

Direct space methods of crystal structure determination from powder diffraction applied to intermetallic compounds

Radovan ČERNÝ^{1*}

¹ *Laboratoire de Cristallographie, Université de Genève, 24 Quai Ernest-Ansermet, CH-1211 Genève 4, Switzerland*

* *Corresponding author. Tel.: + 41-22-3796450; fax: +41-22-3796108; e-mail : Radovan.Cerny@cryst.unige.ch*

Received September 19, 2007; accepted May 20, 2008; available on-line September 10, 2008

Direct space methods of structure determination from powder diffraction of non-molecular compounds (inorganics, extended solids, intermetallic compounds, etc.) are reviewed. They do not need powder pattern decomposition, and are based on a global optimization of a structural model to improve the agreement between the observed and calculated diffraction patterns. The success of the method depends very much on an appropriate modeling of the structure from building blocks. Modeling from larger building blocks improves the convergence of the global optimization algorithm by a factor of up to ten. The correctness of the building block (its rigidity, deformation, bonding distances, and ligand identity) must be examined carefully. Dynamical Occupancy Correction implemented in the direct space program FOX has shown to be useful when merging excess atoms, and even larger building blocks like coordination polyhedra. It also allows to join smaller blocks into larger ones in the case when the connectivity is not *a-priori* evident from the structural model. Available computer programs working in direct space are listed.

Structure solution / Powder diffraction / Inorganic compound / Simulated annealing / Genetic algorithm

1. Introduction: Why are powders more difficult than single crystals?

Powder diffraction using X-rays and neutrons plays a major role in the search for new materials that are not available in the form of single crystals. Moreover, most of the industrial applications of inorganic and organic compounds are in the form of polycrystalline materials (for example metal hydrides for storage and battery applications, alloys and intermetallic compounds in industry, thin films, organic compounds in pharmaceutical industry, etc.). Structure determination from powder diffraction (SDPD) is more difficult than structure determination on single crystals, because the available data are a projection of a three-dimensional diffraction pattern onto one dimension (radial distance from the reciprocal space origin), and consequently the diffraction peaks overlap. The extraction of structure factor amplitudes can be further complicated by broadening (often anisotropic) due to crystal lattice defects. Two alternative solutions of this problem exist when trying to solve a crystal structure from powder data: Either we try to improve the decomposition of the observed powder pattern into individual peaks, or we try to

model the observed pattern as a whole. Consequently the methods of SDPD can be divided into two groups according to the working space, as can be found in [1] and references therein:

- *Reciprocal space methods*: They use procedures developed for single crystal data, like direct methods or Patterson synthesis, and optimized for powder data. They need structure factor amplitudes obtained by powder pattern decomposition.

- *Direct space methods*: Different algorithms for a search in the direct space of structural parameters are used, an agreement factor between the observed and calculated powder diffraction data is evaluated, and the structural model is optimized to improve the agreement.

In this paper we review the second case as they are applied to compounds that do not contain isolated molecules (extended solids), i.e. most of inorganic compounds. Molecular compounds like organics, or hybrids like coordination compounds, are in principle treated by the same approach. However, we do not review here the details of the description of molecules by internal coordinates, use of knowledge of molecule conformation obtained by other methods, active use of organic structure databases, or energy minimization of molecular crystals.

2. Direct space methods

2.1. Some definitions (from A. Le Bail's talk on ESCA-9, Egypt 2004)

Sometimes the "direct space methods" (not to be confused with the direct methods) are called "global optimization methods" or "model building methods", and even sometimes "real space methods". "Direct space" was the definition retained in the pioneering papers. "Direct space" as opposed to "reciprocal space" has an adequate crystallographic structural sense, and should be preferred to "real space", which, opposed to "imaginary", would call to mind both parts of the diffraction factors. "Global optimization" has a large sense and designates the task of finding the absolutely best set of parameters in order to optimize an objective function, a task not at all limited to crystallography.

Under the name "direct space methods" we will not understand methods for the interpretation of electron density or Patterson maps by a search for molecular fragments, even if they work in direct space and use the global optimization algorithm as genetic algorithm. These methods still need structure amplitudes, i.e. decomposition of the powder pattern, which is avoided by the direct space methods considered here.

2.2. History

The first successful attempt to solve a crystal structure by an automatic (not manual!) localization of a building block (rigid molecule) in direct space can be seen in the program RISON [2], which was then modified for powder data as P-RISON [3]. The optimization algorithm used was a constrained least-squares refinement, which was limited to structures not larger than 10 independent atoms and resulted in only approximate atomic positions.

