JOURNAL OF MOLECULAR SPECTROSCOPY **191,** 223–231 (1998)
ARTICLE NO. MS987646

REVIEW

Uncertainty, Sensitivity, Convergence, and Rounding in Performing and Reporting Least-Squares Fits

R. J. Le Roy

Guelph-Waterloo Center for Graduate Work in Chemistry, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada

Received October 22, 1997; in revised form May 14, 1998

This paper describes a procedure for optimal rounding of parameters determined from a linear or nonlinear least-squres fit in order to minimize the number of digits which must be quoted while ensuring that the resulting rounded constants can predict the input data with no significant loss of precision. Related problems concerning nonlinear least-squares convergence and taking account of model dependence of fitted or predicted parameters are also addressed. The recommended rounding procedure is illustrated by applications to electronic band data for the *A–X* system of I_2 and to infrared and microwave data for HF (yielding optimal new Dunham expansion coefficients for ground state HF). An automated version of this sequential rounding procedure has been incorporated in a general subroutine for performing linear or nonlinear least-squares fits.

I. INTRODUCTION

In multiparameter least-squares fits of models to experimental data sets, the scientific literature has seen very inconsistent treatments of the rounding of fitted parameters. When attempting to minimize the clutter of unnecessary digits to be reported, the simplest and most common approach is to round off all parameters at the first, second, or third digit of their uncertainty. However, it is shown below that predictions generated from the resulting constants may differ from the input data by orders of magnitude more than the residual discrepancies associated with the original fit. The alternate approach of blindly reporting all constants to maximum computer precision removes this shortcoming, but requires tediously large numbers of digits to be listed, making such parameter sets inconvenient to use and prone to transcription error by users. Two strategies for addressing this problem are presented and illustrated by applications to infrared and microwave data for HF and to electronic band data for the *A–X* system of I_2. The related problem of estimating the effect of model dependence on the value and uncertainty of a physically significant fitting parameter is also discussed.

Throughout this work, the quality of fit to an experimental data set, and the ability of a given set of parameters to accurately reproduce those data are represented by the "dimensionless standard error" of the fit

$$\bar{\sigma}_f = \left\{ \frac{1}{N-M} \sum_{i=1}^{N} \left[\frac{y_{\rm calc}(i) - y_{\rm obs}(i)}{u(i)} \right]^2 \right\}^{1/2}, \qquad [1]$$

where each of the N experimental data $y_{\rm obs}(i)$ has an uncertainty of $u(i)$, and $y_{\rm calc}(i)$ is the value of datum i predicted by the model of interest.[1] If the experimental uncertainties used are correct, a $\bar{\sigma}_f$ value of <1.0 means that on average, the predictions of the model differ from the input data by less than the associated experimental uncertainties, while a value of (say) 3.6 would mean that on average those predictions differ from experiment by 3.6 times the experimental uncertainties.

Note too that all parameter uncertainties quoted in the present work are 95% confidence limit uncertainties

$$U(p_j) = U_M(p_j) = f_{95}(N-M) \times \sigma_M(p_j), \qquad [2]$$

where $\sigma_M(p_j)$ is the usual one standard error uncertainty in parameter j obtained from an M-parameter fit to N data, and $f_{95}(N-M)$ is the student t factor for 95% confidence in a fit with $N-M$ degrees of freedom (*1, 2*). For an infinitely large data set, these values of $U(p_j)$ would correspond to approximately two standard errors in the parameters.

II. PRESENTATION OF THE PROBLEM

The problem being addressed is illustrated first by application of a standard model to a recent compendium of high-quality infrared and microwave data for HF. Ram *et al.* (*3*) recently reported extensive new hot infrared emission spectra for HF and DF, and combined those data with complementary earlier IR data and the best existing microwave data in least-squares fits which yielded optimal new band constants $\{G_v, B_v, D_v,$ etc.$\}$ for levels $v = 0$–5 of both isotopomers. The

[1] The quantity $\bar{\sigma}_f$ is the square root of the conventional "reduced chi squared" or "chi squared per degree of freedom": $\overline{\chi^2} = \bar{\sigma}_f^2$ (*1*).

0022-2852/98 $25.00

present section illustrates the nature of the problem being addressed by examining the results of a fit of the HF transition frequencies from this data set to the usual energy difference expression $y_{\rm calc}(i) = E_{\rm D}(v_i', J_i') - E_{\rm D}(v_i'', J_i'')$ based on the familiar Dunham model for vibration–rotation energies of a $^1\Sigma$ diatomic molecule (*4–6*)

$$E_{\rm D}(v, J) = \sum_{l,m} Y_{l,m}(v + 1/2)^l [J(J + 1)]^m, \qquad [3]$$

where single ($'$) and double ($''$) primes, respectively, label the quantum numbers for the upper and lower levels of transition i. Note that while all of the data for this case have the same units ($\rm cm^{-1}$), the magnitudes of the experimental uncertainties $u(i)$ are very different for different types of transitions.

Fitting the 326 HF data (*3*) to a 26-parameter version of the above model yields a dimensionless standard error of $\bar{\sigma}_f = 0.927$ and the parameter values and uncertainties listed in the second and third columns of Table 1. This quality of fit is essentially the same as that for the published 34-parameter band-constant fit to these same data, which yielded $\bar{\sigma}_f = 0.957$ (*3*). To facilitate subsequent discussion, the parameter values in column 2 of Table 1 are all quoted to an assumed machine precision of 15 significant digits, although in all cases the uncertainties are clearly sufficiently large that many of the digits listed have no physical significance.

To test the most common approach to the reporting of rounded-off parameters, the set of fitted parameters listed in Table 1 was, in turn, rounded off at the first, second, third, etc. significant digit of its (95% confidence limit) uncertainty, and the resulting rounded constants used to predict the input data and generate a value for $\bar{\sigma}_f$. The predictions generated from constants obtained on rounding after the first digit of the uncertainty (such that, e.g., $Y_{0,6}$ was fixed at 1.1×10^{-19} and $Y_{4,1}$ at $1. \times 10^{-6}$) yielded $\bar{\sigma}_f = 59.62$, which means that on average they would differ from the input data they were determined from by 59.62 times the experimental uncertainties, implying average discrepancies more than 6300% larger than those associated with the original fit! Similarly, the effect of rounding each parameter at the second, third, fourth, etc. significant digit of its uncertainty yielded the results seen in the rest of the first row of Table 2; they show that it is not until the rounding is moved to the fourth significant digit of the uncertainty (such that $Y_{0,6}$ would be reported as 1.0898×10^{-19} and $Y_{4,1}$ as 0.508×10^{-6}) that the value of $\bar{\sigma}_f$ became equivalent to that associated with the unrounded constants. In other words, rounded constants did not adequately reproduce the input data unless $3 \times 26 = 78$ digits of physically insignificant precision were included in their representation.

