computer programs

GIDInd: an automated indexing software for grazing-incidence X-ray diffraction data

Manuel Peter Kainz,[a] Lukas Legenstein,[a] Valentin Holzer,[a] Sebastian Hofer,[a] Martin Kaltenegger,[a] Roland Resel[a] and Josef Simbrunner[b]*

[a]Institute of Solid State Physics, Graz University of Technology, Petersgasse 16, Graz, 8010, Austria, and [b]Division of Neuroradiology, Vascular and Interventional Radiology, Medical University Graz, Auenbruggerplatz 9, Graz, 8036, Austria. *Correspondence e-mail: josef.simbrunner@medunigraz.at

Grazing-incidence X-ray diffraction (GIXD) is a widely used technique for the crystallographic characterization of thin films. The identification of a specific phase or the discovery of an unknown polymorph always requires indexing of the associated diffraction pattern. However, despite the importance of this procedure, only a few approaches have been developed so far. Recently, an advanced mathematical framework for indexing of these specific diffraction patterns has been developed. Here, the successful implementation of this framework in the form of an automated indexing software, named GIDInd, is introduced. GIDInd is based on the assumption of a triclinic unit cell with six lattice constants and a distinct contact plane parallel to the substrate surface. Two approaches are chosen: (i) using only diffraction peaks of the GIXD pattern and (ii) combining the GIXD pattern with a specular diffraction peak. In the first approach the six unknown lattice parameters have to be determined by a single fitting procedure, while in the second approach two successive fitting procedures are used with three unknown parameters each. The output unit cells are reduced cells according to approved crystallographic conventions. Unit-cell solutions are additionally numerically optimized. The computational toolkit is compiled in the form of a MATLAB executable and presented within a user-friendly graphical user interface. The program is demonstrated by application on two independent examples of thin organic films.

1. Introduction

The formation of potentially new polymorphs due to the transition from bulk to thin films is a well know phenomenon (Jones et al., 2016; Gentili et al., 2019). The presence of a substrate surface during the crystallization process can induce crystal structures with new molecular packing motifs. Frequently, the crystallization at surfaces is associated with a strong preferred orientation of the crystallites, which exhibit a well defined crystallographic plane (called the contact or texture plane) parallel to the substrate surface. In many cases, no azimuthal (i.e. in-plane) order of the crystallites is observed owing to the isotropic nature of the substrate surfaces. This type of crystalline orientation is called uniplanar texture (Heffelfinger & Burton, 1960) or fibre texture (Roe & Krigbaum, 1964).

Grazing-incidence X-ray diffraction (GIXD) has become a sophisticated technique for structural characterization of thin films. The primary X-ray beam with the wavevector k_0 and the scattered X-ray beam with the wavevector k determine the scattering vector q by $q = k - k_0$. According to the Laue equation, diffraction occurs if the scattering vector q is equal to a reciprocal lattice vector g. The surface sensitivity in a
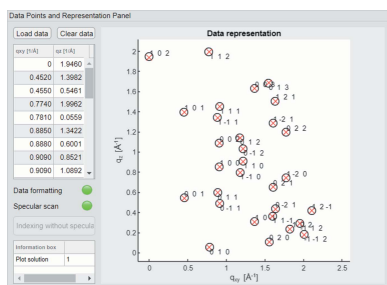
GIXD setup is high, making it particularly well suited for investigations of crystalline properties of thin films. Just as in single-crystal X-ray diffraction (XRD) and powder X-ray diffraction, the crystallographic unit cell is obtained by indexing of the diffraction data. While in case of single-crystal XRD usually all three components of the scattering vector are utilized for the indexing procedure (Pan et al., 2012; Sauter et al., 2004), powder diffraction uses only the length of the scattering vector (Boultif & Louër, 2004; Coelho, 2003). GIXD investigations provide two components of the scattering vector: the in-plane component q_{xy}, aligned parallel to the substrate surface, and the out-of-plane component q_z, perpendicular to the substrate surface. These two components are connected to the total length of the scattering vector by $q^2 \equiv q_{xyz}^2 = q_{xy}^2 + q_z^2$.

A few approaches for processing experimental GIXD data have been developed (Smilgies & Blasini, 2007; Breiby et al., 2008; Hailey et al., 2014; Jiang, 2015; Schrode et al., 2019; Savikhin et al., 2020). Not all of them address the task of indexing with regard to unit-cell determination and even fewer consider the contact plane of the crystallites and the substrate surface as separate parameters necessary to determine. For this reason, the rotation matrix of the thin-film crystallites relative to the substrate surface must be considered (Shmueli, 2006). If the lattice parameters are, however, unknown, both the lattice constants and the rotation matrix need to be determined, which is a much more challenging task.

A significant improvement of the issue is the computational tool Diffraction Pattern Calculator (DPC) (Hailey et al., 2014). It was the first toolkit that incorporated the determination of the unit-cell parameters by simultaneous assignment of the Laue indices. Prior to the unit-cell analysis, several operation parameters have to be defined by the user. Although the toolkit integrates many features, the high amount of necessary input and previous knowledge for indexing could be a limitation of the DPC.

A recent tool addressing the indexing problem of GIXD patterns is the MATLAB-based software package GIWAXS-SIIRkit (Savikhin et al., 2020). Diffraction patterns can be optionally analysed to generate data sets of q_{xy} and q_z. Similarly to our presented algorithm (Simbrunner et al., 2018, 2019), GIWAXS-SIIRkit splits the indexing process into two parts. In the first part, the q_{xy} data are preliminarily indexed by tuples of two Laue indices (hk). This is achieved by evaluation of the in-plane component of the scattering vector. Solutions obtained that way are used to assign indices to the remaining peaks. The program cycles through the index permutations within a range of −8 to 8, forms the difference for every case and thereby searches for a minimum in Δq_{xy}. In the second indexing step, the out-of-plane components of the scattering vectors are evaluated. The result of this algorithm is the unit-cell solution with the lowest total error expressed in Cartesian distances. This routine is limited to problems where the (001) plane is the plane parallel to the substrate surface. If a contact plane with any Miller indices exists, the reduced unit cell can be linearly transformed to a supercell where the (001) plane is the new contact plane.

In our previous work, we exploited a combination of GIXD data together with information acquired by specular X-ray diffraction. In specular XRD geometry, the angle between the primary beam and sample surface is the same as the angle enclosed by the scattered beam and sample surface. The presence of a specular diffraction peak indicates that a defined crystallographic plane (contact plane) of the crystallites is oriented parallel to the substrate. If the contact plane of the crystallographic unit cell with the substrate surface is considered an additional deducible quantity, a novel mathematical treatment for the indexing procedure of the GIXD data is required. Consequently, additional unknown parameters must be determined by the algorithm. A comprehensive formalism that addresses exactly this problem is proposed in recent work (Simbrunner et al., 2018, 2019, 2020). The general (triclinic) case as well as systems of higher symmetry have been dealt with.

In this article a freely available, automatic indexing software called GIDInd is presented, in which our algorithm has been implemented. We have included the option to search for either a monoclinic or a triclinic unit cell without prior knowledge of the specular diffraction peak from specular XRD. However, owing to the amount of computational work required, this method is much more time consuming and the inclusion of the specular scan is recommended.

2. Methods

2.1. Indexing formalism

With GIXD measurements, two individual components of the scattering vector **q** are accessible. Here, every point in the two-dimensional reciprocal-space map is described by a tuple of q_{xy} and q_z. These peak positions (or reflections) are the required input data for the here-proposed formalism as well as for GIDInd. After appropriate substitution, the potential lattice constants a, b, c, α, β and γ are obtained by solving sets of linear systems of equations (LSEs) under variation of the three Laue indices h, k and l. Using the GIXD data only (without the use of a specular diffraction peak at a defined q_z value and $q_{xy} = 0$), this approach is a single-step procedure with every solution yielding one potential set of lattice constants. Owing to the high number of possible combinations of the three Laue indices, this can be computationally challenging, often limited by performance and memory of the available computer resources.