The authors of [4] were one step from being the first ones using a true global optimization algorithm – simulated annealing (SA) – for structure solution from powder data. However, they did not believe in the power of the method: "At present the method is not efficient enough for use in most practical problems of *ab-initio* structure determination". The authors used SA for structure prediction based on the optimization of the crystal potential energy. Hence, the first use of a global optimization algorithm (SA) in the structure solution from powder data is generally attributed to Newsam *et al.* [5], even if the structure solved in the paper was known and small (benzene). Later on the direct space methods of structure solution from powder data developed rapidly, using different algorithms like Monte Carlo (MC) search [6] and genetic algorithm (GA) [7]. An essential step forward was achieved by applying the description of structural blocks by internal coordinates like bond distances, angles and torsion angles [8], allowing a direct stereochemical interpretation and/or constraining of optimized structural parameters. Since then, the list of

programs dealing with direct space methods of structure solution from powder (but also single crystal) X-ray and/or neutron diffraction data continues to grow. For a review see Table 1 and <http://www.cristal.org/> or <http://www.ccp14.ac.uk/>.

2.3. Principles

The direct space methods are based on the location of building blocks in the elementary cell by using random or systematic moves and/or modifications of the blocks, and the comparison of the calculated and observed diffraction patterns and/or other cost functions (CF) such as crystal energy, atomic coordination, etc. (Fig. 1). Based on the "fitness" of the current structural model, decisions are taken how to improve the model. Generally said, it is a global optimization problem of a great complexity, where the algorithm must explore a hypersurface (see Fig. 1), which describes the "cost" of the model as a function of all structural parameters (see chapter 15.6 in [1]), and find its global minimum. A flow chart representing a typical implementation, like in the program FOX [16], of the global optimization approach to the crystal structure solution from powder diffraction data is given in Fig. 2. Two algorithms of global optimization have found larger application in SDPD:

2.4. Simulated annealing and parallel tempering

SA and PT algorithms are both based on MC sampling, earlier known as "statistical sampling" (for a review, see [25]). The first, and now widely used, algorithm of MC sampling is based on the Boltzmann distribution, and is known as Metropolis algorithm [26]. The MC sampling as applied in SDPD is also called Reverse Monte Carlo [27], because the system is modified by random changes under the constraint of observed data, such as the diffraction pattern. A flow chart of the Metropolis algorithm applied to SDPD is given in Fig. 3.

The modification of SA called parallel tempering algorithm (PT) was first used in SDPD by [24]. The principal advantage of the PT algorithm within SDPD as compared to the SA algorithm is its generality for any type of problem; no parameters like annealing rate, starting temperature, are required. The algorithm is also generally able to escape from local minima in the parameter space [16].

2.5. Evolutionary theory – genetic algorithm

GA form a subset of broader classes of global-optimization strategies called *population-based methods*, and *evolutionary algorithms*. The concept of GA follows the old idea of minimizing human efforts in solving difficult scientific and technical problems by learning from Nature. The genetic computation proceeds in the space of (usually binary coded) variables. It mimics the evolution of living organisms, represented by points in this space (trial solutions). In the beginning, a population of individuals (also called

