_\alpha\, r^2}\left[1+g^{(\alpha)}(r)\right] \right. \qquad (41)$$
$$\left. +\, {\rm sg}_\Lambda(e/f)\;\Delta V_\Lambda^{(\alpha)}(r)\;[J(J+1)]^\Lambda \right\} \psi_{v,J}(r) \;=\; E_{v,J}\;\psi_{v,J}(r) \quad ,$$

in which ${\rm sg}_\Lambda(e/f)$ is a dimensionless numerical factor defined by the e/f parity of the level of interest (see below), and the overall Λ-doubling function is defined as [7]

$$\Delta V_\Lambda^{(\alpha)}(r) \;=\; \left(\frac{\hbar^2}{2\mu_\alpha\, r^2}\right)^{2\Lambda} f_\Lambda(r) \quad . \qquad (42)$$

We know of no theoretical predictions regarding the long-range behaviour expected for the λ-doubling function $\Delta V_\Lambda^{(\alpha)}(r)$ other than that provided by the first term on the right-hand side of Eq. (42). We therefore choose to write the mass-independent radial function $f_\Lambda(r)$ as a simple polynomial expansion in the reduced variable $y_{p\Sigma}^{\rm eq}(r)$ of Eq. (5), that is,

$$f_\Lambda(r) \;=\; \sum_{i=0}^{N_\Lambda} w_i^\Lambda\; y_{q\Lambda}^{\rm eq}(r)^i \quad , \qquad (43)$$

in which the expansion coefficients w_i^Λ have units [8] $1/({\rm cm}^{-1})^{2\Lambda-1}$. The fact that $y_q^{\rm eq}(r) \to 1$ as $r \to \infty$ means that $f_\Lambda(r)$ will necessarily approach a finite value in this limit, and hence that at long range $\Delta V_\Lambda^{(\alpha)}(r) \to 0$ as $1/r^{2\Lambda}$.

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

5 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 [5] yields the following effective radial Hamiltonian for an electronic state with $^2\Sigma$ symmetry:

$$\left\{ -\frac{\hbar^2}{2\mu_\alpha}\,\frac{d^2}{dr^2} \;+\; \left[V_{\rm ad}^{(1)}(r) + \Delta V_{\rm ad}^{(\alpha)}(r)\right] \;+\; \frac{N(N+1)\,\hbar^2}{2\mu_\alpha\, r^2}\left[1+g^{(\alpha)}(r)\right] \right. \qquad (44)$$
$$\left. +\, {\rm sg}_\Sigma(e/f;N)\;\Delta V_\Sigma^{(\alpha)}(r) \right\} \psi_{v,J}(r) \;=\; E_{v,J}\;\psi_{v,J}(r) \quad ,$$

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

$$\Delta V_\Sigma^{(\alpha)} \;=\; \left(\frac{\hbar^2}{2\mu_\alpha\, r^2}\right)\; f_\Sigma(r) \quad , \qquad (45)$$

[7] Note that while the derivation of Ref. [5] only addressed the case of Λ-doubling in Π states, the present implementation has been extended to handle states for which $\Lambda > 1$.

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

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

$$f_\Sigma(r) \;=\; \sum_{i=0} \; w_i^\Sigma \; y_{q_\Sigma}^{\rm eq}(r)^i \quad , \tag{46}$$

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 #6 in §9.3) 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), (38), (41) or (44) 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 lower-state eigenvalues 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:

$$\frac{\partial E_{v,J}}{\partial p_j} \;=\; \left\langle \psi_{v,J}(r) \left| \frac{\partial \widehat{\mathcal{H}}}{\partial p_j} \right| \psi_{v,J}(r) \right\rangle \tag{47}$$

These quantities are required 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 subroutines `SCHR` [59–63]. The present version of those routines incorporates several unique features, such as the ability to locate quasibound (or tunneling-predissociation) levels automatically, and calculate both their widths and the partial derivatives of those widths with respect to the potential function parameters [5, 64–66]. This last capability is required for cases in which measured tunneling predissociation level widths are included in the experimental data set being analyzed [5]. Most details and features of the Schrödinger-solver routine `SCHRQ` used by DPotFit are described in the manual for program `LEVEL` [66], 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, each of which 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 #10 of the data file) used in the numerical integration of Eq. (1), (38), (41) or (44). For potentials that are not too steep or too sharply curved, adequate 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

$$\texttt{RH} \;=\; \pi / \left(\texttt{NPN} \times \left[\mu \times \max\{E - V(r)\} / 16.857629205 \right]^{1/2} \right) \tag{48}$$

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

cases it is approximately the potential well depth $\mathfrak{D}_e$), the reduced mass μ is in amu, and the numerical factor is $\hbar^2/2$ expressed in "spectroscopists' units" [amu cm^{-1} Å^2]. A value of NPN that is too small yields unreliable results, while a value that is too large may require excessive computational effort, and require that array dimensions be made inconveniently large. Note that while Eq. (48) is a useful guide, a careful user *should always* examine the effect of different RH values on the calculated band constants written to output channel-7 in order to ensure that the calculation yields results that have an accuracy appropriate for their particular application.

The numerical integration is performed on the interval from RMIN to RMAX (see READ statement #10) using the Numerov algorithm [59, 67]. These bounds must lie sufficiently far into the inner and outer classically-forbidden regions (where $V_{\rm eff}(r) > E_{v,J}$) that the wavefunction has decayed by several orders of magnitude relative to its amplitude in the classically-allowed region. The present version of the code prints warning messages if this decay is not smaller by a factor of at least 10^{-9}; if such warnings are printed, a smaller RMIN or larger RMAX value should be used to remove them and to ensure that the desired accuracy is achieved. However, if RMIN lies too far into the classically-forbidden regions and $[V_J(r) - E]$ becomes extremely large, then the integration algorithm may become numerically unstable for the specified mesh size. If it does, a warning message is printed, and the beginning of the integration range is automatically shifted outward until the problem disappears. However, use of a slightly larger value of RMIN will cause such warning messages to disappear and (marginally) reduce the computational effort. For most diatomic molecules, a reasonable value of RMIN is ca. $0.5 - 0.8$ times the smallest inner turning point for the levels involved in the data set, but for hydrides or other species of low reduced mass, even smaller values may be required.

