# On a Test for Trends in Piecewise Monotonic Data Approximation Method

E. E. Vassiliou and I. C. Demetriou

Abstract—We consider the numerical results of the piecewise monotonic approximation method combined with a test for trends of the residuals that is applied to a data set which is considered difficult to be fitted. If n are the data and k-1 is the number of positions of the joins of the monotonic sections, the method requires  $\mathcal{O}(n^{k-1})$  combinations of positions in order to find an optimal one. The test attempts to distinguish genuine trends from data errors in order to provide automatically an adequate number of monotonic sections. This problem is hard to solve, due to both the inherent combinatorial nature of the piecewise monotonic approximation method and the need to compare the goodness of competing fits. We show that the test for trends by controlling an upper bound of the variance of the fit allows progressively improved approximations. Our results expose some critical questions and suggest a subject for future research.

Index Terms—approximation, data smoothing, divided difference, dynamic programming, least squares, piecewise monotonic, Raman spectrum, turning point

#### I. INTRODUCTION

ET { $\phi_i$  : i = 1, 2, ..., n} be a sequence of measurements of function values { $f(x_i)$  : i = 1, 2, ..., n}, where f(x) is a real function of one variable, and where the abscissae are in the strictly ascending order  $x_1 < x_2 < \cdots < x_n$ . We assume that due to errors of measurement the sequence of the first differences

$$\{\phi_{i+1} - \phi_i: \quad i = 1, 2, \dots, n-1\}$$
(1)

has far more sign changes than the sequence of the first differences of the function values  $f(x_i)$ . The number of sign changes in the sequence (1) may suggest to smooth the data, particularly if it is known that f(x) is composed of a smaller number of monotonic sections than that shown by the first differences (1). Demetriou and Powell [5] proposed and studied a data smoothing method by imposing a prescribed number, say k - 1, of sign changes on the first differences of the smoothed values  $\{y_i : i = 1, 2, \ldots, n\}$ . These values allow at most k sections of monotonic components alternately increasing and decreasing. There is no loss of generality to assume that the first monotonic section of the best fit is increasing. Ideally, k - 1 is the number of sign changes in the first derivative of f(x), but the user can try many values of k that may be suitable.

Vassiliou and Demetriou [10] have extended this method by applying a test on the residuals for trends in order to identify automatically an adequate value for k. The test has been motivated by a test of data trends by Powell [8] on curve fitting by splines. The underlying assumption is that the residuals are random variables with the same probability density function, which has mean zero. This assumption is useful to our theoretical analysis, but in practice our knowledge is restricted to just the data at hand, which indicates the difficulty of the problem we have addressed.

Our test for trends includes a parameter,  $\eta$  say, that forces some smoothness in the fit by controlling the number of monotonic sections k. Therefore, changing  $\eta$  allows different numbers of monotonic sections in the derived fit.

In this paper we consider the details of one calculation that considers best approximations to real data from a Raman spectrum in order to identify peaks. Peak finding is an important problem of spectroscopy of continuous interest (see, for example, Gunther [6]; also, [2] and [4] are two applications of the piecewise monotonic method on peak finding). The data show many peaks of various sizes at irregular positions, sudden changes, and great variability. The complexity of the underlying physical laws make this a good test of the efficacy of the extended piecewise monotonic approximation method. We consider ranges of values of  $\eta$  that satisfy an inclusion relation, each range giving the same k. Thus, a sequence of values of  $\eta$  is generated that detects progressively monotonic sections with diminishing importance on the fit.

Piecewise monotonic approximation by Demetriou and Powell [5] is a combinatorial optimization problem that requires about  $\mathcal{O}(n^{k-1})$  combinations of positions of sign changes in the first differences of the smoothed values in order to identify an optimal combination. The problem has been solved by developing a dynamic programming algorithm that requires only  $\mathcal{O}(kn^2)$  computer operations in the least squares case. Software has also been developed by Demetriou [1] that is about 1800 Fortran lines including comments, which gives an idea of the size of the probelm. The inclusion of the test for trends by Vassiliou and Demetriou [10] reorganizes the iterations of the dynamic programming algorithm at the expense of only  $\mathcal{O}(kn)$  computer operations, and is also supported by software development that is necessary to apply the test for trends.