A specular diffraction peak provides data in a region that is not accessible with conventional GIXD experiments. This additional peak at $q_{xy} = 0$ allows splitting the determination of the lattice constants into two consecutive parts. The presence of a specular peak reveals that the orientation of the crystallographic cell on the substrate surface can be expressed by a contact plane, which is a defined net plane of the crystal denoted by Miller indices. In this way, it is not necessary to compute and evaluate an entire set of lattice parameters for every single combination of integer numbers that constitute the Laue indices of diffraction. It is rather possible to

Table 1
Summary of the used equations for total length of the reciprocal lattice vectors g_{xyz}, the magnitude of the reciprocal lattice vector associated with the specular diffraction peak g_{spec}, and the in-plane component g_{xy} and out-of-plane component g_z of the reciprocal lattice vector **g** using the Laue indices hkl and the Miller indices (uvw) of the contact plane together with the lattice constants of the direct lattice (a, b, c, α, β, γ) and the reciprocal lattice parameters (a^*, b^*, c^*, α^*, β^*, γ^*).

$$g_{xyz}^2 = h^2\left(\frac{2\pi}{a\sin\gamma}\right)^2 + k^2\left(\frac{2\pi}{b\sin\gamma}\right)^2 - 2hk\left(\frac{2\pi}{a\sin\gamma}\right)\left(\frac{2\pi}{b\sin\gamma}\right)\cos\gamma + (ha^*\cos\beta^* + kb^*\cos\alpha^* + lc^*)^2$$

$$g_{xyz}^2 = h^2 a^{*2} + k^2 b^{*2} + l^2 c^{*2} + 2hka^*b^*\cos\gamma^* + 2hla^*c^*\cos\beta^* + 2klb^*c^*\cos\alpha^*$$

$$g_{spec}^2 = u^2\left(\frac{2\pi}{a\sin\gamma}\right)^2 + v^2\left(\frac{2\pi}{b\sin\gamma}\right)^2 - 2uv\left(\frac{2\pi}{a\sin\gamma}\right)\left(\frac{2\pi}{b\sin\gamma}\right)\cos\gamma + (ua^*\cos\beta^* + vb^*\cos\alpha^* + wc^*)^2$$

$$g_{xy}^2 g_{spec}^2 = \left(\frac{2\pi}{a\sin\gamma}\right)^2\left[u^2 g_{xy}^2 + (hg_{spec} - ug_z)^2\right] + \left(\frac{2\pi}{b\sin\gamma}\right)^2\left[v^2 g_{xy}^2 + (kg_{spec} - vg_z)^2\right]$$
$$- 2\left(\frac{2\pi}{a\sin\gamma}\right)\left(\frac{2\pi}{b\sin\gamma}\right)\cos\gamma\left[uvg_{xy}^2 + (hg_{spec} - ug_z)(kg_{spec} - vg_z)\right] - (hv - ku)^2\left(\frac{2\pi}{a\sin\gamma}\right)^2\left(\frac{2\pi}{b\sin\gamma}\right)^2\sin^2\gamma$$

$$g_z = \left[hu\left(\frac{2\pi}{a\sin\gamma}\right)^2 + kv\left(\frac{2\pi}{b\sin\gamma}\right)^2 - (hv + ku)\left(\frac{2\pi}{a\sin\gamma}\right)\left(\frac{2\pi}{b\sin\gamma}\right)\cos\gamma + (ua^*\cos\beta^* + vb^*\cos\alpha^* + wc^*)(ha^*\cos\beta^* + kb^*\cos\alpha^* + lc^*)\right]$$
$$\left/ \left[u^2\left(\frac{2\pi}{a\sin\gamma}\right)^2 + v^2\left(\frac{2\pi}{b\sin\gamma}\right)^2 - 2uv\left(\frac{2\pi}{a\sin\gamma}\right)\left(\frac{2\pi}{b\sin\gamma}\right)\cos\gamma + (ua^*\cos\beta^* + vb^*\cos\alpha^* + wc^*)^2\right]^{1/2}\right.$$

$$g_z g_{spec} = hua^{*2} + kvb^{*2} + lwc^{*2} + (hv + ku)a^*b^*\cos\gamma^* + (hw + lu)a^*c^*\cos\beta^* + (kw + lv)b^*c^*\cos\alpha^*$$

intervene at an earlier point and to apply sorting and filtering to the numerical solutions, thereby keeping the computational effort at a much lower level.

2.1.1. Formalism using a specular diffraction peak. The indexing formalism used in this first part of the work considers the contact plane [with Miller indices (uvw)] of the sample with respect to the substrate surface as a separate set of integer parameters. The consequence of such a defined crystallographic plane is the (possible) occurrence of a specular diffraction peak. By systematic variation of the integer numbers and additional incorporation of the position of a specular diffraction peak, the above-mentioned separation of the lattice parameter calculation becomes feasible. For the comprehensive analytical derivation of the mathematical framework we refer to earlier work (Simbrunner et al., 2019, 2018). Important for this algorithm is the accessibility of the individual components of the reciprocal lattice vectors. The equations for the components are adopted without further analytic derivation and are summarized in Table 1.

For every given diffraction peak, the Laue condition must be fulfilled for the total length of the scattering vector with $q_{xyz} = g_{xyz}$, and consequently for the components $q_{xy} = g_{xy}$ and $q_z = g_z$. This must also be valid for the specular diffraction peak with $q_{spec} = g_{spec}$. We can therefore relate every pair (q_{xy}, q_z) to a point in reciprocal space (g_{xy}, g_z), if the Laue indices and the unit-cell parameters are known. Conversely, the unit-cell parameters must be derivable for a given set of Laue indices and for a known point (q_{xy}, q_z).

Thus, in the first part of the indexing routine, the Miller indices of the contact plane, i.e. the integers u and v, as well as the Laue indices h and k of the experimentally observed diffraction peaks are assigned by systematically varying the integer variables and by calculating three lattice parameters a,

b and γ. In a subsequent step (i) w of the contact plane, (ii) the Laue indices l of the diffraction peaks, and (iii) the lattice constants c, α and β are determined. This is schematically shown in Fig. 1.

Our algorithm builds upon the following mathematical expression:

$$\begin{aligned}
g_{xy}^2 g_{spec}^2 = &\ z_a^2\left[u^2 g_{xy}^2 + (hg_{spec} - ug_z)^2\right] \\
&+ z_b^2\left[v^2 g_{xy}^2 + (kg_{spec} - vg_z)^2\right] \\
&- 2z_a z_b\cos\gamma\left[uvg_{xy}^2 + (hg_{spec} - ug_z)(kg_{spec} - vg_z)\right] \\
&- (hv - ku)^2 z_a^2 z_b^2\sin^2\gamma,
\end{aligned} \tag{1}$$

where g_{spec} is the magnitude of the reciprocal lattice vector associated with the specular diffraction peak, $z_a = 2\pi/(a\sin\gamma)$ and $z_b = 2\pi/(b\sin\gamma)$. In equation (1), which is valid for any contact plane (uvw), the real-space parameters are reduced to a, b and γ. The unknown integers are the Miller indices u and v and the Laue indices h and k for every observed Bragg peak. Here, the first indexing step contains the determination of sets of (u, v, a, b, γ) with potential values for (h, k). We will follow this choice of order throughout this work. Equation (1) is symmetric and analogous expressions are valid for (u, w, a, c, β) with (h, l) and (v, w, b, c, α) with (k, l).

Equation (1) facilitates the mathematical analysis, where the integer variables can be varied and only three real-space unknowns have to be calculated. Therefore, when indexing GIXD patterns, the acquisition of a specular diffraction peak is of considerable help. The unknown real-space parameters a, b and γ can be calculated from the q_{xy} and q_z values of three independent Bragg peak series. This can be achieved

analytically by employing appropriate mathematical substitutions to obtain linearly solvable equations (see Appendix A).

Following this procedure, several sub-sets with solutions of (u, v, a, b, γ) are obtained. These are used to subsequently index the measured components q_{xy}. By computing the vector components $g_{xy}(h_i, k_i)$ [cf. equation (1)] and searching for the smallest absolute differences, tuples of Laue indices can be assigned to the experimental diffraction peaks. A set of (u, v, a, b, γ) is then sorted and evaluated with respect to its summed root-mean-square deviation Δq_{xy}. This is one of the core tasks of GIDInd and is elaborated in Section 3. The final solutions are reduced because of the restrictions imposed by Niggli's scalar-product criterion (Niggli, 1928) and by optional user-selected boundary conditions on a, b and γ.

The remaining lattice constants α, β and c together with the third Miller index w of the contact plane (uvw) can be acquired in a similar way using the expressions for the out-of-plane component g_z of the reciprocal-lattice vector and for the specular component g_{spec} from Table 1. More mathematical details are given in Appendix B.

This algorithm, when applied in a reasonable range for the Miller indices (uvw) of the contact plane and the Laue indices hkl of all observed Bragg peaks, allows us to derive numerically possible sets of parameters in the form of $(u, v, w, a, b, c, \alpha, \beta, \gamma)$. Using these sets, the full data (namely every tuple of q_{xy} and q_z) can be indexed and the deviations not only in q_{xy} but also in q_z and especially in the total length q_{xyz} are

Figure 1
Process workflow of the indexing algorithm realized with GIDInd. The markers on the left branch indicate the split processing of the peak positions (q_{xy}, q_z) for derivation of the lattice parameters $[a, b, \gamma]$ followed by $[\alpha, \beta, c]$ of the unit cell when a specular diffraction peak is available. The right branch demonstrates the single fitting procedure when working without a specular scan.

evaluable. The here-proposed indexing software is used to apply these equations in a highly automated manner, thereby deriving the best-fitting lattice constants for a given set of Bragg peaks from GIXD experiments.