The above difficulties tend to increase with the number of free parameters required by the model. This is illustrated by the results for an analogous 47-parameter Dunham expansion fit to a 9552-line data set for the A state of I_2, seen in the third line of Table 2. Details regarding the data set and the associated Dunham constants for this system are found in Ref. (*7*). For this case, full consistency (to three decimal places) with the $\bar{\sigma}_f$ value yielded by the unrounded parameters from the original fit is not achieved until eight physically insignificant digits are retained for each of the 47 parameters.[2] Note that modest deviations of the "final" (those for $y > 8$) $\bar{\sigma}_f$ values in Table 2 from unity are to be expected, as absolute experimental uncertainties $u(i)$ are difficult to determine accurately, and different types of models may generally be expected to give slightly different $\bar{\sigma}_f$ values for the same data set.

For fits to nonlinear models, such as near-dissociation expansions (NDE) expressions for the level energies (*7–12*) or models for the potential energy function itself (*13–16*), the situation can be particularly difficult. For example, an equivalent quality NDE fit to the same 9552-line data set for A state I_2 mentioned above requires only half as many parameters as the Dunham expansion fit (row 3 of Table 2) discussed above. However, the last row in Table 2 shows that in fits to the NDE model, full convergence to the $\bar{\sigma}_f$ value yielded by the unrounded parameters is still not achieved until the rounding is performed at the 11th significant digit of the parameter uncertainties. The analogous results for a fit of the 326-line HF data set of Ram *et al.* (*3*) to a 14-parameter model consisting of a "modified Morse oscillator" function for the radial potential (*16*) plus a polynomial expansion for the nonadiabatic centrifugal correction function are seen in the second row of Table 2. Here, full (three decimal place) convergence to the $\bar{\sigma}_f$ value for the unrounded parameters was not achieved until rounding was performed at the sixth significant digit of the uncertainty. Thus, it seems clear that an uncertainty-based approach to parameter rounding is at best tedious, as the number of significant digits which must be retained may only be determined by trial and error.

III. A PARTIAL SOLUTION: THE "SENSITIVITY" OF A FITTED PARAMETER

A basic solution to the above problem was presented by Watson some 20 years ago (*17*) but seems to have largely escaped notice and certainly has largely escaped application. The total parameter uncertainties yielded by a proper least-squares fitting procedure have two sources: the first is the actual sensitivity of the predicted value of each datum $y_{\rm obs}(i)$ to the value of the given parameter p_j, a quantity determined by the partial derivatives $\partial y_{\rm calc}(i)/\partial p_j$ required as input to the least-squares procedure, and the second is the interparameter correlation which allows the effect of a change in the value of one parameter to be (at least) partially compensated for by a correlated change in one or more other parameter(s). In all cases except linear fits to truly orthogonal functions (such as orthogonal polynomials), the latter is usually the dominant source of the overall parameter uncertainty. On the other hand,

[2] Note that the results reported in Ref. (*7*) do not include all of these $47 \times 8 = 376$ nonphysical digits, because of an approximate manual application of the type of sequential rounding procedure described below.

TABLE 1
Results of Fitting HF Infrared and Microwave Transition Frequencies to Eq. [3]

