

# Some further considerations in powder diffraction pattern indexing with the dichotomy method

Daniel Louër,<sup>1,a)</sup> and Ali Boultif<sup>2</sup>

<sup>1</sup>Retired from Centre National de la Recherche Scientifique and Université de Rennes 1, France

<sup>2</sup>Faculté des Sciences Exactes, Département de Physique, Laboratoire de Cristallographie, Université de Constantine 1, 25000 Constantine, Algeria

(Received 28 August 2014; accepted 24 September 2014)

Some improvements have been introduced in the current computer program for powder diffraction pattern indexing using the dichotomy algorithm. The resulting version, DICVOL14, includes optimizations and extension of scanning limits for triclinic cases, a detailed review of the input data from the indexing solutions, cell centering tests and a new approach for zero-point offset evaluation. The performance of the new version is illustrated with many examples, such as triclinic cases with long axes and dominant zones. Some important parameters in pattern indexing based on the dichotomy algorithm are commented upon, e.g. the precision of data and spurious lines. © 2014 International Centre for Diffraction Data. [doi:10.1017/S0885715614000906]

Key words: powder pattern indexing, dichotomy method, DICVOL, DICVOL14, X-ray diffraction

## I. INTRODUCTION

The use of a dichotomy algorithm for indexing powder diffraction patterns was introduced by Louër and Louër (1972) with the first computer program written for orthorhombic and higher symmetry systems. Later, with the advances in computer technology, extensions to monoclinic (Louër and Vargas, 1982) and triclinic systems (Boultif and Louër, 1991, 2004) were reported. In this method, the position  $2\theta_i$  of the diffraction lines and the absolute uncertainty  $\Delta(2\theta)_i$  on their measure, namely  $2\theta_i \pm \Delta(2\theta)_i$ , are used for the search of indexing solutions. This approach presents the advantage of avoiding a propagation of errors on a random basis, provided experimental errors do not exceed the absolute uncertainty  $\Delta(2\theta)_i$ . Nowadays, more complex cases are reported in powder diffraction studies, particularly with the use of high-resolution powder diffractometers with synchrotron X-rays. It is therefore appropriate to introduce some optimization in the analysis, in particular for triclinic powder data, as well as new facilities, in the current versions of the computer program: DICVOL04 (Boultif and Louër, 2004) and its expanded version DICVOL06 (Louër and Boultif, 2007). These features are described in the present study and some aspects of importance in the use of the dichotomy algorithm are discussed. The performance of the new version, DICVOL14, is illustrated with many powder diffraction data sets.

## II. THE DICHOTOMY METHOD AND DICVOL14

Powder pattern indexing is based on the quadratic equation obtained by squaring the reciprocal lattice vectors  $\mathbf{d}_{hkl}^*$  ( $=1/d_{hkl}$ ), i.e.  $Q_{hkl} (=1/d_{hkl}^2) = f(a, b, c, \alpha, \beta, \gamma, h, k, l)$ , where  $a, b, c, \alpha, \beta$ , and  $\gamma$  are the linear and angular parameters of the unit cell and  $hkl$  the Miller indices. The dichotomy method

is based on the variation in direct space (for monoclinic and higher symmetry systems) of the lengths of cell edges and the interaxial angle of monoclinic cells by means of finite sections (0.4 Å and 5°, respectively), followed by a progressive reduction of these intervals with a dichotomy procedure. For the triclinic system a specific strategy is used in order to reduce calculation times. The algorithm is applied in  $Q$ -space to the powder constants,  $Q_A, Q_B, \dots, Q_F$  of the quadratic equation (for details, see Boultif and Louër, 1991), and the first two lines in the input data are constrained by selected  $hkl$  indices. The successive dichotomy process is applied until the intervals become very narrow. Whatever the symmetry, at each dichotomy level a pattern is calculated in the form of intervals  $[Q_-(hkl), Q_+(hkl)]$  for each line. They are compared with the observed data,  $Q_i \pm \Delta Q_i$ , and the solution is retained if the following conditions are verified:

$$Q_-(hkl) - \Delta Q_i \leq Q_i \leq Q_+(hkl) + \Delta Q_i$$

Optimizations and some heuristic knowledge are introduced to improve the convergence of the tree-type process, e.g. the search of mathematical solutions with smallest cell volumes from scanning the volume space by shells of  $400 \text{ \AA}^3$ . For the triclinic symmetry, the first volume shell is estimated from the relation between the number of diffraction lines and the volume of the triclinic cell (see Smith, 1977). The reliability of the solutions is checked with the de Wolff (1968) figure of merit (FoM), whose definition shows that the higher accuracy of data and the more complete the pattern, the larger is the FoM.