Section II gives some details of the calculation that are needed later. Section III considers the details of an experiment with the method of Section II on the mentioned Raman spectrum data. Section IV gives some concluding remarks.

### II. AN OUTLINE OF THE METHOD

In this section we include basic ideas of the piecewise monotonic method [5] that are necessary to apply this method together with our test for trends.

Therefore we let k be a positive number that is smaller than n, and the piecewise monotonic method seeks a vector

Manuscript received April 25, 2021; revised June 24, 2021

E. E. Vassiliou is an Assistant Professor of Department of Financial and Business Engineering, University of the Aegean, 41 Kountouriotou Street, Chios 82100, Greece. e-mail: e.vassiliou@aegean.gr

I. C. Demetriou is a Professor of Department of Economics, National and Kapodistrian University of Athens, 1 Sofokleous and Aristidou Street, Athens 10559, Greece. e-mail: demetri@econ.uoa.gr

Proceedings of the World Congress on Engineering 2021 WCE 2021, July 7-9, 2021, London, U.K.

y in  $\mathbb{R}^n$  that minimizes the sum of squares

$$\Phi(\underline{y}) = \sum_{i=1}^{n} (\phi_i - y_i)^2 \tag{2}$$

subject to the piecewise monotonicity constraints

$$\begin{cases} y_{t_{j-1}} \le y_{t_{j-1}+1} \le \dots \le y_{t_j}, \ j \text{ odd} \\ y_{t_{j-1}} \ge y_{t_{j-1}+1} \ge \dots \ge y_{t_j}, \ j \text{ even} \end{cases}$$

$$(3)$$

where  $\{t_j : j = 1, 2, ..., k - 1\}$  are integers that satisfy the conditions

$$1 = t_0 \le t_1 \le \dots \le t_k = n. \tag{4}$$

While k is provided by the user, the integers  $\{t_j : j = 1, 2, ..., k-1\}$  are variables of the minimization calculation together with the components  $\{y_i : i = 1, 2, ..., n\}$ . There are about  $O(n^{k-1})$  combinations of positions of the variables  $\{t_j : j = 1, 2, ..., k-1\}$  for solving this problem, which makes it a formidable calculation. Demetriou and Powell have developed a dynamic programming method that generates the required fit in only  $O(kn^2)$  computer operations. The calculation takes account of the following two properties of the solution that depend on the problem and that do not occur generally.

The first property is that the optimal fit interpolates the data at  $\{t_j : j = 1, 2, ..., k - 1\}$  giving

$$y_{t_i} = \phi_{t_i}.$$

The second property is that the optimal fit when  $k \ge 2$  consists of separate optimal monotonic components that occur between successive integers  $\{t_j\}$ . Therefore the calculation of the optimal fit reduces to dividing the data into at most k disjoint sets of adjacent data and solving a monotonic calculation for each set. The division into suitable sets is achieved by the mentioned dynamic programming method. In order to provide a brief description of it, we introduce the notation

$$\alpha(t_{j-1}, t_j) = \min_{y_{t_{j-1}} \le y_{t_{j-1}+1} \le \dots \le y_{t_j}} \sum_{i=t_{j-1}}^{t_j} (y_i - \phi_i)^2, \ j \text{ odd},$$

for the monotonic increasing section on the interval  $[t_{j-1}, t_j]$ , and similarly the notation  $\beta(t_{j-1}, t_j)$  for the monotonic decreasing section on the interval  $[t_{j-1}, t_j]$ . It follows that if the values  $\{t_j : j = 1, 2, ..., k-1\}$  are optimal and k is odd, say, then the least value of the objective function (2) is the expression (see Demetriou [3])

$$\Phi(\underline{y}) = \alpha(t_0, t_1) + \beta(t_1, t_2) + \alpha(t_2, t_3) + \dots + \alpha(t_{k-1}, t_k).$$