name	p_j(*direct fit*)	$U(p_j)$	$S(p_j)$	p_j(SRR − *A*)	p_j(SRR − *B*)
$Y_{1,0}$	$4.13838769424422\times10^{+03}$	2.7×10^{-03}	8.2×10^{-07}	$4.13838758\times10^{+03}$	$4.13838707\times10^{+03}$
$Y_{2,0}$	$-8.99457582531803\times10^{+01}$	2.8×10^{-03}	2.9×10^{-07}	$-8.994564\times10^{+01}$	$-8.994511\times10^{+01}$
$Y_{3,0}$	$9.25492342337343\times10^{-01}$	1.2×10^{-03}	6.7×10^{-08}	9.25441×10^{-01}	9.2521×10^{-01}
$Y_{4,0}$	$-1.56656918089349\times10^{-02}$	2.2×10^{-04}	1.2×10^{-08}	-1.5657×10^{-02}	-1.5613×10^{-02}
$Y_{5,0}$	$-4.28496084501599\times10^{-04}$	1.5×10^{-05}	2.1×10^{-09}	-4.29×10^{-04}	-4.32×10^{-04}
$Y_{0,1}$	$2.09537278608473\times10^{+01}$	7.4×10^{-06}	2.4×10^{-08}	$2.09537273\times10^{+01}$	$2.09537256\times10^{+01}$
$Y_{1,1}$	$-7.93420037882570\times10^{-01}$	2.3×10^{-05}	8.0×10^{-09}	-7.934188×10^{-01}	-7.934125×10^{-01}
$Y_{2,1}$	$1.09694206726486\times10^{-02}$	2.2×10^{-05}	2.0×10^{-09}	1.09683×10^{-02}	1.09622×10^{-02}
$Y_{3,1}$	$-2.36014594895663\times10^{-04}$	8.5×10^{-06}	4.2×10^{-10}	-2.3557×10^{-04}	-2.3314×10^{-04}
$Y_{4,1}$	$-5.08252824082059\times10^{-07}$	1.4×10^{-06}	7.9×10^{-11}	-5.8×10^{-07}	$-1.\times10^{-06}$
$Y_{5,1}$	$-6.43974172476720\times10^{-07}$	8.8×10^{-08}	1.5×10^{-11}	-6.4×10^{-07}	-6.14×10^{-07}
$Y_{0,2}$	$-2.14896287807783\times10^{-03}$	4.3×10^{-08}	3.6×10^{-11}	-2.148963×10^{-03}	-2.148964×10^{-03}
$Y_{1,2}$	$5.86009180266378\times10^{-05}$	5.6×10^{-08}	1.4×10^{-11}	5.860×10^{-05}	5.8593×10^{-05}
$Y_{2,2}$	$-1.16133283615565\times10^{-06}$	3.6×10^{-08}	4.5×10^{-12}	-1.1593×10^{-06}	-1.1540×10^{-06}
$Y_{3,2}$	$-1.37969914146843\times10^{-08}$	9.1×10^{-09}	1.1×10^{-12}	-1.44×10^{-08}	-1.6×10^{-08}
$Y_{4,2}$	$-5.55187139617819\times10^{-09}$	7.6×10^{-10}	2.5×10^{-13}	-5.5×10^{-09}	-5.35×10^{-09}
$Y_{0,3}$	$1.66120587655227\times10^{-07}$	1.6×10^{-10}	2.7×10^{-14}	1.66129×10^{-07}	1.66133×10^{-07}
$Y_{1,3}$	$-4.68494491579844\times10^{-09}$	8.5×10^{-11}	1.6×10^{-14}	-4.691×10^{-09}	-4.685×10^{-09}
$Y_{2,3}$	$-2.80091413349209\times10^{-11}$	2.8×10^{-11}	6.0×10^{-15}	-2.9×10^{-11}	$-3.\times10^{-11}$
$Y_{3,3}$	$-2.62209552349185\times10^{-11}$	3.3×10^{-12}	1.9×10^{-15}	-2.6×10^{-11}	-2.58×10^{-11}
$Y_{0,4}$	$-1.56192833279561\times10^{-11}$	2.6×10^{-13}	2.0×10^{-17}	-1.5633×10^{-11}	-1.5635×10^{-11}
$Y_{1,4}$	$3.17041606264502\times10^{-13}$	7.1×10^{-14}	1.4×10^{-17}	3.27×10^{-13}	3.2×10^{-13}
$Y_{2,4}$	$-6.27986411598560\times10^{-14}$	8.2×10^{-15}	5.9×10^{-18}	-6.3×10^{-14}	-6.3×10^{-14}
$Y_{0,5}$	$1.49282838412224\times10^{-15}$	1.9×10^{-16}	1.5×10^{-20}	1.50×10^{-15}	1.50×10^{-15}
$Y_{1,5}$	$-1.16258830570504\times10^{-16}$	2.3×10^{-17}	1.1×10^{-20}	-1.2×10^{-16}	-1.17×10^{-16}
$Y_{0,6}$	$-1.08978648236473\times10^{-19}$	5.1×10^{-20}	1.1×10^{-23}	-1.1×10^{-19}	-1.1×10^{-19}

the ability to accurately back-calculate the input data clearly depends on the former. Thus, an approach to rounding based on these partial derivatives would seem appropriate.

Watson reasoned that the effect of rounding on predictions of a model might reasonably be allowed to be as large as $0.1\times\bar{\sigma}_f$, which implied in turn that for a fit to M parameters, the effect of rounding off any one parameter might be allowed to be as large as $(0.1/M)\bar{\sigma}_f$. For a many-parameter fit, this is of course a worst-case estimate, since it neglects the fact that the effects of rounding multiple parameters would usually tend to have some degree of cancellation. In any case, his criterion suggests the introduction of a quantity labeled here as the "parameter sensitivity,"

$$S(p_j) \equiv (0.1/M)\bar{\sigma}_f\left[\frac{1}{N}\sum_{i=1}^{N}\left\{\frac{1}{u(i)}\frac{\partial y_{\mathrm{calc}}(i)}{\partial p_j}\right\}^2\right]^{-1/2}, \quad [4]$$

TABLE 2
Results of Rounding off Constants Yielded by Fits to Various Models

Model	No. data	No. param.	$\bar{\sigma}_f$ after rounding at y–th digit of parameter uncertainties							
			$y=1$	2	3	4	6	8	10	12
Dunham fit to HF(X)	326	26	59.62	1.413	0.987	0.927	0.927	0.927	0.927	0.927
Potential fit to HF(X)	326	14	5832.	494.3	106.7	6.468	1.055	1.055	1.055	1.055
Dunham fit to I_2 (A)	9552	47	$>10^6$	167073.	17966.	2667.	22.86	1.335	1.329	1.329
NDE fit to I_2 (A)	9552	26	204010.	97130.	42684.	$>10^6$	4754.	26.60	1.407	1.383

which is the magnitude of the change in parameter p_j whose effect on the predictions of the model would increase $\bar{\sigma}_f$ by no more than $(0.1/M)\bar{\sigma}_f$. For the Dunham fit to the HF data set described in Section II, the values of these parameter sensitivities are shown in column 4 of Table 1. They are typically 2–5 orders of magnitude smaller than the associated overall parameter uncertainties and clearly provide a more selective and systematic basis on which to perform rounding. The fact that the ratios of the uncertainty to the sensitivity range over several orders of magnitude for the various parameters clearly demonstrates that any parameter rounding scheme based on the uncertainties alone is fundamentally unsound. Note that $S(p_j)$ is very simply related to the uncertainty in the value of that parameter for the case of a one-parameter fit with all other parameters held fixed at their optimum values:[3]

$$S(p_j) = (0.1\sqrt{N}/M)U_1(p_j)/f_{95}(N-M) = (0.1\sqrt{N}/M)\sigma_1(p_j). \quad [5]$$

For iterative nonlinear least-squares fits, the set of $\{S(p_j)\}$ values also provides an optimum convergence criterion. Requiring the iterative procedure to continue until the incremental change in each free parameter is smaller than the associated $S(p_j)$ value would certainly ensure that parameter changes in additional cycles would be too small to have any significant effect on the predictions of the model. This is much more stringent than conventional criteria such as requiring those changes to be smaller than the total (correlated) parameter uncertainties $\sigma_M(p_j)$, or requiring that $\bar{\sigma}_f$ (or $\overline{\chi^2}$) approaches some asymptote. Its only drawback is that in highly correlated many-parameter fits on computers with finite precision, numerical noise may prevent this limit from ever being reached. On the other hand, as long as each parameter change Δp_j is smaller than $U_M(p_j)$, one can assert that physical convergence has been achieved, and in cases where numerical noise prevents full convergence to the $S(p_j)$ limit, it will eventually give rise to small increases in $\bar{\sigma}_f$ or in the parameter changes in adjacent iterations, which would indicate that the iterative procedure should be stopped. A general iterative-fit convergence criterion which encompasses this case proceeds as follows: (i) attempt to cycle until all parameter changes $\{\Delta p_j\}$ are smaller than the associated sensitivities ($|\Delta p_j| < S(p_j)$, $\forall j$), while monitoring the size of $\bar{\sigma}_f$ and of $\max_j|\Delta(p_j)/U(p_j)|$ from cycle to cycle; (ii) if the sensitivity criterion is not (yet) satisfied, but $\bar{\sigma}_f$ or $\max_j|\Delta(p_j)/U(p_j)|$ increases from one cycle to the next (indicating numerical noise) and (say) $\max_j|\Delta(p_j)/U(p_j)| < 0.01$, simply stop iterating, as further changes will not affect the parameters within their physical uncertainties.