Compared to DICVOL04/06, the new version of the computer program, DICVOL14, includes: (i) an optimization of filters in the final stages of the convergence of the successive dichotomy process, which are adapted to the higher precision obtained with synchrotron X-rays; (ii) an optimization for the triclinic system of the step length for scanning the powder constants, as well as an extension of the scanning limits for cases

<sup>a)</sup>Author to whom correspondence should be addressed. Electronic mail: [daniel.louer@free.fr](mailto:daniel.louer@free.fr)

with a long axis; (iii) a new approach for zero-point offset evaluation; and (iv) a detailed review of the input data from the resulting unit cells. In addition to the visualization on the computer screen of the progress of a calculation, two output files are created. The main file contains the solution(s) found in the course of the indexing process, and the second one with the extension “ord” (already introduced, in a shorter form, in the DICVOL06 version) gives the solutions ranked according to their FoM,  $M_{20}$ . In this file, all calculated lines, together with the observed lines, are listed for the top five mathematical solutions. Further results from cell-centering tests on the final indexing output listings are shown and the probable Bravais-type lattice is derived. The review of the powder data can be used for checking that all input observed lines are indexed and a careful analysis of systematic extinctions can give space group information. The determination of the correct space group from powder data is rarely unambiguous so that, as stated by McCusker and Baerlocher (2002) “it is important to keep the uncertainty of the assumed space group in mind in subsequent steps” of a structure determination.

### III. DICHOTOMY ALGORITHM AND DATA QUALITY

#### A. Effect of the uncertainty on peak positions

The absolute uncertainty  $\Delta(2\theta)_i$  on the peak position is an important parameter in powder pattern indexing. Too large selected errors  $\Delta(2\theta)_i$  would introduce instabilities arising from more possible theoretical lines within the error interval, they would lengthen computing time and they would increase the risk of missing the correct solution. Clearly, the higher the precision, the narrower the absolute error can be set, and easier is the pattern indexing. Generally, for data collected with laboratory X-rays and  $\text{CuK}\alpha$  radiation a standard experimental error is  $\pm 0.03^\circ (2\theta)$  (default value in DICVOL14 for monoclinic and higher symmetry); however, it is desirable to reduce it for the triclinic symmetry, e.g.  $\pm 0.025^\circ (2\theta)$  (default value in DICVOL14) and, also, when large cell axes are expected or if the density of lines is high (e.g. large-cell volume, presence of spurious lines, short wavelength). For a shorter wavelength  $\lambda$ , as frequently used with synchrotron X-rays, the default value in DICVOL14 is set proportional to the ratio  $\lambda/\lambda_{\text{CuK}\alpha}$ . The default values must, however, be critically considered according to the diffraction line broadening. It is worth considering the absolute error on the individual lines with respect to their width and resolution in the diffraction pattern. The instrumental resolution function (IRF) of the diffractometer used to collect the data (see, for example, Louër and Langford, 1988) is the reference for minimum line broadenings. In practice, the measured full-width at half-maximum (FWHMs) can be used for estimating the error intervals  $\pm\Delta(2\theta)$  on peak positions. As a rule of thumb an error interval in the range  $1/2$ – $1/3$  FWHM is a reasonable (starting) estimation for separated lines. [It can be noted that for non-uniform FWHMs, e.g. in case of anisotropic line broadening, line overlap or low-counting statistics, the absolute error  $\Delta(2\theta)_i$  can be set for each individual line  $i$ .]

With the advent of high resolution with synchrotron X-rays and the use of short wavelengths, data collection with step sizes as small as  $0.001^\circ (2\theta)$  are reported. This is illustrated in Figure 1 with data collected at the 11-BM

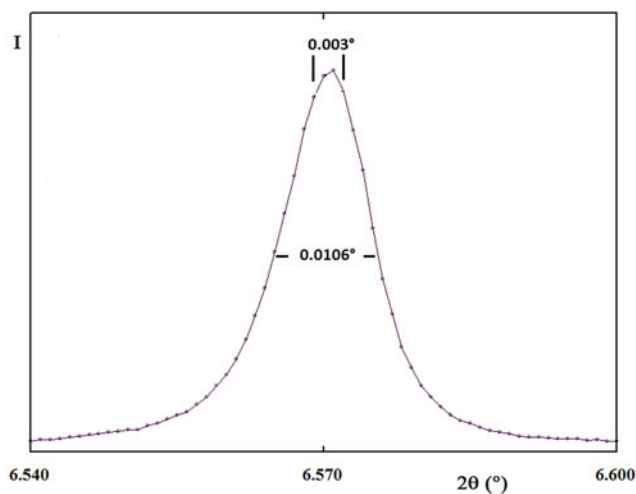


Figure 1. Example of diffraction line profile in the powder diffraction pattern of tetracaine hydrochloride collected with synchrotron X-rays,  $\lambda = 0.45882 \text{ \AA}$  (11-BM synchrotron beam line, Argonne Advanced Photon Source), step size  $0.001^\circ (2\theta)$  (Papoular *et al.*, 2011). The FWHM is  $0.0106^\circ (2\theta)$  and the  $0.003^\circ$  interval corresponds to an absolute error  $\Delta(2\theta)$  of  $0.0015^\circ (2\theta)$ .

synchrotron beam line (Argonne Advanced Photon Source) (Papoular *et al.*, 2011). It shows a line profile in the pattern of triclinic tetracaine hydrochloride ( $\lambda = 0.45882 \text{ \AA}$ ), collected with steps of  $0.001^\circ (2\theta)$ , with a FWHM =  $0.0106^\circ (2\theta)$ . The angular interval shown at the top of the line profile corresponds to  $0.003^\circ (2\theta)$ , i.e.  $\pm 1.5$  steps or  $\pm 0.0015^\circ (2\theta)$ , close to  $1/3$  FWHM. When using this small uncertainty on input peak positions the solution is obtained in 0.1 s with DICVOL14.

