NRC Publications Archive Archives des publications du CNRC

SVDdiagnostic, a program to diagnose numerical conditioning of Rietveld refinements

Mercier, P. H. J.; Le Page, Y.; Whitfield, P. S.; Mitchell, L. D.

This publication could be one of several versions: author's original, accepted manuscript or the publisher's version. /
La version de cette publication peut être l'une des suivantes : la version prépublication de l'auteur, la version
acceptée du manuscrit ou la version de l'éditeur.

For the publisher's version, please access the DOI link below. / Pour consulter la version de l'éditeur, utilisez le lien
DOI ci-dessous.

Publisher's version / Version de l'éditeur:

<https://doi.org/10.1107/S0021889806009915>

Journal of Applied Crystallography, 39, June 3, pp. 458-465, 2006-06-01

NRC Publications Record / Notice d'Archives des publications de CNRC:

<https://nrc-publications.canada.ca/eng/view/object/?id=76ee2e28-04cb-4c8c-a154-be455d92b44f>

<https://publications-cnrc.canada.ca/fra/voir/objet/?id=76ee2e28-04cb-4c8c-a154-be455d92b44f>

Access and use of this website and the material on it are subject to the Terms and Conditions set forth at

<https://nrc-publications.canada.ca/eng/copyright>

READ THESE TERMS AND CONDITIONS CAREFULLY BEFORE USING THIS WEBSITE.

L'accès à ce site Web et l'utilisation de son contenu sont assujettis aux conditions présentées dans le site

<https://publications-cnrc.canada.ca/fra/droits>

LISEZ CES CONDITIONS ATTENTIVEMENT AVANT D'UTILISER CE SITE WEB.

Questions? Contact the NRC Publications Archive team at

PublicationsArchive-ArchivesPublications@nrc-cnrc.gc.ca. If you wish to email the authors directly, please see the
first page of the publication for their contact information.

Vous avez des questions? Nous pouvons vous aider. Pour communiquer directement avec un auteur, consultez la
première page de la revue dans laquelle son article a été publié afin de trouver ses coordonnées. Si vous n'arrivez
pas à les repérer, communiquez avec nous à PublicationsArchive-ArchivesPublications@nrc-cnrc.gc.ca.

<http://irc.nrc-cnrc.gc.ca>

SVDdiagnostic, a program to diagnose numerical conditioning of Rietveld refinements

NRCC-50566

Mercier, P.H.J.; Le Page, Y.; Whitfield, P.S.; Mitchell, L.D.

A version of this document is published in / Une version de ce document se trouve dans:
Journal of Applied Crystallography, v. 39, no. 3, June 2006, pp. 458-465 doi:
[10.1107/S0021889806009915](http://dx.doi.org/10.1107/S0021889806009915)

The material in this document is covered by the provisions of the Copyright Act, by Canadian laws, policies, regulations and international agreements. Such provisions serve to identify the information source and, in specific instances, to prohibit reproduction of materials without written permission. For more information visit <http://laws.justice.gc.ca/en/showtdm/cs/C-42>

Les renseignements dans ce document sont protégés par la Loi sur le droit d'auteur, par les lois, les politiques et les règlements du Canada et des accords internationaux. Ces dispositions permettent d'identifier la source de l'information et, dans certains cas, d'interdire la copie de documents sans permission écrite. Pour obtenir de plus amples renseignements : <http://lois.justice.gc.ca/fr/showtdm/cs/C-42>

National Research
Council Canada

Conseil national
de recherches Canada

Canada

***SVDdiagnostic*, a program to diagnose numerical conditioning of Rietveld refinements**

Patrick H. J. Mercier, Yvon Le Page, Pamela S. Whitfield and Lyndon D. Mitchell

Copyright © International Union of Crystallography

Author(s) of this paper may load this reprint on their own web site provided that this cover page is retained. Republication of this article or its storage in electronic databases or the like is not permitted without prior permission in writing from the IUCr.