In conclusion, a minimal recommended alternative to conventional (unreliable) uncertainty-based parameter rounding is simply to round off each parameter at the first significant digit of the parameter sensitivity $S(p_j)$. For the four cases considered in Table 2, to three decimal places, rounding done using this criterion had no effect on the ability of the rounded constants to reproduce the values of $\bar{\sigma}_f$ from the original fits![4] This is a much more systematic and robust scheme than traditional uncertainty-based approaches to rounding and can reliably be used in all cases. Moreover, calculation of $S(p_j)$ only involves quantities already required for the least-squares procedure, so it can readily be incorporated into any standard least-squares fitting routine.[5, 6]

A residual shortcoming of this sensitivity-based parameter

[3] The ratio $S(p_j)/U_M(p_j)$ is also simply $(0.1\sqrt{N}/M)/f_{95}(N-M)$ times the "freedom parameter" $F = \sigma_1(p_j)/\sigma_M(p_j)$ proposed by Femenias (*18*) as a measure of the overall lack of dependence of a given parameter on all the other free parameters of the fit. This F parameter is in turn simply the inverse square root of a measure of the degree of interparameter correlation introduced earlier by Watson *et al.* (*19*).

[4] While better than might have been expected, this three-decimal-place stability in $\bar{\sigma}_f$ should not be surprising for two reasons. The first is the fact that the magnitude of the actual change on rounding a particular parameter will range from zero to a maximum of $S(p_j)/2$, and the second is the fact that the 10% maximum change in $\bar{\sigma}_f$ associated with our definition of $S(p_j)$ allows for the hypothetical worst case in which the effects of all truncations add constructively, with no partial cancellation.

[5] See, e.g., subroutine packages "llsqf" and "llsqfvl" obtainable from subdirectory pub/leroy/fitting on our anonymous ftp site. An associated set of notes on least-squares methodologies and the use of these packages may also be found on this site (*20*), or obtained from the author on request.

[6] To obtain this material by anonymous ftp, begin by using the command "ftp theochem.uwaterloo.ca" to connect to our computer. The response to the userid prompt should be "anonymous" and that to the password prompt should be the caller's name or e-mail address. The desired subdirectory may then be accessed with the command "cd pub/leroy/⟨subdirectory name⟩," and the desired file(s) copied with the "get ⟨filename⟩" command.

rounding scheme is that it is still necessary to report a substantially larger number of significant digits than would be required by the actual physical uncertainties in our knowledge of the parameters. Thus, one is still open to excessive opportunity for transcription errors in preparation, publication, or use of those reported quantities. For example, although applying this sensitivity-based rounding scheme to the HF data set did give a reliable set of rounded parameters, and the requisite number of digits required to fully reproduce the quality of fit is reduced from the (assumed) machine precision value of $26 \times 15 = 390$ to a more manageable 175, the latter is still tediously large.

IV. A FURTHER IMPROVEMENT: SEQUENTIAL ROUNDING AND REFITTING

A simple means of minimizing the above problem has been introduced by Tromp, Le Roy, and Tellinghuisen (*21, 22, 12*), and described by Tellinghuisen (*23, 24*). It is based on the observation that the difference between $\sigma_1(p_j)$ and $\sigma_M(p_j)$ represents the degree to which correlated changes in other parameters can compensate for rounding off parameter j. The essential idea is simple: in turn round off each parameter and repeat the fit to the remaining parameters while holding the rounded value(s) fixed, allowing the interparameter correlation to compensate for the effect of the parameter rounding. This approach can lead to a significant reduction in the numbers of significant digits which must be reported, with essentially no loss in the quality of the representation of the original data set (*12, 21–23*). Tellinghuisen's illustration (*23*) involves a manual trial-and-error means of selecting where each parameter should be rounded off. However, the present section presents and illustrates a straightforward systematic algorithm for doing this.

The recommended procedure is simply to round off the least significant parameter yielded by a given fit at the first significant digit of its (95% confidence limit) uncertainty $U(p_j)$, and then to repeat the fit to the remaining parameters with this last one held fixed at that rounded value. This rounding step can be easily automated, and the procedure iteratively repeated until the last cycle involves a fit to only one parameter. If for this last free parameter $S(p_1) < \sigma_1(p_1)$, its rounding should be at the first significant digit of its sensitivity $S(p_j)$, rather than of its uncertainty, since there would be no other free parameters to absorb the effect of its rounding; however (see Eq. [5]), this could only occur for small data sets. Because of the systematic lowering of correlation as more and more parameters are held fixed, the predicted uncertainties in all of the remaining free parameters will tend to decrease as this procedure proceeds, and it is this reduced "local" uncertainty which should be used in subsequent rounding steps. However, the uncertainties reported at the end of the procedure should be those from the original fit in which all parameters were free.

An apparent source of uncertainty in the above procedure is the fact that there is no commpletely unique order in which to perform the rounding and refitting. The simplest approach would be to design the overall parameter ordering in the model such that the "least significant" (i.e., likely to be most uncertain) parameters are listed last, and to sequentially round from last to first. In the case of the present Dunham expansion fits to Eq. [3], this means rounding the parameters as ordered in Table 1 from bottom to top. For models that involve (sets of) power series, such as the Dunham-type expansion considered here, this might seem the most natural order in which to perform such rounding. On the other hand, it would seem even better to always select to round off the remaining free parameter with the highest relative uncertainty. These two (and other possible) orderings would in general give slightly different final sets of rounded parameters. However, their differences should be smaller than the physical uncertainties $U(p_j)$, and the overall quality of the predictions would be unaffected.