2.1.2. Formalism without the use of a specular diffraction peak. In the triclinic case, the following expressions for the total length of the reciprocal-lattice vector $g_{xyz} = (g_{xy}^2 + g_z^2)^{1/2}$ and the out-of-plane component g_z (see Table 1) are used:

$$g_{xyz}^2 = h^2 a^{*2} + k^2 b^{*2} + l^2 c^{*2} + 2hka^*b^* \cos \gamma^* \\ + 2hla^*c^* \cos \beta^* + 2klb^*c^* \cos \alpha^*, \quad (2)$$

where $a^* = 2\pi bc \sin \alpha / V$, $b^* = 2\pi ac \sin \beta / V$, $c^* = 2\pi ab \times \sin \gamma / V$, $\cos \alpha^* = (\cos \beta \cos \gamma - \cos \alpha)/(\sin \beta \sin \gamma)$, $\cos \beta^* = (\cos \alpha \cos \gamma - \cos \beta)/(\sin \alpha \sin \gamma)$ and $\cos \gamma^* = (\cos \alpha \cos \beta - \cos \gamma)/(\sin \alpha \sin \beta)$ are the reciprocal cell parameters and V is the unit-cell volume, which can be explicitly written as

$$V = abc(1 - \cos^2 \alpha - \cos^2 \beta - \cos^2 \gamma + 2 \cos \alpha \cos \beta \cos \gamma)^{1/2}, \quad (3)$$

and

$$g_z = h \left(\frac{u}{g_{\text{spec}}} a^{*2} + \frac{v}{g_{\text{spec}}} a^*b^* \cos \gamma^* + \frac{w}{g_{\text{spec}}} a^*c^* \cos \beta^* \right) \\ + k \left(\frac{v}{g_{\text{spec}}} b^{*2} + \frac{u}{g_{\text{spec}}} a^*b^* \cos \gamma^* + \frac{w}{g_{\text{spec}}} b^*c^* \cos \alpha^* \right) \\ + l \left(\frac{w}{g_{\text{spec}}} c^{*2} + \frac{u}{g_{\text{spec}}} a^*c^* \cos \beta^* + \frac{v}{g_{\text{spec}}} b^*c^* \cos \alpha^* \right). \quad (4)$$

In a first step, equation (4) is built up as overdetermined LSEs by choosing the q_z of four reflections and varying the Laue indices h_i, k_i and l_i for $i = 1$–4. The solutions with the smallest residual error are further processed and expanded by adding two additional reflections and varying their Laue indices. Using equations (2) and (4), the resulting LSEs with the smallest residual errors are considered as tentative solutions. By continuously adding the q_{xyz} and q_z of further reflections and by varying the associated Laue indices, the overdetermined LSEs are expanded and the best solutions are determined. Requiring considerable computational work, this procedure is of course time consuming. Therefore, the Laue indices of the first six equations, which must be linearly independent, can only be varied in a small interval. For this reason, the reflections with the smallest q_{xyz} are chosen first. Restrictions on the final lattice parameters are imposed by the Niggli criteria (Niggli, 1928) and by optionally chosen boundary conditions.

The orientation parameters of the final unit cell are determined by using equation (4) (see Appendix C).

In the monoclinic lattice, equations (2) and (4) reduce to

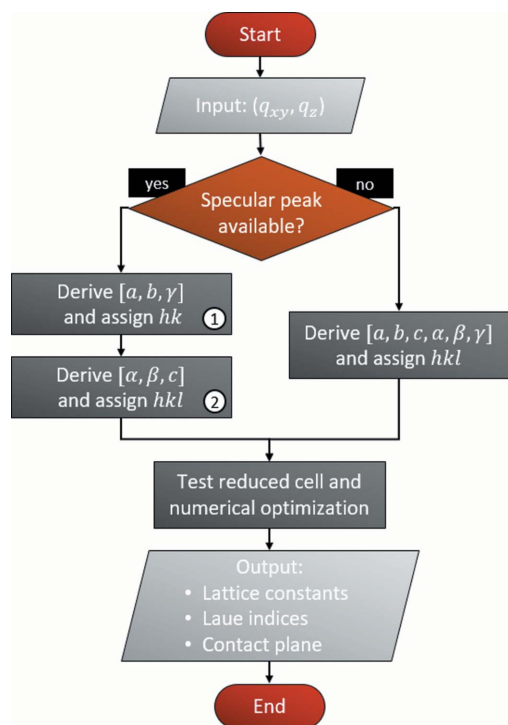$$g_{xyz}^2 = h^2 a^{*2} + k^2 b^{*2} + l^2 c^{*2} - 2hla^*c^* \cos \beta \quad (5)$$

and

$$g_z = h\left(\frac{u}{g_{spec}}a^{*2} - \frac{w}{g_{spec}}a^*c^*\cos\beta\right) + k\frac{v}{g_{spec}}b^{*2}$$
$$+ l\left(\frac{w}{g_{spec}}c^{*2} - \frac{u}{g_{spec}}a^*c^*\cos\beta\right). \quad (6)$$

Then, of course, our algorithm results in less computational work and is less time consuming. Therefore, we offer the option – in a first attempt – to check if the data are compatible with a monoclinic system.

2.2. Common final steps

In both implemented algorithms, that using the specular diffraction peak and that not using the specular diffraction peak, the final solutions for the lattice constants are optimized by performing first-order corrections (Simbrunner *et al.*, 2020). Furthermore, it is checked whether the final unit cells correspond to the reduced cells (Buerger, 1957). Redundant solutions due to the symmetry of the equations are removed.

3. Indexing software: *GIDInd*

The programmatic implementation of the algorithm is based on MATLAB and provided as a standalone desktop application. The peak positions given in the form of positive tuples of (q_{xy}, q_z) are the required input for the program. A proven and efficient tool to derive the diffraction peaks from experimental GIXD data is the software package *GIDVis* (Schrode *et al.*, 2019). For *GIDInd*, the input data have to be provided within a formatted `.xls` (or `.xlsx`) file. Each peak position is listed in a row with numerical values of q_{xy} in the first column and the corresponding values for the out-of-plane components q_z in the second column. If applicable, the specular peak must be included directly in the uploaded file in the row where $q_{xy} = 0$. The uploaded file should contain only numeric floating-point numbers with a dot for decimal separation. If the formatting is accepted and at least one specular diffraction peak is included, the data are added to the q map within the graphical user interface (GUI). Multiple indication icons and status messages are provided to guide the operator through the indexing process. If no specular information is available within the list, a subroutine is enabled addressing this case.

3.1. Indexing using a specular diffraction peak

The indexing procedure and the calculation of the real-space lattice parameters is split into two consecutive parts if the specular peak is utilized (see Section 2.1.1). The internal architecture of the first part, namely derivation of a, b and γ upon successive assignment of the Laue indices h and k to q_{xy} data, is outlined in Fig. 2. For setting up the LSEs [equation (12), Appendix A], 'start values' are specified by the routine. These are linearly independent sets of three Bragg peaks from the provided reciprocal-space map. That has a direct impact on the input file containing the Bragg peaks: for indexing GIXD data with *GIDInd* using a specular diffraction peak, at least four independent peaks, with one line having the entry

$q_{xy} = 0$, are required. The approach elaborated in Appendix D has been shown to be successful in checking the (in)dependence. After these initialization steps, the matrices for the LSEs are constructed. For every combination of integers (u, v, h, k) in a specified range, a quadratic array is set up. The pre-adjusted limit for permuting the Miller indices is ± 2. The option to restrict the solutions to a (001) contact plane or to any other desired combination of fixed (uvw) is also provided to the operator. The default limit for h and k is ± 3, but this is adjustable if required. This choice is reasonable, as the 'start sets' to compute a, b and γ are generated using the lowest q values first, and the Laue indices are typically the lowest in this region. This assumption is in agreement with other reported programs addressing the indexing problem of GIXD patterns (Savikhin *et al.*, 2020; Hailey *et al.*, 2014). The evaluation of every LSE is based on MATLAB's symbolic matrix left division for every matrix with a determinant unequal to zero. The substitutions given in Appendix A are applied to derive the sub-sets with solutions of (u, v, a, b, γ). For machines with available multicore processors, these calculations are executed in parallel computing algorithms using parallelized loops.

Crystallographic restrictions, formulated as Niggli criteria (Niggli, 1928), as well as user-defined limitations to the lattice constants allow a first confinement of the possible solutions at this point. Non-real or negative numerical results are eliminated and only solutions with $60 \leq \gamma \leq 120°$ (Hahn, 2005) are considered for the subsequent indexing procedure. Additionally, the lattice constants a and b are restricted to the range
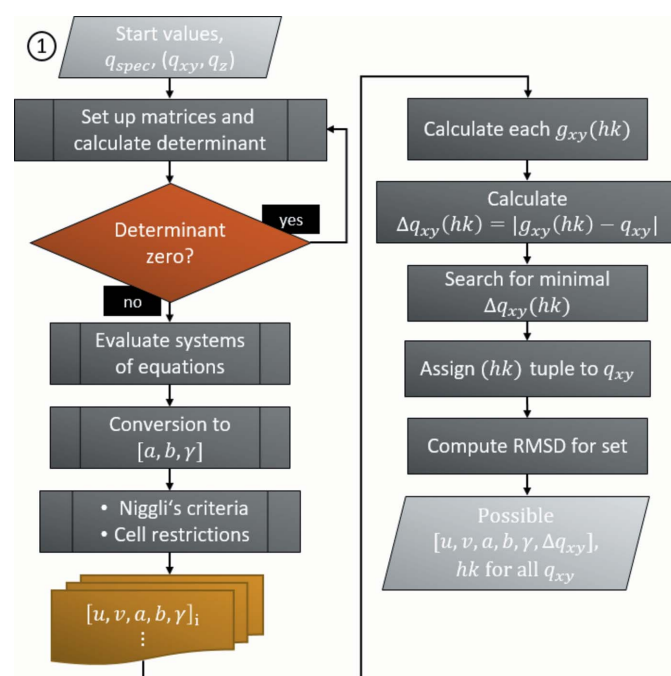

Figure 2
Process diagram describing the determination of the lattice parameters $[a, b, \gamma]$ using the diffraction peaks (q_{xy}, q_z) and the specular peak q_{spec}. After comparison with calculated values for the component of the reciprocal-lattice vector $g_{xy}(hk)$, the whole branch results in sub-sets containing the three lattice parameters, preliminarily assigned pairs of Laue indices h and k, two Miller indices u and v, and an RMSD for the set expressed as Δq_{xy}.

between 3 and 60 Å. Adjustment of an upper limit at ∼30 Å is recommended, to keep computational effort low.