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

6.2 Calculating Second Virial Coefficients

For species formed from ground-state rare-gas atoms, measured virial and transport property coefficients may be used together with spectroscopic data in a combined analysis to determine a potential energy function for the species in question. DPOTFIT is currently able to fit to a combination of virial coefficients with all types of spectroscopic data [68]. Second virial coefficients are calculated in the usual fashion, from a sum of a classical term with first and second quantum corrections [69], that is,

$$B_2(T) \;=\; B_{\rm cl}(T) \;+\; \left(\frac{h^2}{2\mu}\right) B_{\rm Q}^{\rm I}(T) \;+\; \left(\frac{h^2}{2\mu}\right)^2 B_{\rm Q}^{\rm II}(T) \;+\; \ldots \tag{49}$$

in which the classical term may be written as

$$B_{\rm cl}(T) \;=\; -\,2\pi\, N_{\rm A} \int_0^\infty \left[e^{-V(r)/{\rm k_B T}} - 1\right] r^2 \; dr \quad , \tag{50}$$

with N_A the Avogadro's number, k_B the Boltzmann constant, and m the atomic mass. The first quantum correction, $B_Q^I(T)$, is given by

$$B_Q^I(T) \;=\; 2\pi N_A \left(\frac{1}{48\pi^2\,(k_B T)^3}\right) \int_0^\infty e^{-V(r)/k_B T} \left(\frac{dV(r)}{dr}\right)^2 r^2\, dr \quad , \tag{51}$$

while the second quantum correction, $B_Q^{II}(T)$, is obtained as

$$\begin{aligned} B_Q^{II}(T) \;=\; -2\pi N_A \left(\frac{1}{1920\pi^4\,(k_B `T)^4}\right) \int_0^\infty e^{-V(r)/k_B T} \Bigg[\left(\frac{d^2V(r)}{dr^2}\right)^2 \;+\; \frac{2}{r^2}\left(\frac{dV(r)}{dr}\right)^2 \\ +\; \frac{10}{9k_B T}\,\frac{1}{r}\left(\frac{dV(r)}{dr}\right)^3 \;-\; \frac{5}{36(k_B T)^2}\left(\frac{dV(r)}{dr}\right)^4 \Bigg] r^2\, dr \quad . \end{aligned} \tag{52}$$

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

7 Fitting Strategies

7.1 Initial Trial Parameters

In a DPF treatment of experimental data, the observables – the transition energies or tunneling lifetimes – are not linear functions of the parameters of the radial functions characterizing the effective Hamiltonian. As in any non-linear least-squares procedure, it is essential to have a set of realistic initial trial values of all fitting 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 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 $V_{ad}^{(1)}(r)$ has been obtained.

For the potential energy function $V_{ad}^{(1)}(r)$ (or $V_{CN}(r)$) itself, the problem of determining initial trial parameters is somewhat more challenging, and a number of strategies have been used. From a conventional preliminary analysis of the data, it is usually fairly straightforward to obtain a good estimate of the potential minimum position r_e, which is a central parameter in all of the model potentials. It is usually 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 potential models. However, 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 set of interest to determine analytic level energy expressions, such as Dunham expansions or near-dissociation expansions [70], and then use the resulting functions to generate a pointwise RKR potential for that state [71]. 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 [72].

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 then when preliminary optimized values of those parameters have been determined, they may replace 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) also to be free. This sort of stepwise procedure is often necessary if the fit to determine the main supporting potential is to be stable. However, once a converged value of 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 #6, DPotFit allows a user to limit the range of data to be utilized in a given fit without having the 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

If one is performing a fit to data involving more than one electronic state, it is often‘ necessary to utilize a stepwise procedure – initially optimizing parameters for one state at a time – before proceeding to the 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 usually the only source of information about the upper vibrational levels of a given state, so they cannot be ignored when one is attempting to obtain a good description of that state.

The best way to treat this problem is then 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: the extension to allow fits to

band constants is introduced here in Version 2.0. These options are invoked by giving the 'potential-type' parameter PSEL that is input via READ #8 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 §9.3). 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 to 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.

8 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. The values of the relevant physical constants occur in the program as the single factor $\hbar^2/(2\,\mu) = 16.857629205/\mu$ [cm^{-1} Å^2] (with μ in amu) appearing in the effective radial Schrödinger equations (1), (38), (41) or (44). This numerical constant is based on the 2010 CODATA recommended physical constant values [73], while the atomic isotope masses stored in subroutine MASSES were taken from the 2003 compilation of Ref. [74].

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

$$\texttt{DRMSD} \;\equiv\; \overline{dd} \;=\; \left\{ \frac{1}{N} \sum_{i=1}^{N} \left[\frac{y_i^{\mathrm{calc}} - y_i^{\mathrm{obs}}}{u_i} \right]^2 \right\}^{1/2} \quad , \tag{53}$$

or by the related dimensionless standard error σ_f, defined as $\texttt{DSE} \equiv \bar{\sigma}_f = \overline{dd}\,\sqrt{N/(N-M)}$. This data weighting allows observables with very different magnitudes and very different absolute uncertainties (e.g., microwave *vs.* electronic band head data) 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. However, the occurrence of converged values larger than unity may simply reflect the fact that the experimental uncertainties assigned to the data were overly optimistic.

In addition to reporting the 95% confidence limit (approximately 'two-sigma') uncertainty in each fitted parameter, DPOTFIT follows the approach of Ref. [75] by also always listing the associated "parameter sensitivity" (identified as PS in the output). This quantity is defined (see Eq. (4) of Ref. [75]) as the magnitude of the largest change in the given parameter whose effect on the predictions of the model could increase $\bar{\sigma}_f$ by a maximum of $(0.1/M)\,\bar{\sigma}_f$. This 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. [75], 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 $\overline{\mathrm{dd}}$.

Another feature of DPOTFIT is its implementation (via subroutine NLLSSRR) of the automated "sequen-

tial rounding and refitting" (SRR) procedure of Ref. [75], 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 set of trial fits required to determine manually the optimum values of $r_{\rm ref}$ and N_β. Application of this SRR procedure is turned on or off by the value of the flag `IROUND` that is set by the user in READ #5 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. [75], in most cases setting `IROUND` $= \pm 1$ yields a maximum degree of rounding without significant loss of precision, but in some cases it may be necessary to set $|$`IROUND`$| > 1$ (see §9.3).

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. [76] 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. [76] and replaces the normal least-squares data weights $w_i = 1/(u_i)^2$ by the '*robust*' weights $w_i^{\rm rob} = 1/\left[(u_i)^2 + (y_i^{\rm calc} - y_i^{\rm obs})^2/3\right]$. Because the latter depend on the then-current degree of agreement of the data with the model, fits of this type are repeated iteratively, with the parameter values and the weights being updated in each cycle until self-consistency is achieved. As a result, *robust* fits require substantially more computer time than normal fits do. Moreover, the fact that *robust* weighting reduces the effect of large $[y_i^{\rm calc} - y_i^{\rm obs}]$ values on `DSE` and $\overline{dd}$ makes it more difficult to interpret differences in those quantities obtained from fits to different versions of a model, and it may tend to obscure the presence of systematic discrepancies that indicate shortcomings of the model, rather than of the data. However, our (limited) experience with this option indicates that it can facilitate identifying 'bad' data as well as local and/or systematic discrepancies from a model, when one examines the $[y_i^{\rm calc} - y_i^{\rm obs}]$ results in the Channel-8 output file.