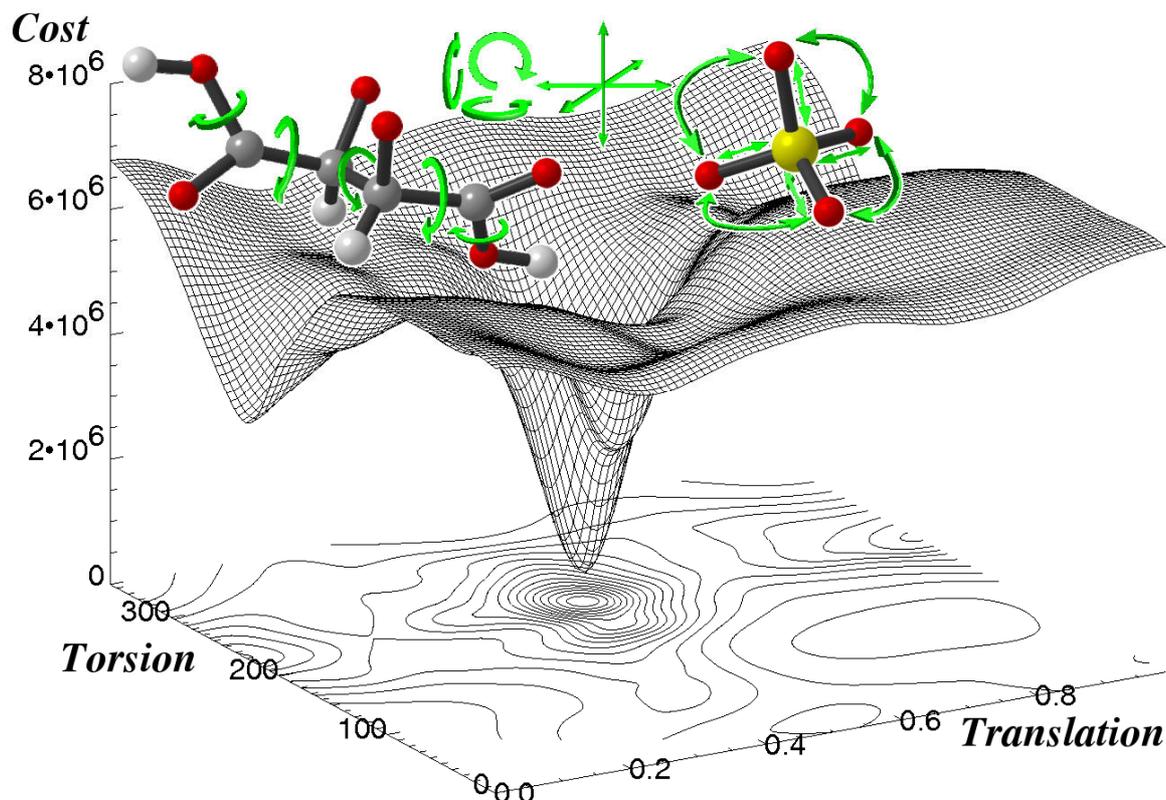


Fig. 1 Solving a structure *ab-initio* in direct space implies describing the structure through a number (N) of Degrees of Freedom (DoF): translation and rotation of the molecule or polyhedron, and internal DoF like torsion angles, bond length and bond angles. These parameters must then be randomly changed in order to find the minimal cost (usually the best agreement between the calculated and experimental powder pattern). This corresponds to exploring a N -dimensional hypersurface until the global minimum is found. The surface represented here corresponds to a 2D cut of the hypersurface corresponding to the variation of one torsion angle and one translation.

chromosomes, agents...), which may represent trial solutions of the optimization task, is generated. Next generations are successively created using simplified principles of plant or animal (Darwinian) evolution. The calculation is terminated by application of a suitable stop condition. The basic genetic operators used in the formation of each new population include selection, crossover and mutation. GA was first used for SDPD in [7] and [12]. A flow chart of the genetic algorithm applied to SDPD is given in Fig. 4.

2.6. Modeling a non-molecular structure

Inorganic samples generally require a more complex model building than molecular crystals. They can be built up from different building blocks such as coordination polyhedra, monoatomic layers or structural sheets of finite thickness (see [28]). The easiest description is by coordination polyhedra. Once

the type of polyhedra present is known from literature or experience, the choice of the number of each polyhedron to use must be made, taking into account (i) how many atoms are expected per unit cell, (ii) whether some atoms are expected to fall on special positions, (iii) whether different building blocks share some atoms.

The final choice is not trivial, in particular due to the presence of special positions. The use of a dynamical occupancy correction (DOC) [16] simplifies the problem, as it is no longer necessary to manually adjust the occupancy of atoms falling in a special position, and also makes two identical elements fully overlapping half-occupied to be “seen” by the diffraction as a single atom. DOC has proved to be very powerful in cases where the exact composition of the studied compound is *a-priori* not known exactly, e.g. for metal hydrides obtained by hydrogen