During this part of the routine, every experimental peak position gets preliminarily equipped with four potential pairs of h and k. (An unambiguous assignment of the Laue indices is not possible at this point owing to symmetry considerations of the equations provided.) The lower and upper limits are set to ±6 by default and adjustable to a maximum of ±8. This program parameter can be differentiated from the maximum cap for h and k for solving the LSEs, as not only the lowest values out of the q map are considered now, but all input diffraction peaks.

For a given sub-set of (u, v, a, b, γ), the program computes $g_{xy}(hk)$ for every single combination of hk. The four tuples of hk yielding the smallest absolute differences $\Delta q_{xy}(hk)$ are assigned to a chosen value of q_{xy}. This is done for each of the N input Bragg peaks and the 'quality' of the set is derived by computing the summed root-mean-square deviations (RMSD) with

$$\text{RMSD}|_{(u,v,a,b,\gamma)} = \left(\frac{1}{N}\sum_{n=1}^{N} \Delta q_{xy_n}\right)^{1/2}$$
$$= \left[\frac{1}{N}\sum_{n=1}^{N}\left(g_{xy_n} - q_{xy_n}\right)^2\right]^{1/2}. \quad (7)$$

This algorithm applied to all possible sub-sets will result in two outputs: (i) a list of diffraction peaks with assigned pairs of potential Laue tuples and (ii) the corresponding list of partial solutions (u, v, a, b, γ). This serves as input for the second part of the indexing routine. As demonstrated in Fig. 3, the deri-
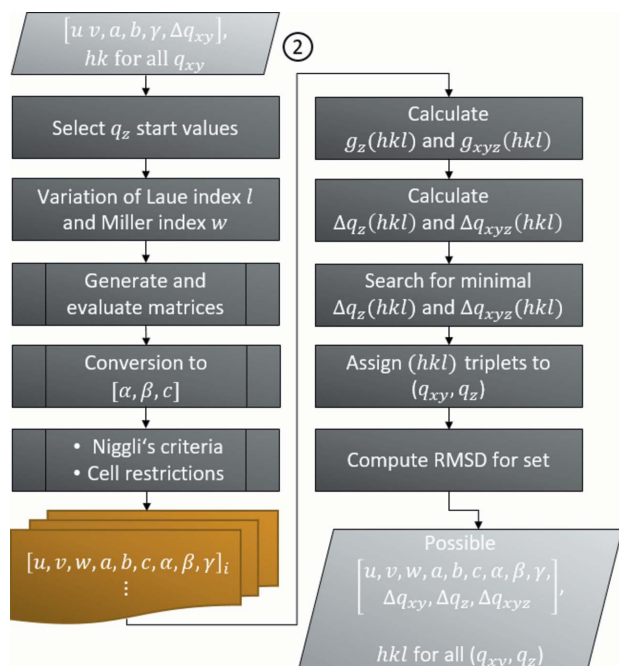
vation of the remaining parameters follows a similar process architecture. In a first step, 'start values' out of the q_z data are selected. For generation of the LSEs [equation (22), Appendix B], the three lowest values which are not multiples of each other are considered. The limit for variation of the Laue index l is a program parameter and thus adjustable by the user. By default, l is varied in the range between −6 and +6 and adjustable to a maximum of ±8. If not specifically defined through a GUI input, the cap of w is internally adopted to the maximal occurring value of u and v.

The program generates one LSE for every possible numerical combination. The number of systems to be solved is therefore directly dependent on the range of the indices and on how many previously acquired solutions (sub-sets) are included in the second round. This can also be controlled from the GUI. The so-obtained LSEs can again be solved using MATLAB's matrix left division function, which returns the least-squares solution to each overdetermined system of equations. The norm of the residuals is used as a sorting quantity at this point. Using the previously defined substitutions (see Appendix B), the three remaining cell parameters are deduced for each system. The routine presents the resulting sets in the form of $(u, v, w, a, b, c, \alpha, \beta, \gamma)$ together with the list of peak positions and the tentatively assigned tuples of h and k. Again, the (purely numerical) output of lattice constants can be gradually confined and reduced by the application of user restrictions and by continuous inquiries regarding the fulfillment of Niggli's criteria. Only those sets passing these tests will continue to be considered for full indexing in the second part.

During the first indexing step, every Bragg peak is equipped only with preliminarily assigned pairs of Laue indices h and k. At this point of the indexing routine, the third Laue index l remains to be determined for every single reflection. This could be done by again varying integer numbers in a certain range and evaluating every single possible combination regarding the overall deviation of the diffraction pattern. However, the equations from Table 1 allow the value for l to be determined analytically, as all other formerly unknown parameters are now available. This option is preferred to minimize computational time and enhance the overall performance of the program. For a fixed set of $(u, v, w, a, b, c, \alpha, \beta, \gamma)$, the value of l is acquired for each reflection and each combination of h_i, k_i where $i = 1$–4. It is subsequently rounded to the next integer number. By applying exactly the same equations again, every value for $g_z(hkl)$ and $g_{xyz}(hkl)$ can be computed. The program searches for the smallest absolute differences $\Delta q_z(hkl)$ and $\Delta q_{xyz}(hkl)$ for every combination. This time, the triplet hkl with the smallest deviation in g_{xyz} is assigned to the reflection and settled as the final set of Laue indices. As already introduced in equation (7), the 'quality' of the as-derived individual solutions is again assessed using the summed root-mean-square deviation.

With all Laue indices derived for all provided Bragg peaks, each single set of lattice parameters undergoes a numerical optimization routine based on first-order corrections (Simbrunner et al., 2020). This subroutine computes correction

Figure 3
Process diagram containing the branch for derivation of the preliminary lattice constants with the associated RMSDs Δq_{xy}, Δq_z and Δq_{xyz} before reduction and optimization is applied. In addition to the first step, the out-of-plane component $g_z(hkl)$ and the total length of the reciprocal-lattice vector $g_{xyz}(hkl)$ are used for assignment of the Laue indices h, k and l.

terms for the six lattice constants a, b, c, α, β and γ in order to minimize the summed RMSDs in the component q_z and in the total length q_{xyz} of the scattering vectors. Even though the received solutions face minimal numerical errors with regard to the experimental diffraction pattern, these unit-cell solutions are not necessarily unique and superlattices can appear. Each solution is therefore checked to see if it corresponds to the reduced cell (the unit cell based on the three shortest non-coplanar lattice vectors) and, if not, adjusted accordingly. If possible, similar and reoccurring results are combined by comparison of their Laue indices. Those exhibiting the smallest errors are presented in an ascending order to the user through the GUI of *GIDInd*. The results are provided to the operator in the form of numbered lists with the six lattice parameters a, b, c, α, β and γ, the three Miller indices of the contact plane u, v and w, the volume V of the cell, and the four summed errors in q_{xy}, q_z, q_{xyz} and q_{spec}. Using the GUI, the simulated diffraction patterns for every solution can be calculated to be compared with the experimental input data. The *GIDInd* interface offers options to generate compiled output files (`.xlsx` format) with all information of a specific solution and with all Laue indices for the corresponding Bragg peaks. The basic functionality of the GUI will be demonstrated with two samples in Section 4. For a detailed explanation we refer the reader to the user's manual, available with the links provided in Section 5.

3.2. Indexing without using a specular diffraction peak

If no specular peak is included, a subprogram is enabled. The interface of the subroutine offers several adjustments for the indexing procedure. These possible adjustments are of particular interest in this case, as the computation time and the use of memory are the liming factors when working without a specular diffraction peak. In particular, the possibility to restrict the potential Laue indices is crucial in order to make the derivation of solutions possible. The cap of the Laue indices h and k for building the LSEs is ± 1 and ± 2, respectively. The third index l is varied between -2 and $+2$. The subprogram is able to assign integers in the range between -6 and $+6$ upon subsequent indexing. An additional computational reduction of effort when deriving a solution with this routine is the possible choice between a monoclinic cell and a triclinic cell. As already implied by the indexing formalism in Section 2.1.2, the assumption of a monoclinic unit cell facilitates the mathematical treatment. Using the triclinic approach, the six lattice constants a, b, c, α, β and γ are derived simultaneously. The minimum number of peak positions is therefore also raised to six. In addition, these peaks must be linearly independent, which is not necessarily the case when deriving data from GIXD patterns. In the monoclinic approach, only four independent peak positions have to be provided. Note that the indexing routine without incorporation of the specular scan has no automation for checking the (in)dependence of the Bragg peaks. The use of the subroutine is demonstrated in Section 4.1.