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 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 F'90 and F'95 compilers) the use of a Fortran-77 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 utility routine `arrsizes.h`, which is supplied with the program. If this file resides in the same directory as the source code when the program is being compiled, FORTRAN compilers will automatically incorporate it into relevant subroutines at compilation time through FORTRAN '`include`' statements in the code. Parameters defined in this way include the maximum number of isotopologues being consideredi, `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 file `arrsizes.h` before compiling the code, and assign these parameters values appropriate 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 § 9.2.

The program writes standard output to Channel-6 plus 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 (aproximately 'two-sigma') uncertainties, and presents a summary of the $[y_i^{\rm calc} - y_i^{\rm obs}]$ 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 $\{[y_i^{\rm calc} - y_i^{\rm obs}]\}$ and $\{[y_i^{\rm calc} - y_i^{\rm obs}]/u_i\}$ values for all data utilized in the fit.

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

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

Channel-11 contains the radial exponent function $\beta(r)$ for the EMO, MLR or DELR model potentials,

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

Channels 14 & 15 contain the non-adiabatic BOB radial strength functions $\widetilde{R}_{\rm na}^{\rm A}(r)$ and $\widetilde{R}_{\rm na}^{\rm B}(r)$, respectively, and

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

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

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

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

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

```
rdpot filename
```

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

This program is written for serial execution, and the amount of computer memory required is ???? HELP, Toby !!!!!

8.3 Generating Sets of Predicted Transition Energies

`DPotFit` may also be employed to generate a set of predicted data $\{y_i^{\rm calc}\}$ from a given (fixed) set of input parameters. This option is invoked by setting the value of the input variable in READ #2, which is normally the name of the file containing the experimental data, to be `MAKEPRED`. In this case the program will use READ #33 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 Data File Structure and Input Parameter Definitions

9.1 The Experimental Data File

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 collected into separate vibrational bands (or fluorescence series, or sets of photo-association spectroscopy (PAS) binding energies, or tunneling predissociation level widths), each characterized by the upper- and lower-state vibrational quantum numbers $v' =$ `VP` and $v'' =$ `VPP`, respectively, by the two-alphanumeric-character labels `LABLP` and `LABLPP` (enclosed between single quotes; e.g. '`X0`') that have been chosen to label the particular upper and lower electronic states, and by the (integer) mass numbers `MN1` and `MN2` of the atoms forming that particular isotopologue. 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 $y_i^{\rm obs} =$ `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
       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 = 'FS'` 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.

Another 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 = 'PA'`. 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.

`PA`-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 if EMO, MLR or DELR functions are used for the potential energy function. If this is not the case, DPOTFIT simply treats a `PA`-type data-group as an ordinary fluorescence series whose upper level is a free fitting parameter.

Another type of experimental data that the program can use is a set of tunneling predissociation level widths. Again, the band-type input data structure is retained, but in this case the input parameter value `LABLP = 'WI'` 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 cm^{-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 *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 = 'VV'`. 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
       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.

A final type of experimental data that can be used in a fit is a set of measured second virial coefficients [68]. In this case, one of the isotopologues specified in READ #4 below should correspond to mass numbers `MN(1,IISTP)` = `MN(2,IISTP)` = 0 to allow 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 = 'VR'`. 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:

```
       COUNT= 1
    10 IBAND= IBAND+ 1
#1     READ(4,*,END=20)  VP(IBAND), VPP(IBAND), LABLP, LABLPP, MN1, MN2
       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 $\mathrm{cm}^3\,\mathrm{mol}^{-1}$) 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 [68] from a combination of microwave data [77], high-resolution vacuum laser spectroscopy [68], and virial coefficient data [78, 79, 80]. 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    'X0'  'X0'      40   128          % v' v" LABLP LABLPP MN1 MN2
    3 +1  2 +1  0.19375109  3.3D-08        % J' p' J" p" FREQ UFREQ8
    4 +1  3 +1  0.25830956  3.3D-08        % MW data of Jaeger et al. (1993)
    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    'X0'  'X0'      40   129          % v' v" LABLP LABLPP MN1 MN2
    3 +1  2 +1  0.19339688  3.3D-08        % J' p' J" p" FREQ UFREQ8
    4 +1  3 +1  0.25783738  3.3D-08        % MW data of Jaeger et al. (1993)
    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    'X0'  'X0'      40   130         % v' v" LABLP LABLPP MN1 MN2
    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'  'X0'      40   132         % v' v" LABLP LABLPP MN1 MN2
    1 +1  0 +1  77239.1655  0.0012       % J' p' J" p" FREQ UFREQ
    2 +1  1 +1  77239.2240  0.0012       % VUV data of Piticco & Merkt
    3 +1  2 +1  77239.2817  0.0010       %    from [JMS 264, 83 (2010)]
    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    'VR'  'X0'       0    0     % v' v" LABLP LABLPP MN1 MN2
   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
```

9.2 The Channel–5 'Instruction' Input File: Specifying the Model and the Fit

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

 #8       READ(5,*) PSEL(ISTATE), VLIM(ISTATE), BOBCN(ISTATE), 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)
 #9              READ(5,*) VTST,(NBC(I,IISTP,ISTATE),IISTP= 1,NISTP)
```