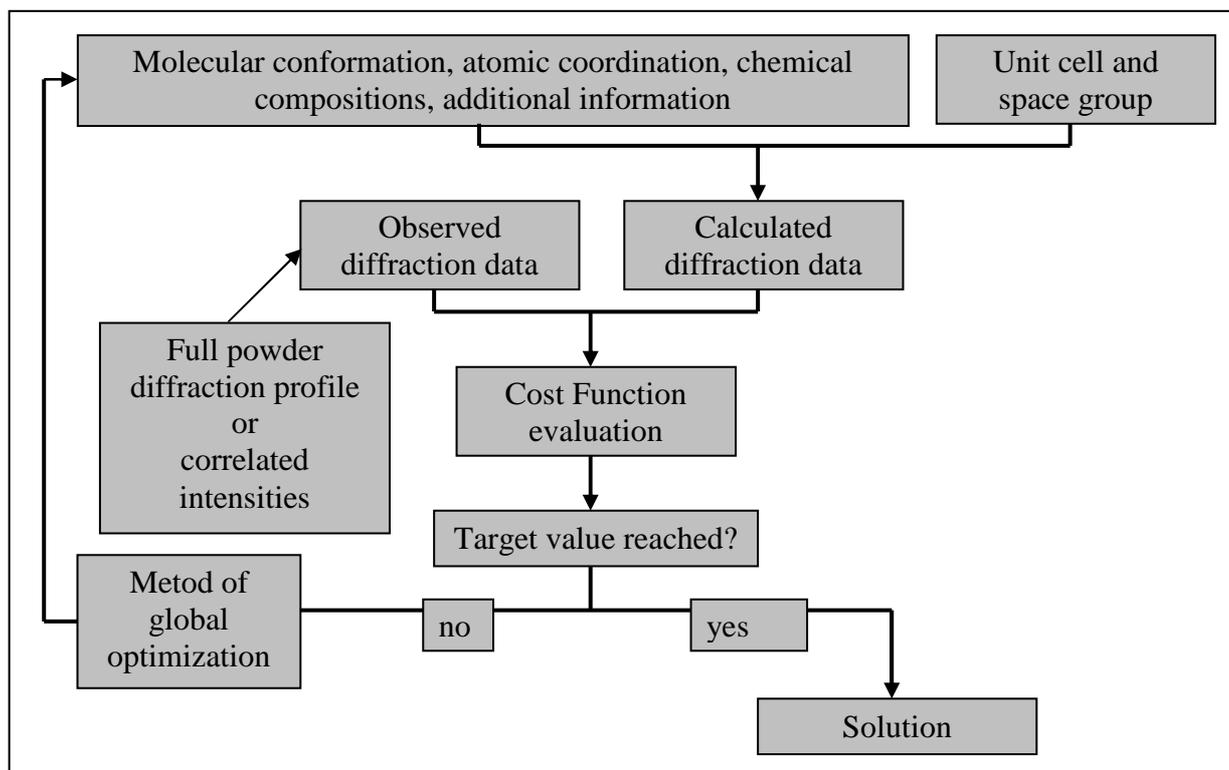
Table 1 List of available computer programs that use direct space methods for SDPD.

Program	Access	GO	CF	Reference	www
DASH	C	SA	P	[11]	www.ccdc.cam.ac.uk
EAGER (former GAPSS)	A	GA	WP	[12]	www.cardiff.ac.uk/chemy/staff/harris.html
ENDEAVOUR	C	SA	I+E	[13]	www.crystalimpact.com
ESPOIR	O	MC	L	[14]	www.cristal.org
FOCUS	O		I+TS	[15]	www.crystal.mat.ethz.ch
FOX	O	SA(PT)	WP,I,AC	[16]	objcryst.sf.net
GEST	O	GA	I	[17]	crystallography.zhenjie.googlepages.com/GEST.html
OCTOPUS	A	MC	WP	[18]	www.cardiff.ac.uk/chemy/staff/harris.html
ORGANA	A	MC(E)	I+E	[19]	
POSSUM	A	DE	WP	[20]	www.chem.bham.ac.uk/staff/tremayne.shtml
POWDERSOLVE	C	MC	WP	[21]	www.accelrys.com
PSSP	O	SA	L		powder.physics.sunysb.edu/programPSSP/pssp.html
SAFE	A	SA	WP+SE	[22]	www.crystal.mat.ethz.ch
SA	A	SA	WP	[8]	ch-www.st-andrews.ac.uk/staff/pgb/group
TOPAS	C	SA	I,WP,E	[23]	members.optusnet.com.au/~alancoelho
ZEFSAII	O	MC(B)	I+AC	[24]	www.mwdeem.rice.edu/zefsaII

Access: C = Commercial with academic prices, O = Open access, A = contact the authors

GO = Global Optimization : MC = Monte Carlo, MC(B) = Biased Monte Carlo, MC(E) = Energy guided Monte Carlo, SA = MC+Simulated Annealing, PT = MC + Parallel Tempering, GA = Genetic Algorithm, DE = Differential Evolution

CF = CostFunction : P = Pawley [9], L = Le Bail [10], I = Integrated intensities, WP = Whole Pattern, E = potential Energy, SE = Structure Envelopes, AC = Atomic Coordination, TS = Topology Search