#### B. Error magnitude and computing time

The magnitude of the absolute error  $|\Delta(2\theta)_i|$  has an effect on the search of indexing solutions with the dichotomy algorithm and, consequently, on computing times (CPU times). This is also dependent on the lattice symmetry since the density of lines increases according to the number of unit cell parameters. It has already been shown that DICVOL91 is tolerant of imprecision (Louër, 1992). Simulated powder data of orthorhombic and monoclinic materials with a random error were used in order to assess the behavior of the program (see Figure 2 in Louër, 1992). Successful indexing was obtained with random errors as large as  $0.10^\circ (2\theta)$  and  $0.07^\circ (2\theta)$ , respectively. In the same review, the effect of zero-point errors (see Figure 3 in Louër, 1992) has also been studied. For the two examples the correct solution was found up to  $0.30^\circ (2\theta)$  and  $0.15^\circ (2\theta)$  zero-angle offset, respectively. This tolerance to random and systematic errors was also noted by Shirley (1980). Some powder data collected with laboratory and synchrotron X-rays are used to further illustrate the effect of the magnitude of the error parameter.

##### (i) Standard powder diffraction data

The powder data were collected with laboratory monochromatic X-rays ( $\lambda = 1.540596 \text{ \AA}$ ) using the Debye–Scherrer (D–S) geometry for low-absorbing materials and Bragg–Brentano (B–B) geometry for materials with heavy atoms. The absolute error  $|\Delta(2\theta)|$  on powder data input in

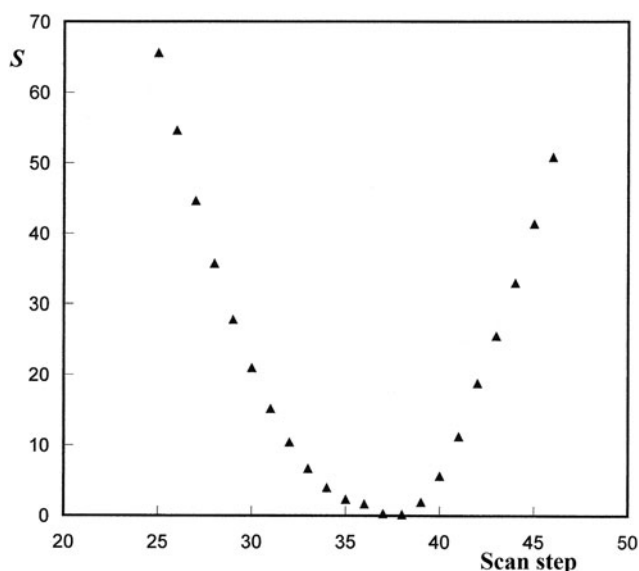


Figure 2. Search of the zero-point offset with the scanning approach based on the pairs of harmonic lines detected in the powder data set.  $S$  is the residual minimized with a least-squares analysis.

DICVOL14 was progressively increased until the program failed (the number  $N$  of lines for the indexing search was set to 20). The results are as follows:

- Tetragonal urea (D–S), Powder Diffraction File (PDF) entry 55-0142 (ICDD, 2014), only one solution is found up to  $|\Delta(2\theta)| = 0.70^\circ$ , CPU times are lower than 1 s until  $|\Delta(2\theta)| = 0.20^\circ$ .
- Orthorhombic theophylline (D–S, PDF entry 55-1653), only one solution is found up to  $|\Delta(2\theta)| = 0.19^\circ$ , CPU times are lower than 2s.
- Monoclinic barium titanate oxalate (B–B, Louër *et al.*, 1990), the maximum volume in DICVOL14 has been set to  $3000 \text{ \AA}^3$ , the solution is found up to  $|\Delta(2\theta)| = 0.15^\circ$ , CPU times are lower than 1 s until  $|\Delta(2\theta)| = 0.05^\circ$ .
- Triclinic uranium uranyl phosphate (B–B, PDF 48-1227), the solution is found up to  $|\Delta(2\theta)| = 0.06^\circ$ , CPU times are lower than 60 s until  $|\Delta(2\theta)| = 0.04^\circ$ .

It appears again that indexing with the dichotomy algorithm (DICVOL14) is generally not too sensitive to imprecision. The tolerance  $|\Delta(2\theta)|$  decreases from high to low symmetry, i.e. for the selected examples  $0.70^\circ$  ( $2\theta$ ) (tetragonal),  $0.50^\circ$  (orthorhombic),  $0.15^\circ$  (monoclinic), and  $0.06^\circ$  (triclinic). As expected, a side effect of increasing  $|\Delta(2\theta)|$  is the increase of CPU times, which remain small for high-symmetry cases but become significantly greater for low-symmetry cells, particularly the triclinic case.

