# betaFIT: A computer program to fit pointwise potentials to selected analytic functions

Robert J. Le Roy [a],*, Asen Pashov [b]

[a] Department of Chemistry, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1
[b] Department of Physics, Sofia University, bul. J. Bourchier 5, 1164 Sofia, Bulgaria

## ARTICLE INFO

## ABSTRACT

This paper describes program **betaFIT**, which performs least-squares fits of sets of one-dimensional (or radial) potential function values to four different types of sophisticated analytic potential energy functional forms. These families of potential energy functions are: the Expanded Morse Oscillator (EMO) potential [J Mol Spectrosc 1999;194:197], the Morse/Long-Range (MLR) potential [Mol Phys 2007;105:663], the Double Exponential/Long-Range (DELR) potential [J Chem Phys 2003;119:7398], and the "Generalized Potential Energy Function (GPEF)" form introduced by Šurkus et al. [Chem Phys Lett 1984;105:291], which includes a wide variety of polynomial potentials, such as the Dunham [Phys Rev 1932;41:713], Simons–Parr–Finlan [J Chem Phys 1973;59:3229], and Ogilvie–Tipping [Proc R Soc A 1991;378:287] polynomials, as special cases. This code will be useful for providing the realistic sets of potential function shape parameters that are required to initiate direct fits of selected analytic potential functions to experimental data, and for providing better analytical representations of sets of ab initio results.

© 2016 Elsevier Ltd. All rights reserved.

## 1. Introduction

Potential energy functions provide a ubiquitous means for summarizing what we know about a molecule. However, while it is most convenient and useful to have these functions in analytic form with very high orders of smoothness and differentiability, they are most often determined in pointwise form, either empirically, by utilizing the semiclassical Rydberg–Klein–Rees (RKR) procedure [1–4], or by employing ab initio methods. Moreover, because of the sophisticated form of the successful modern potential functional forms, it is a non-trivial task to obtain the sensible preliminary sets of parameters values that are required in order to enable initiation of the type of complex non-linear least-squares 'Direct Potential Fit'

(DPF) procedure that has become an increasingly common means for analysing experimental data in recent years [5–15]. To help address these problems, this paper describes a computer program that has been developed for taking a set of such radial or 1-D potential function points and fitting it to one of the four families of analytic functional forms that are differentiable and smooth to infinite order. Three of the four forms have the physically significant well depth, $\mathfrak{D}_e$, and equilibrium distance, $r_e$, as central defining parameters, and display sensible short- and long-range extrapolation behaviours outside the range of the set of input points. The fourth family, which is included mainly for historical reasons, has $r_e$ as a central parameter but, as for all polynomial functions, its extrapolation behaviour is completely unpredictable.

Another important application of this code will be as a tool for providing better analytic functions to represent sets of ab initio potential energies. Much recent work on weakly bonded systems has been based on the use of potential forms that are represented as a simple sum of

* Corresponding author.
E-mail addresses: leroy@UWaterloo.ca (R.J. Le Roy),
pashov@phys.uni-sofia.bg (A. Pashov).

damped inverse-power dispersion terms with a single repulsive exponential term whose exponent coefficient and pre-factor are an increasingly complicated clutters of distance-dependent terms [16–22]. The potential function forms offered in **betaFIT** provide a more natural way of combining two key physically significant parameters, the well depth $\mathfrak{D}_e$ and the equilibrium distance $r_e$, with a variable number of potential shape parameters, in forms that merge naturally to the correct inverse-power-sum long-range limit.

## 2. Potential function forms

Program **betaFIT** will fit a set of read-in potential-function points to one of the four families of analytic functional forms. The first two may be thought of as generalizations of familiar simple two- and three-parameter potentials that allow the resulting functions to represent not only conventional single-well potentials, but also potentials that have double-minima or 'shelf' behaviour. The third type of potential form (the DELR function) was originally developed to deal with potentials possessing a barrier that protrudes above the asymptote, but it should also be able to represent normal single-minimum, double-minimum, or shelf-state potentials. The fourth family of potential functions considered herein consists of polynomials expanded about the potential minimum (or equilibrium distance), for which different definitions of the radial expansion variable provide a variety of particular historically significant forms.

As the potential functional forms described below are all expressed in terms of non-standard dimensionless radial expansion variables, the presentation begins with some remarks about those radial variables.

### 2.1. The radial expansion variable

Although potential energy functions are defined on the infinite radial domain $r \in [0, \infty)$, it is not advisable to represent them as explicit functions of the internuclear distance $r$ itself, because all polynomial functions of $r$ approach $\pm \infty$ as $r \to \infty$. The first three types of potential function described below are therefore based around an exponential term whose exponent coefficient is a slowly varying function of distance that is expressed in terms of radial variables having the forms

$$y_p^{r_e}(r) = y_p(r; r_e) \equiv \frac{r^p - r_e^p}{r^p + r_e^p}, \tag{1}$$

and/or

$$y_q^{\mathrm{ref}}(r) = y_q(r; r_{\mathrm{ref}}) \equiv \frac{r^q - r_{\mathrm{ref}}^q}{r^q + r_{\mathrm{ref}}^q}, \tag{2}$$

in which the powers $p$ and $q$ are small positive integers, $r_e$ is the equilibrium internuclear distance, and $r_{\mathrm{ref}}$ is a positive reference distance (usually set greater than $r_e$) that is chosen as the expansion centre. These variables are particular versions of the "GPEF" expansion variable discussed in Section 2.5, particular forms of which had been introduced by a number of authors [23–27]. Note that the
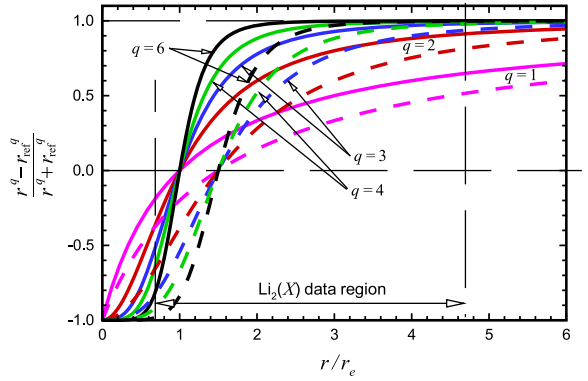


**Fig. 1.** (Adapted from Fig. 2 of Ref. [28]) Illustrative plots of the expansion variables of Eqs. (1) and (2) showing the 'data range' associated with an analysis for the $X^1\Sigma_g^+$ state of Li$_2$ [28]. The solid curves show the variables $y_q^{r_e}(r) = y_q(r; r_e)$ for a range of $q$ values, while the dashed curves show analogous $y_q^{\mathrm{ref}}(r) = y_q(r; r_{\mathrm{ref}})$ plots for $r_{\mathrm{ref}} = 1.5\, r_e$. (For interpretation of the references to color in this figure, the reader is referred to the web version of this paper.)

potential form of Section 2.3 requires two versions of the variable of Eq. (2), $y_p^{\mathrm{ref}}(r)$ and $y_q^{\mathrm{ref}}(r)$, that are defined by different values of the power on the radial variable.

The nature of these variables is illustrated in Fig. 1 for a range of values of the power $q$ [28]. Because $y_p^{r_e}(r)$ and $y_{\{p/q\}}^{\mathrm{ref}}(r)$ approach finite limits both as $r \to 0$ and as $r \to \infty$, functions of these variables will also approach finite values in these limits. At the same time, the fact that $y_p^{r_e}(r) \propto (r - r_e)$ and $y_q^{\mathrm{ref}} \propto (r - r_{\mathrm{ref}})$ 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 onto the finite interval $y_{\{p/q\}}(r) \in [-1, +1)$ greatly facilitates the imposition of proper theoretical constraints on the behaviour of the potential function both at long range, and in the very short-range region, and it means that functions defined as finite power series in $y_q^{\mathrm{ref}}(r)$ will not have singularities at either very small or very large values of $r$.