4. Application examples

In addition to the calculation of lattice parameters, different, although related, tools are provided when working with *GIDInd*. The main window consists of five sub-windows (panels): 'Data Points and Representation Panel', 'Add Crystal Panel', 'Indexing Panel', 'Error Panel' and 'Results Panel'. For indexing without using a specular diffraction peak, a subprogram is opened for controlling the indexing procedure. The functionality of the indexing program is demonstrated here on two different thin-film samples. In both cases, the peak positions were extracted from GIXD maps using the GIXD analysis tool *GIDVis* (Schrode *et al.*, 2019).

### 4.1. Pentacenequinone on highly oriented pyrolytic graphite

6,13-Pentacenequinone (PQ) is an example of an organic semiconductor material with the chemical formula $C_{22}H_{12}O_2$. GIXD data of a PQ thin-film sample, grown on highly oriented pyrolytic graphite, are chosen to demonstrate the applicability of *GIDInd*. A specular diffraction peak is provided at $q_z = 1.9460$ Å^{-1}. In total, a set with 74 Bragg reflections is provided, of which 30 peak positions are taken for the calculation of the lattice parameters. Upon indexing of this data set, the program is applied with initial settings only, to emphasize the fast and easy-to-use approach.

The data are uploaded via the 'Data Points and Representation Panel' and, as the formatting is accepted and the specular diffraction peak is included within the set, the routine adds the peak positions to the 'Data representation' graph right away. A screenshot of the panel is shown in Fig. 4(*a*). Black markers are used to indicate (uploaded) experimental peaks and diffraction patterns simulated with the derived unit-cell parameters are printed with red cross markers. The GUI guides the operator through the indexing procedure via dialogue windows, indication icons and continually updated process bars. After successful execution of the first indexing part, the user gets possible results presented within a panel at the bottom of the GUI. The routine allows sorting of the obtained sets of (u, v, a, b, γ) either according to their summed RMSDs (Δq_{xy}) or by the areas of the parallelograms formed through a, b and γ. Both options can be useful; however, a general recommendation for sorting the sub-sets cannot be given. It is usually expedient to search for the smallest cells with a reasonable deviation. A rough indication for a good match is $\Delta q_{xy} \simeq 0.01$ Å^{-1} and below, but this error varies from sample to sample and depends also on the quality of the provided diffraction data. Using the indexing routine several times will certainly help the user to develop intuition about deviation values yielding the highest matching quality. For the presented example of PQ, the sets are sorted by ascending parallelogram areas, as the smallest appearing areas also exhibit the least deviations in q_{xy}. A possible strategy for the second indexing part is monitoring the quality of the derived unit-cell solutions using the 'Error map', shown in the screenshot in Fig. 4(*b*). The user can stop the program manually when the summed RMSD in the total length of the scattering vector (Δq_{xyz}) converges to a minimum. As long as

the operator does not intervene, the program continues processing the earlier derived sub-sets. A defined range for the sub-sets that should be considered can be specified via the GUI. We emphasize again that the mathematical solutions are not unique. Various unit-cell solutions may be able to reproduce a matching diffraction pattern.

Here, the program is ended after nine different sets of (u, v, a, b, γ) are processed, as the error converged to a minimum value of $\Delta q_{xyz} = 0.0015$ Å^{-1}. The unit-cell solution exhibiting the least deviations with $\Delta q_{xy} = 0.0010$ Å^{-1} and $\Delta q_z = 0.0022$ Å^{-1} is given explicitly with $a = 5.056$, $b = 8.076$, $c = 8.871$ Å, $\alpha = 91.54$, $\beta = 93.03$, $\gamma = 94.14°$ and a volume of $V = 360.8$ Å^3. The derived Miller indices of the contact plane are $(uvw) = (102)$. This solution agrees with the recently reported new polymorph of the PQ crystal (Simbrunner *et al.*, 2018) within the expected range of uncertainty.

The same GIXD data are used for verifying the indexing algorithm without incorporating the specular diffraction peak. When uploading a file without a specular peak listed, the subroutine 'Indexing without specular scan' is unlocked. A screenshot of the control panel is shown in Fig. 5(*a*). The number of available and independent peaks is not always sufficient for the triclinic approach (at least six independent Bragg reflections are required). Therefore, the user can choose between the cell types 'monoclinic' and 'triclinic'. As in the case where the specular peak is used, the caps for the Laue indices can be set separately. Restrictions on the real-space lattice constants and the volume are also provided to filter the output solutions. Once this sub-routine starts, it processes the input in a single fitting procedure to derive sets of $(a, b, c, \alpha, \beta, \gamma)$. The results, if available, are embedded in the GUI and can then be compared graphically or exported as mentioned
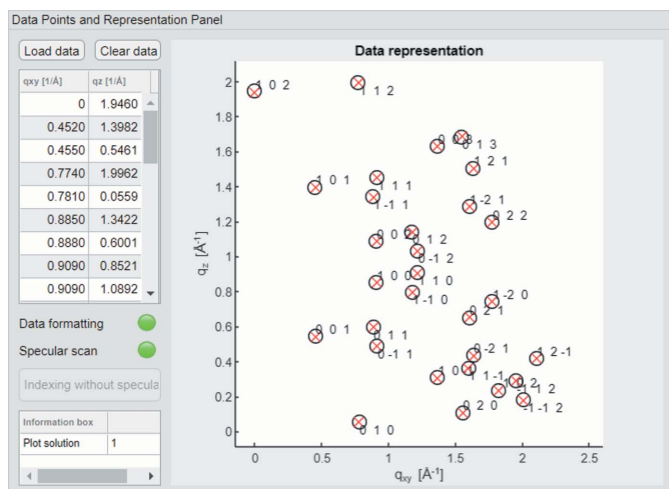
(*a*)
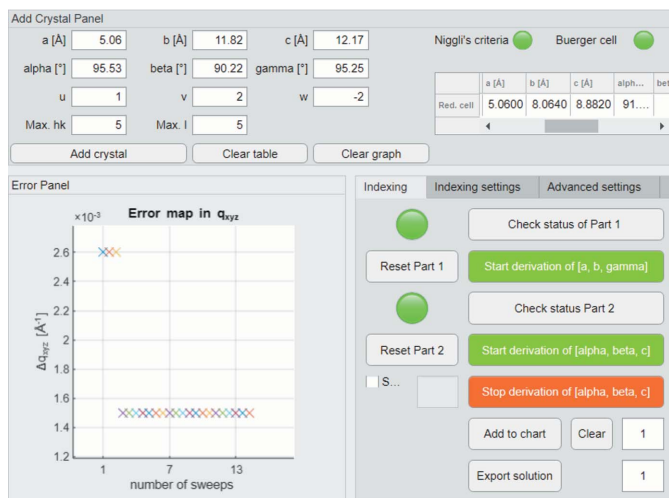
(*b*)

Figure 4
Selected screenshots of the graphical user interface of *GIDInd* after indexing. (*a*) 'Data Points and Representation Panel' with a selection of 30 peak positions for indexing of the PQ data set (plotted in black). The pattern printed in red represents the simulated diffraction pattern using the best solution with respect to the error in q_{xyz}. (*b*) Cut-out section of the interface containing the 'Add Crystal Panel', the 'Error Panel' and the main control panel labelled 'Indexing'.
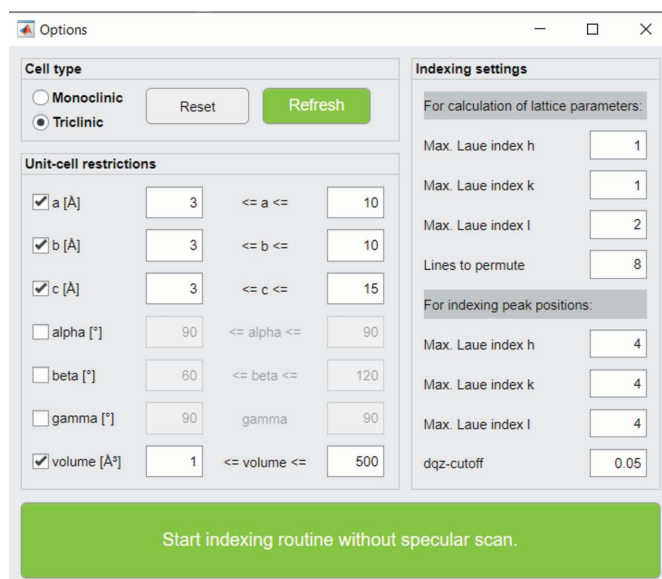
(*a*)
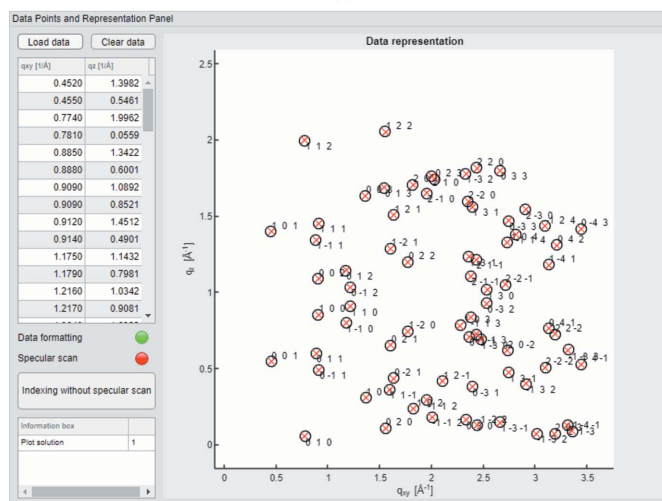
(*b*)