Application of sequential rounding and refitting in fits to the HF data set described above yielded the sets of rounded parameters seen in the last two columns of Table 1. The results in the second to last column ("*SRR—A*") were obtained on rounding the 26 parameters from last to first as listed in the table, in the order: 26, 25, 24, 23, . . . etc. The results in the last column ("SRR—*B*") were obtained by selecting for rounding at each stage the remaining parameter with the largest relative uncertainty; in the present case this means rounding them off in the order: 10, 19, 26, 15, 22, 23, 25, 5, . . . etc. While the parameter values obtained in these two cases are usually not exactly the same, the differences between them, and between them and the original unrounded parameters of column 2 are always much smaller than the parameter uncertainties, and both rounded sets give equivalently excellent representations of the data. In particular, the $\bar{\sigma}_f$ values for predictions yielded by the two rounded parameter sets (0.929 and 0.930, respectively) are only *ca.* 0.15% larger than the value from the original fit (0.927). [7] The key result, however, is the fact that both sets of results consist of more than 1/3 fewer significant digits (113) than the number (175) associated with sensitivity-rounded parameters implied by the results of column 4. In any case, either of the equivalent sets of results in the last two columns of Table 1 may be taken as a current "best" set of single-isotope Dunham $\{Y_{l,m}\}$ constants for this species.[8]

The results descrided above clearly demonstrate that these "sequential rounding and refitting" (SRR) schemes achieve a considerable economy of representation with essentially no loss of accuracy! Similarly, for the other three cases considered in Table 2, application of these SRR procedures yielded parameter sets comprising less than 2/3 as many digits as were required by the sensitivity-based rounding criterion (63 *vs.* 97, 232 *vs.* 403, and 152 *vs.* 255, respectively), while having virtually no effect (to three decimal places) on $\bar{\sigma}_f$.

[7] Of course, even this slight increase in $\bar{\sigma}_f$ could be removed by performing the sequential rounding at the second significant digit of each uncertainty.

[8] Ram *et al.* (*3*) only reported band constants $\{G_v, B_v, D_v,$ etc.$\}$ for the observed levels of the individual isotopomers and not Dunham expansion coefficients.

There seems to be little to choose between the two schemes for sequential rounding considered above. One often requires one or two more significant digits than the other, but the identity of the more economical representation changes from case to case. The only generalization which seems evident is that the second approach (at each stage, rounding and fixing the remaining parameter with the largest relative uncertainity) will be more likely than the first to occasionally round one or more of the parameters to exactly zero. That approach tends to require more iterative fitting cycles, and hence slightly more CPU time, but this is not a significant expense. In any case, either procedure may be selected by a user of the automated subroutine for performing SRR fits developed in the present work.[9]

Tellinghuisen has described and applied a procedure similar to that presented here, although he gave no explicit prescription for performing the rounding, and he did not discuss the question of parameter ordering (*23*). In subsequent work he also showed that for a linear least-squares fit (where the observables are strictly linear functions of the parameters), this sequential rounding could be effected without actually performing the repeated fits, since the partial derivatives $\{\partial y_{\mathrm{calc}}(i)/\partial p_j\}$ are not affected by the rounding (*24*). In that case his correlated rounding scheme gives exactly the same results as sequential rounding and refitting, if the parameters are treated in the same order. However, his approach may be more complicated to implement, especially in a program set up to perform fits to both linear and nonlinear models, as his correlated approach is not reliable for the latter case.

The one weakness of Tellinghuisen's second paper (*24*) is his suggestion that although his correlated adjustment procedure is technically not correct for nonlinear fits ". . . the formal inequivalence seems unlikely to be a practical concern in more realistic adjustments of parameters for non-linear fits". This is an optimistic assessment which will only be valid when either the parameter changes due to rounding are "small" or the axis of the confidence ellipsoid in parameter space has little curvature. In many cases this will not be true, and no shortcuts to the full sequential rounding and refitting are possible. This in turn raises a significant practical problem, since each subsequent (nonlinear) fit to the reduced parameter set requires a realistic set of initial trial parameters, and a simple combination of the last rounded parameter with the (unrounded) other fitted parameters may not be adequate. An illustration of this problem and a means of dealing with it are presented below.

V. "SRR" IMPLEMENTATION: CORRECTING TRIAL PARAMETERS IN SEQUENTIAL ROUNDING FOR NONLINEAR FITS

In the application of sequential rounding and refitting to linear least-squares fits, there is no need to be concerned about optimizing the trial values of the remaining parameters, since with a linear model, fitting to converge on a parameter correction is equivalent to fitting to determine the parameters themselves. This is not the case for nonlinear fits, and if the parameter sensitivity $S(p_j)$ is much smaller than the overall parameter uncertainty $U(p_j)$, rounding that constant can massively change the predictions of the model. As a result, even though such changes can be fully compensated for by correlated changes in other parameters, predictions generated on simply combining that rounded parameter with the other fitted parameters from the preceeding cycle may be sufficiently implausible that the nonlinear fitting procedure fails. This has been found to be the case in fits to near-dissociation expansions for vibrational energies and rotational constants, where an accurate fit may require use of a moderately high-order rational polynomial or exponent polynomial (*7, 10, 12, 25*).