Earlier work employing this type of exponent radial variable was based on the single variable $y_p^{r_e}(r)$ of Eq. (1) [29–32]. Later work, however, showed that use of an expansion variable $y_q^{\mathrm{ref}}(r)$ centred at a distance $r_{\mathrm{ref}} > r_e$ can, with no loss of accuracy, lead to much more compact and robust potential function expressions than could otherwise be obtained [28,33,34]. This is readily explained by the differences between the solid and dashed curves associated with particular values of $q$, shown in Fig. 1. In particular, for $q = 4$ the solid green curve for $y_4^{r_e}(r)$ is flat and lies very close to its upper limit value of $+1$ over a substantial fraction of the 'data region'. Consequently, low-order expansions in that variable would have difficulty providing accurate representations of properties that vary significantly across this domain. In contrast, the dashed green curve shows that the corresponding variable, $r_4^{\mathrm{ref}}(r) = y_4^{1.5\,r_e}(r) = y_4(r;\ r_{\mathrm{ref}} = 1.5\,r_e)$, undergoes significant change across the whole data region, and hence may be expected to provide a more robust and compact description of properties that vary across this whole domain.

The expansion variable of Eq. (1) is a special case of a more general expansion variable introduced by Šurkus et al. [27] (see Section 2.5), who used it to represent the potential energy function as a simple power series in $y_p^{r_e}(r)$, with a recommended value of $p$ based upon that of the asymptotically dominant inverse-power term in the long-range intermolecular potential for the state of interest. However, in other potential forms (see Sections 2.2–2.4) this power does *not* define the limiting long-range behaviour: rather, its introduction there was based on findings [12,35–38] that an expansion in powers of $y_q^{r_e}(r)$ using some (small) integer power of $q > 1$ (say, $q = 2 - 4$), rather than $q = 1$, greatly reduces the propensity for the resulting expansion to extrapolate badly outside the radial interval to which the data are most sensitive. This same concern with attaining physically reasonable extrapolation behaviour has also been part of the motivation for introducing the option of setting $r_{\text{ref}}$ at a (user-selected) value differing from the conventional molecular expansion centre of $r_e$.

In any case, values of $r_{\text{ref}}$ and $q$ (and $p$, for the MLR form of Section 2.3) 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. [30,33].

## 2.2. The Expanded Morse Oscillator (EMO) potential function

The first potential function form considered here is the *Expanded Morse Oscillator* or EMO function [39], which has the form of a Morse potential [40] in which the exponent coefficient varies with distance. Two similar functions of this type were introduced by other authors: the *Generalized Morse Oscillator* or 'GMO' potential function of Coxon and Hajigeorgiou [41], and the *Modified Morse Oscillator* or 'MMO' potential function of Dulick and co-workers [42]. However, because of its simpler structure and better extrapolation behaviour, only the EMO form is considered here.

An EMO potential has the form

$$V_{\text{EMO}}(r) = \mathfrak{D}_e \left[ 1 - e^{-\beta(r) \cdot (r - r_e)} \right]^2, \tag{3}$$

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

$$\beta(r) = \beta_{\text{EMO}}(y_q^{\text{ref}}(r)) = \sum_{i=0}^{N_\beta} \beta_i y_q^{\text{ref}}(r)^i. \tag{4}$$

While the earlier work based on this form used $y_q^{r_e}(r)$ as the expansion variable in Eq. (4), rather than $y_q^{\text{ref}}(r)$, the latter is employed here because allowing $r_{\text{ref}} \neq r_e$ introduces considerable additional flexibility. In a previous version of this code [43], this exponent polynomial was broken into two, with different polynomial orders being used for $r > r_e$ and $r < r_e$. However, the flexibility provided by using $y_q^{\text{ref}}(r)$ as the expansion variable rather than $y_q^{r_e}(r)$ has made this complication unnecessary, and allowed its high-order-derivative discontinuities at $r = r_e$ to be eliminated.

The EMO potential form has been used successfully in a number of demanding data analyses involving both normal single-well potentials [38,39,44,45] and a state whose potential function has an additional 'ripple' [37]. However,

the fact that $[\mathfrak{D}_e - V_{\text{EMO}}(r)]$ dies off exponentially at large $r$ means that it provides a poor representation for the long-range parts of *all* potential energy functions, a deficiency that is particularly troublesome for electronic states for which experimental data extends fairly close to the dissociation limit. This shortcoming is addressed by the potential function forms described the next two subsections.

## 2.3. The Morse/Long-Range (MLR) potential function

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

$$V(r) \simeq \mathfrak{D} - \frac{C_{m_1}}{r^{m_1}} - \frac{C_{m_2}}{r^{m_2}} - \frac{C_{m_3}}{r^{m_3}} - \cdots, \tag{5}$$

in which the powers $m_1$, $m_2$, …, etc., are determined by the nature of the atoms into which the given molecular state dissociates [46–51], and the coefficients $C_{m_i}$ may often be calculated ab initio. For cases in which realistic estimates of one or more of the long-range potential coefficients $C_{m_i}$ are available, it is always desirable to use a potential form with the limiting behaviour of Eq. (5), especially if the available experimental information extends fairly close to the dissociation limit $\mathfrak{D}$. These considerations instigated the development of the *Morse/Long-Range* or MLR potential energy functional form [28–30,33],

$$V_{\text{MLR}}(r) = \mathfrak{D}_e \left\{ 1 - \frac{u_{\text{LR}}(r)}{u_{\text{LR}}(r_e)} e^{-\beta(r) \cdot y_p^{r_e}(r)} \right\}^2, \tag{6}$$

in which $\mathfrak{D}_e$ is the well depth, $r_e$ the equilibrium internuclear distance, the exponent coefficient $\beta(r) = \beta_{\text{MLR}}(y_{\{p/q\}}^{\text{ref}}(r))$ is a relatively slowly varying function of $r$, while $u_{\text{LR}}(r)$ incorporates the attractive limiting long-range form imposed by theory for a particular case, namely,

$$u_{\text{LR}}(r) = \frac{C_{m_1}}{r^{m_1}} + \frac{C_{m_2}}{r^{m_2}} + \cdots + \frac{C_{m_{\text{last}}}}{r^{m_{\text{last}}}} = \sum_{i=1}^{\text{last}} \frac{C_{m_i}}{r^{m_i}}, \tag{7}$$

and $u_{\text{LR}}(r_e)$ is its value at $r_e$. The potential energy function of Eq. (6) is an extended form of what had been called the 'Morse/Lennard-Jones' or MLJ potential function form [52–54], which corresponds to a case in which the sum in Eq. (7) consists of a single inverse-power term, for which case the pre-exponential factor in Eq. (6) reduces to $(r_e/r)^{m_1}$.

A defining feature of the MLR function is the fact that

$$\lim_{r \to \infty} \beta_{\text{MLR}}(r) \equiv \beta_\infty = \ln\{2\,\mathfrak{D}_e/u_{\text{LR}}(r_e)\}. \tag{8}$$

Since the pre-factor to the exponential term in Eq. (6) is equal to 1 at the equilibrium distance, $r_e$, and $y_p^{r_e}(r) \to +1$ as $r \to \infty$, at large distances Eq. (6) becomes

$$V_{\text{MLR}}(r) \simeq \mathfrak{D}_e - u_{\text{LR}}(r) + \frac{1}{4\mathfrak{D}_e} [u_{\text{LR}}(r)]^2 \simeq \mathfrak{D}_e$$
$$- \sum_{i=1}^{\text{last}} \frac{C_{m_i}}{r^{m_i}} + \frac{1}{4\mathfrak{D}_e} \left( \sum_{i=1}^{\text{last}} \frac{C_{m_i}}{r^{m_i}} \right)^2. \tag{9}$$

This indicates that although the MLR function does take on the form of Eq. (7) at long range, if $2 m_1 \leq m_{\text{last}}$, the quadratic term in Eqs. (6) and (9) would change the long-range behaviour specified by Eq. (7) by adding additional terms with powers $2m\_i \leq m_{\text{last}}$. This problem is most serious for cases in which $m_1 = 3$ [28], but can also cause

difficulties for systems with larger $m_1$ values [15]. However, it is resolved here by the fact that, as necessary, the code internally modifies the input $C_{m_i}$ coefficients, and/or adds additional terms, so as to precisely cancel the effect of those non-physical quadratic long-range terms, as was illustrated by the discussion of the coefficients $C_6^{\text{adj}} = C_6 + (C_3)^2/4\mathfrak{D}_e$ and $C_9^{\text{adj}} = C_3\, C_6^{\text{adj}}/2\mathfrak{D}_e$ for the $m_1 = 3$ case in Ref. [28].