Figure 5
(*a*) 'Options' panel of the sub-routine 'Indexing without specular scan'. (*b*) 'Data Points and Representation Panel' showing the simulated diffraction pattern (red) of a unit-cell solution derived with 74 input peak positions (black) without specular peak.

above. With the here-provided 74 reflections the following unit cell results: $a = 5.053$, $b = 8.076$, $c = 8.8671$ Å, $\alpha = 91.55$, $\beta = 93.08$, $\gamma = 94.15°$, $V = 360.0$ Å^3 with $\Delta q_{xyz} = 0.0028$ Å^{-1}, $\Delta q_{xy} = 0.0028$ Å^{-1} and $\Delta q_z = 0.0017$ Å^{-1}. Within the range of uncertainty, this is the same result as derived above. The corresponding simulated diffraction pattern is shown in Fig. 5(b). With this framework, a distinct contact plane cannot be specified with integer Miller indices. The rotation angles ψ and ϕ to calculate the rotation matrix (as described in Appendix C) are stored in the generated output file. Although the results appear to be the same in the case of PQ, we report significant drawbacks in high memory occupation and time consumption if the routine is used without a specular diffraction peak. However, this application can be of advantage for the case that no specular scan is available.

4.2. Acetylsalicylic acid on thermally oxidized Si(100)

If the initial settings (as used earlier for PQ) do not lead to sufficient solutions, a different approach for the derivation of the unit-cell parameters can be helpful. For demonstration, GIXD data of an acetylsalicylic acid (ASS) thin film on Si(100) are evaluated. We derived ten diffraction peaks from the reciprocal-space map and, additionally, a specular diffraction peak is used for the indexing procedure with *GIDInd*. First, a run with initial settings is performed to find the tendencies according to which the lattice parameters converge and to define reasonable ranges of the errors. After this survey run, the edge lengths of the unit cell a, b and c and especially the volume show a clear trend towards particular values. We therefore use the 'Indexing settings' tab of *GIDInd* to restrict certain lattice parameters for the solution sets, as shown in Fig. 6(a). The screenshot of the 'Error panel' in Fig. 6(b) was taken during the survey run and shows a population of solutions around $\Delta q_{xyz} = 0.005$ Å^{-1} and below. With the applied restrictions, we can immediately identify a unit cell which generates a matching, and therefore promising, diffraction pattern [Fig. 6(c)]. The corresponding unit-cell solution is shown and highlighted in Fig. 6(d) and explicitly given with the parameters $a = 6.594$, $b = 11.378$, $c = 11.418$ Å, $\alpha = 95.29$, $\beta = 90.28$ and $\gamma = 90.07°$. The volume of the cell is $V = 852.98$ Å^3, the total-length deviation is $\Delta q_{xyz} = 0.0043$ and we find the plane $(uvw) = (020)$ to be parallel to the substrate. The solution found here agrees with the known polymorph of ASS at ambient temperature (Wheatley, 1964; Kim *et al.*, 1985). Note again that Niggli's criteria are implemented and the output solutions are reduced cells. The order of listing the parameters can therefore deviate from that stated in the literature. The
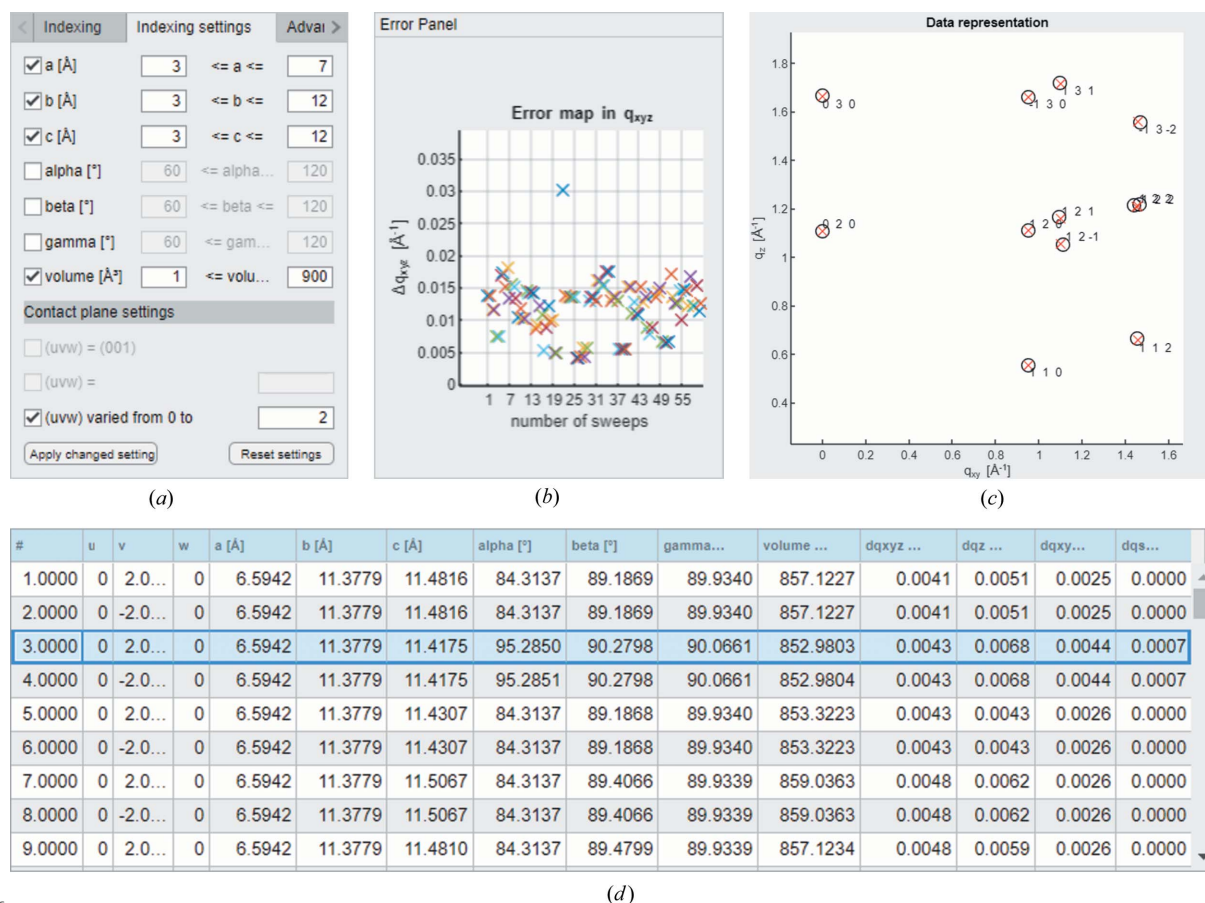

(a)　　　　　　　　(b)　　　　　　　　(c)

#	u	v	w	a [Å]	b [Å]	c [Å]	alpha [°]	beta [°]	gamma...	volume ...	dqxyz ...	dqz ...	dqxy...	dqs...
1.0000	0	2.0...	0	6.5942	11.3779	11.4816	84.3137	89.1869	89.9340	857.1227	0.0041	0.0051	0.0025	0.0000
2.0000	0	-2.0...	0	6.5942	11.3779	11.4816	84.3137	89.1869	89.9340	857.1227	0.0041	0.0051	0.0025	0.0000
3.0000	0	2.0...	0	6.5942	11.3779	11.4175	95.2850	90.2798	90.0661	852.9803	0.0043	0.0068	0.0044	0.0007
4.0000	0	-2.0...	0	6.5942	11.3779	11.4175	95.2851	90.2798	90.0661	852.9804	0.0043	0.0068	0.0044	0.0007
5.0000	0	2.0...	0	6.5942	11.3779	11.4307	84.3137	89.1868	89.9340	853.3223	0.0043	0.0043	0.0026	0.0000
6.0000	0	-2.0...	0	6.5942	11.3779	11.4307	84.3137	89.1868	89.9340	853.3223	0.0043	0.0043	0.0026	0.0000
7.0000	0	2.0...	0	6.5942	11.3779	11.5067	84.3137	89.4066	89.9339	859.0363	0.0048	0.0062	0.0026	0.0000
8.0000	0	-2.0...	0	6.5942	11.3779	11.5067	84.3137	89.4066	89.9339	859.0363	0.0048	0.0062	0.0026	0.0000
9.0000	0	2.0...	0	6.5942	11.3779	11.4810	84.3137	89.4799	89.9339	857.1234	0.0048	0.0059	0.0026	0.0000

(d)

Figure 6
Selection of screenshots of the *GIDInd* interface. (a) 'Indexing settings' panel showing the restrictions on the potential lattice constants a, b and γ and on the volume of the crystallographic cell. (b) 'Error panel' monitoring the deviation in q_{xyz} during execution of the indexing routine. (c) Input peak positions (black) of an ASS thin film together with the simulated diffraction peaks (red) calculated with the solution highlighted in the 'Result Panel' in (d).

equivalence of the obtained solutions can easily be proven using known symmetry relations (Simbrunner *et al.*, 2018).