Further, we introduce the notation  $\gamma(m, t) = \min_{\underline{z} \in Y(m,t)} \sum_{i=1}^{t} (z_i - \phi_i)^2$  for any integers  $m \in [1, k]$  and  $t \in [1, n]$ , where Y(m, t) is the set of *t*-vectors  $\underline{z}$  with *m* monotonic sections. Therefore in order to calculate  $\gamma(k, n)$ , which is the least value of the objective function (2), the calculation begins with the values m = 1 and

$$\gamma(1,t) = \alpha(1,t), \text{ for } t = 1, 2, \dots, n.$$
 (5)

Then, as  $m = 2, 3, \ldots, k$ , it proceeds by applying the formulae

$$\gamma(m,t) = \begin{cases} \min_{\substack{1 \le s \le t}} \left[ \gamma(m-1,s) + \alpha(s,t) \right], \ m \text{ odd} \\ \min_{\substack{1 \le s \le t}} \left[ \gamma(m-1,s) + \beta(s,t) \right], \ m \text{ even,} \end{cases}$$
(6)

and storing  $\tau(m, t)$ , which is the value of s that minimizes the right hand term of expression (6), for t = 1, 2, ..., n. In this way,  $\gamma(k, n)$  can be found in  $O(kn^2)$  computer operations. At the end of the calculation, m = k occurs and the value  $\tau(k, n)$  is the integer  $t_{k-1}$ . Hence, by setting  $t_0 = 1$  and  $t_k = n$ , the sequence of optimal values  $\{t_j: j = 1, 2, ..., k - 1\}$  is obtained by the backward formula

$$t_{j-1} = \tau(j, t_j), \text{ for } j = k, k-1, \dots, 2,$$
 (7)

and the components of the relevant optimal approximation are obtained by independent monotonic approximation calculations between successive  $\{t_j\}$ .

As already stated, in this paper we take the point of view that k is unknown, which makes the piecewise monotonic problem even harder to solve. So our algorithm combines the dynamic programming method described above with a test for trends between the monotonic sections (3) in a way that increases m in formulae (6) up to an adequate value of k.

Our algorithm is given the data  $\{\phi_i : i = 1, 2, ..., n\}$ , while the user may specify a value for a positive parameter  $\eta$  (for example  $\eta = 1$ ), which will be explained later. Initially, the algorithm sets m = 1, makes the assignments (5), obtains the components  $\{y_i : i = 1, 2, ..., n\}$  of the best approximation that gives  $\gamma(1, n)$ , and applies the test. If the test fails, then an iterative procedure starts. On each iteration, m is increased by one, the best approximation to the data that has m monotonic sections is calculated, and the residuals are tested for trends between successive  $t_j$ . If trends are found, another iteration starts so as to add one more monotonic section. Otherwise the iteration sets k to the current value of m and stops.

The test for trends by Vassiliou and Demetriou is briefly described as follows. We let  $\{y_i : i = 1, 2, ..., n\}$  be the components of the best approximation with m sections. Next, we form the residuals

$$e_i = y_i - \phi_i, \quad i = t_{j-1}, t_{j-1} + 1, \dots, t_j,$$
 (8)

and calculate the quantity

$$R_j = \sum_{i=t_{j-1}+1}^{t_j} e_{i-1}e_i.$$
(9)

We suppose that there is a trend if

$$R_j \ge \eta C_j,\tag{10}$$

where  $\eta$  is a positive parameter whose initial value is provided by the user, and where

$$C_j = \sqrt{t_j - t_{j-1}} \sum_{i=t_{j-1}}^{t_j} e_i^2 / (t_j - t_{j-1} + 1).$$
(11)

The current approximation is considered to be satisfactory if (10) fails to be satisfied for all j such that  $1 \le j \le m$ . Otherwise m is increased by one and another iteration is commenced. The iterations continue until no more trends are indicated in the fit.