In order to use the MLR potential form, it is clearly necessary to know appropriate values for the powers $\{m_i\}$ and to have realistic estimates for the coefficients $\{C_{m_i}\}$. If no realistic estimate for the leading (smallest-power) coefficient $C_{m_1}$ is available, then the MLR form has no significant advantages over the simpler EMO function, and the latter will likely be more 'robust'. However, if the leading long-range term is the only one for which a reliable estimate of the coefficient ($C_{m_1}$) is available, experience suggests that more realistic extrapolation behaviour will be achieved if a second term with some plausible ad hoc estimate of $C_{m_2}$ is also included.

The exponent coefficient function $\beta(r)$ must both asymptotically approach the value $\beta_\infty$ defined by Eq. (8), and be sufficiently flexible to provide an accurate description of the shape of the potential energy well. To this end, it has normally been represented as a constrained polynomial expansion (PE) in the variable $y_q^{\text{ref}}(r)$, in which a second radial variable $y_p^{\text{ref}}(r)$ acts as a switching function that constrains $\beta(r)$ to approach the specified asymptotic values of $\beta_\infty$ [12,29,35,36,55]:

$$\beta(r) \equiv \beta_{\text{PE}-\text{MLR}}(y_{\{p/q\}}^{\text{ref}}(r))$$
$$\equiv y_p^{\text{ref}}(r)\beta_\infty + \left[1 - y_p^{\text{ref}}(r)\right]\sum_{i=0}^{N_\beta}\beta_i y_q^{\text{ref}}(r)^i. \quad (10)$$

Although most early work with this model was performed with $q = p$, it has since been shown that use of a separate power $q$ in the power-series portion of Eq. (10), usually with $q < p$, will often lead to more compact and robust expressions for the potential energy function [28,33,34].

The limiting long-range behaviour of the exponential term in Eq. (6),

$$e^{-\beta(r)\cdot y_p^{re}(r)} \simeq e^{-\beta_\infty}\left(1 + \frac{A_p}{r^p} + \frac{A_{p,q}}{r^{p+q}} + \cdots\right)$$
$$= \frac{u_{\text{LR}}(r_e)}{2\,\mathfrak{D}_e}\left(1 + \frac{A_p}{r^p} + \frac{A_{p,q}}{r^{p+q}} + \cdots\right), \quad (11)$$

must also be taken into account [28,30]. This expression shows that this exponential factor will change the long-range behaviour from that specified by Eq. (7) unless the integer $p$ defining the radial variable of Eqs. (1) and (6) and the switching function variable in Eq. (10) is set to a value greater than the difference between the largest and smallest (inverse) powers of the terms included in the chosen definition of $u_{\text{LR}}(r)$, i.e., $p > (m_{\text{last}} - m_1)$ [28,33]. In practise, it may also be desirable to set $p$ equal to the difference between the leading power $m_1$ and the power of the first long-range term predicted by theory that is *not* included in the chosen definition of $u_{\text{LR}}(r)$, $p = (m_{\text{next}} - m_1)$, as this would cause the leading contribution of Eq. (11) to the long-range potential to have the same $1/r^{m_{\text{next}}}$ radial behaviour as that first missing term.

The above condition will often require the use of relatively large values of $p$ (e.g., $p \geq 6$), which means that the radial variables $y_p^{re}(r)$ and $y_p^{\text{ref}}(r)$ will be fairly flat and close to their common upper limits over a relatively wide range of $r$ (see Fig. 1). However, the use of a separate smaller power $q$ to define the expansion variable $y_q^{\text{ref}}(r)$, together with the ability to vary the location of the (common) expansion centre $r_{\text{ref}}$, ensure the presence of adequate functional flexibility for the exponent coefficient function $\beta_{\text{PE}-\text{MLR}}(r)$, and as shown by Eq. (11), the choice of $q$ does not affect the leading deviation from the limiting long-range behaviour arising from the exponential term. Note, however, that experience suggests that for very small values of $q$ such as 1 or 2, $\beta_{\text{PE}-\text{MLR}}(r)$ is more likely to misbehave in the extrapolation region at very short distances, and/or in the interval between the 'data region' and the long-range region where the potential finally takes on the form $\mathfrak{D}_e - u_{\text{LR}}(r)$ [35,36]. At the same time, for too large values of $q$ the variable $y_q^{\text{ref}}(r)$ will lie very close to its limiting upper bound (of $+1$) over an increasingly wide range of values of $r$ (see Fig. 1), and the order of polynomial required to attain a satisfactory fit to the input points may become inconveniently large [28]. Thus, there is no definitive suggestion regarding choice of a value for $q$, so use of the MLR-PE form will normally require experimentation for each case.

### 2.3.1. A more general definition of $u_{LR}(r)$: inclusion of damping functions

While all potential energy functions take on the limiting form of Eq. (5) for 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 that contribute to this expression. This led a number of groups to propose models for representing such 'damping' behaviour [56–60]. Moreover, Eq. (9) shows that for an MLR potential the quadratic term in Eq. (6), which defines the short-range repulsive wall of an MLR potential, contains the factor $[u_{\text{LR}}(r)]^2$. Thus, at very short range the strength of this term would grow as $1/r^{2m_{\text{last}}}$, which is much steeper than the typical exponential-type behaviour expected for the repulsive wall of a normal potential energy function. These concerns instigated the implementation of an (optional) alternate definition for the long-range potential incorporated in Eq. (6), namely,

$$u_{\text{LR}}(r) = D_{m_1}(r)\frac{C_{m_1}}{r^{m_1}} + D_{m_2}(r)\frac{C_{m_2}}{r^{m_2}} + \cdots + D_{m_{\text{last}}}(r)\frac{C_{m_{\text{last}}}}{r^{m_{\text{last}}}}$$
$$= \sum_{i=1}^{\text{last}} D_{m_i}(r)\frac{C_{m_i}}{r^{m_i}}, \quad (12)$$

in which the $D_m(r)$ are referred to as "damping functions". Program **betaFIT** allows use of the two families of damping functions introduced in Ref. [33], specifically, a generalized version of a function introduced by Douketis et al. [58],

$$D_m^{\text{ds}(s)}(r) = \left\{1 - \exp\left(-\frac{b^{\text{ds}}(s)(\rho\, r)}{m} - \frac{c^{\text{ds}}(s)(\rho\, r)^2}{m^{1/2}}\right)\right\}^{m+s}, \quad (13)$$

and a generalized version of the Tang–Toennies damping function [59,61]

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

For both these damping function forms $r$ has units Å, while $\rho$ is a dimensionless, system-dependent range-scaling factor. Moreover, both these families of functions have the same limiting short-range behaviour [33],

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

Hence, incorporation of Eq. (12) into the MLR functional form would cause the limiting very short-range behaviour of the repulsive wall to be proportional to $r^{2s}$. This means that the value of $s$ must be set $\leq 0$ in order to prevent the potential from turning over and going to zero as $r \to 0$.