Fortunately, it is a straightforward matter to predict corrections to the remaining free parameters which will compensate in an optimum way for the effect of rounding off and fixing a particular parameter. As is shown in the Appendix, if the change in parameter m due to rounding is δp_m, the optimum compensating adjustment to (remaining free) parameter j is[10]

$$\Delta p_j = (\mathbf{C})_{j,m}\, \delta p_m\, U(p_j)/U(p_m), \qquad [6]$$

where $\mathbf{C}$ is the correlation matrix for the fit, which is readily accessible as output from any reasonable least-squares fitting routine.[5,9] This expression assumes that the axis of the confidence ellipsoid in parameter space is locally linear, and is implicit in Tellinghuisen's correlated rounding procedure for fits to linear models. While not exact for nonlinear fits, it provides excellent initial trial parameters for proceeding with fits following even very drastic rounding [e.g., rounding of an exponent polynomial expansion parameter for which $S(p_m)/\delta p_m < 10^{-10}$]. A general least-squares fitting routine ("nllssrr") which utilizes Eq. [6] while applying the present recommended sequential rounding and refitting scheme may be obtained from the author by anonymous ftp.[9]

VI. CALCULATING UNCERTAINTIES IN QUANTITIES PREDICTED FROM THE PARAMETERS OF A MODEL

For the sake of completeness, this section reviews a topic which is not directly concerned with parameter rounding, but with the closely related subject of determining the uncertainties in the values of properties predicted from parameters determined by a fit. For example, what is the uncertainty in a given RKR potential energy function turning point due to the uncertainties in the vibration–rotation constants used to calculate it, or what is the uncertainty in the value of the vibrational

[9] Subroutine "nllssrr" may be obtained by anonymous ftp from subdirectory "pub/leroy/fitting" on the computer "theochem.uwaterloo.ca."

[10] I am grateful to Professor M. E. Thompson for bringing to my attention both the simple relationship of Eq. [6] and its derivation for the more general case in which several parameters are rounded simultaneously (summarized in the Appendix).

constant $\omega_e x_e$ associated with a near-dissociation expansion (NDE) expression for vibrational energies, or what are the uncertainties in the values of an analytic intermolecular potential function determined from a direct fit to experimental transition frequencies? While answers to these particular questions have been reported (*25–27*), the generality of those answers does not seem to be widely appreciated, with the result that values of derived (spectroscopic) properties are often reported without being accompanied by proper estimates of the associated uncertainties. The objective here is simply to remind the reader of the straightforward procedure for calculating such uncertainties, and to adjure them to always do so.

For a property $F(\{p_j\})$ defined by a set of M parameters $\{p_j\}$ determined from a least-squares fit yielding parameter uncertainties $\{U(p_j)\}$ and correlation matrix **C**, the associated uncertainty in $F(\{p_j\})$ is (*28, 29*)

$$U(F) = \sqrt{\mathbf{D}^t \cdot \mathbf{C} \cdot \mathbf{D}}, \qquad [7]$$

where **D** is an $M \times 1$ column matrix of elements $U(p_j)(\partial F/\partial p_j)$, for $j = 1 - M$. This expression may, of course, also be written in terms of the variance–covariance matrix **V** instead of **C**.[11] One or the other of these matrices, together with the correlated one standard error (66% confidence limit) or 95% confidence limit parameter uncertainties, is returned by any reasonable least-squares fitting subroutine package (*20*).[5]

In many cases the application of this procedure is quite straightforward, as the property of interest is an explicit analytic function of the parameters. For example, direct "Hamiltonian" fits of spectroscopic transition frequencies to analytic potential energy functions are becoming increasingly common (*13–16*), and the potential parameters determined in this way can be readily used to predict the uncertainties in the resulting potential function in the extrapolation region beyond the range spanned by the data (*27*). However, this has rarely been done. For cases in which the dependence is not direct or cannot be expressed in closed form, the programming may be a little more complex, but straightforward application of error propagation techniques allows proper uncertainties to be generated. This is the case for the determination of RKR turning point uncertainties, where the turning points, and hence also their partial derivatives with respect to the various spectroscopic parameters, are algebraically defined in terms of numerical quadratures (*26*). However, while fits to such turning points are often used in attempts to determine dissociation limits and long-range potential constants, their associated uncertainties and their correlations (*28–32*) are usually [inappropriately (*33*)] ignored.

[11] The elements of **C** and **V** are interrelated as (*2, 28, 29*): $(\mathbf{C})_{a,b} = (\mathbf{V})_{a,b}/\sqrt{(\mathbf{V})_{a,a}(\mathbf{V})_{b,b}}$.

VII. PARAMETER UNCERTAINTIES AND MODEL DEPENDENCE

The final topic addressed here is one which is usually ignored in polite (scientific) society. It concerns the problem of estimating the values of and real total uncertainty in physically interesting quantities determined from fits of experimental data to particular models. If the quantity of interest is one of the explicit parameters of the model, the normal least-squares procedure itself yields an estimate of its uncertainty within that model, while if it is a derived quantity, the procedure of Section VI does the same. However, this takes no account of the fact that the choice of the analytic form of the model itself is a further source of uncertainty. The nature of the problem is illustrated here by consideration of the particular problem of fits to expressions for the vibrational energy levels of a diatomic molecular ion, but the concerns are quite general.

Consider the problem of using a fit to the observed energies $G(v)$ of the lower vibrational levels of a diatomic molecular ion to determine optimum estimates of its dissociation energy $\mathcal{D}$ and of the effective vibrational index at dissociation $v_{\mathcal{D}}$. Theory tells us that the limiting near-dissociation behavior of the pattern of vibrational levels is given by

$$G_{\lim}(v) = \mathcal{D} - X_0(4)(v_{\mathcal{D}} - v)^4, \qquad [8]$$

where $X_0(4)$ is a known constant determined by the polarizability of the uncharged fragment atom yielded on dissociation (*34–37*). Since the observed levels generally do not lie in the limiting region where this behavior is quantitatively obeyed, the data are actually fitted to a more general "near-dissociation expansion"

$$G_{\mathrm{NDE}}(v) = \mathcal{D} - X_0(4)(v_{\mathcal{D}} - v)^4 \times \mathcal{F}(v_{\mathcal{D}} - v) \qquad [9]$$

which combines the limiting behavior of Eq. [8] with some empirical expansion function $\mathcal{F}(v_{\mathcal{D}} - v)$ which takes account of the deviation of the observed lower (small v) levels from the limiting form. However, there is no single "true" or "best" form for this empirical deviation function, and the dispersion among the estimates of the values and uncertainties in the fitted values of the physically interesting parameters $\mathcal{D}$ and $v_{\mathcal{D}}$ yielded by different models is usually much greater than the properly correlated uncertainties in the values obtained from a fit to any one model (*9, 38, 25*).