With the possibility of restricting the parameters, the user can specifically scan regions and use potential initial guesses if available. In addition to restricting the parameters, the Miller indices can be set to pre-defined values. In some cases, it can be useful to first search for solutions that exhibit a (001) lattice plane. This can give a first impression of the possible solutions and, overall, help to save time when indexing the GIXD data with *GIDInd*.

Application of the indexing routine and the accompanying memory usage did not lead to any computational restrictions for the investigated samples. If the program parameters are kept in the ranges as demonstrated, the program should not face any programmatic problems. Nevertheless, no claim is made to completeness. Neither should the impression arise that the program has no limitations or bottlenecks, nor that the obtained solutions are unambiguous.

5. Availability

The program is provided in the form of an executable file (`.exe`) and as a MATLAB application (`.mlapp` file), written and tested in MATLAB version R2019b, Update 5, on Windows OS. It is released under the terms of the GNU General Public License, either version 3 of the licence or any later version. The software can be used as a standalone executable file (MATLAB Runtime required) and as a `.mlapp` file (valid MATLAB licence required) together with the functions provided in the zip-compiled folder, available on the web sites stated below. The executable can be used right away with the MATLAB Runtime installed, which can be downloaded from The Mathworks Inc. (https://mathworks.com/products/compiler/matlab-runtime.html) free of charge. For individual adaptions and potential further development, all source codes are provided. To access the program, codes and further instructions, visit https://www.if.tugraz.at/amd/GIDInd/ and https://github.com/m-kainz/GIDInd. More program details, further tutorials, and additional help regarding operation and use of *GIDInd* can be found in a separate documentation file, available on the web sites.

APPENDIX A
Mathematical procedure for analytically determining the cell parameters *a*, *b* and *γ*

For analytically determining the unit-cell parameters *a*, *b* and *γ*, it is convenient to introduce the parameters Z_a^2, Z_b^2 and X_γ with the substitutional relations

$$Z_a^2 = z_a^2\left(1 - v^2 z_b^2 \sin^2\gamma / g_{\text{spec}}^2\right), \tag{8}$$

$$Z_b^2 = z_b^2\left(1 - u^2 z_a^2 \sin^2\gamma / g_{\text{spec}}^2\right), \tag{9}$$

$$X_\gamma = z_a z_b\left(\cos\gamma - uv z_a z_b \sin^2\gamma / g_{\text{spec}}^2\right). \tag{10}$$

Note that Z_a^2 and Z_b^2 are always positive. Using these substitutions, equation (1) can be linearized and rewritten as

$$1 = Z_a^2\left[\frac{1}{g_{xy}^2}\left(h - u\frac{g_z}{g_{\text{spec}}}\right)^2 + \frac{u^2}{g_{\text{spec}}^2}\right]$$
$$+ Z_b^2\left[\frac{1}{g_{xy}^2}\left(k - v\frac{g_z}{g_{\text{spec}}}\right)^2 + \frac{v^2}{g_{\text{spec}}^2}\right]$$
$$- 2X_\gamma\left[\frac{1}{g_{xy}^2}\left(h - u\frac{g_z}{g_{\text{spec}}}\right)\left(k - v\frac{g_z}{g_{\text{spec}}}\right) + \frac{uv}{g_{\text{spec}}^2}\right]. \tag{11}$$

The parameters Z_a^2, Z_b^2 and X_γ can be determined from three independent Bragg peak series by solving the following LSE:

$$\begin{bmatrix} f_{11}(h_1) & f_{12}(k_1) & f_{13}(h_1, k_1) \\ f_{21}(h_2) & f_{22}(k_2) & f_{23}(h_2, k_2) \\ f_{31}(h_3) & f_{32}(k_3) & f_{33}(h_3, k_3) \end{bmatrix}\begin{pmatrix} Z_a^2 \\ Z_b^2 \\ X_\gamma \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \tag{12}$$

where

$$f_{i1}(h_i) = \frac{1}{g_{xyi}^2}\left(h_i - u\frac{g_{zi}}{g_{\text{spec}}}\right)^2 + \frac{u^2}{g_{\text{spec}}^2}, \tag{13}$$

$$f_{i2}(k_i) = \frac{1}{g_{xyi}^2}\left(k_i - v\frac{g_{zi}}{g_{\text{spec}}}\right)^2 + \frac{v^2}{g_{\text{spec}}^2}, \tag{14}$$

$$f_{i3}(h_i, k_i) = -2\left[\frac{1}{g_{xyi}^2}\left(h_i - u\frac{g_{zi}}{g_{\text{spec}}}\right)\left(k_i - v\frac{g_{zi}}{g_{\text{spec}}}\right) + \frac{uv}{g_{\text{spec}}^2}\right], \tag{15}$$

and *i* = 1, 2 and 3. For obtaining *a*, *b* and *γ* from equations (8)–(10) the following identity is helpful:

$$\Lambda = z_a^2 z_b^2 \frac{\sin^2\gamma}{g_{\text{spec}}^2} = \frac{Z_a^2 Z_b^2 - X_\gamma^2}{g_{\text{spec}}^2 - \left(u^2 Z_a^2 + v^2 Z_b^2 - 2uv X_\gamma\right)}. \tag{16}$$

From equations (8)–(10) and (16), the following expressions can be derived:

$$z_a = \left(Z_a^2 + v^2\Lambda\right)^{1/2}, \tag{17}$$

$$z_b = \left(Z_b^2 + u^2\Lambda\right)^{1/2} \tag{18}$$

and

$$\cos\gamma = \frac{X_\gamma + uv\Lambda}{z_a z_b}. \tag{19}$$

APPENDIX B
Mathematical procedure for analytically determining the cell parameters *c*, *α* and *β*

For setting up the LSE, the expressions for g_z and g_{spec} in Table 1 can be transformed to:

$$h_i a^* \cos \beta^* + k_i b^* \cos \alpha^* + l_i c^*$$

$$= \frac{g_{z_i} g_{\text{spec}} - h_i u z_b^2 - k_i v z_b^2 + (h_i v + k u) z_a z_b \cos \gamma}{\left(g_{\text{spec}}^2 - u^2 z_a^2 - v^2 z_a^2 + 2uv z_a z_b \cos \gamma\right)^{1/2}} := g_{z,\text{T}_i},$$

$$(20)$$

for three linearly independent peaks with indices $i = 1, 2$ and 3, and

$$ua^* \cos \beta^* + vb^* \cos \alpha^* + wc$$

$$= \left(g_{\text{spec}}^2 - u^2 z_a^2 - v^2 z_a^2 + 2uv z_a z_b \cos \gamma\right)^{1/2} := g_{\text{spec},\text{T}}. \quad (21)$$

Then, it is again possible to build an LSE of the form

$$\begin{pmatrix} u & v & w \\ h_1 & k_1 & l_1 \\ h_2 & k_2 & l_2 \\ h_3 & k_3 & l_3 \end{pmatrix} \begin{pmatrix} \kappa \\ \lambda \\ \rho \end{pmatrix} = \begin{pmatrix} g_{\text{spec},\text{T}} \\ g_{z,\text{T}_1} \\ g_{z,\text{T}_2} \\ g_{z,\text{T}_3} \end{pmatrix}. \quad (22)$$

Under variation of the Miller index w and the Laue indices l_1, l_2 and l_3, the terms for $g_{\text{spec},\text{T}}$ and g_{z,T_i} have to be computed for every previously derived set of (u, v, a, b, γ). These overdetermined systems are solvable for κ, λ and ρ. The parts containing the unknown reciprocal lattice constants a^*, b^*, c^*, α^* and β^* are fully replaced by $\kappa = [-2\pi/(a \sin \gamma)]\delta$, $\lambda = [-2\pi/(b \sin \gamma)]\mu$ and $\rho = 2\pi/(c \sin \epsilon)$, $\delta = (\cos \beta - \cos \alpha \cos \gamma)/(\sin \gamma \sin \epsilon)$, $\mu = (\cos \alpha - \cos \beta \cos \gamma)/(\sin \gamma \sin \epsilon)$, and $\sin \epsilon = \sin \gamma/(\sin^2 \gamma + \delta^2 + \mu^2 + 2\delta\mu \cos \gamma)^{1/2}$ (Truger et al., 2016). Then, the remaining lattice parameters can be calculated by using the following expressions:

$$\cos \alpha = \frac{\mu + \delta \cos \gamma}{\left(\sin^2 \gamma + \delta^2 + \mu^2 + 2\delta\mu \cos \gamma\right)^{1/2}}, \quad (23)$$

$$\cos \beta = \frac{\delta + \mu \cos \gamma}{\left(\sin^2 \gamma + \delta^2 + \mu^2 + 2\delta\mu \cos \gamma\right)^{1/2}} \quad (24)$$

and

$$c = \frac{2\pi}{\rho \sin \epsilon}. \quad (25)$$

APPENDIX C
Determining the parameters of the rotation matrix

The orientation parameters of the final unit cell are determined by using equation (4), which can be alternatively written as $g_z = ha_z^* + kb_z^* + lc_z^*$:

$$\begin{pmatrix} a^{*2} & a^* b^* \cos \gamma^* & a^* c^* \cos \beta^* \\ a^* b^* \cos \gamma^* & b^{*2} & b^* c^* \cos \alpha^* \\ a^* c^* \cos \beta^* & b^* c^* \cos \alpha^* & c^{*2} \end{pmatrix} \begin{pmatrix} u_q \\ v_q \\ w_q \end{pmatrix} = \begin{pmatrix} a_z^* \\ b_z^* \\ c_z^* \end{pmatrix}, \quad (26)$$

where $u_q = u/g_{\text{spec}}$, $v_q = v/g_{\text{spec}}$ and $w_q = w/g_{\text{spec}}$. By building the ratios of these factors, the Miller indices of the contact plane may be found.