The parameter  $\eta$  deserves our attention. Inequality (10) indicates that the termination of the Vassiliou and Demetriou algorithm with an adequate approximation depends on the

value of the parameter  $\eta$ . This parameter controls the sensitivity of the method with respect to the magnitude of the identified turning points  $y(t_i)$ . Larger values of  $\eta$  reveal only dominant turning points of the fit, while smaller values of  $\eta$  detect in addition turning points of minor importance to the fit, which are not detected with the larger values of  $\eta$ . Thus, the user, instead of having a direct control on k as it happens in the method of Demetriou and Powell, has an indirect control on it through the parameter  $\eta$ , because the ratio  $R_i/(\eta\sqrt{t_i-t_{i-1}})$  provides an upper bound on the variance of the fit on the section  $[t_{i-1}, t_i]$  as we can derive from equations (10) and (11). Two main advantages of this approach are the following. First, an adequate value of k for identifying simultaneously all the turning points with almost the same importance is automatically determined. Second, the computation benefits as no redundant values of k are attempted. In particular, provided that the user knows the experimental accuracy, he may well take it into account to stop when the error in the fit is only due to this accuracy.

# III. ANALYSIS BY AN APPLICATION ON A DIFFICULT DATA SET

This section presents an example that illustrates the efficiency of our method for identifying the peaks of a Raman spectrum sample. Peak finding is an important application of the piecewise monotonic approximation method, and the complexity of the underlying structure of the spectral data makes it an excellent test for our method. We use the descriptive term "turning point" for the value  $y(t_j)$  at the integer variable  $t_j$ . To comply with the terminology of the specialists we refer to  $y(t_j)$  as a peak. The data set named "Artesunate" was downloaded form RRUFF [9], which is a project website containing an integrated database of Raman spectra, X-ray diffraction and chemistry data for minerals. Our choice of the data set is due to its contamination of different levels of variation and the many isolated peaks of different magnitude it contains.

The Artesunate datafile contains n = 2374 pairs of data points. The first coordinate represents the Raman shift  $(cm^{-1})$  and the second coordinate represents the intensity, feeding our algorithm with values  $\{x_i : i = 1, 2, ..., n\}$  and  $\{\phi_i : i = 1, 2, ..., n\}$  respectively. The data are too many to be presented in these pages, but their main characteristics are illustrated in Fig. 1. Specifically, across the data, we can distinguish ranges of small variability, many noticeable peaks of different magnitude, sharp increases and sharp decreases as well. We do not know the underlying physical law, and we make no assumption about the nature of any underlying function f(x).

We seek turning points that reveal simultaneously all the peaks in the data with the same importance with respect to the reduction of the residuals after employing the test for trends. To this end, we fed the data to our computer program by trying different values of  $\eta$  in the range [1, 50] with step 0.1. The final choice of the range of values that was picked for this experimentation is due to a preparatory experimental outcome indicating that this choice can capture piecewise monotonic fits from 1 to 488 monotonic sections. In the second case there are on average five data points in a monotonic section. This level of the experimental extend was considered appropriate, but for illustration purposes only

piecewise monotonic fits with 2 to 48 monotonic sections are presented in these pages. Again, in the second case there are on average 51 data points per monotonic section.

We illustrate our findings by presenting in Table I the turning points and their positions by piecewise monotonic fits to the Artesunate data file for nine successive ranges of values  $\eta$ . Table I consists of a duplex of columns, each duplex having 49 rows, one for each turning point of the fit when k = 48 including also the end point indices  $t_0$  and  $t_{48}$ . The left hand duplex gives  $j, t_i, x_{t_i}$  and  $\phi_{t_i}$ . The right hand duplex contains nine columns of ranges of values of  $\eta$  and the corresponding values of k obtained from the calculation. Each row presents the positions of the turning points  $(t_i)$ for each distinct optimal fit that was derived for all values of tested  $\eta$  in the given intervals. At first, we found that identical piecewise monotonic fits were derived for many different successive values of  $\eta$  until a piecewice fit was produced automatically providing a significant shift to the value of k. For example, for all  $\eta \in [29.7, 44.3]$  a piecewise monotonic fit with k = 4 monotonic sections is obtained, where the turning points occur at positions 493 (peak), 1082 (trough) and 1978 (peak) as indicated by the times signs in the column labeled "[29.7, 44.3]". However, when the value of  $\eta$  was further reduced by 0.1, a different piecewise monotonic fit with k = 10 monotonic sections was obtained, where the turning points occur at positions 493 (peak), 561 (trough), 609 (peak), 681 (trough), 781 (peak), 1082 (trough), 1573 (peak), 1829 (trough), 1978 (peak) as again indicated by the times signs in the column labeled "[21.9, 29.6]". The sum of squares of residuals and the maximum absolute residual of the fits associated with these columns is added at the bottom of Table I and visualized in Fig. 5 to further evaluate the importance of the non captured peaks of the resultant fit. For example, when  $\eta \in [29.7, 44.3]$  these quantities are equal to  $\gamma(k = 4, n) = 0.65 \times 10^6$  and  $D = 1.01 \times 10^2$  respectively.