```
                ENDDO
              ENDIF
#10         READ(5,*) RMIN(ISTATE), RMAX(ISTATE), RH(ISTATE)
            IF(PSEL(ISTATE).EQ.0) THEN
#11             READ(5,*) NPT, NUSE, IR2, ILR, NCN, CNN
#12             READ(5,*) RFACT, EFACT, VSHIFT
#13             READ(5,*) (XI(I), YI(I), I= 1,NTP)
                EXIT
                ENDIF
            IF((PSEL(ISTATE).GE.2).AND.(PSEL(ISTATE).LE.5)) THEN
      c** For an MLR or DELR or HPP or TT-type potential ...
#14             READ(5,*) NCMM(ISTATE), rhoAB(ISTATE), IDF(ISTATE, IDSTT(ISTATE)
                DO  m= 1,NCMM(ISTATE)
#15                 READ(5,*) MMLR(m,ISTATE), CmVAL(m,ISTATE), IFXCm(m,ISTATE)
                    ENDDO
                ENDIF
      c** For a GPEF potential ...
#16           IF(PSEL(ISTATE).EQ.6) READ(5,*) AGPEF(ISTATE), BGPEF(ISTATE)
      c** For all  PSEL > 0  cases ...
#17           READ(5,*) DE(ISTATE), IFXDE(ISTATE)
#18           READ(5,*) RE(ISTATE), IFXRE(ISTATE)
#19           READ(5,*) APSE(ISTATE), Nbeta(ISTATE), nPB(ISTATE), nQB(ISTATE), RREF(ISTATE)
            DO  I= 0, Nbeta(ISTATE))
#20             READ(5,*) BETA(I,ISTATE), IFXBETA(I,ISTATE)
                ENDDO
#21           READ(5,*) NUA(ISTATE), NUB(ISTATE), pAD(ISTATE), qAD(ISTATE)
      c** Adiabatic potential energy BOB function parameters for atom-A
            IF(NUA(ISTATE).GE.0) THEN
                DO  I= 0, NUA(ISTATE)
#22                 READ(5,*) UA(I,ISTATE), IFXUA(I,ISTATE)
                    ENDDO
#23             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)
#24                 READ(5,*) UB(I,ISTATE), IFXUB(I,ISTATE)
                    ENDDO
#25             READ(5,*) UBinf, IFXUBinf
                ENDIF
#26           READ(5,*) NTA(ISTATE), NTB(ISTATE), pNA(ISTATE), qNA(ISTATE)
      c** Non-adiabatic centrifugal BOB function parameters for atom-A
            IF(NTA(ISTATE).GE.0) THEN
                DO  I= 0, NTA(ISTATE)
#27                 READ(5,*) TA(I,ISTATE), IFXTA(I,ISTATE)
                    ENDDO
#28             READ(5,*) TAinf, IFXTAinf
                ENDIF
      c** Non-adiabatic centrifugal BOB function parameters for atom-B
            IF(NTB(ISTATE).GE.0) THEN
                DO  I= 0, NTB(ISTATE)
#29                 READ(5,*) TB(I,ISTATE), IFXTB(I,ISTATE)
                    ENDDO
#30             READ(5,*) TBinf, IFXTBinf
                ENDIF
            IF(IOMEG(ISTATE).NE.0) THEN
      c** Lambda-doubling or 2\Sigma splitting expansion parameters
#31             READ(5,*) NwCFT(ISTATE), Pqw(ISTATE), efREF(ISTATE)
                IF(NwCFT(ISTATE).GE.0) THEN
                    DO  I= 0, NwCFT(ISTATE)
#32                     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
#33             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*********************************************************************
```

9.3 Definitions and Descriptions for the Channel-5 Input File Data

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`. Setting `LPRINT`$=0$ yields no internal printout except for convergence-failure warning messages; this is the recommended choice when there are no problems with the fits. If `LPRINT`< 0, print converged unrounded parameters; if `LPRINT`≥ 1, also print converged rounded parameters (when `IROUND`$\neq 0$); if `LPRINT`≥ 2, also print parameter changes on each rounding step; if `LPRINT`≥ 3, also report parameter convergence criterion satisfied; if `LPRINT`≥ 4, also print convergence test on each fitting cycle; if `LPRINT`≥ 5, also print the change and uncertainty for each parameter in each fitting cycle.

`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.

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.

#4 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.

#5 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= 'WI') in the data file; otherwise, set NOWIDTHS $= 0$.

IROUND: Setting (integer) IROUND $\neq 0$ causes the "sequential rounding and refitting" procedure of Ref. [75] 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 § 5.1 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 print 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.

#6. READ(5,*) SLABL(s), IOMEG(s), VMIN(s), VMAX(s), JTRUNC(s), EFSEL(s)

SLABL is a two-alphanumeric-character label for this electronic state, enclosed in single quotes (e.g., 'X0', 'a1', 'f+', ...), 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 $= -2$.

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.

#7. 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.

#8. READ(5,*) PSEL(s), VLIM(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 [66].

If PSEL$=1$ use the Expanded Morse Oscillator (EMO) form of §2.2.

If PSEL$=2$ use the Morse/Long-Range (MLR) potential form of §2.3.

If PSEL$=3$ use the Double-Exponential/Long-Range (DELR) form of §2.4.

If PSEL$=4$ use the Tiemann/Hannover polynomial of §2.5.

If PSEL$=5$ use the Tang-Toennies type potential of Eq. (31) in §2.6.

If PSEL$=6$ use Seto's modification [9] of Šurkus' Generalized Potential Energy Function (GPEF) [18], as described in §2.7. Its parameters p, a_S, and b_S are input via READs #16 and 19.

- 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$=-2$ 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, #9–32, are ignored.

If 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 #9 is used to input the number of such constants for each state, and the remainder of the READs for this state, #10–32, 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.

BOBCN specifies the manner in which the BOB mass-scaling is to be done.

If BOBCN$=0$, use the mass scaling of Eqs. (3) and (4) with isotopologue #1 as the reference isotopologue.

If BOBCN$=1$, combine use of the clamped-nuclei reference potential with mass scaling factors $m_e/M_A^{(\alpha)}$, of Eqs. (39) and (40).

`OSEL`: if (integer) `OSEL` > 0, write every `OSEL`'th point of the final potential energy, its exponent coefficient function $\beta(r)$, and any BOB radial function arrays to Channels 10-16; if `OSEL` ≤ 0, omit such output. Smaller (positive) values of `OSEL` yield larger output files written on a finer radial mesh.

If `PSEL` $= -1$, loop over vibrational levels `VMIN` to `VMAX`, reading in the vibrational quantum number `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 `PSEL` $\neq -1$, ignore READ #9.

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

For all cases for which `PSEL` ≥ 0, read READ #10.

#10. `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 `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. (48) in §3.

If `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 `PSEL` $\neq 0$, skip READs #11 – 13.