Scoles and Thakkar and co-workers [58] suggested that such damping functions be defined by reference to the accurately known damping of the leading dispersion terms in the long-range non-bonding interaction of two ground-state H atoms. Least-squares fits to those ab initio H-atom damping function results [62] yielded values of the system-independent parameters $b^{\text{ds}}(s)$, $c^{\text{ds}}(s)$, and $b^{\text{tt}}(s)$ for a range of values of the limiting short-range power '$s$' of Eq. (15) [33], and those values are built into program **betaFIT**. In Ref. [58] it was also suggested that the system-dependent scaling parameter for molecule AB be defined as $\rho \equiv \rho^{\text{AB}} = [2\,\rho^{\text{A}}\rho^{\text{B}}/(\rho_d^{\text{A}}+\rho^{\text{B}})]$, with $\rho^{\text{A/B}}$ defined in terms of the ratio of the ionization potential of the atom of interest to that for ground-state atomic hydrogen, viz., $\rho^A \equiv \left(I_p^{\text{A}}/I_p^{\text{H}}\right)^{2/3}$. Note, however, that other system-dependent scaling factors have also been proposed [63].

Program **betaFIT** allows the user to choose to define $u_{\text{LR}}(r)$ in terms of either Eq. (7) or (12), and to use either Eq. (13) or (14) to represent the damping functions, but it also requires the user to specify a value for the limiting short-range power, $s$ and the range-scaling parameter $\rho$. Note that Ref. [33] recommended the use of the Douketis-type damping function (13) with $s = -1$.

### 2.3.2. Treatment of interstate coupling near the asymptote: states of alkali homodimers dissociating to $X(nS) + X(nP)$

The above discussion of the MLR form focused upon use of the inverse-power sums of Eq. (7) or (12) to represent the long-range potential. However, there is no need to restrict $u_{\text{LR}}(r)$ to these forms, and in applications to states of $\text{Li}_2$ dissociating to the $\text{Li}(^2_{1/2}P + \text{Li}_2(^2_{1/2}P)$ asymptote, $u_{\text{LR}}(r)$ has been represented by one of the roots of a diagonalization arising from two-state [28,64] or three-state [34,65] coupling near that asymptote. In particular, the theory of interstate mixing presented in Refs. [66,67] shows that (omitting damping) the expressions for the $2 \times 2$ roots may be written in closed form as

$$u_{\text{LR}}^{\text{A}-\text{F}}(r) = -\frac{A_{\text{so}}}{2} + \frac{C_3^\Sigma + C_3^\Pi}{2\,r^3} + \frac{C_6^\Sigma + C_6^\Pi}{2\,r^6} + \frac{C_8^\Sigma + C_8^\Pi}{2\,r^8}$$
$$\pm \frac{1}{2}\left\{ \left( \frac{C_3^\Sigma - C_3^\Pi}{3\,r^3} + \frac{C_6^\Sigma - C_6^\Pi}{3\,r^6} + \frac{C_8^\Sigma - C_8^\Pi}{3\,r^8} - A_{\text{so}} \right)^2 \right.$$

$$\left. + 8\left( \frac{C_3^\Sigma - C_3^\Pi}{3\,r^3} + \frac{C_6^\Sigma - C_6^\Pi}{3\,r^6} + \frac{C_8^\Sigma - C_8^\Pi}{3\,r^8} \right)^2 \right\}^{1/2}, \qquad (16)$$

in which $A_{\text{so}}$ is the (positive) separation between the $^2P_{1/2}$ and $_2P_{3/2}$ atomic spin–orbit level energies, and $\{C_m^\Sigma/C_m^\Pi\}$ are the long-range coefficients associated with the relevant coupled states [28]. Use of the '+' sign with the square root term yields the correct long-range tails for the $A^1\Sigma_u^+$ states of homonuclear alkali dimers, while use of the '−' sign yields those for the $0_u^+$ component of their $b^3\Pi_u$ states [66,67].

The analogous expressions for the long-range tails of the $1^3\Sigma_g^+$ or $B^1\Pi_u$ state alkali dimer potentials involve use of, respectively, the lowest and middle roots of the $3 \times 3$ coupling matrix of Ref. [67]. The implementation in program **betaFIT** is based on the description of that result presented in Eq. (7) of Ref. [65] [28,65], with the extension that the present implementation of these special cases allows the inclusion of damping functions with all of the six $\{C_m^\Sigma/C_m^\Pi\}$ coefficients of the $2 \times 2$ cases for $m = 3$, 6, and 8, and all of the analogous nine $\{C_m^{\Sigma^{1/3}\Pi}\}$ coefficients for the $3 \times 3$ cases. The ease with which it can incorporate these special cases further illustrates the capabilities of the MLR functional form.

### 2.3.3. A spline-pointwise representation of the MLR exponent coefficient $\beta(r)$

Most existing applications of the MLR potential have been based on use of the constrained polynomial expansion of Eq. (10) to represent the exponent coefficient $\beta(r)$. However, while preliminary work on its application to double-minimum and shelf-state potentials has shown some promise [68,69], it has not yet been *demonstrated* to be capable of providing practical, compact and accurate representations of potential functions for such cases. In contrast, a novel approach introduced by Pashov and co-workers, in which the potential is defined as a cubic spline through a set of points whose ordinate values are the parameters of the model, has proven to be remarkably successful for such cases [9–11]. However, that approach requires the use of a relatively large set of points/parameters (typically $\gtrsim 50$) to define a potential function accurately. Moreover, both short-range extrapolations and the theoretically predicted inverse-power long-range tail may only be attached in an ad hoc manner, and the resulting function lacks high-order analytic continuity at all 50 or so spline points, as well as at the long- and short-range extrapolation-connection points.

A way of combining these approach that is now under investigation [70] is the 'Spline-Exponent-MLR' (SE-MLR) function, which uses the Pashov 'spline-pointwise' approach to define the exponent coefficient $\beta(r)$ in the MLR potential function form of Eq. (6). In particular, $\beta(r)$ is defined as a 'natural' cubic spline function passing through $\beta(r_j) = \beta(y_q^{\text{ref}}(r_j))$ values at a specified set of $y_q^{\text{ref}}(r_j)$ values, and those $\beta(r_j)$ values become the parameters that define the shape of the potential. Following the original Pashov approach [9,71], the exponent coefficient function is

written as

$$\beta(r) = \beta_{SE-MLR}(r) = \sum_{k=1}^{N_\beta} S_k(y_q^{ref}(r))\beta_k, \qquad (17)$$

in which the 'spline basis functions' $S_k\left(y_p^{ref}(r)\right)$ are completely defined by the chosen mesh of values of $y_q^{ref}(r_j)$. In this case, the connection to the long-range behaviour of Eq. (9) is provided by the constraint of fixing $\beta_{N_\beta} = \beta_\infty$. A straightforward application of the chain rule of calculus yields the partial derivatives required for the least-squares fit procedure; for example,

$$\frac{\partial V(r)}{\partial \beta_k} = 2\,\mathfrak{D}_e\left\{1 - \frac{u_{LR}(r)}{u_{LR}(r_e)}e^{-\beta(r)\cdot y_p^{r_e}(r)}\right\}$$
$$\times \left(\frac{u_{LR}(r)}{u_{LR}(r_e)}e^{-\beta(r)\cdot y_p^{eq}(r)}\right)y_p^{r_e}(r)S_k\left(y_q^{ref}(r)\right). \qquad (18)$$

Thus, although the least-squares problem is still non-linear, some computational simplifications occur because the $S_k(y_p^{ref}(r))$ functions do not depend on the parameter values $\{\beta_k\}$.