The best approximation with k = 4 is obtained for  $\eta \in [29.7, 44.3]$  and its corresponding peaks are illustrated in Fig. 1. The plot shows the two most important peaks across all the data points that were identified by the method. Here, we assume that there is some differentiation in the importance of the peaks, which can be distinguished by picking an alternative value of  $\eta$  within the range [44.4, 48.6] (see, Table I).

The next three figures show the results of the next steps of the process with k = 10, 15 and 26 corresponding to  $\eta \in [21.9, 29.6], [20.8, 21.8]$  and [16.1, 20.7] respectively, in order to give more emphasis to sets of turning points that imply fits with less important monotonic trends. The piecewise monotonic approximation with k = 10 was calculated giving nine turning points and having sum of squares of residuals equal to  $0.13 \times 10^6$  as we see in Table I. Fig. 2 displays the data and the fit. The new fit maintained the peaks of the fit presented in Fig. 1, while three extra peaks were automatically detected in-between the old ones and highlighted in Fig. 2. It can be seen that the importance of the new peaks is relatively lower compared to the two peaks identified in the previous step (Fig. 1). The coordinates of the turning points can be seen in Table I.

The piecewise monotonic approximation with k = 15 gave five extra turning points which enhance both the left hand side part and the right hand side part of the fit that occurs in

# Proceedings of the World Congress on Engineering 2021 WCE 2021, July 7-9, 2021, London, U.K.

[44.4, 48.6][16.1, 20.7] [13.8, 16.0][12.2, 13.7][11.1, 12.1] [9.0, 11.0][29.7]44.3[21.9, 29.6][20.8, 21.8] $\eta \in k =$ 2 10 1526 30 32 36 48 φ.  $\mathbf{4}$  $\frac{w_{l_j}}{142.26}$  $\frac{\varphi \iota_{j}}{0.00}$ × × × × × 180.31 82 33.47× × × × × 131 203.940.22× × × × 2 3 4 5 6 7 8 9 × 202 238.1736.46× × × × 279  $275\ 29$ 0.01 × × × × 346 307.59 90.76 × × × 318.20368 57.88× 393 330.2584.78 × 402 334.59 63.75 433 349.54142.4× × × ××× 10 455 360.14113.38 × × 11 493 378.46271.16× × × × 12 535 398.71 × 0.53542 402.09 19.09 ×× 13 14 15 561 411.250.04 × × 609 434.39 159.17 × × × × × × × × × 16 17 632 445 48 64 47 × × × × × × × ×× 655 105.24 456.57 18 19 681 469.10 0.11 × × × 101.44 781 517.31× × × × × × × 838 544.79 20 21 22 23 24 25 26 27 28 28 30 31 32 33 34 35 36 37 38 0.14× × ×× 912 580.4718.24951 599.276.42× × × 966 606.50 21.94× × × 662.43 1082 0.19 × × × × × 1222 729.93 18.29×× 1319 776.69 0.241352 792.60 12.87 ×× 1404 817.67 0.04 × × × 1458 21.82 ×××× 843.71 × 1496 862.03 899.15 2.441573 65.39 × × 1684 952.66 0.15× × 1716 968.09 30.89 × × × × × × 1753 985.932.6422.521776 997.02 1829 1022.57 0.01× × × × × × × × 1059.21 1905 1917  $61.55 \\ 37.15$ 1065.00× 39 40 41 1978 1094.41 327.15 × × × × × × ××× × × 2003 1106.46166.14× × × × 2028 1118.51 246.08 × × × × 2075 1141.17 65.46× 42 43 44 45 × 2095 1150.8188.30 × × × 2188 1195.650.22 × 2201 12.741201.92× 46 1219.270.02 × × × Х × × × × 47 2295 1247.2433.53 × × × 48 2374 0.00 1285.32  $\gamma(k, n)$ 0.317  $0.65_{6}$  $0.13_{6}$ 0.665  $0.22_{5}$  $0.15_{5}$  $0.13_{5}$  $1.00_{4}$  $0.51_{4}$ D2.39 $1.01_{2}$  $4.23_{1}$  $2.40_{1}$  $1.76_{1}$  $1.35_{1}$ 1.351.09 $7.76_{0}$ 