Although the results reported here demonstrate the efficiency of the dichotomy algorithm, it is important to keep in mind that wide uncertainty limits on peak positions increase the risk of generating wrong solutions, particularly as soon as the number of cell parameters increases. Then, data quality (noted by minimal errors on measured Bragg angles) remains a recommended practice in powder pattern indexing. Moreover, incompleteness of the input data may be an additional reason for failure or long computing times. For triclinic symmetry examples, incompleteness in the input data is reflected in the parameter  $N_{\text{calc}}$  of the Smith and Snyder FoM

(1979) reported as “ $F_N = \text{value} [\Delta(2\theta), N_{\text{calc}}]$ ”. Each solution found by DICVOL14 is characterized by  $F_{20}$ . A high value of the number of calculated lines  $N_{\text{calc}}$ , until the 20th observed line, indicates that many lines are probably missing in the input data set, and then the solution must be further examined thoroughly, e.g. from the review of the entire data set ( $N_{\text{total}}$  lines) reported in the file “.ord”. Additionally, impurity lines may also cause indexing failure. The user must approach indexing with discernment and critical thinking when impurities may be present in a sample.

### (ii) High-resolution powder diffraction data

With synchrotron X-ray sources narrow line profiles are observed for materials with negligible intrinsic line broadening. This is illustrated in Figure 1. The precision on measured peak positions is high and the absolute error is small. An immediate consequence in pattern indexing is a significant increase of the magnitude of the FoM of solutions obtained from synchrotron data compared with laboratory data (see, for example, Cernik and Louër, 1993). The effect of the magnitude of  $|\Delta(2\theta)|$  on the indexing with DICVOL14 of the powder data of triclinic tetracaine hydrochloride [ $a = 7.4060(6) \text{ \AA}$ ,  $b = 8.5713(7) \text{ \AA}$ ,  $c = 13.710(1) \text{ \AA}$ ,  $\alpha = 106.252(6)^\circ$ ,  $\beta = 90.767(7)^\circ$ ,  $\gamma = 98.705(9)^\circ$ ,  $V = 824.49(11) \text{ \AA}^3$ ,  $M_{20} = 204$ ,  $F_{20} = 2073(0.0004, 22)$ ] can be summarized as follows: the solution is obtained for errors in the range  $0.0015^\circ$ – $0.015^\circ$  ( $2\theta$ ), with CPU time varying from 0.07 to 15 s, respectively. Only the correct solution ( $M_{20} \approx 200$ ) is found up to an error of  $0.005^\circ$  ( $2\theta$ ). With greater errors the correct solution is still found, but additional cells with lower  $M_{20}$  values are also proposed.

### C. Zero-point offset evaluation from a scanning approach

An optional analysis of the zero-angle offset in the input data was proposed in DICVOL04. Its interest was essentially for uncalibrated data with significant zero shift. Indeed, for small shifts [e.g.  $\lesssim 0.05^\circ$  ( $2\theta$ )], the extra degree of freedom (for the mixed error parameter, including the true zero-point shift and specimen-surface displacement in B–B optics) in the refinement process of the lattice parameters is generally efficient enough to solve the problem. A benefit of refining the “zero-shift” is an increase of the FoMs (see, Louër and Boulton, 2006).

For greater zero-point shifts the method used in DICVOL04 was based on the reflection-pair method reported by Dong *et al.* (1999). In the context of pattern indexing, some ambiguous results were often found with DICVOL04, leading to long CPU times. In the new version DICVOL14, a scanning approach has been preferred. The procedure is also based on the use of higher orders of reflections, if they can be identified unambiguously. In the present approach, the input  $2\theta_i$  positions are artificially shifted by a zero-point varying in an interval  $[-0.30^\circ, +0.30^\circ$  ( $2\theta$ )] with a scan step width of  $0.005^\circ$ . For each step, the effect on the different line orders is calculated through the quantity  $\Delta m$  ( $=m-d_i/d_j$ , where  $d_i$  and  $d_j$  are the  $d$ -spacings of two line orders for a particular step, and  $m$  is the integer corresponding to the ratio of the  $d$ -spacings). All identified pairs of lines with an integer ratio are considered. The zero-point is then estimated from a least-squares analysis,

by minimizing the residual  $S [= \Sigma(\Delta m/m)^2]$ . The sum is taken over pairs of harmonics. As shown in Figure 2, for a particular example, the minimum of the residual is located between steps 37 and 38, i.e. the zero-shift  $[(37 + 38)/2 \times \text{“step width”}]$  is  $0.1875^\circ (2\theta)$ .

The scanning approach for zero-point shift estimation is stable and a unique solution is obtained, provided unambiguous higher orders of lines at low angles are identified. The method has been tested successfully on many experimental powder data sets. In particular, all 30 triclinic powder data reported in the whole NBS monograph 25 (1963–1985) were artificially degraded by a zero shift in the range  $0.15^\circ$ – $0.30^\circ (2\theta)$ . These data were subsequently analyzed with DICVOL14. Twenty shifted patterns were correctly indexed.