Experience suggests that when using the SE-MLR form, the radial variable $y_q^{ref}(r)$ used to define $\beta(r)$ should normally be defined by a value of $r_{ref}$ that is significantly greater than $r_e$, in order to ensure that the chosen mesh of points samples the full range of $y_q^{ref}(r)$ values fairly uniformly [70]. This consideration is illustrated in Fig. 2, which compares the SE-MLR exponent coefficient functions determined from fits to an extensive data set for the $X^1\Sigma_g^+$ state of Ca$_2$ [30,70,72] that was carried out using $r_{ref} = r_e$ (blue round points, curve, and lines), and $r_{ref} = 1.5\,r_e$ (red square points, curve, and lines). In each case two points, equally spaced in $y_q^{ref}(r)$, were placed at $r < r_e$, and 13 points, equally spaced in $y_q^{ref}(r)$ were placed at $r > r_e$, while one more point was placed at $r = r_e$ and the fixed point $\beta_{N_\beta} = \beta_\infty$ placed at $y_q^{ref} = 1$ defines the $r \to \infty$ limit. From the distributions of points for the two cases shown in Fig. 3, it will be intuitively clear that for the same number of spline points, it will be easier to obtain an accurate description of $\beta(r)$ for this case using points based on $r_{ref} = 1.5\,r_e$, as they are distributed more evenly across
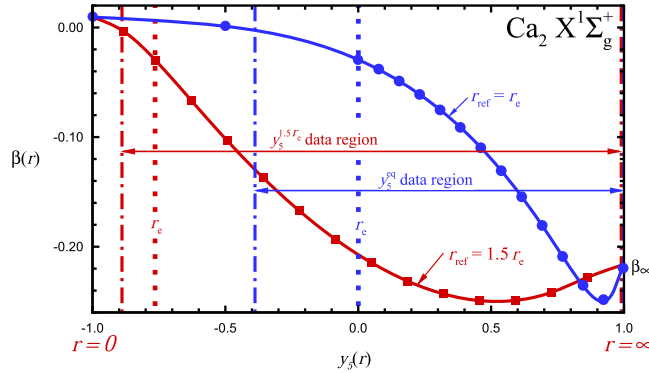


**Fig. 2.** Comparison of $\beta(y_q^{ref}(r))$ functions for ground-state Ca$_2$ determined from DPFs of spectroscopic data to SE-MLR models based upon $r_{ref} = r_e$ (blue round points and curve), and $r_{ref} = 1.5\,r_e$ (red square points and curve) (figure taken from Ref. [70]). (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)
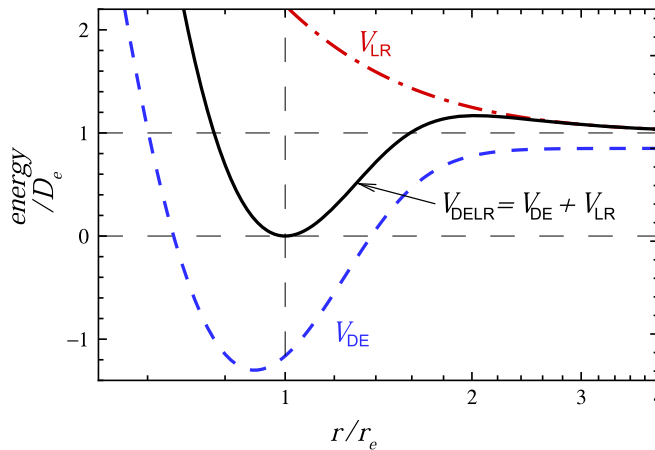


**Fig. 3.** Schematic illustration of the form of a DELR potential function for representing a barrier-state potential (solid curve) and its components (dashed and dot-dash curves) (adapted from Fig. 1 of Ref. [12]).

the $y_q(r)$ domain. Further work on developing and testing this approach is now under way.

### 2.4. The Double-Exponential/Long-Range (DELR) model potential

A need for a flexible analytic potential function form with a barrier that protrudes above the potential asymptote for distances $r > r_e$ instigated the development of the *Double-Exponential/Long-Range* (DELR) potential function model [12,35]. This function is written as [73]:

$$V_{\text{DELR}}(r) = A\, e^{-2\,\beta(r)\cdot(r-r_e)} - B\, e^{-\beta(r)\cdot(r-r_e)} + \mathfrak{D}_e - u_{\text{LR}}(r), \quad (19)$$

in which the exponent coefficient $\beta(r)$ is defined by the same simple power series in $y_q^{r_{\text{ref}}}(r)$ that was used for the EMO potential, Eq. (4), while the long-range tail function $u_{\text{LR}}(r)$ may be represented by any one of the functions discussed in Section 2.3. A schematic illustration of the nature of a DELR potential function for which $u_{\text{LR}}(r)$ is negative (repulsive) is shown in Fig. 3.

The only current published application of this form [12] set $r_{\text{ref}} = r_e$ and used different power-series orders for $r \leq r_e$ and for $r > r_e$ in order to prevent unphysical extrapolation behaviour at small $r$. However, subsequent experience with the MLR form has led us to believe that this $\beta(r)$ function should be treated as a single simple polynomial in the new expansion variable $y_p^{r_{\text{ref}}}(r)$, and that placing the expansion centre $r_{\text{ref}}$ at some distance greater than $r_e$ would ensure stable short-range behaviour: experience to date validates this assumption. In this model, the presence of a repulsive leading term in the additive long-range function $u_{\text{LR}}(r)$ (which, in principle, may be attractive or repulsive) would introduce a potential function barrier at a distance beyond the main well, as is seen in Fig. 3.

The pre-exponential coefficients $A$ and $B$ in Eq. (19) are defined in terms of the overall well depth $\mathfrak{D}_e$ (relative to the potential asymptote), the position of the potential minimum, $r_e$, and the properties of the exponent coefficient function $\beta(r)$ and the long-range tail function $u_{\text{LR}}(r)$ at that minimum:

$$A = \mathfrak{D}_e - u_{\text{LR}}(r_e) - u'_{\text{LR}}(r_e)/\beta(r_e), \quad (20)$$

$$B = 2\,\mathfrak{D}_e - 2\,u_{\text{LR}}(r_e) - u'_{\text{LR}}(r_e)/\beta(r_e), \quad (21)$$

in which $u'_{\text{LR}}(r_e) \equiv [du_{\text{LR}}(r)/dr]_{r=r_e}$. If $u_{\text{LR}}(r) = 0$ the DELR potential becomes the EMO function of Eq. (3). However, other choices of $u_{\text{LR}}(r)$ allow it to represent the outer wall of a potential function with a barrier [12], a multi-term attractive inverse-power long-range potential, or even the outer wall of a double-minimum or shelf-state potential [73].

In general, $u_{\text{LR}}(r)$ should be represented by the damped long-range expansion of Eq. (12), rather than by the undamped expansion of Eq. (7), since the short-range singularities of attractive inverse-power terms in the latter would cause the potential energy function to turn over at small distances. Moreover, the damping-function short-range parameter must be set as $s \geq 0$ if the higher-order $C_{m_i}$ coefficients are positive (attractive), since otherwise the negative singularities implied by Eq. (15) would cause the potential wall to turn over at small $r$. Of course, other

damping function expressions [57,60] or entirely different types of expressions for $u_{\text{LR}}(r)$ could be used in the DELR potential form.

### 2.5. The Generalized Potential Energy Function (GPEF) of Šurkus

The fourth family of potential functions implemented in **betaFIT** is a generalization of the familiar Dunham polynomial potential [23] that Šurkus et al. [27] introduced and named the *Generalized Potential Energy Function* (GPEF). Using the modified (but equivalent) expression for the Šurkus radial expansion variable devised by Seto [74], the GPEF potential form is written as

$$V_{\text{GPEF}}(r) = c_0\, z_q^2 \left[ 1 + \sum_{m=1}^{N} c_m\, z_q^m \right], \quad (22)$$

with

$$z_q \equiv \frac{(r^q - r_e^q)}{(a_S\, r^q + b_S\, r_e^q)}. \quad (23)$$

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

- setting $q = 1$, $a_S = 0$ and $b_S = 1$ yields the Dunham expansion variable [23]: $z = (r - r_e)/r_e$;
- setting $q = 1$, $a_S = 1$ and $b_S = 0$ yields the Simons–Parr–Finlan (SPF) expansion variable: [24]: $z = (r - r_e)/r$;
- setting $q = 1$, $a_S = b_S = 0.5$ yields the Ogilvie–Tipping (OT) expansion variable [26]: $z = 2(r - r_e)/(r + r_e)$;
- setting $a_S = b_S = 1$ yields the expansion variable of Eq. (1);
- setting $q = 1$, $a_S = 1$ and fixing $b_S$ at some small value yields the type of asymmetric expansion variable used by Tiemann and coworkers [72,75].