**Fig. 2** A typical flow chart of a direct space method as applied to SDPD, case of the program FOX [16].

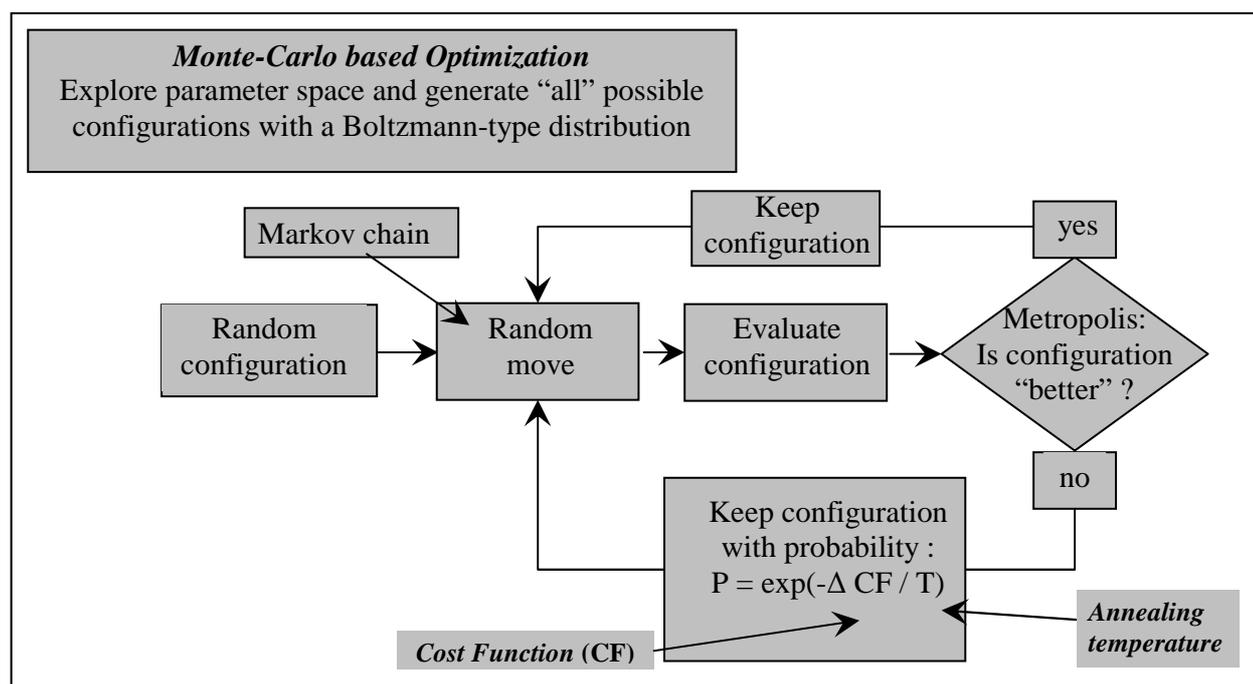


Fig. 3 Flow chart of the Metropolis algorithm [26] of simulated annealing as applied to SDPD.

absorption in a metallic matrix (for examples see [29]), or for even larger building blocks like coordination polyhedra [30].

Theoretically, this means that it is possible to add more atoms than initially deemed necessary, expecting the DOC to artificially “merge” the excess atoms. In practice, adding too many atoms will slow down the optimization, since more atoms have to fall into the correct position to find the correct structure. This also implies that if some atoms are shared between polyhedra, a choice should be made on where the bridging atoms should be used.

A nice example where careful model building helped considerably can be found in [31]; the structure solution for $\text{Al}_2(\text{CH}_3\text{PO}_3)_3$ was done using two different models: (i) $2\text{Al}+3\text{CH}_3\text{PO}_3$, and (ii) $\text{AlO}_4 + \text{AlO}_5 + \text{CH}_3\text{-P}$. Both models avoid the inclusion of oxygen atoms in both the AlO_x polyhedra and the phosphonate molecule, to keep the number of independent oxygen atoms equal to 9. However, model (i) made it possible to find the correct solution in $750 \cdot 10^3$ trials, while (ii) required $6.5 \cdot 10^6$ trials. This can be explained by the presence in model (i) of equivalent building blocks, which reduced the conformation space to search. Moreover, in model (ii) the AlO_4 and AlO_5 blocks are not only independent, but they are very similar, which can easily create false minima, which can slow down the optimization.

For more details and guidelines for the modeling of crystal structures for direct space methods see [29] and [32].

2.7. FOX: Free Objects for Xtallography

At the end of the last century the direct space methods were developing intensively in the field of molecular crystals. Significantly less activity was found in the domain of non-molecular (inorganic) crystals. However, the idea of constructing the crystal structure from well defined building blocks, like the molecules in the case of molecular crystals, can be applied also to non-molecular crystals such as extended solids or framework structures. This was the main idea behind FOX [16,33], which has become a user-friendly tool for solving not only non-molecular but also molecular structures from powder diffraction data. FOX is open-source software, released under the GNU General Public License. It can be downloaded from <http://objcryst.sf.net/Fox>. Precompiled versions are available for Windows and MacOS X.

Any crystal structure can be described in FOX as a combination of scattering objects, which can be independent atoms, molecules, polyhedra, or molecular fragments. These were originally described using Z-matrices, as for molecules, to keep a uniform description for all building blocks in FOX. The description was later changed to a restraint-based approach [33] to avoid some pitfalls of the Z-matrices, however, with less benefit for the non-molecular compounds.