The rotation matrix $\mathbf{R}$ can be expressed as

$$\mathbf{R}(\psi, \Phi) = \begin{bmatrix} \begin{array}{c} \cos^2 \psi \\ + \cos \Phi \sin^2 \psi \end{array} & \begin{array}{c} \cos \psi \sin \psi \\ \times (1 - \cos \Phi) \end{array} & -\sin \psi \sin \Phi \\[2ex] \begin{array}{c} \cos \psi \sin \psi \\ \times (1 - \cos \Phi) \end{array} & \begin{array}{c} \sin^2 \psi \\ + \cos \Phi \cos^2 \psi \end{array} & \cos \psi \sin \Phi \\[2ex] \sin \psi \sin \Phi & -\cos \psi \sin \Phi & \cos \Phi \end{bmatrix}, \quad (27)$$

where ψ and ϕ are the rotation angles. Then the following relations are valid (Simbrunner et al., 2018):

$$\cos \psi = \frac{(u_q/a) \cos \gamma - v_q/b}{\left[(u_q/a)^2 + (v_q/b)^2 - 2(u_q/a)(v_q/b) \cos \gamma\right]^{1/2}}, \quad (28)$$

$$\sin \psi = \frac{u_q/a}{\left[(u_q/a)^2 + (v_q/b)^2 - 2(u_q/a)(v_q/b) \cos \gamma\right]^{1/2}}, \quad (29)$$

$$\cos \Phi = (u_q a^* \cos \beta^* + v_q b^* \cos \alpha^* + w_q c^*)$$
$$/\left[u_q^2 z_a^2 + v_q^2 z_b^2 - 2u_q v_q z_a z_b \cos \gamma \right.$$
$$\left. + \left(u_q a^* \cos \beta^* + v_q b^* \cos \alpha^* + w_q c^*\right)^2\right]^{1/2}. \quad (30)$$

For ϕ, the following relations can also be derived (Simbrunner et al., 2021):

$$\cos \Phi = \tan \Omega_a \cot \psi$$
$$= \tan \Omega_a \left(\cot \gamma - \frac{v_q}{u_q} \frac{a}{b} \frac{1}{\sin \gamma}\right) \text{ if } u_q \neq 0, \quad (31)$$

where

$$\cos \Omega_a = \frac{1}{\left[1 - \left(2\pi u_q/a\right)^2\right]^{1/2}}$$
$$\times \frac{(u_q/a) \cos \gamma - (v_q/b)}{\left[(u_q/a)^2 + (v_q/b)^2 - 2(u_q/a)(v_q/b) \cos \gamma\right]^{1/2}}, \quad (32)$$

and

$$\cos \Phi = \tan \Omega_b \cot(\gamma - \psi)$$
$$= \tan \Omega_b \left(\cot \gamma - \frac{u_q}{v_q} \frac{b}{a} \frac{1}{\sin \gamma}\right) \text{ if } v_q \neq 0, \quad (33)$$

where

$$\cos \Omega_b = \frac{1}{\left[1 - \left(2\pi v_q/b\right)^2\right]^{1/2}}$$
$$\times \frac{(u_q/a) - (v_q/b) \cos \gamma}{\left[(u_q/a)^2 + (v_q/b)^2 - 2(u_q/a)(v_q/b) \cos \gamma\right]^{1/2}}. \quad (34)$$

If $u_q = v_q = 0$, the contact plane is (001) (unrotated system).

APPENDIX *D*
Determining the start values for the linear systems of equations

To assess whether or not numerical values are multiples of one another, the following algorithm is used: For various pairs out of N, the total number of input Bragg peaks, the quotient of two different points q_{xy_i} and q_{xy_k} is computed, where $q_{xy_i} > q_{xy_k}$ and $i, k \in (1, N)$. If the absolute difference between the quotient and its nearest integer neighbour is below a certain numerical limit (*e.g.* $\epsilon_c = 0.1$), the two values are considered as linearly dependent. A pair q_{xy_i} and q_{xy_k} where

$$\left| \frac{q_{xy_i}}{q_{xy_k}} - \left| \frac{q_{xy_i}}{q_{xy_k}} \right| \right| > \epsilon_c \qquad (35)$$

are considered to be numerically independent and a triplet of such independent values is referred to here as a 'start set'. The routine switches through a specified number of q_{xy} values, permutes the lines in every possible way and thus creates a series of different sets. Initially, these sets are chosen independently from any user-defined numerical limit. The number of points included, however, can be adopted via the user interface. If there is at least one possible, independent combination of three Bragg peaks, it can be found with the algorithm.

References

Boultif, A. & Louër, D. (2004). *J. Appl. Cryst.* **37**, 724–731.
Breiby, D. W., Bunk, O., Andreasen, J. W., Lemke, H. T. & Nielsen, M. M. (2008). *J. Appl. Cryst.* **41**, 262–271.
Buerger, M. J. (1957). *Z. Kristallogr.* **109**, 42–60.
Coelho, A. A. (2003). *J. Appl. Cryst.* **36**, 86–95.
Gentili, D., Gazzano, M., Melucci, M., Jones, D. & Cavallini, M. (2019). *Chem. Soc. Rev.* **48**, 2502–2517.
Hahn, T. (2005). Editor. *International Tables for Crystallography*, Vol. A, *Space-Group Symmetry*. Heidelberg: Springer.
Hailey, A. K., Hiszpanski, A. M., Smilgies, D.-M. & Loo, Y.-L. (2014). *J. Appl. Cryst.* **47**, 2090–2099.
Heffelfinger, C. J. & Burton, R. L. (1960). *J. Polym. Sci.* **47**, 289–306.
Jiang, Z. (2015). *J. Appl. Cryst.* **48**, 917–926.
Jones, A. O. F., Chattopadhyay, B., Geerts, Y. H. & Resel, R. (2016). *Adv. Funct. Mater.* **26**, 2233–2255.
Kim, Y., Machida, K., Taga, T. & Osaki, K. (1985). *Chem. & Ph. Bull.* **33**, 2641–2647.
Niggli, P. (1928). *Handbuch der Experimentalphysik*, Vol. 7, Part 1. Leipzig: Akademische Verlagsgesellschaft.
Pan, Q., Guo, P., Duan, J., Cheng, Q. & Li, H. (2012). *Chin. Sci. Bull.* **57**, 3867–3871.
Roe, R. J. & Krigbaum, W. R. (1964). *J. Chem. Phys.* **40**, 2608–2615.
Sauter, N. K., Grosse-Kunstleve, R. W. & Adams, P. D. (2004). *J. Appl. Cryst.* **37**, 399–409.
Savikhin, V., Steinrück, H.-G., Liang, R.-Z., Collins, B. A., Oosterhout, S. D., Beaujuge, P. M. & Toney, M. F. (2020). *J. Appl. Cryst.* **53**, 1108–1129.
Schrode, B., Pachmajer, S., Dohr, M., Röthel, C., Domke, J., Fritz, T., Resel, R. & Werzer, O. (2019). *J. Appl. Cryst.* **52**, 683–689.
Shmueli, U. (2006). Editor. *International Tables for Crystallography*, Vol. B, *Reciprocal Space*, 1st online ed. Chester: International Union of Crystallography.
Simbrunner, J., Hofer, S., Schrode, B., Garmshausen, Y., Hecht, S., Resel, R. & Salzmann, I. (2019). *J. Appl. Cryst.* **52**, 428–439.
Simbrunner, J., Schrode, B., Domke, J., Fritz, T., Salzmann, I. & Resel, R. (2020). *Acta Cryst.* A**76**, 345–357.
Simbrunner, J., Schrode, B., Hofer, S., Domke, J., Fritz, T., Forker, R. & Resel, R. (2021). *J. Phys. Chem. C*, **125**, 618–626.
Simbrunner, J., Simbrunner, C., Schrode, B., Röthel, C., Bedoya-Martinez, N., Salzmann, I. & Resel, R. (2018). *Acta Cryst.* A**74**, 373–387.
Smilgies, D.-M. & Blasini, D. R. (2007). *J. Appl. Cryst.* **40**, 716–718.
Truger, M., Roscioni, O. M., Röthel, C., Kriegner, D., Simbrunner, C., Ahmed, R., Głowacki, E. D., Simbrunner, J., Salzmann, I., Coclite, A. M., Jones, A. O. & Resel, R. (2016). *Cryst. Growth Des.* **16**, 3647–3655.
Wheatley, P. J. (1964). *J. Chem. Soc.* pp. 6036.