For $a_S \neq 0$, this function always asymptotically approaches a finite limit with a $1/r^q$ functional behaviour. Thus, if appropriate constraints are applied to the coefficients, such an expansion may in principle be required to have the theoretically predicted limiting long-range behaviour of Eq. (5) [27,76]. However, apart from the relatively simple case in which $q$ is set equal to the power of the leading long-range term in Eq. (5) [27], such constraints have proven to be too unwieldy for practical use.

## 3. Methodology

### 3.1. Preliminary estimates of non-linear parameters

The first three of the potential function forms described above are based on an exponential term whose exponent coefficient $\beta(r)$ is expanded as a power series in the radial variable $y_q^{\text{ref}}(r)$ of Eq. (2). Such potential functions are non-linear functions of the expansion coefficients $\beta_i$, so an approximate method must be devised to obtain the realistic initial trial values for these parameters that are required in order to initiate the non-linear least-squares

fits. This is done in **betaFIT** by first fitting to a linearized version of the potential function equation. For example, Eq. (3) for the EMO potential may be re-arranged to yield the following expressions for the exponent in the exponential term:

$$(r - r_e)\beta(r) = -\ln\left\{1 \pm \sqrt{\frac{V(r) - \text{VMIN}}{\mathfrak{D}_e}}\right\} = (r - r_e)\sum_{i=0}^{N_\beta}\beta_i\left[y_q^{\text{ref}}(r)\right]^i. \tag{24}$$

The $+$ and $-$ signs in the logarithmic term correspond, respectively, to the inner $(r < r_e)$ and outer branches of the potential well. Values of the logarithmic term in Eq. (24) are defined by the input data points, and the resulting expression is a linear function of the expansion parameters $\beta_i$. Thus, if plausible initial values of the physical parameters $r_e$ ($\equiv$ Re), $\mathfrak{D}_e$ ($\equiv$ De) and the energy of the potential minimum VMIN are provided in the input data file, a simple linear fit to Eq. (24) will yield realistic values of the $\beta_i$ parameters for use in the subsequent direct non-linear fits to Eq. (3).

Although the algebra is somewhat more complicated, essentially the same approach is employed to determine initial estimates of the expansion parameters for a PE-MLR potential. If Eq. (10) is used to represent the exponent coefficient function, the exponent in exponential term of Eq. (6) may be written as:

$$y_p^{r_e}(r)\beta(r) = -\ln\left\{\left(1 \pm \sqrt{\frac{V(r) - \text{VMIN}}{\mathfrak{D}_e}}\right)\frac{u_{\text{LR}}(r_e)}{u_{\text{LR}}(r)}\right\}$$
$$= y_p^{r_e}(r)\left\{y_p^{\text{ref}}(r)\beta_\infty + [1 - y_p^{\text{ref}}(r)]\sum_{i=0}^{M_\beta}\beta_i y_q^{\text{ref}}(r)^i\right\}. \tag{25}$$

in which the $\pm$ signs in the logarithmic term have the same significance as in Eq. (24). As was the case for the EMO potential, this logarithmic term is a simple linear function of the $\beta_i$ exponent expansion parameters.

For the case of an SE-MLR potential, in which the exponent coefficient function $\beta(r)$ is represented by a spline function passing through a specified set of $y_q^{\text{ref}}(r_i)$ values, this linearization step consists of simply using spline interpolation through the exponent function defined by the first line of Eq. (25) to determine their values at the chosen grid of $\{y_q^{\text{ref}}(r_i)\}$ values, and dividing the results by the associated values of $\{y_p^{r_e}(r_i)\}$ to obtain the desired $\{\beta_i = \beta(r_i)\}$ values. Note, however, that this step must avoid placing one of those grid point too close to the distance $r = r_e$, since at that point $y_p^{r_e}(r) = 0$, and both sides of Eq. (25) vanish. The current version of **betaFIT** addresses this problem simply by defining the grid point closest to $r_e$ to be $y_q^{\text{ref}}(r_e) - 0.5$.

For the DELR potential the situation is a little more complicated, since in the analog of Eq. (24), which is

$$\beta(r)(r - r_e) = -\ln\left\{\frac{B}{2A} \pm \sqrt{\frac{V(r) - \text{VMIN} - \mathfrak{D}_e + u_{\text{LR}}(r)}{A} + \left(\frac{B}{2A}\right)^2}\right\}$$
$$= (r - r_e)\sum_{i=0}^{N_\beta}\beta_i\left[y_q^{\text{ref}}(r)\right]^i, \tag{26}$$

the logarithmic term depends upon the value of $\beta(r_e)$ (and hence upon the values of the $\beta_i$ expansion parameters),

through the definitions of the exponential term prefactors $A$ and $B$ (see Eqs. (20) and (21)). For this reason it is necessary to apply the fit to Eq. (26) iteratively in order to generate a self-consistent value of $\beta(r_e)$ and set of trial $\beta_i$ values. However, experience indicates that virtually any non-zero initial trial value of $\beta(r_e)$ will serve as an adequate starting point, so this apparent problem presents no practical difficulty.

Since GPEF potential functions are simple polynomials, their fits are only non-linear with respect to the equilibrium distance $r_e$, and hence there is no need for any preliminary estimates of their $c_m$ expansion parameter, and the fits will converge rapidly for any realistic initial trial value of $r_e$.

### 3.2. Uncertainties and parameter rounding

In any least-squares fit, each datum should be weighted by the inverse square of its uncertainty. The current version of **betaFIT** allows all input potential points to be equally weighted by a common read-in uncertainty UNC, but if the value input for UNC is set $\leq 0.0$, the program will read in a separate uncertainty $u_i$ for each datum $i$. The quality of a fit of an $M$-parameter model to $N_d$ input data, which yields the predicted quantities $\{y_{\text{calc}}(i)\}$, is indicated by the value of the dimensionless-root-mean-square-deviation,

$$\text{DRMSD} \equiv \overline{dd} = \left\{\frac{1}{N}\sum_{i=1}^{N_d}\left[\frac{y_{\text{calc}}(i) - y_{\text{obs}}(i)}{u(i)}\right]^2\right\}^{1/2}, \tag{27}$$

or by the dimensionless standard error $\text{DSE} \equiv \overline{\sigma}_f = \overline{dd}\sqrt{N_d/(N_d - M)}$.

Note that if one is using **betaFIT** to attempt to obtain an accurate analytic representation of a set of ab initio points, it is extremely important to have realistic estimates of their uncertainties in order to prevent the fit from being totally dominated by one region or another: for example by focussing on and fitting accurately to the points on a very steep short-range wall where the energies are very large, or focussing on a long-range region where there may be many points, but the interaction energy is very small.

In addition to reporting the 95% confidence limit (approximately 'two-$\sigma$') uncertainty in each fitted parameter, **betaFIT** follows the approach of Ref. [77] by always also listing the associated "parameter sensitivity" (identified as PS in the output). This quantity is defined (see Eq. (4) of Ref. [77]) as the magnitude of the largest change in the given parameter whose effect on the predictions of the model could increase $\overline{\sigma}_f$ by a maximum of $(0.1/M)\overline{\sigma}_f$. In other words, the parameter sensitivity indicates the degree to which any particular fitted parameter value may be rounded off while having no significant effect (within the uncertainties) on the ability of the resulting parameter set to predict the input data accurately. For the illustrative cases considered in Ref. [77], rounding off all parameters at the first significant digit of their *sensitivity* had no significant effect on the values of $\overline{\sigma}_f$ or $\overline{dd}$ to three significant digits.