Any CF used in addition to the diffraction data can be valuable to find the correct structure, either to find the global minimum, or to disfavor unsound configurations and thus reduce the overall parameter

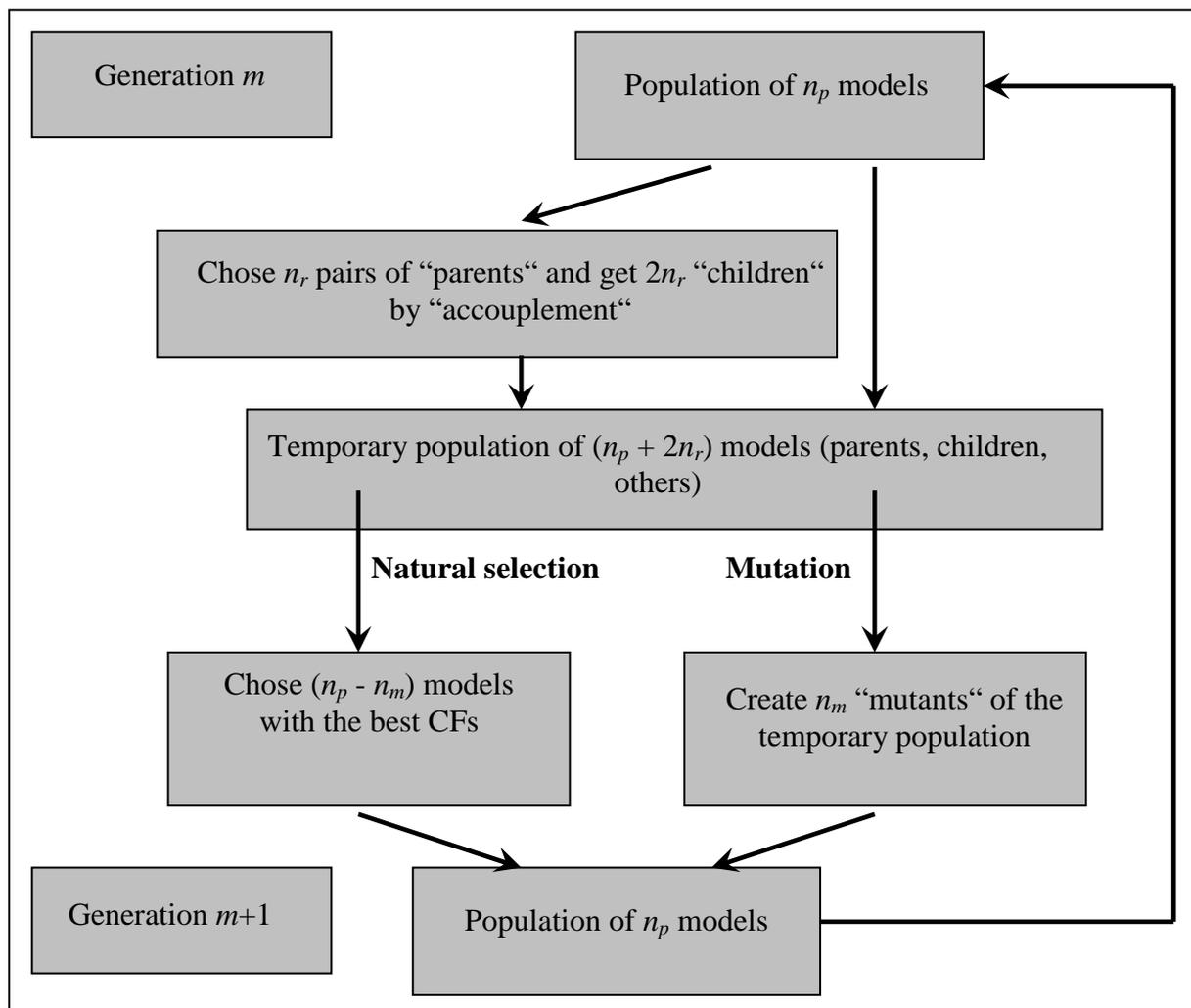


Fig. 4 A general flow chart of the genetic algorithm as applied to SDPD.

space to be sampled. Because of the non-uniqueness of the energetic description of atomic interaction in crystal structures we have preferred to implement in FOX a simple anti-bump (AB) CF that adds a penalty when two atoms are closer than a minimum distance. This minimum distance can be input by the user for each pair of atom types. For identical elements, this function also allows DOC to merge the atoms (when the distance tends toward zero), so that for identical atom types that completely overlap, the penalty decreases to zero. CF based on the **bond valence sum** [4] available in FOX seems to be a good option for testing the validity of structural models, and identifying the local minima in the parameter space.