The powder data of acetylcysteine reported in the PDF entry 46-1988 (ICDD, 2014), with the quality “i”, are characterized by a low FoM  $F_{30} = 20(0.031, 49)$ . They can be used to illustrate the usefulness of the scanning method implemented in DICVOL14. The unit cell is triclinic ( $a = 5.897 \text{ \AA}$ ,  $b = 6.548 \text{ \AA}$ ,  $c = 5.108 \text{ \AA}$ ,  $\alpha = 103.93^\circ$ ,  $\beta = 102.19^\circ$ ,  $\gamma = 96.34^\circ$ ,  $V = 184 \text{ \AA}^3$ ). With  $0.03^\circ (2\theta)$  as absolute error on peak positions, an unsatisfactory solution ( $V = 477 \text{ \AA}^3$ ), with a poor FoM [ $F_{20} = 17(0.018, 65)$ ], is found in the second volume domain (352 s). If the parameter for searching a zero-point offset is activated, a zero-point error  $-0.125^\circ (2\theta)$  is suggested and the correct solution [ $M_{20} = 33$ ,  $F_{20} = 51(0.016, 25)$ ] is found in 0.3 s.

#### IV. TESTING DICVOL14

The compilation of the FORTRAN program (MS-DOS version) and all applications of DICVOL14 have been performed with a Dell XPS 15 laptop equipped with an Intel processor 2.4 GHz with Turbo Boost up to 3 GHz. The executable program can be downloaded from the web site (<http://www.cdifx.univ-rennes1.fr/progs/dicvol14/DICVOL14.zip>).

##### A. General tests

DICVOL14 has been tested from powder data reported in the NBS Monograph 25 (Natl Bur. Stand., 1963–1985), namely 70 experimental and one calculated patterns reported in section 20 (Natl Bur. Stand., 1984) [absolute uncertainty  $\Delta(2\theta) = 0.03^\circ$ ], as well as the 30 patterns of triclinic compounds (including seven calculated patterns) reported in the

whole NBS Monograph 25 [absolute uncertainty  $\Delta(2\theta) = 0.025^\circ$ ]. [For the powder data of  $\text{Li}_4\text{Th}_7(\text{MoO}_4)_{16}$  and  $\text{Li}_4\text{Th}_7(\text{WO}_4)_{16}$  the parameter “VOLMAX” was set to 3500  $\text{Å}^3$ .] The patterns were all correctly indexed, except for the data of triclinic  $\text{KIO}_3$  (Section 15,  $a = 7.708 \text{ \AA}$ ,  $b = 7.222 \text{ \AA}$ ,  $c = 7.689 \text{ \AA}$ ,  $\alpha = 109.25^\circ$ ,  $\beta = 108.96^\circ$ ,  $\gamma = 109.37^\circ$ ,  $V = 355.9 \text{ \AA}^3$ ). This phase is described as a “pseudo-monoclinic perovskite-type” and only a triclinic cell with half the correct volume was obtained with DICVOL14. For the 100 indexed patterns, the average computing times were less than 1 s for orthorhombic and higher-symmetry data sets and 4 s for monoclinic cases, except for one data set ( $\text{Eu}_2\text{O}_3$ ) indexed in 80 s. Thirteen powder data sets of triclinic compounds were indexed in less than 5 s, eight in the range 5–30 s, six in the range 30–160 s, and the last two data sets in 360 and 690 s. The program has also been successfully tested with experimental powder data using the density and formula weight of the material as input information. As shown by Louër and Vargas (1982), substantial reduction of computing times are obtained when the density information is used, since only limited volume domains of the direct parameter-space are explored. The powder data of pharmaceutical materials listed in Boultif and Louër (2004; Table 2) were also indexed.

##### B. Triclinic examples

To improve further the triclinic section of DICVOL14, additional powder diffraction data of materials with a triclinic cell from the PDF, with quality “C”, “\*”, and “i”, have been used to check the behavior of the program. Among them, 40 examples with peculiarities were selected according to the presence of a dominant zone (several successive first lines with a common zero Miller index, No. “0” in Table I), to long edges with several higher orders of the first line as first lines ( $N_{\text{harm}}$  in Table I) of the data set, or having close linear parameters. They were successfully indexed; some representative examples are listed in Table I. From the input data the presence of harmonic lines as first lines in the pattern ( $N_{\text{harm}}$ ) can be easily detected, and it is the indication of a long axis with respect to the two others. In practice, with such experimental powder data, it is necessary to act on some input control parameters to have a chance to identify the correct solution, i.e. the number of lines  $N (>20)$  used for searching solutions, the absolute error  $\Delta(2\theta)$  and the

TABLE I. Examples of triclinic powder data with a dominant zone or long axis indexed with DICVOL14. Data quality (PDF, ICDD, 2014): “C” for calculated, “\*” for precise data, and “i” for indexed data. The linear cell parameters are given, No. “0” is the number of successive first lines with a common zero index,  $N_{\text{harm}}$  is the number of successive harmonic lines at lowest angles,  $N$  is the number of lines used for the indexing search,  $N_{\text{imp}}$  is the number of allowed impurity lines,  $N_{\text{calc}}$  is the number of calculated lines up to the 20th observed line,  $\Delta(2\theta)$  the input absolute error in DICVOL14 (i.e.  $0.002^\circ$  for calculated patterns and  $0.025^\circ$  for experimental patterns), and CPU is the computing time.