Another feature of **betaFIT** is its implementation of the (optional) automated "sequential rounding and refitting" (SRR) procedure of Ref. [77], 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 is turned on or off by the value of the flag `IROUND` that is set by the user in the input data file. One would normally turn this flag off (by setting `IROUND=0`) for preliminary analyses, and only turn it on when one wishes to generate a final parameter set to report and distribute. However, for the type of problem addressed by **betaFIT**, the difference in absolute computation time for these two cases is very small. As was discussed in Ref. [77], in many cases setting $IROUND = \pm 1$ yields an excellent degree of rounding without significant loss of precision (the first author prefers $IROUND = -1$). However, in **betaFIT** applications, it is often necessary to set $|IROUND| > 1$ to obtain a rounded parameter set that reproduces the observables to the desired degree of precision.

### 3.3. Input/output conventions, and program execution

The array dimensions assumed by **betaFIT** are set by the (identical) 'PARAMETER' statements located near the beginning of the Main program and of subroutines `DYIDPJ` and `Lkoef`, at lines #16, 1274 and 1753, respectively, in the source code, and may be changed by a user, as needed. The current settings allow for up to 1001 input potential function points and up to 50 expansion parameters, while the long-range tail of Eq. (7) or (12) may include up to 25 terms.

Program **betaFIT** requires one input data file, which it reads in 'free-format' form on Channel-5, and it writes its main output to Channel-6. A compact supplementary listing of the results of each fit in a form that facilitates its being incorporated into an input instruction file for program **PotFit** [78] is also written to Channel-7. If desired, listings of values of the resulting functions may be written to Channel-8. The structure of the input data file and the definitions of and options for the various input quantities are presented in Appendices A and B of the Supplementary Material accompanying this paper.

The source code was written in the FORTRAN language, and has been tested using the SUN FORTRAN-77 and FORTRAN-90, and COMPAQ compilers, as well as the public domain GFORTRAN compilers.

If one is executing **betaFIT** in a UNIX or Linux operating system environment, it may be convenient to do so using a shell (named, say, `rbeta`) such as that shown below, which may be stored in the system or user's `bin` directory:

```
#UNIX shell 'beta' to execute the compiled
# program betaF.x with input data file $1.5
# and output to $1.6, $1.7, & $1.8
time/home/rleroy/betaFIT/code/betaF.x < $1.5 > $1.6
mv fort.7 $1.7 >& /dev/null
mv fort.8 $1.8 >& /dev/null
```

in which `userpath` is a path specifying the location of the executable file `betaF.x` on the user's computer. This shell allows the program to be executed using the input file 'molec.5' with the simple command:

```
rbeta molec
```

in which `molec` is the particular name assigned by the user

to an input data file named `molec.5` that contains a list of the potential points being fitted to and the user's instructions regarding what fits are to be performed. The root component of this filename, `molec`, may be chosen arbitrarily by the user. In this case, the standard output to Channel-6 will be written to file `molec.6` and the Channels-7 and 8 output to files `molec.7` and `molec.8`, respectively, in the same directory as the `molec.5` Channel-5 input data file.

## 4. Concluding remarks

The problem that led to the creation of program **betaFIT** was the need for the plausible sets of initial trial parameters that are required to initiate direct potential fits (DPFs) to large bodies of experimental data using tools such as program **dPotFit** [78,79]. However, **betaFIT** may also prove useful for providing fully analytic functions to represent sets of high-quality ab initio potential energy values, since having an effective analytic model to represent such data can be as important as the ab initio results themselves. Extensions of the present methodology to the representation of ab initio results for multi-dimensional atom plus linear molecule and linear-molecule plus linear-molecule interactions have already proved their worth [80–83].

## Appendix A. Supplementary data

Supplementary materials associated with this paper consist of: a text file containing the full standalone FORTRAN source code for program **betaFIT**, a text file containing the sample data files that are presented in Appendix C, and a document containing: (i) Appendix A, a structured listing of the READ statements in the program, (ii) Appendix B full definitions of all of the parameters and quantities read in the input data file, and descriptions of the associated program options, and 1 (iii) Appendix C, commented listings of illustrative sample input and output data files for various cases. Anyone who wishes to be Registered with the author as a user of this code, eligible to be sent any future bug fixes or updates, should fill in the online form at the www address http://scienide2.uwaterloo.ca/~rleroy/betaFIT16.

Supplementary data associated with this paper can be found in the online version at http://dx.doi.org/10.1016/j.jqsrt.2016.03.036.