TABLE ILEFT FOUR COLUMNS: TURNING POINTS IN THE ARTESUNATE SPECTRUM BY A BEST FIT WITH k = 48 Monotonic Sections. Right NineCOLUMNS: THE TURNING POINT POSITIONS OF THE OPTIMAL FIT, FOR  $\eta$  Ranging in a Way That Gives  $k \in \{2, 4, 10, 15, 26, 30, 32, 36, 48\}$ ,<br/>ARE INDICATED BY THE TIMES SIGN

Fig. 2 with two extra peaks. The fit with k = 15 is presented in Fig. 3, the additional peaks are highlighted and the sum of squares of residuals is equal to  $0.65 \times 10^5$ . As expected, the importance of the two extra peaks is relatively lower compared to the peaks identified in the previous step (Fig. 2).

One more run with k = 26 gave 11 extra turning points that enhanced the fit of Fig. 3 with six more peaks, as we see in Fig. 4. Now the sum of squares of residuals is equal to  $0.22 \times 10^5$ . Again, the significance of the newly identified peaks seems to be even less important compared to the significance of the peaks that were detected in the previous step (Fig. 3).

Clearly a visual comparison of Figs. 1 to 4 reveals the differences of the final fits to the given Raman spectrum with respect to the ranges of values of  $\eta$ . It is noticeable that for a specific range of values of  $\eta$  (see, for instance Fig. 1 where  $\eta \in [29.7, 44.3]$ ) the method succeeds at identifying ranges of data with peaks that have about the same magnitude. On the other hand, different ranges of values of  $\eta$  imply different number of monotonic sections for the resultant fit, urging the user to derive as an appropriate value of k the one that reveals a new set of peaks that are significantly different than those

in the previous attempt. As the values of  $\eta$  are reduced, the method detects subtle monotonic trends in the data, which are not detected in the previous steps.

We see in Table I that the sum of squares of residuals decreased from  $0.31 \times 10^7$  down to  $0.51 \times 10^4$  as we move from one interval of values of  $\eta$  to the next one, which in turn increased the value of k from k = 2 to k = 48. Analogously, we see a gradual reduction in the values of the maximum absolute residual as k increased due to changing the value of  $\eta$ , which indicates that the best fit comes closer to the data. The sum of squares of residual values and the maximum absolute residual values across k are both displayed in Fig. 5 giving one vertical axis of values for each case.

As already mentioned, adjusting the value of the parameter  $\eta$  within a range of values, provides control to the method with respect to the change of k in order to gradually identify peaks of relatively similar importance. However, by decreasing  $\eta$  and thus increasing k, piecewise monotonic approximation has the freedom to add more turning points, making the sum of the squares of residuals progressively smaller. As can be seen in Fig 5, a piecewise monotonic approximation with more than k = 26 monotonic sections does not provide any noticable further improvement, neither to the



Fig. 1. Detected peaks (circles) by a best piecewise monotonic fit with k = 4 to 2374 data points (plus signs) of the Artesunate Raman spectrum. The solid line illustrates the best fit. The Raman shifts and the intensities are given in the x-axis and the y-axis, respectively.