PDF entry	Cell edges (Å)	$V (\text{Å}^3)$	No. “0”	$N_{\text{harm}}$	$N$	$N_{\text{imp}}$	$N_{\text{calc}}$	$\Delta(2\theta) (^\circ)$	CPU time (s)
89-4351 “C”	4.2/20.9/20.9	1839	19	0	35		40	0.002	93
85-1301 “C”	7.1/5.4/40.6	1553	16	5	30		31	0.002	0.5
78-1721 “C”	5.4/7.1/26.6	1011	9	3	30		27	0.002	0.2
72-1930 “C”	16.6/17.7/4.0	1150	23	0	35		23	0.002	1.5
53-1564 “*”	5.8/13.6/17.0	1281	6	0	25		50	0.025	2884
53-1533 “*”	4.3/4.8/28.3	558	10	5	25	1	67	0.025	505
52-2154 “*”	5.8/8.6/14.9	615	5	0	20		46	0.025	27
51-2118 “*”	14.9/7.1/4.3	440	8	2	25		27	0.025	92
47-2269 “i”	9.8/11.6/20.2	2229	4	0	20		48	0.025	81

TABLE II. Binary powder data sets.  $V$  is the cell volume given for both compounds,  $\Delta(2\theta)$  is the absolute error,  $N$  is the number of observed lines used for indexing the phase with the highest number of lines, and  $N_{\text{imp}}$  is the number of allowed unindexed lines. Lattices are denoted as follows: (T) tetragonal, (H) hexagonal, (O) orthorhombic, (M) monoclinic, and (Tr) triclinic. The lattice in bold characters corresponds to the solution found by DICVOL14.

Binary patterns	$V$ ( $\text{\AA}^3$ )	$\Delta(2\theta)$ ( $^\circ$ )	$N$	$N_{\text{imp}}$	CPU time (s)
MgSiO <sub>3</sub> <sup>a</sup> (M) + Fe(NbO <sub>3</sub> ) <sub>2</sub> <sup>a</sup> (O)	416/413	0.01	35	15	605
BaTi <sub>2</sub> O <sub>7</sub> <sup>a</sup> (M) + BaTiO <sub>3</sub> <sup>a</sup> (H)	610/396	0.01	35	15	1650
SrAl <sub>2</sub> O <sub>4</sub> <sup>a</sup> (M) + KSr <sub>2</sub> (NbO <sub>3</sub> ) <sub>3</sub> <sup>a</sup> (T)	384/612	0.01	35	15	135
Zn <sub>3</sub> (VO <sub>4</sub> ) <sub>2</sub> <sup>a</sup> (O) + YbF <sub>3</sub> <sup>a</sup> (O)	585/187	0.01	35	15	9
CsH <sub>5</sub> (PO <sub>4</sub> ) <sub>2</sub> <sup>a</sup> (M) + CrWO <sub>4</sub> <sup>a</sup> (M)	801/251	0.01	35	15	410
Na <sub>4</sub> P <sub>4</sub> O <sub>12</sub> ·4H <sub>2</sub> O <sup>b</sup> (Tr) + ZrTiO <sub>4</sub> <sup>a</sup> (O)	370/133	0.01	35	15	2
SrZr <sub>4</sub> (PO <sub>4</sub> ) <sub>6</sub> <sup>a</sup> (H) + KSr <sub>2</sub> (NbO <sub>3</sub> ) <sub>3</sub> <sup>a</sup> (T)	1531/612	0.015	35	15	1
$\alpha$ -ZnMoO <sub>4</sub> <sup>b</sup> (Tr) + CsH <sub>5</sub> (PO <sub>4</sub> ) <sub>2</sub> <sup>a</sup> (M)	520/801	0.01	35	15	4140
C <sub>20</sub> H <sub>26</sub> ClNO <sub>3</sub> <sup>b</sup> (Tr) + Cd <sub>2</sub> P <sub>2</sub> O <sub>7</sub> <sup>b</sup> (Tr)	986/266	0.01	35	15	2
Ce(NO <sub>3</sub> ) <sub>3</sub> ·6H <sub>2</sub> O <sup>b</sup> (Tr) + MgSiO <sub>3</sub> <sup>a</sup> (M)	604/416	0.01	35	15	6890
Piracetam I <sup>c</sup> (M) + Piracetam II <sup>c</sup> (Tr)	726/347	0.01	40	20	640
Piracetam III <sup>c</sup> (M) + Piracetam II <sup>c</sup> (Tr)	684/347	0.01	35	15	100

<sup>a</sup>Natl Bur. Stand. (1984).

<sup>b</sup>Natl Bur. Stand. (1963–1985).

<sup>c</sup>ICDD (2014).

tolerance for spurious lines  $N_{\text{imp}}$  (which include impurity lines and measured line positions with an error exceeding the input absolute error).