An "averaging-over-models" procedure for addressing this problem was introduced in Refs. (*9, 38, 25*). In particular, if a fit to a particular model "*k*" which has an overall (dimensionless) standard error of $\bar{\sigma}_f(k)$ yields an estimate P_k of the physical quantity P, with (95% confidence limit) uncertainty $U(P_k)$, then the recommended best estimate for the value of this quantity is

$$\bar{P} = \sum_k (w_k^P P_k) / \sum_k w_k^P, \quad [10]$$

and for the associated 95% probability uncertainty in $\bar{P}$ is

$$\bar{U}(\bar{P}) = 3\left(\frac{\Sigma_k w_k^P\{(P_k - \bar{P})^2 + [U_M(P_k)/f_{95}(N-M)]^2\}}{\Sigma_k w_k^P}\right)^{1/2}, \quad [11]$$

where the (unnormalized) weight given to the value of (and uncertainty in) parameter P yielded by model k is

$$w_k^P = 1/\{U_M(P_k)[1 + \bar{\sigma}_f(k)]\}^2. \quad [12]$$

These particular definitions of w_k^P and $\bar{U}(\bar{P})$ differ slightly from those used in earlier work (*9*, *38*, *25*). The second factor in the denominator of Eq. (12) is included to discriminate against models which may give a poor overall fit to the data, but for which a relatively low degree of interparameter correlation may give rise to relatively small parameter uncertainties; it also presumes that the data uncertainties $\{u(i)\}$ are sufficiently realistic that for "good" fits the magnitude of $\bar{\sigma}_f(k)$ is close to unity.

The sums over k in Eqs. [10] and [11] should include as wide a range of (realistic) models as possible, but there will clearly always be a practical limit to the range of models which may be considered. However, use of this type of approach with some range of models will always give more realistic estimates of the true parameter values and their uncertainties than would the conventional approach of considering only a single model and relying solely on the values and uncertainties it yields. This particular procedure has been incorporated into a general nonlinear least-squares fitting program for fitting near-dissociation expansions to diatom vibrational level energies, which may be obtained from the author on request.[12]

VIII. CONCLUSIONS AND RECOMMENDATIONS

In conclusion, the present work shows that at the very least, the number of significant digits quoted for parameters determined in a least-squares fit should be based on rounding at the first significant digit of the Watson (*17*) parameter sensitivity of Eq. [4]. Considerable further contraction may be achieved by applying the sequential rounding and refitting procedure of Section IV, or for linear fits, Tellinghuisen's correlated rounding procedure (*24*). Present experience suggests that this will yield a further $\gtrsim\frac{1}{3}$ reduction in the number of significant digits which must be reported. For nonlinear fits, application of this approach will require, or at least be facilitated by, use of the procedure of Section V for estimating corrections to initial trial parameters used in subsequent cycles of the sequential rounding procedure. A general least-squares fitting subroutine for application to linear or nonlinear models which (optionally) implements this SRR procedure may be obtained from the author by anonymous ftp.[6,9] Finally, the reported values of any derived quantities calculated from parameters determined from such fits should always be accompanied by proper correlated uncertainties, calculated as described in Section VI, and whenever possible, the effect of model dependence should be taken into account when reporting the overall uncertainty in physically interesting parameters. One approach to this latter problem is described in Section VII.

[12] Fortran program "gvnde" for fitting experimental vibrational energies to Eq. [9] and performing the averaging over models represented by Eqs. [10]–[11] (*25*) may be obtained by sending a request by e-mail to "leroy@UWaterloo.ca."

APPENDIX. CORRECTING SOME PARAMETERS TO COMPENSATE FOR CHANGES TO OTHERS

In a normal linear (or nonlinear) least-squares problem involving uncorrelated data, the parameter (or parameter change) column vector $\mathbf{p}$ is determined as the solution of the matrix equation

$$\mathbf{p} = \mathbf{V}(\mathbf{X}^t \cdot \bar{\mathbf{y}}), \quad [13]$$

where $\bar{\mathbf{y}}$ is the $N \times 1$ column vector of reduced data $(\bar{\mathbf{y}})_i = y_{\text{obs}}(i)/u(i)$, $\mathbf{X}$ is the $N \times M$ matrix of weighted partial derivatives $(\mathbf{X})_{i,j} = [1/u(i)][\partial y_{\text{calc}}(i)/\partial p_j]$, and $\mathbf{V} = (\mathbf{X}^t \cdot \mathbf{X})^{-1}$. If the parameter vector is now divided into two component vectors, $\mathbf{p} = \begin{pmatrix}\mathbf{p}_a \\ \mathbf{p}_b\end{pmatrix}$, where $\mathbf{p}_b$ is the set of parameters which are to be rounded to have values $\mathbf{p}'_b$, then we can similarly partition $\mathbf{X} = (\mathbf{X}_a\ \mathbf{X}_b)$, and the new least-squares estimate for parameters $\mathbf{p}_a$ when parameters $\mathbf{p}_b$ are fixed at the (rounded) values $\mathbf{p}'_b$ is

$$\mathbf{p}'_a = (\mathbf{X}_a^t \cdot \mathbf{X}_a)^{-1}\, \mathbf{X}_a^t(\bar{\mathbf{y}} - \mathbf{X}_b \cdot \mathbf{p}'_b). \quad [14]$$

On partitioning the $\mathbf{V}$ matrix analogously, Eq. [13] may be rewritten as

$$\begin{pmatrix}\mathbf{p}_a \\ \mathbf{p}_b\end{pmatrix} = \begin{pmatrix}\mathbf{V}_{a,a} & \mathbf{V}_{a,b} \\ \mathbf{V}_{b,a} & \mathbf{V}_{b,b}\end{pmatrix}\begin{pmatrix}\mathbf{X}_a^t \cdot \bar{\mathbf{y}} \\ \mathbf{X}_b^t \cdot \bar{\mathbf{y}}\end{pmatrix} = \begin{pmatrix}\mathbf{V}_{a,a} \cdot \mathbf{X}_a^t \cdot \bar{\mathbf{y}} + \mathbf{V}_{a,b} \cdot \mathbf{X}_b^t \cdot \bar{\mathbf{y}} \\ \mathbf{V}_{b,a} \cdot \mathbf{X}_a^t \cdot \bar{\mathbf{y}} + \mathbf{V}_{b,b} \cdot \mathbf{X}_b^t \cdot \bar{\mathbf{y}}\end{pmatrix}. \quad [15]$$