Fig. 2. As in Fig. 1, but k = 10. The extra peaks as compared to Fig. 1 are indicated by circles.

sum of squares of residuals nor to the maximum absolute residual. Therefore, a piecewise monotonic approximation with k = 26 monotonic sections can be considered as an adequate choice that has revealed the peaks and in-between trends that seem to have real significance.

#### IV. CONCLUSION

Piecewise monotonic approximation provides the optimal solution in only  $O(kn^2)$  computer operations to a challenging combinatorial problem that requires k monotonic sections to n data. We have given attention to an extension of the method, where the residuals are tested for trends in order to obtain automatically an adequate value of k.

We applied the method to the peak estimation problem of a Raman spectrum. The data set was chosen, because the complexity of the underlying physical laws and its variability makes it a hard test for our method. We considered the behaviour of the residual mean square by decreasing systematically the value of a parameter  $\eta$ , which resulted to

ISBN: 978-988-14049-2-3 ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online) increasing values of k. The main advantage of this approach is that the user has a direct control on the accuracy of the approximation.

We saw that the mean square error decreased consistently at first and finally levelled off to a fairly constant value, which made clear where to stop. The example of the this paper drew attention to interesting questions on the test for trends that deserves further study. The given calculations have helped us to comprehend the behavior of the parameter  $\eta$  on a difficult data set and gives promise for future research.

## REFERENCES

- I. C. Demetriou, "Algorithm 863: L2WPMA, a Fortran 77 package for weighted least-squares piecewise monotonic data approximation," *ACM Trans. Math. Softw.*, Vol. 33, No 1, pp. 1–19, 2007.
- [2] I. C. Demetriou, "Peak estimation of a spectrum from noisy measurements by least squares piecewise monotonic data approximation," *Lecture Notes in Engineering and Computer Science: Proceedings* of The International Multiconference of Engineers and Computer Scientists 2018 (editors: S. I. Ao, O. Castillo, C. Douglas, D. D. Feng, A. M. Korsunsky), 14–16 March, 2018, Hong Kong, pp. 51–56, 2018.



Fig. 3. As in Fig. 1, but k = 15. The extra peaks as compared to Fig. 2 are indicated by circles.



Fig. 4. As in Fig. 1, but k = 26. The extra peaks as compared to Fig. 3 are indicated by circles.



Fig. 5. Sum of squares of residuals (solid line - primary axis) and the maximum absolute residual (dot line - secondary axis) across k.

[3] I. C. Demetriou, "A decomposition theorem for the least squares piecewise monotonic data approximation problem," *Approximation* 

and Optimization - Algorithms, Complexity and Applications (editors: Ioannis C. Demetriou and Panos M. Pardalos), Springer Optimization and Its Applications, Vol. 145, Springer Nature, Switzerland, pp. 119–134, 2019.

- [4] I. C. Demetriou and I. N. Perdikas, "Peak estimation of univariate spectra by the best L1 piecewise monotonic approximation method," *WSEAS Transactions on Systems and Control*, Vol. 15, pp. 477–489, 2020.
- [5] I. C. Demetriou and M. J. D. Powell, "Least squares smoothing of univariate data to achieve piecewise monotonicity," *IMA J. of Numerical Analysis*, Vol. 11, pp. 411–432, 1991.
- [6] H. Gunther, NMR Spectroscopy: Basic Principles, Concepts and Applications in Chemistry, 3rd ed., Chichester, U. K.: J. Wiley and Sons, 2013.
- [7] C. van Eeden, "Maximum likelihood estimation of ordered probabilities," *Indag. Math.*, Vol. 18, pp. 444–455, 1956.
- [8] M. J. D. Powell, "Curve fitting be splines in one variable," *Numerical Approximation to Functions and Data* (editor: J. G. Hayes), The Institute of Mathematics and its Applications, The Athlone Press, 1970.
- [9] RRUFF Project, Department of Geosciences, University of Arizona. Available online: https://rruff-2.geo.arizona.edu/artesunate/display= default/D120022, Database, January 2021.
- [10] E. Vassiliou and I. C. Demetriou, "An adaptive algorithm for least squares piecewise monotonic data fitting," *Computational Statistics* and Data Analysis, Vol. 49, pp. 591–609, 2005.