The powder data of *n*-docosane (PDF entry 53–1533 “\*”) can be used for illustration. The parameters of the reduced cell are  $a = 4.28 \text{ \AA}$ ,  $b = 4.82 \text{ \AA}$ ,  $c = 28.26 \text{ \AA}$ ,  $\alpha = 86.41^\circ$ ,  $\beta = 85.95^\circ$ ,  $\gamma = 73.61^\circ$ , and  $V = 557 \text{ \AA}^3$ . The first five lines are  $00l$  lines ( $l = 1-5$ ), consequently the value for  $N$  was increased to 25. With the default value for the absolute error and no spurious lines [ $\Delta(2\theta) = 0.025^\circ$ ,  $N_{\text{imp}} = 0$ ], a first calculation (81 s) with DICVOL14 is unsuccessful. A second run with a tolerance for one spurious line [ $\Delta(2\theta) = 0.025^\circ$ ,  $N_{\text{imp}} = 1$ ] gives the correct reduced cell [ $M_{20} = 14$  and  $F_{20} = 22(0.014, 67)$ ] in 510 s. The value of  $N_{\text{calc}} = 67$  indicates that many possible lines are missing from the input data and the indexing listing shows that one line among the 25 lines used for indexing has a discrepancy ( $2\theta_{\text{calc}} - 2\theta_{\text{obs}} = 0.042^\circ 2\theta$ ).

### C. Impurity lines, binary patterns

Powder pattern indexing is a demanding application of the powder method. This is a consequence of the ill posed mathematical problem as noted in this statement by Vand and Johnson (1968) “we cannot solve the indexing problem by the method of ordinary algebra, as no matter how many (quadratic) equations we take, we have always more unknowns than knowns”. The dichotomy algorithm is a means to circumvent this problem. Generally, data quality must be as high as possible and it is desirable to avoid impurity lines, because they introduce additional unknown parameters and methodological instabilities. However, for a few two-phase patterns, some successful results have been reported with DICVOL04 (Lou er and Boultif, 2006). A sample of 15 powder data sets of binary-mixture patterns have been used to test the efficiency of the parameter “number of spurious lines” in DICVOL14. Data were taken from the NBS Monograph 25 (1963–1985) and from the PDF (ICDD, 2014) for the piracetam polymorphs. The individual data sets were added to simulate mixing and DICVOL14 was used for searching an indexing solution. Clearly, only the solution of the phase with higher density of lines could be expected. A small error  $\Delta(2\theta)$

(e.g.,  $0.01^\circ 2\theta$ ) was used to reduce instabilities and to limit the number of mathematical solutions. For 12 data sets, a correct indexing solution was obtained (Table II). For two binary patterns of triclinic/monoclinic and triclinic/triclinic phases, many solutions were proposed; however, the correct solution was ranked at the fourth position over 32 and first over 37 solutions, respectively. For one mixture of patterns of triclinic phases the program failed because the first line belonged to one phase and the second line to the second phase (which contradicts the strategy used for the triclinic symmetry).

The use of DICVOL14 with binary mixture data sets must be considered with caution. Only in some particular circumstances, for example in the course of solid-state syntheses or mixture of polymorphs, such application could be helpful to identify the content of the pattern. This is illustrated with the pharmaceutical compound piracetam. In the course of the structure determination study of form I from powder diffraction data (Lou er *et al.*, 1995) and stability study of piracetam polymorphs (C olin *et al.*, 1996), mixtures of polymorphs II and III, on one hand, and I and II, on the other hand, were observed. Powder data of these two binary mixtures were simulated from the powder data of the pure phases collected with a D–S setup, namely form I (monoclinic, PDF entry 54-2432), form II (triclinic, entry 54-1783) and form III (monoclinic, entry 54-1784). The monoclinic cells of form I and form III in mixtures I/II and III/II, respectively, were found with DICVOL14 (Table II). As already noted, this application of DICVOL14 remains marginal. Obviously, the main objective of DICVOL14 is to index single-phase data. It is always desirable to identify spurious lines, for example with the help of search/match methods for phase identification (for a review, see Langford and Lou er, 1996) or other appropriate technique, e.g. temperature- or time-dependent X-ray powder diffraction in the cases of Advil tablet (Boultif and Lou er, 2004) and piracetam (C olin *et al.*, 1996).

### V. CONCLUSION

The new version DICVOL14 of the program for indexing powder diffraction patterns with the dichotomy algorithm is the successor of DICVOL04. Some improvements and

optimizations have been introduced, particularly for the study of triclinic phases, generally the most difficult case to solve with the dichotomy algorithm. It must be noted that the major contribution to the program series was done in DICVOL04, which remains the base reference. With our long experience from hundreds of patterns, including many cases for which indexing opened the door to subsequent *ab initio* structure determination, it appeared useful to introduce in a new version some facilities and to comment the use of some control parameters.

## ACKNOWLEDGEMENTS

The authors thank Dr. Robert Papoular (IRAMIS/Léon Brillouin Laboratory, CEA/CEN-Saclay, France) for providing the powder diffraction data of tetracaine hydrochloride and for helpful discussions, and Dr Thierry Roisnel (Centre de Diffractométrie X, Institut des Sciences Chimiques, Université de Rennes 1, France) for his assistance for the web site.