Since its release in 2001, FOX [33] has been quite often used for solving non-molecular structures from powder diffraction data (for a review see [29]). The complexity of the structures solved by FOX ranges from 2 to 34 independent atoms found *ab-initio*. Decreasing the DoF by modeling the structure with larger building blocks was one of the reasons of using FOX for SDPD of non-molecular compounds. In

many cases tetrahedral and octahedral units were successfully used. A list of structures solved with FOX can be found on the FOX wiki <http://objcryst.sf.net/Fox/FoxBiblioStructures>.

Future developments of FOX do not depend only on its authors, but, since FOX is available as an open source program, also on any user who decides to make modifications. One of the “most wanted” features for FOX has long been the ability to index powder patterns, without having to use a separate package. This would make FOX a *complete structure solution package*, able to proceed from a raw powder pattern to a structural model. The only remaining step not handled by FOX would be a full (publication-ready) least squares refinement. Two indexing algorithms are now implemented: the first one is a new version of the successful *dichotomy-in-volume* algorithm [34] specifically written for FOX, and the second one, which is entirely new, uses a differential evolution algorithm. Both algorithms are quite fast, evaluating typically 100 000 unit cells/s for a default search (3–20 Å).

2.8. Perspectives of FOX

Deep intervention in FOX, like the introduction of profile fitting or grid computing [35], done by the users of a program dedicated originally to structure solution, illustrates what can be done from a well-documented and available open-source library. Our own activity concentrates on introducing in FOX an analysis of disordered and weakly crystallized materials by Pair Distribution Function modeling [36], which seems to be a promising way for the structural characterization of nano-materials [37]. We strongly encourage all users to contribute actively to the next evolution of the program.

3. Conclusion

Structure determination of non-molecular compounds from powder diffraction data has undergone an intensive development in the last 25 years. Reciprocal space methods have been applied and optimized to work with lower quality data obtained from powder diffraction patterns (like the programs SIRPOW-EXPO [38], DOREES-POWSIM [39] and XLENS [40]), with no important difference when applied to molecular or non-molecular crystals. Following the pioneering work by [4], and mainly by [5], the direct space method has rapidly evolved, and continues still to be developed, as a user-friendly tool for SDPD of non-molecular crystals. The main principles are the same as for molecular crystals, however, it was necessary to develop some additional tools for the treatment of special crystallographic positions, sharing of atoms between different building blocks such as coordination polyhedra, and for a correct optimization of disordered atomic positions. The current (known) limits of direct space methods are around 30-100 independent atoms. The success depends on the quality of the diffraction data, but even more on the amount of additional chemical information (knowledge about structural building blocks) injected into the structure solution process. The same is true for the reciprocal space methods. However, the use of additional information such as atomic coordination, interatomic distances, angles, is easier and natural when working in direct space, the space where this information comes from. Among the current challenges and prospects of SDPD one can mention active modeling of preferred orientation, active evolution of the structural model during the optimization, improvement of the optimization algorithm, and speeding up of the calculations. All these developments may proceed towards an automatic SDPD connected with structure prediction. The actual state of knowledge, however, still requires active interaction of an experienced crystallographer.

Acknowledgements

The author wants to thank all users of FOX, and especially those who have kindly provided details of

their work when solving the crystal structures. The discussion with Yuri Andreev from the University of St. Andrews on the principles of direct space methods is highly appreciated. The review of Genetic Algorithm principle relies much on the discussion with Wojciech Paszkowicz from the Polish Academy of Sciences in Warsaw, which is highly appreciated.