#11. `READ(5,*) NTP, NUSE, IR2, ILR, NCN, CNN`

NTP: gives the number of turning point pairs to be input via READ #13 to define the potential.

NUSE: Specifies how the interpolation is to be done. If `NUSE` > 0, use `NUSE`–point piecewise polynomials; if `NUSE` ≤ 0, perform cubic spline interpolation. For highly precise and smooth input points, such as those generated from an RKR calculation, `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 (`NUSE` $= 4$) or splines (`NUSE` ≤ 0) are usually best.

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 `IR2` > 0 causes this to be done (normally recommended). A comparison between results obtained with this option turned on *vs.* off (setting `IR2` ≤ 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 `ILR` $= -1$, 0, or 1 is often most appropriate; however, if the outer turning points extend moderately close to the dissociation limit (at `VLIM`), one should set `ILR` ≥ 2 and specify the theoretically appropriate value of `NCN` (≥ 1), and if it is available, also input an estimate of `CNN` (see below).

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

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

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

For `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 `ILR` ≥ 4, fit outermost `ILR` turning points to a sum of `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 `ILR` ≥ 2, `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 `ILR` ≤ 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} $\text{Å}^{\texttt{NCN}}$] rather than to be determined from a fit to the outermost turning points.

#12. `READ(5,*) RFACT, EFACT, VSHIFT`
#13. `READ(5,*) (XI(I), YI(I), I= 1,NTP)`

RFACT & EFACT : are multiplicative factors required to convert units of the `NTP` input turning point distances `XI`(i) and energies `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 `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 `VLIM`.

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

If $\texttt{PSEL} = 2-5$ (for an MLR, DELR, HPP or TT-type potential), specify the long-range potential energy tail function $u_{\text{LR}}(r)$. For $\texttt{PSEL} \neq 2-5$, skip READs #14 & 15.

#14. `READ(5,*) NCMM(`s`), rhoAB(`s`),IDF(`s`), IDSTT(`s`)`
`DO i= 1, NCMM(`s`)`
#15. `READ(5,*) MMLR(i,`s`), CmVAL(i,`s`), IFXCm(i,`s`)`
`ENDDO`

`NCMM(`s`)` is the number of inverse-power terms included in the definition of $u_{\text{LR}}(r)$ in Eq. (11) or (16).

If $\texttt{rhoAB} \leq 0.0$, the long-range potential is the pure inverse-power sum of Eq. (11).

If $\texttt{rhoAB} > 0.0$, include damping functions in $u_{\text{LR}}(r)$, and $\texttt{rhoAB} = \rho \equiv \rho_{\text{AB}}$ is the system-dependent scaling factor of Eqs. (19). In this case:

- $\texttt{IDF} = 2s$ specifies the limiting short-range behaviour of the damping function to be $D_m(r)/r^m \propto r^s = r^{\texttt{IDF}/2}$.
- If $\texttt{IDSTT} > 0$, use the generalized Douketis-type damping functions of Eq. (17) (recommended). Allows any integer $-4 \leq \texttt{IDF} \leq 0$ ($\texttt{IDF} = -2$ or -1 recommended).
- If $\texttt{IDSTT} \leq 0$, use the generalized Tang-Toennies-type damping functions of Eq. (18). Allows even integers $-4 \leq \texttt{IDF} \leq 4$.

`MMLR(i,`s`)` is the power and `CmVAL(i,`s`)` the coefficient of the i'th contribution to $u_{\text{LR}}(r)$ in Eq. (11) or (16), C_{m_i}/r^{m_i}. Positive `CmVAL` values yield attractive potential energy terms.

`IFXCm(i,`s`)` controls whether $C_{m_i} = \texttt{CmVAL(i,}s\texttt{)}$ is to be fitted ($\texttt{IFXCm} \leq 0$) or held fixed ($\texttt{IFXCm} = 1$) 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. Setting $\texttt{IFXCm} > 1$ causes this C_m coefficient to be constrained to equal the value of parameter #`IFXCm`, which has a smaller parameter # index than its own. E.g. to fix C_m for state `ISTATE` equal to the leading C_m value for the first electronic state ($\texttt{ISTATE} = 1$), set $\texttt{IFXCm}(s) = 3$ (since r_{e} and $\mathfrak{D}_{\text{e}}$ are parameters #1 and 2, respectively). This allows one to fit simultaneously to potentials for two states with this C_m coefficient allowed to vary subject to the constraint that the values for the two states always remain the same.

For the special case of the states of Li_2 dissociating to the Li($2S$)+Li($2P$) asymptote, to use the Aubert-Frécon 2×2 or 3×3 diagonalization long-range form [38], set $\texttt{NCMM} = 6$ with $\texttt{MMLR}(i) = 3,\ x,\ 6,\ 6,\ 8,$ & 8, for which the input values of `CmVAL`(i) are C_3^{Σ}, A_{so}, C_6^{Σ}, C_6^{Π}, C_8^{Σ} and C_8^{Π}, for $i = 1 \cdots 6$, respectively, and

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

If PSEL = 6, read in parameters defining the expansion variable in the GPEF potential of Eq. (32); otherwise, skip READ #16.

#16. READ(5,*) AGPEF(s), BGPEF(s)

In the GPEF radial expansion variable of Eq. (32): $a_S = \texttt{AGPEF}$ and $b_S = \texttt{BGPEF}$, while $p = \texttt{nPB}$ is input via READ #19.

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

#17. READ(5,*) DE(s), IFXDE(s)

#18. READ(5,*) RE(s), IFXRE(s)

DE $\equiv \mathfrak{D}_e$ is the dissociation energy. This is a dummy variable for the case of a GPEF potential.

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 integers specifying the form used for the exponent-coefficient function $\beta(r)$, the order of the polynomial defining the potential function $V_{\rm ad}^{(1)}(r)$ (or $V_{\rm CN}(r)$), and the nature of the radial variables of Eqs. (5) and (6).

#19. READ(5,*) APSE(s), Nbeta(s), nPB(s), nQB(s), RREF(s)

For PSEL = 2,

- Setting APSE(s) ≤ 0 invokes use of the constrained-polynomial expansion of Eq. (14) for the exponent-coefficient function $\beta(r) = \beta^{\rm PE}(r)$ of the MLR potential.
- Setting APSE(s) > 0 invokes use of the 'Pashov-Spline' expression for the MLR exponent coefficient $\beta^{\rm SE}(r)$ described in §2.3.3 [44, 72] (this option is not yet fully implemented).

For PSEL = 1–3, Nbeta is the order of the potential function exponent-coefficient polynomial expansion of Eq. (8) or (14).

For PSEL = 4, Nbeta is the order of the "X–representation" polynomial expansion of Eq. (28).

For PSEL = 5, set Nbeta = 0 and omit READs #21-32.

For PSEL = 6, Nbeta is the order of the GPEF polynomial expansion of Eq. (32).

nPB $= p$, nQB $= q$ and RREF $= r_{\rm ref}$ define the radial variables of Eqs. (5), (6), (14) and (32).

For PSEL > 0, loop over the range $i = 0$ to $N_\beta = \texttt{Nbeta}(s)$ and read initial trial values of the exponent expansion parameters $\beta_i = \texttt{BETA(i)}$ of Eqs. (8), (14) or (28) or of the coefficients $c_i = \texttt{BETA(i)}$ of the GPEF potential of Eq. (32). Skip this READ statement if PSEL ≤ 0.