## References

[1] Rydberg R. Z Phys 1931;73:376.
[2] Klein O. Z Phys 1932;76:226.
[3] Rydberg R. Z Phys 1933;80:514.
[4] Rees ALG. Proc Phys Soc (Lond) 1947;59:998.
[5] Coxon JA. J Mol Spectrosc 1992;152:274.
[6] Samuelis C, Tiesinga E, Laue T, Elbs M, Knöckel H, Tiemann E. Phys Rev A 2000;63:012710.
[7] Brühl R, Zimmermann D. J Chem Phys 2001;114:3035.
[8] Seto JY, Le Roy RJ, Vergès J, Amiot C. J Chem Phys 2000;113:3067.
[9] Pashov A, Jastrzębski W, Kowalczyk P. Comput Phys Commun 2000;128:622.
[10] Pashov A, Jastrzębski W, Kowalczyk P. J Chem Phys 2000;113:6624.
[11] Pashov A, Jastrzębski W, Jaś niecki W, Bednarska V, Kowalczyk P. J Mol Spectrosc 2000;203:264.
[12] Huang Y, Le Roy RJ. J Chem Phys 2003;119: 7398 erratum: Huang Y, Le Roy RJ. J Chem Phys 2007;126:169904.
[13] Grochola A, Jastrzebski W, Kowalczyk P, Crozet P, Ross AJ. Chem Phys Lett 2003;372:173.
[14] Coxon JA, Hajigeorgiou PG. J Phys Chem 2006;A 110:6261.
[15] Cho Y-S, Le Roy RJ. J Chem Phys. 2016;144:024311.
[16] Slavíček P, Kalus R, Paška P, Odvásková I, Hobza P, Malijevský A. J Chem Phys 2003;119:2102.
[17] Hanni M, Lantto P, Runeberg N, Jokisaari J, Vaara J. J Chem Phys 2004;121:5908.
[18] Hellmann R, Bich E, Vogel E. Mol Phys 2007;105:3013.
[19] Patkowski K, Špirko V, Szalewicz K. 326;1382:2009.
[20] Patkowski K, Szalewicz K. J Chem Phys 2010;133:094304.
[21] Jäger B, Hellmann R, Bich E, Vogel E. Mol Phys 2009;107: 2181 erratum: Jäger B, Hellmann R, Bich E, Vogel E. Mol Phys 2010;108:105.
[22] Waldrop JM, Song B, Patkowski K, Wang X. J Chem Phys 2015;142:204307.
[23] Dunham JL. Phys Rev 1932;41:721.
[24] Simons G, Parr RG, Finlan JM. J Chem Phys 1973;59:3229.
[25] Thakkar AJ. J Chem Phys 1975;62:1693.
[26] Ogilvie JF. Proc R Soc (Lond) 1981;A 378:287.
[27] Šurkus AA, Rakauskas RJ, Bolotin AB. Chem Phys Lett 1984;105:291.
[28] Le Roy RJ, Dattani N, Coxon JA, Ross AJ, Crozet P, Linton C. J Chem Phys 2009;131:204309.
[29] Le Roy RJ, Huang Y, Jary C. J Chem Phys 2006;125:164310.
[30] Le Roy RJ, Henderson RDE. Mol Phys 2007;105:663.
[31] Salami H, Ross AJ, Crozet P, Jastrzebski W, Kowalczyk P, Le Roy RJ. J Chem Phys 2007;126:194313.
[32] Shayesteh A, Henderson RDE, Le Roy RJ, Bernath PF. J Phys Chem A 2007;111:12495.
[33] Le Roy RJ, Haugen CC, Tao J, Li H. Mol Phys 2011;109:435.
[34] Dattani N, Le Roy RJ. J Mol Spectrosc 2011;268:199.
[35] Huang Y. Determining analytical potential energy functions of diatomic molecules by direct fitting [M.Sc. thesis]. Department of Chemistry, University of Waterloo; 2001.
[36] Le Roy RJ, Huang Y. J Mol Struct (Theochem) 2002;591:175.
[37] Le Roy RJ, Appadoo DRT, Anderson K, Shayesteh A, Gordon IE, Bernath PF. J Chem Phys 2005;123:204304.
[38] Le Roy RJ, Appadoo DRT, Colin R, Bernath PF. J Mol Spectrosc 2006;236:178.
[39] Lee EG, Seto JY, Hirao T, Bernath PF, Le Roy RJ. J Mol Spectrosc 1999;194:197.
[40] Morse PM. Phys Rev 1929;34:57.
[41] Coxon JA, Hajigeorgiou PG. J Mol Spectrosc 1990;139:84.
[42] Hedderich HG, Dulick M, Bernath PF. J Chem Phys 1993;99:8363.
[43] (a) Le Roy RJ. betaFIT 2.1. University of Waterloo Chemical Physics Research Report CP-666; 2013. (b) Le Roy RJ. betaFIT 2.0. CP-665; 2009. (c) Le Roy RJ. phiFIT 1.1. CP-663R; 2006. (d) Le Roy RJ. phiFIT 1.0, CP-663; 2006. See also ⟨http://leroy.uwaterloo.ca/programs/⟩.
[44] Seto JY, Morbi Z, Charron F, Lee SK, Bernath PF, Le Roy RJ. J Chem Phys 1999;110:11756.
[45] Walji S, Sentjens KM, Le Roy RJ. J Chem Phys 2015;142:044305.
[46] Margenau H. Rev Mod Phys 1939;11:1.
[47] Hirschfelder JO, Curtiss CF, Bird RB. Molecular theory of gases and liquids. New York: Wiley; 1964.
[48] Hirschfelder JO, Meath WJ. In: Hirschfelder JO, editor. Intermolecular forces. Advances in chemical physics, vol. 12. New York: Interscience; 1967. p. 3–106 [chapter 1].
[49] Le Roy RJ, Bernstein RB. J Chem Phys 1970;52:3869.
[50] Meath WJ. Am J Phys 1972;40:21.
[51] Le Roy RJ. In: Barrow RN, Long DA, Millen DJ, editors. Molecular spectroscopy, vol. 1. London: Chemical Society of London; 1973. p. 113–76 specialist periodical report 3.
[52] Hajigeorgiou PG, Le Roy RJ. In: 49th Ohio State University international symposium on molecular spectroscopy, Columbus, Ohio, paper WE04; 1994.
[53] Hajigeorgiou PG, Le Roy RJ. J Chem Phys 2000;112:3949.
[54] In the early work [53] that potential was called a *Modified Lennard-Jones* (MLJ) oscillator. However, as omitting the exponential term yields a Lennard-Jones($2m_1, m_1$) potential while omitting the $(r_e/r)^{m_1}$ term yields a Morse-like potential function, its algebraic form suggests that *Morse/Lennard-Jones* is a more appropriate name for it.
[55] In earlier versions of the model an exponential switching function was used [52,85,86,53,8,87], but that form involves greater complexity and has no significant advantages.
[56] Hepburn J, Scoles G, Penco R. Chem Phys Lett 1975;36:451.
[57] Koide A, Meath WJ, Allnatt AR. Chem Phys 1981;58:105.
[58] Douketis C, Scoles G, Marchetti S, Zen M, Thakkar AJ. J Chem Phys 1982;76:3057.
[59] Tang KT, Toennies JP. J Chem Phys 1984;80:3726.
[60] Douketis C, Hutson JM, Orr BJ, Scoles G. Mol Phys 1984;52:763.
[61] The original Tang–Toennies damping function was the normalized incomplete gamma function of order $(m+1)$, which corresponds to the case $s = +1$ [59]. while the original Douketis et al. [58] function corresponds to $s=0$.
[62] Kreek H, Meath WJ. J Chem Phys 1969;50:2289.
[63] Lu JX, Marlow WH. Phys Rev Lett 1995;74:1724.
[64] Gunton W, Semczuk M, Dattani NS, Madison KW. Phys Rev A 2013;88:062510.
[65] Semczuk M, Li X, Gunton W, Haw M, Dattani NS, Witz J, et al. Phys Rev A 2013;87:052505.
[66] Martin F, Aubert-Frécon M, Bacis R, Crozet P, Linton C, Magnier S, et al. Phys Rev A 1997;55:3458.
[67] Aubert-Frécon M, Hadinger G, Magnier S, Rousseau S. J Mol Spectrosc 1998;188:182.
[68] Huang Y, Le Roy RJ. In: 55th Ohio State University international symposium on molecular spectroscopy, Columbus, Ohio, paper RA01; 2000.
[69] Seto JY, Le Roy RJ. In: 55th Ohio State University international symposium on molecular spectroscopy, Columbus, Ohio, paper RA02; 2000.
[70] Tao J, Le Roy RJ, Pashov A. In: 65th Ohio State University international symposium on molecular spectroscopy, Columbus, Ohio, paper MG07; 2010.
[71] Ivanova M, Stein A, Pashov A, Knöckel H, Tiemann E. J Chem Phys 2011;134:024321.
[72] Allard O, Samuelis C, Pashov A, Knöckel H, Tiemann E. Eur Phys J D 2003;26:155.
[73] Note that the sign convention for the long-range term $u_{LR}(r)$ in the DELR potential has been changed from that of Refs. [35,12] in order to make it consistent with that used for the MLR potential form.
[74] Seto JY. *Direct fitting of analytic potential functions to diatomic molecule spectroscopic data* [M.Sc. thesis]. Department of Chemistry, University of Waterloo; 2000.
[75] Knöckel H, Bodermann B, Tiemann E. Eur Phys J D 2004;28:199.
[76] Ogilvie JF. J Chem Phys 1988;88:2804.
[77] Le Roy RJ. J Mol Spectrosc 1998;191:223.
[78] Le Roy RJ. J Quant Spectrosc Radiat Transf 2016;000:0000. **dPotFit:** A computer program to fit diatomic molecule spectral data to potential energy functions.
[79] (a) Le Roy RJ, Seto J, Huang Y. DPotFit 2.0: a computer program for fitting diatomic molecule spectra to potential energy functions. University of Waterloo Chemical Physics Research Report CP-667; 2013. (b) Le Roy RJ, Seto J, Huang Y. DPotFit 1.2. CP-664; 2007. (c) Le Roy RJ, Seto J, Huang Y. DPotFit 1.1. CP-662R; 2006. See also ⟨http://leroy.uwaterloo.ca/programs/⟩.
[80] Li H, Le Roy RJ. Phys Chem Chem Phys 2008;10:4128.
[81] Li H, Blinov N, Roy P-N, Le Roy RJ. J Chem Phys 2009;130:144305.
[82] Li H, Roy P-N, Le Roy RJ. J Chem Phys 2010;132:214309.
[83] Wang L, Xie D, Le Roy RJ, Roy P-N. J Chem Phys 2013;139:034312.

[84] Press WH, Flannery BP, Teukolsky S, Vetterling WT. Numerical recipes. The art of scientific computing. Cambridge: Cambridge University Press; 1986.
[85] Coxon JA, Colin R. J Mol Spectrosc 1997;181:215.
[86] Coxon JA, Hajigeorgiou PG. J Mol Spectrosc 1999;193:306.
[87] Coxon JA, Hajigeorgiou PG. J Mol Spectrosc 2000;203:49.