- Boultif, A. and Louër, D. (1991). "Indexing of powder diffraction patterns for low symmetry lattices by the successive dichotomy method," *J. Appl. Crystallogr.* **24**, 987–993.
- Boultif, A. and Louër, D. (2004). "Powder pattern indexing with the dichotomy method," *J. Appl. Crystallogr.* **37**, 724–731.
- Céolin, R., Agafonov, V., Louër, D., Dzyabchenko, V. A., Toscani, S., and Cense, J. M. (1996). "Phenomenology of polymorphism, III: *p*, *T* diagram and stability of paracetam polymorphs," *J. Solid State Chem.* **122**, 186–194.
- Cernik, R. J. and Louër, D. (1993). "Indexing unit cells from synchrotron X-ray powder diffraction data," *J. Appl. Crystallogr.* **26**, 277–280.
- de Wolff, P. M. (1968). "A simplified criterion for the reliability of a powder pattern indexing," *J. Appl. Crystallogr.* **1**, 108–113.
- Dong, C., Wu, F., and Chen, H. (1999). "Correction of zero shift in powder patterns using the reflection-pair method," *J. Appl. Crystallogr.* **32**, 850–853.
- ICDD (2014). PDF-4+ 2014 (Database), edited by Soorya Kabekkodu, International Centre for Diffraction Data, Newtown Square, PA, USA.
- Langford, J. I. and Louër, D. (1996). "Powder diffraction," *Rep. Progr. Phys.* **59**, 131–234.
- Louër, D. (1992). "Automatic indexing: procedures and applications," in *Accuracy in Powder Diffraction II*, edited by E. Prince and J. K. Stalick, NIST Spec. Pub. No. 846 (US Department of Commerce, Gaithersburg, MD), pp. 92–104.
- Louër, D. and Boultif, A. (2006). "Indexing with the successive dichotomy method, DICVOL04," *Z. Kristallogr. Suppl.* **23**, 225–230.
- Louër, D. and Boultif, A. (2007). "Powder pattern indexing with the dichotomy algorithm," *Z. Kristallogr. Suppl.* **26**, 191–196.
- Louër, D. and Langford, J. I. (1988). "Peak shape and resolution in conventional diffractometry with monochromatic X-rays," *J. Appl. Crystallogr.* **21**, 430–437.
- Louër, D. and Louër, M. (1972). "Méthode d'essais et erreurs pour l'indexation automatique des diagrammes de poudre," *J. Appl. Crystallogr.* **5**, 271–275.
- Louër, D. and Vargas, R. (1982). "Indexation automatique des diagrammes de poudre par dichotomies successives," *J. Appl. Crystallogr.* **15**, 542–545.
- Louër, D., Boultif, A., Gotor, F. J., and Criado, J. M. (1990). "X-ray powder diffraction analysis of barium titanyl oxalate tetrahydrate," *Powder Diffr.* **5**, 162–164.
- Louër, D., Louër, M., Dzyabchenko, V. A., Agafonov, V., and Céolin, R. (1995). "Structure of a metastable phase of paracetam from X-ray powder diffraction using the atom–atom potential method," *Acta Crystallogr. B* **51**, 182–187.
- McCusker, L. B. and Baerlocher, Ch. (2002). "Chemical information and intuition in solving crystal structures," in *Structure Determination from Powder Data*, edited by W. I. F. David, K. Shankland, L. B. McCusker, and Ch. Baerlocher (IUCr/OUP, Oxford), pp. 307–323.
- Natl Bur. Stand. (US) (1963–1985). "Standard X-ray Diffraction Powder Patterns," Monogr. No. 25, Sections 2–21.
- Natl Bur. Stand. (US) (1984). "Standard X-ray Diffraction Powder Patterns," Monogr. No. 25, Section 20.
- Papoular, R., Toby, B. H., and Agafonov, V. N. (2011). "A winning pair to study solid pharmaceuticals: the 11-BM synchrotron beamline at the APS and the expo 2009 software," Abstract, 2011 ICDD Spring Meetings, Newtown Square (PA). *Powder Diffr.* **26**, 193.
- Shirley, R. (1980). "Data accuracy for powder indexing," in *Accuracy in Powder Diffraction*, edited by S. Block and C. R. Hubbard, NBS Spec. Pub. No. 567 (US Department of Commerce, Gaithersburg, MD), pp. 361–380.
- Smith, G. S. (1977). "Estimating the unit-cell volume from one line in a powder diffraction pattern: the triclinic case," *J. Appl. Crystallogr.* **10**, 252–255.
- Smith, G. S. and Snyder, R. L. (1979). " $F_N$ : a criterion for rating powder diffraction patterns and evaluating the reliability of powder-pattern indexing," *J. Appl. Crystallogr.* **12**, 60–55.
- Vand, V. and Johnson, G. G. (1968). "Indexing of X-ray powder patterns. Part I. The theory of the triclinic case," *Acta Crystallogr. A* **24**, 543–546.