#20. READ(5,*) (BETA(I,s), IFXBETA(I,s), I=0, Nbeta(s))

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

IFXBETA(I,s) controls whether β_i is to be fitted (IFXBETA ≤ 0) or held fixed (IFXBETA$' > 0$).

Now, read in parameters specifying the Born-Oppenheimer breakdown radial functions. First, for the effective adiabatic functions $\widetilde{S}_{\rm ad}^A$ and $\widetilde{S}_{\rm ad}^B$ of Eqs. (3) and (34).

#21. READ(5,*) NUA(ISTATE), NUB(ISTATE), pAD(ISTATE), qAD(ISTATE)

`NUA` $= N_{\rm ad}^{\rm A}$ and `NUB` $= N_{\rm ad}^{\rm B}$ are the (integer) orders of the polynomial in $y_{q_{\rm ad}}(r)$ used to define the effective adiabatic BOB radial function of Eq. (34) for atoms A and B, respectively. If `NUA` < 0 or `NUB` < 0, omit the adiabatic BOB terms for that atom; i.e., skip READ s #22 and #23, or READ s #24 and #25, respectively.

• *Note*: for a chemically homonuclear molecule (`AN`(1) = `AN`(2)), set `NUB` $= -1$, as the parameters for atom A are used for both atoms.

`pAD`(s) $= p_{\rm ad}$ and `qAD`(s) $= q_{\rm ad}$ define the nature of the radial variables $y_{p_{\rm ad}}(r)$ and $y_{q_{\rm ad}}(r)$ appearing in Eq. (34). In cases for which `NCMM`$(s) > 0$ normally set `pAD`(s) = `MMLR`$(1, s)$.

If `NUA`$(s) \geq 0$, loop over $i = 1$ to `NUA`(s) while reading in initial trial values of parameters `UA`$(i, s) = u_i^{\rm A}$ of Eq. (34) and integer parameter `IFXUA`(i, s) to control whether $u_i^{\rm A}$ is to be fitted (`IFXUA`$(i, s) \leq 0$) or held fixed (`IFXUA`$(i, s) > 0$).

#22. `READ(5,*) UA(I,`s`), IFXUA(I,`s`), I=0, NUA(`s`)`

If `NUA`$(s) \geq 0$, read `UAinf` $= u_\infty^{\rm A}$, the limiting asymptotic value of the atom-A radial BOB function $\widetilde{S}_{\rm ad}^{A}(R)$, and integer parameter `IFXUAinf` that controls whether `UAinf` is to be fitted (`IFXUAinf` ≤ 0) or held fixed (`IFXUAinf` > 0).

#23. `READ(5,*) UAinf, IFXUAinf`

If `NUB`$(s) \geq 0$, loop over $i = 1$ to `NUB`(s) while reading in initial trial values of parameters `UB`$(i, s) = u_i^{\rm B}$ of Eq. (34) and integer parameter `IFXUB`(i, s) to control whether $u_i^{\rm B}$ is to be fitted (`IFXUB`$(i, s) \leq 0$) or held fixed (`IFXUB`$(i, s) > 0$).

#24. `READ(5,*) (UB(I,`s`), IFXUB(I,`s`), I=0, NUB(`s`))`

If `NUB`$(s) \geq 0$, read `UBinf` $= u_\infty^{\rm B}$, the limiting asymptotic value of the atom-B radial BOB function $\widetilde{S}_{\rm ad}^{\rm B}(R)$, and integer parameter `IFXUBinf` that controls whether `UBinf` is to be fitted (`IFXUBinf` ≤ 0) or held fixed (`IFXUBinf` > 0).

#25. `READ(5,*) UBinf, IFXUBinf`

Now, read in parameters defining the non-adiabatic centrifugal BOB radial functions $\widetilde{R}_{\rm na}^{\rm A}$ and $\widetilde{R}_{\rm na}^{\rm B}$ of Eqs. (4) and (35).

#26. `READ(5,*) NTA(ISTATE), NTB(ISTATE), pNA(ISTATE), qNA(ISTATE)`

`NTA` $= N_{\rm na}^{\rm A}$ and `NTB` $= N_{\rm na}^{\rm B}$ are the (integer) orders of the polynomial in $y_{q_{\rm na}}(r)$ used to define the nonadiabatic centrifugal BOB radial function of Eq. (35) for atoms A and B, respectively. If `NTA` < 0 or `NTB` < 0, omit the centrifugal BOB terms for that atom; i.e., skip READ s #27 and #28, or READ s #29 and #30, respectively.

• *Note*: for a chemically homonuclear molecule (`AN`(1) = `AN`(2)), set `NTB` $= -1$, as the parameters for atom A are used for both atoms.

`pNA`(s) $= p_{\rm na}$ and `qNA`(s) $= q_{\rm na}$ define the nature of the radial variables $y_{p_{\rm na}}(r)$ and $y_{q_{\rm na}}(r)$ appearing in Eq. (35).

If `NTA`$(s) \geq 0$, loop over $i = 1$ to `NTA`(s) while reading in initial trial values of parameters `TA`$(i, s) = t_i^{\rm A}$ of Eq. (34), and integer parameter `IFXTA`(i, s) to control whether $t_i^{\rm A}$ is to be fitted (`IFXTA`$(i, s) \leq 0$) or held fixed (`IFXTA`$(i, s) > 0$).

#27. `READ(5,*) TA(I,`s`), IFXTA(I,`s`)`

If `NTA`$(s) \geq 0$, read `TAinf` $= t_\infty^{\rm A}$, the limiting asymptotic value of the atom-A radial BOB function $\widetilde{R}_{\rm na}^{\rm A}(R)$, and the integer parameter `IFXTAinf` which controls whether `TAinf` is to be fitted (`IFXTAinf` ≤ 0) or held fixed (`IFXTAinf` > 0).

#28. `READ(5,*) TAinf, IFXTAinf`

If `NTB`$(s) \geq 0$, loop over $i = 1$ to `NTB`(s) while reading in initial trial values of parameters `TB`$(i, s) = t_i^{\rm B}$ of Eq. (35), and integer parameter `IFXTB`(i, s) to control whether $t_i^{\rm B}$ is to be fitted (`IFXTB`$(i, s) \leq 0$) or held fixed (`IFXTB`$(i, s) > 0$).

#29. `READ(5,*) (TB(I,`s`), IFXTB(I,`s`), I=0, NTB(`s`))`

If `NTB(`s`)` ≥ 0, read `TBinf` $= t_\infty^{\rm B}$, the limiting asymptotic value of the atom-B radial BOB function $\widetilde{R}_{\rm na}^{\rm B}(R)$, and the integer parameter `IFXTBinf` which controls whether `TBinf` is to be fitted (`IFXTBinf` ≤ 0) or held fixed (`IFXTBinf` > 0).

#30. `READ(5,*) TBinf, IFXTBinf`

Finally, if the electronic angular momentum quantum number `IOMEG(`s`)` is not zero, read in parameters defining the radial Λ-doubling or $^2\Sigma$ spin-splitting radial strength function of Eqs. (43) or (46), respectively.

#31. `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 #32.

`Pqw` is the integer q_Λ or q_Σ defining the radial expansion variable $y_{p_\Lambda}^{\rm eq}(r)$ in Eq. (43) or $y_{p_\Sigma}^{\rm eq}(r)$ in Eq. (46).

`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`)` ≥ 0, loop over $i = 1$ to `NwCFT(`s`)` while reading in initial trial values of parameters `wCFT(`i,s`)` $= w_i$ of Eq. (43) or (46), and integer parameter `IFXwCFT(`i,s`)` to control whether thia s w_i is to be fitted (`IFXwCFT(`i,s`)` ≤ 0) or held fixed (`IFXwCFT(`i,s`)` > 0).

#32. `READ(5,*) wCFT(I,`s`), IFXwCFT(I,`s`)`

To Generate predictions from a set of fixed system parameters.

Loop over READ #33 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.

#33. `READ(5,*) VP(IBAND),VPP(IBAND), LABLP,LABLP, MN1,MN2, PP,PPP, 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 two-alphanumeric-character names (enclosed in single quotes, as in '`X0`') 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`.

9.4 Illustrative Sample Input Data Files

This program has been successfully applied to a wide variety of problems [7, 8, 28, 81, 4, 5, 20, 21, 13, 14, 15, 16, 10, 68, 11, 12, 29, 82]. The following subsection briefly describes two of these cases and presents the associated Channel-5 instruction input data file.

9.4.1 Determining the Potential Energy Function for the Ground $X\,^1\Sigma_g^+$ State of Ca_2

This is the case described in Refs. [14] and [11], 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 [31, 32] 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 ss a free parameter.