Making use of matrix identities expressing the matrix blocks $\mathbf{V}_{\alpha,\beta}$ in terms of the products $(\mathbf{X}_\alpha^t \cdot \mathbf{X}_\beta)$ and their inverses (for α and $\beta = a$ or b) then yields

$$\mathbf{p}'_a - \mathbf{p}_a = \mathbf{V}_{a,b} \cdot \mathbf{V}_{b,b}^{-1}(\mathbf{p}'_b - \mathbf{p}_b). \quad [16]$$

For the case in which parameter subset $\mathbf{p}_b$ consists of only one

parameter, the elements of this difference vector are given by Eq. [6].[10]

ACKNOWLEDGMENTS

I am pleased to acknowledge extensive discussions over many years with Professors Joel Tellinghuisen and Mary E. Thompson, who have helped shape my understanding of these topics, and I would like to thank Professor Tellinghuisen, Professor Peter Bernath, Dr. James Watson, and Dr. Michael Goldstein for helpful comments on the manuscript. I am also grateful to Mr. Jenning Seto for providing the results of the potential fits to the HF data set.

REFERENCES

1. J. R. Taylor, "An Introduction to Error Analysis—The Study of Uncertainties in Physical Measurements," University Science Books, Mill Valley, CA, 1982.
2. W. H. Press, B. P. Flannery, S. Teukolsky, and W. T. Vetterling, "Numerical Recipies. The Art of Scientific Computing," Cambridge University Press, Cambridge, UK, 1986.
3. R. S. Ram, Z. Morbi, B. Guo, K.-Q. Zhang, P. F. Bernath, J. Vander Auwera, J. W. C. Johns, and S. P. Davis, *Astrophys. J. Supp. Ser.* **103,** 247 (1996).
4. J. L. Dunham, *Phys. Rev.* **41,** 713 (1932).
5. J. L. Dunham, *Phys. Rev.* **41,** 721 (1932).
6. G. Herzberg, "Spectra of Diatomic Molecules," Van Nostrand, Toronto, Canada, 1950.
7. D. R. T. Appadoo, R. J. Le Roy, P. F. Bernath, S. Gerstenkorn, P. Luc, J. Vergès, J. Sinzelle, J. Chevillard, and Y. D'Aignaux, *J. Chem. Phys.* **104,** 903 (1996).
8. A.-R. Hashemi-Attar, C. L. Beckel, W. N. Keepin, and S. A. Sonnleitner, *J. Chem. Phys.* **70,** 3881 (1979).
9. R. J. Le Roy and W.-H. Lam, *Chem. Phys. Lett.* **71,** 544 (1980).
10. J. W. Tromp and R. J. Le Roy, *Can. J. Phys.* **60,** 26 (1982).
11. C. L. Beckel, R. B. Kwong, A.-R. Hashemi-Attar, and R. J. Le Roy, *J. Chem. Phys.* **81,** 66 (1984).
12. J. W. Tromp and R. J. Le Roy, *J. Mol. Spectrosc.* **109,** 352 (1985).
13. R. J. Le Roy and J. van Kranendonk, *J. Chem. Phys.* **61,** 4750 (1974).
14. W. M. Kosman and J. Hinze, *J. Mol. Spectrosc.* **56,** 93 (1975).
15. J. A. Coxon and P. G. Hajigeorgiou, *Can. J. Phys.* **94,** 40 (1992).
16. H. G. Hedderich, M. Dulick, and P. F. Bernath, *J. Chem. Phys.* **99,** 8363 (1993).
17. J. K. G. Watson, *J. Mol. Spectrosc.* **66,** 500 (1977).
18. J.-L. Femenias, *J. Mol. Spectrosc.* **144,** 212 (1990).
19. J. K. G. Watson, S. C. Foster, A. R. W. McKellar, P. Bernath, T. Amano, F. S. Pan, M. W. Crofton, R. S. Altman, and T. Oka, *Can. J. Phys.* **62,** 1875 (1984).
20. R. J. Le Roy, "A Practical Guide to Least-Squares Fitting," University of Waterloo Chemical Physics Research Report CP-628R, 1998.
21. J. W. Tromp and R. J. Le Roy, *in* "37th Ohio State University Symposium on Molecular Spectroscopy." Paper RF14, The Ohio State University, Columbus, Ohio, 1982.
22. J. Tellinghuisen, *Can. J. Phys.* **62,** 1933 (1984).
23. J. Tellinghuisen, *J. Mol. Spectrosc.* **137,** 248 (1989).
24. J. Tellinghuisen, *J. Mol. Spectrosc.* **173,** 308 (1995).
25. R. J. Le Roy, *J. Chem. Phys.* **101,** 10217 (1994).
26. J. Tellinghuisen, *J. Mol. Spectrosc.* **141,** 258 (1990).
27. C. E. Chuaqui, R. J. Le Roy, and A. R. W. McKellar, *J. Chem. Phys.* **101,** 39 (1994).
28. W. E. Deming, "Statistical Adjustment of Data," Dover, New York, 1964.
29. D. L. Albritton, A. L. Schmeltekopf, and R. N. Zare, *in* "Molecular Spectroscopy: Modern Research" (K. Rao, Ed.), Vol. 2, Chap. 1, pp. 1–67, Academic Press, New York, 1976.
30. R. T. Birge, *Phys. Rev.* **40,** 207 (1932).
31. J. Tellinghuisen, *J. Mol. Spectrosc.* **165,** 255 (1994).
32. J. Tellinghuisen, *J. Mol. Spectrosc.* **179,** 299 (1996).
33. M. R. Davies, J. C. Shelley, and R. J. Le Roy, *J. Chem. Phys.* **94,** 3479 (1991).
34. R. J. Le Roy and R. B. Bernstein, *Chem. Phys. Lett.* **5,** 42 (1970).
35. R. J. Le Roy and R. B. Bernstein, *J. Chem. Phys.* **52,** 3869 (1970).
36. R. J. Le Roy, *in* "Molecular Spectroscopy" (R. Barrow, Ed.), Vol. 1, Specialist Periodical Report 3, pp. 113–176, Chemical Society of London, London, 1973.
37. P. F. Bernath, "Spectra of Atoms and Molecules," Oxford University Press, Oxford, New York, 1995.
38. K. J. Jordan, R. H. Lipson, N. A. McDonald, and R. J. Le Roy, *J. Phys. Chem.* **96,** 4778 (1992).