References

- [1] W.I.F. David, K. Shankland, L.B. McCusker, Ch. Baerlocher, *Structure Determination from Powder Diffraction Data*, IUCr Monographs on Crystallography 13, Oxford University Press, 2002.
- [2] R. Bianchi, C.M. Gramaccioli, T. Pilati, M. Simonetta, *Acta Crystallogr. A* 37 (1981) 65-71.
- [3] N. Masciocchi, R. Bianchi, P. Cairati, G. Mezza, T. Pilati, A. Sironi, *J. Appl. Crystallogr.* 27 (1994) 426-429.
- [4] J. Pannetier, J. Bassas-Alsina, J. Rodriguez-Carvajal, V. Caignaert, *Nature* 346 (1990) 343-345.
- [5] J.W. Newsam, M.W. Deem, C.M. Freeman, Accuracy in Powder Diffraction II, *MIST Spec. Publ.* 846 (1992) 80-91.
- [6] K.D.M. Harris, M. Tremayne, P. Lightfoot, P.G. Bruce, *J. Am. Chem. Soc.* 116 (1994) 3543-3547.
- [7] K. Shankland, W.I.F. David, T. Csoka, *Z. Kristallogr.* 212 (1997) 550-552.
- [8] Y.G. Andreev, P. Lightfoot, P.G. Bruce, *J. Appl. Crystallogr.* 30 (1997) 294-305.
- [9] G.S. Pawley, *J. Appl. Crystallogr.* 14 (1981) 357-361.
- [10] A. Le Bail, H. Duroy, J.L. Fourquet, *Mater. Res. Bull.* 23 (1988) 447-452.
- [11] W.I.F. David, K. Shankland, N. Shankland, *Chem. Commun.* (1998) 931-932.
- [12] B.M. Kariuki, H. Serrano-González, R.L. Johnston, K.D.M. Harris, *Chem. Phys. Lett.* 280 (1997) 189-195.
- [13] H. Putz, J.C. Schön, M. Jansen, *J. Appl. Crystallogr.* 32 (1999) 864-870.
- [14] A. Le Bail, *Mater. Sci. Forum* 378-381 (2001) 65-70.
- [15] R.W. Grosse-Kunstleve, L.B. McCusker, Ch. Baerlocher, *J. Appl. Crystallogr.* 30 (1997) 985-995.
- [16] V. Favre-Nicolin, R. Černý, *J. Appl. Crystallogr.* 35 (2002) 734-743.
- [17] Feng Zhenjie, Dong Cheng, *J. Appl. Crystallogr.* 40 (2007) 583-588.
- [18] M. Tremayne, B.M. Kariuki, K.D.M. Harris, *Angew. Chem. Int. Ed. Engl.* 36 (1997) 770.
- [19] V. Brodski, R. Peschar, H. Schenk, *J. Appl. Crystallogr.* 38 (2005) 688-693.
- [20] C.C. Seaton, M. Tremayne, *Chem. Commun.* (2002) 880.

- [21] G.E. Engel, S. Wilke, O. König, K.D.M. Harris, F.J.J. Leusen, *J. Appl. Crystallogr.* 32 (1999) 1169-1179.
- [22] S. Brenner, L.B. McCusker, Ch. Baerlocher, *J. Appl. Crystallogr.* 35 (2002) 243-252.
- [23] A.A. Coelho, *J. Appl. Crystallogr.* 33 (2000) 899-908.
- [24] M. Falcioni, M.W. Deem, *J. Chem. Phys.* 110 (1999) 1754-1766.
- [25] M.E.J. Newman, G.T. Barkema, *Monte Carlo Methods in Statistical Physics*, Clarendon Press, Oxford, 1999.
- [26] N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, E. Teller, *J. Chem. Phys.* 21 (1953) 1087-1092.
- [27] R.L. McGreevy, L. Pusztai, *Mol. Simul.* 1 (1988) 359-367.
- [28] G. Ferraris, E. Makovicky, S. Merlino, *Crystallography of modular materials*, Oxford University Press, 2004.
- [29] R. Černý, V. Favre-Nicolin, *Powder Diffr.* 20 (2005) 359-365.
- [30] H. Sarp, R. Černý, *Eur. J. Mineral.* 17 (2005) 367-373.
- [31] M. Edgar, V.J. Carter, D.P. Tunstall, P. Grewal, V. Favre-Nicolin, P.A. Cox, P. Lightfoot, P.A. Wright, *Chem. Commun.* 8 (2002) 808-809.
- [32] R. Černý, V. Favre-Nicolin, *Z. Kristallogr.* 222 (2007) 105-113.
- [33] V. Favre-Nicolin, R. Černý, *Z. Kristallogr.* 219 (2004) 847-856.
- [34] A. Boultif, D. Louër, *J. Appl. Crystallogr.* 37 (2004) 724.
- [35] R. Černý, V. Favre-Nicolin, J. Rohlíček, M. Hušák, Z. Matěj, R. Kužel, *CPD Newsletter* 35 (2007) 16-18.
- [36] T. Egami, S.J.L. Billinge, *Underneath the Bragg-Peaks: Structural Analysis of Complex Materials*, Pergamon Press, Oxford, 2003.
- [37] P. Juhas, D.M. Cherba, P.M. Duxbury, W.F. Punch, S.J.L. Billinge, *Nature* 440 (2006) 655-658.
- [38] A. Altomare, R. Caliandro, M. Camalli, C. Cuocci, C. Giacovazzo, A.G.G. Moliternia, R. Rizzi, *J. Appl. Crystallogr.* 37 (2004) 1025-1028.
- [39] J. Jansen, R. Peschar, H. Schenk, *J. Appl. Crystallogr.* 25 (1992) 237-243.
- [40] J. Rius, *Powder Diffr.* 14 (1999) 267-273.

Proceeding of the X International Conference on Crystal Chemistry of Intermetallic Compounds, Lviv, September 17-20, 2007.