```
 20 20  0  1  1  0  0          % AN(1) AN(2) CHARGE NISTP NSTATES LPRINT PRINP
 '../Ca2_X_FS.4'               % Name of PotFit input data file
 'N7q4C6rr6_85'                % Writefile
  40  40                       % MN1(1) MN2(1)
 1.d0 1 -0  0  30  0           % UCUTOFF NOWIDTHS IROUND ROBUST CYCMAX uBv
 'X0' 0  0 38 300 0            %(1) SLABL IOMEG VMIN VMAX JTRUNC EFSEL
 2   0.d0   0  20              % PSEL VLIM BOBCN OSEL
 3.00 99.0  0.0005d0           % RMIN RMAX RH

 3  0.5870d0  -2  1            % NCMM rhoAB IDF 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)

  1.102080181693D+03  0        % De  IFXDe
  4.277807094884D+00  0        % Re  IFXRe

 -1  7   5  4   6.85D+00       % APSE Nbeta nPB nQB RREF
 -2.044539573798D-01  0        % BETA 0  IFXBETA 0
 -1.400127193907D-01  0        % BETA 1  IFXBETA 1
  4.725673963595D-02  0        % BETA 2  IFXBETA 2
  4.978784800505D-02  0        % BETA 3  IFXBETA 3
  5.665002150789D-02  0        % BETA 4  IFXBETA 4
  6.977372610379D-02  0        % BETA 5  IFXBETA 5
  3.617615587106D-02  0        % BETA 6  IFXBETA 6
  5.342138733727D-02  0        % BETA 7  IFXBETA 7

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

9.4.2 Determining Potential Energy and BOB Functions for the Ground $X\,^2\Sigma^+$ State of MgH and MgD

This is the case described in Ref. [82], 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 M3LR(12) (i.e., $N_\beta = 12$) potential energy function, as well as adiabatic (potential energy) and non-adiabatic (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. 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
'X0'  -1  0 -11  99  0          % SLABL IOMEG VMAX JTRUNC   EFSEL
```

```
  11 10 10 15 11 11            % VMAX(istp), ISTP= 1,6
  2 0.000   0    20            % PSEL VLIM BOBCN OSEL
 0.60  99.50  0.0025           % RMIN RMAX RH
 3 0.81 -2  1                  % NCMM rhoAB IDF 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

 -1 12   5  4   2.74D+00       % APSE Nbeta nPB nQB 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   6  4                  % NUA NUB  pAD qAD
  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                  % NTA NTB  pNA qNA
  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 0 20                       % PSEL VLIM BOBCN OSEL
'A2' 0 0 99 99 0                   SLABL IOMEG VMAX JRUNC    EFSEL
 -2  0 0 20
'B0' 0 0 99 99 0                   SLABL IOMEG VMAX JRUNC    EFSEL
 -2  0 0 20
```


Acknowledgements

We are grateful to Professor George McBane of Grand Valley State University for helpful comments and suggestions, and to Professor F.R.W. McCourt of the University of Waterloo for a thorough, critical reading of the manuscript.

References

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

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

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

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

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

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

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

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

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

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

[11] 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).

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

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

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

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

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

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

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

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

[20] 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).

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

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

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

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

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

[26] 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.

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

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

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

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

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

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

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

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

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

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

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

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

[44] J. Tao, R. J. Le Roy, and A. Pashov, in *65th Ohio State University International Symposium on Molecular Spectroscopy* (Columbus, Ohio, 2010), paper MG07.

[45] M. Ivanova, A. Stein, A. Pashov, H. Knöckel, and E. Tiemann, J. Chem. Phys. **134**, 024321 (2011).

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

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

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

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

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

[51] K. T. Tang and J. P. Toennies, **106**, 1645 (2008).

[52] J. L. Dunham, Phys. Rev. **41**, 721 (1932).

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

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

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

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

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

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

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

[60] J. Cashion, J. Chem. Phys. **39**, 1872 (1963).

[61] 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).

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

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

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

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

[66] R. J. Le Roy, LEVEL 8.0: *A Computer Program for Solving the Radial Schrödinger Equation for Bound and Quasibound Levels*, University of Waterloo Chemical Physics Research Report CP-663 (2007); see `http://leroy.uwaterloo.ca/programs/`.

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

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

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

[70] R. J. Le Roy, `DParFit 3.3`: *A Computer Program for Fitting Multi-Isotopologue Diatomic Molecule Spectra*, University of Waterloo Chemical Physics Research Report CP-660 (2005); see `http://leroy.uwaterloo.ca/programs/`.

[71] R. J. Le Roy, `RKR1 2.0`: *A Computer Program Implementing the First-Order RKR Method for Determining Diatomic Molecule Potential Energy Curves*, University of Waterloo Chemical Physics Research Report CP-657 (2003); see `http://leroy.uwaterloo.ca/programs/`.

[72] R. J. Le Roy, `betaFIT 2.1`: *A Computer Program to Fit Potential Function Points to Selected Analytic Functions*, University of Waterloo Chemical Physics Research Report CP-666 (2013); see `http://leroy.uwaterloo.ca/programs/`.

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

[74] G. Audi, A. H. Wapstra, and C. Thibault, Nucl. Phys. **A 729**, 337 (2003).

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

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

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

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

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

[80] 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).

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

[82] R. Henderson, A. Shayesteh, J. Tao, C. Haugen, P. F. Bernath, and R. Le Roy, J. Chem. Phys. **XXX**, xxx (2013).

Appendix A. 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 [74] 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` [66], 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. [75].

`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` [66].

`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.

Appendix B. Sample Program Output

Channel-6 output file yielded by the fit to data for ground-state Ca_2 described in § 9.4.1.

```
 Input data for  1  isotopomer(s)
 *******************************
     Isotopomer       Mass of atom-1   Mass of atom-2    Reduced mass
 -----------------    --------------   --------------   --------------
 Ca( 40) - Ca( 40)     39.962590980     39.962590980     19.981295490

Use experimental data input file:  ../Ca2_X_FS.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=  30 cycles

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

State X0 represented by an MLR(p= 5, q= 4) potential defined in terms of
========  exponent coefficient:  beta(R)= betaINF*y5 +(1-y5)*Sum{beta_i*y4^i}
     using exponent power series order  7  with radial variables:
       with radial variable:   y_{p,q} = (R^q - 6.850000^q)/(R^q + 6.850000^q)
   uLR inverse-power terms incorporate DS-type damping with   rhoAB= 0.5870000
       defined to give very short-range  Dm(r)*Cm/r^m  behaviour   r^{-1.0}
       Dm(r)= [1 - exp(- 3.30(rhoAB*r)/m - 0.423(rhoAB*r)^2/sqrt{m})]^{m-1.0}
                                              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= -0.1975595686

  Parameter    Initial Value    Uncertainty   Sensitivity
   VLIM      0.000000000000D+00       --            --
     De      1.102081323362D+03    0.0D+00       0.0D+00
     Re      4.277802505981D+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)  -2.023830181671D-01    0.0D+00       0.0D+00
  beta( 1)  -1.434281568915D-01    0.0D+00       0.0D+00
  beta( 2)   4.117395913512D-02    0.0D+00       0.0D+00
  beta( 3)   4.994173725634D-02    0.0D+00       0.0D+00
  beta( 4)   5.796606380251D-02    0.0D+00       0.0D+00
  beta( 5)   7.791736140275D-02    0.0D+00       0.0D+00
  beta( 6)   4.095935204535D-02    0.0D+00       0.0D+00
  beta( 7)   5.104832694584D-02    0.0D+00       0.0D+00
  beta_INF  -1.975595686440D-01       --            --
  C11{eff}  -1.103784928085D+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

After Cycle # 1:   DRMSD= 2.61426D+06    test(PS)= 2.6D+02   test(PU)= 4.3D+00
After Cycle # 2:   DRMSD= 2.30588D+01    test(PS)= 1.4D-08   test(PU)= 2.4D-10
Full  2-cycle convergence:  Max{|change/sens.|}= 1.4D-08 < 1   DRMSD=  2.31D+01

 3553 data fit to  180 param. yields  DRMS(devn)= 2.30588D+01  tst(PS)= 1.4D-08

After Cycle # 1:   DRMSD= 2.30588D+01    test(PS)= 6.5D+03   test(PU)= 2.9D+00
After Cycle # 2:   DRMSD= 6.19836D-01    test(PS)= 7.9D+03   test(PU)= 2.2D-01
After Cycle # 3:   DRMSD= 6.20546D-01    test(PS)= 1.4D+02   test(PU)= 9.0D-02
After Cycle # 4:   DRMSD= 6.19003D-01    test(PS)= 5.7D+01   test(PU)= 3.7D-02
After Cycle # 5:   DRMSD= 6.18751D-01    test(PS)= 2.3D+01   test(PU)= 1.5D-02
After Cycle # 7:   DRMSD= 6.18703D-01    test(PS)= 3.7D+00   test(PU)= 2.4D-03
After Cycle # 8:   DRMSD= 6.18702D-01    test(PS)= 1.5D+00   test(PU)= 9.7D-04
After Cycle # 9:   DRMSD= 6.18702D-01    test(PS)= 6.1D-01   test(PU)= 3.9D-04

Full  9-cycle convergence:  Max{|change/sens.|}= 6.1D-01 < 1   DRMSD= 6.19D-01

================================================================
For fit of  191  parameters to  3553  transitions,  DSE= 0.63603356

================================================================
 The following  180 Fluorescence Series Origins were determined
----------------------------------------------------------
 ( v', J', p'; ISTP)    T(value)    Uncertainty  Sensitivity
----------------------------------------------------------
```

```
  (   5, 67, +1; 1)    18847.271208    5.4D-03     2.6D-05
  (   6, 97, +1; 1)    19247.655858    5.8D-03     2.2D-05
  (   2,161, +1; 1)    19666.763087    8.6D-03     3.2D-05

..... omit 172 lies to save space .....

  (   9,121, +1; 1)    19892.615893    7.9D-03     7.2D-05
  (  12,119, +1; 1)    20206.419041    6.7D-03     2.8D-05
  (  12, 91, +1; 1)    19895.676097    6.3D-03     5.0D-05
  (  11,125, +1; 1)    20172.567006    8.0D-03     7.1D-05
  (  12, 63, +1; 1)    19665.007946    5.3D-03     2.1D-05
 ================================================================================

 State X0 represented by an MLR(p= 5, q= 4) potential defined in terms of
 ========  exponent coefficient:  beta(R)= betaINF*y5 +(1-y5)*Sum{beta_i*y4^i}
      using exponent power series order  7  with radial variables:
        with radial variable:   y_{p,q} = (R^q - 6.850000^q)/(R^q + 6.850000^q)
    uLR inverse-power terms incorporate DS-type damping with   rhoAB= 0.5870000
        defined to give very short-range  Dm(r)*Cm/r^m  behaviour   r^{-1.0}
        Dm(r)= [1 - exp(- 3.30(rhoAB*r)/m - 0.423(rhoAB*r)^2/sqrt{m})]^{m-1.0}
                                              C6= 1.0452922D+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= -0.1972792053

   Parameter      Final Value    Uncertainty   Sensitivity
    VLIM      0.000000000000D+00       --          --
      De      1.102080181693D+03    5.4D-03       3.2D-06
      Re      4.277807094884D+00    1.0D-05       6.7D-09
     C 6      1.045292227971D+07    6.5D+04       8.8D-01
     C 8      3.060800000000D+08       --          --
     C10      8.344000000000D+09       --          --
   beta( 0)  -2.044539573798D-01    9.6D-04       9.2D-09
   beta( 1)  -1.400127193907D-01    5.6D-03       1.7D-08
   beta( 2)   4.725673963595D-02    7.2D-03       2.6D-08
   beta( 3)   4.978784800505D-02    7.2D-03       3.7D-08
   beta( 4)   5.665002150789D-02    8.8D-03       5.0D-08
   beta( 5)   6.977372610379D-02    1.6D-02       6.5D-08
   beta( 6)   3.617615587106D-02    2.2D-02       8.3D-08
   beta( 7)   5.342138733727D-02    1.2D-02       1.0D-07
   beta_INF  -1.972792052503D-01       --          --
   C11{eff}  -1.087599128775D+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   -0.02681   0.654
   6  97  e    37      0     21   7.6D-03  1.5D-02    0.02953   0.493
   2 161  e    16      1      9   8.8D-03  5.0D-02   -0.00048   0.347
   8  23  e    17      0     14   9.9D-03  1.5D-02    0.03919   0.603

..... omit 172 lies to save space .....

  12 119  e    21      0     13   7.4D-03  1.5D-02   -0.02231   0.489
  12  91  e    12      0      9   1.1D-02  1.5D-02   -0.17082   0.743
  11 125  e     9      0      8   1.2D-02  1.5D-02    0.18483   0.951
  12  63  e    40      0     21   8.3D-03  1.8D-02    0.00723   0.616
 ----------------- For these  3553 lines, overall:   -0.00413   0.619

 ================================================================
 Fit of   193 param to  3553 data:  DRMS(devn)= 0.618702
 ================================================================
```