SVDdiagnostic, a program to diagnose numerical conditioning of Rietveld refinements

Patrick H. J. Mercier,^{a*} Yvon Le Page,^a Pamela S. Whitfield^a and Lyndon D. Mitchell^b

^aInstitute for Chemical Process and Environmental Technology (ICPET), National Research Council of Canada, Ottawa, Ontario, Canada K1A 0R6, and ^bInstitute for Research in Construction (IRC), National Research Council of Canada, Ottawa, Ontario, Canada K1A 0R6. Correspondence e-mail: patrick.mercier@nrc-cnrc.gc.ca

Singular value decomposition (SVD) of the matrix of normal equations is used here both passively to assess numerical stability, and actively to troubleshoot problem refinements, singular or not. Such systems can then either be cured by rank reduction or solved with arbitrary-precision arithmetic carrying a number of digits known to be sufficient. SVD analysis provides objective information about such required rank reduction or number of digits. Pre-conditioning of the normal matrix is seen to decrease its condition number by many orders of magnitude in actual cases, illustrating its great practical usefulness. The methods and tools developed here have general applicability to diagnose problems with least squares, in particular ill-conditioned Rietveld refinements. Crystal-chemical and standard refinements described in the work by Mercier *et al.* [*J. Appl. Cryst.* (2006), **39**, 369–375] are shown to have similar numerical stability. The program *SVDdiagnostic* is freely available at <http://www.tothcanada.com>.

© 2006 International Union of Crystallography
Printed in Great Britain – all rights reserved

1. Introduction

Within the scope of a continuing study involving parameterization of the apatite structure type (Mercier *et al.*, 2005, 2006), we needed to assess the relative numerical stability of crystallographic *versus* crystal-chemical least-squares Rietveld refinements. The corresponding undertaking turned out to be ‘the tip of the iceberg’ with respect to numerical analysis in relation to inversion of square, symmetrical crystallographic least-squares matrices. We ended up with a logically disjoint manuscript that one of the referees wisely suggested be split into two homogeneous parts referring to each other. In this part, we lay the mathematical base for evaluation of numerical stability by singular value decomposition (SVD) in §2, implement it in §3, discuss it in §4 and draw conclusions in §5.

2. Numerical stability for matrix inversion

Authoritative literature about the analysis of numerical stability in matrix inversion based on SVD exists in books about numerical analysis, but we found very little crystallographic work where SVD is used, in line with remarks about SVD by Watkin (1994). Consequently, we developed the corresponding tools needed for assessing the conditioning of crystallographic least-squares matrices by SVD. The development below may seem too detailed to mathematically versed experts because the corresponding items are well known to mathematicians specialized in least squares. Our goal here is mostly to inform Rietveld practitioners and fellow crystallographers in a self-contained way on an apparently neglected topic of utmost importance for their everyday work, without being overly complex.

2.1. Condition number and error propagation in matrix inversion

Round-off error, truncation error, error propagation and numerical stability are standard topics in numerical analysis, discussed for example in ch. 1 of *Numerical Recipes in Fortran: the Art of Scientific Computing* (Press *et al.*, 1994), and references therein. ‘Round-off error’ relates to the limited number of bits in the representation of the mantissa of real numbers in a computer. ‘Truncation error’ refers to the effect of the round-off error on the sum of two numbers A and B with widely different magnitudes: the significant digits of the small number B trail those of the large number A and become chopped by the round-off. If, after the above truncation of $A + B$, a number with value close to A is subtracted from $A + B$, the result will be wrong by significantly more than the round-off error. Re-use of this imprecise number in further calculations causes ‘error propagation’. ‘Numerically stable’ algorithms avoid as much as possible the creation, and especially re-use of such imprecise numbers.

Inversion of square matrices can be performed in a number of ways. The most popular methods are Gauss elimination, LU decomposition, Cholesky decomposition, conjugate gradients, eigen decomposition, *etc.* The inverse of a non-singular matrix is unique. Those derivations are mathematically exact, meaning that elements of $\mathbf{M}^{-1}$ are calculable from the elements of $\mathbf{M}$ through a pre-defined sequence of additions, subtractions, multiplications and divisions with no iterations-until-convergence required in the extraction of an element of $\mathbf{M}^{-1}$. All those methods should therefore give identical results for $\mathbf{M}^{-1}$, but in fact they don’t, mostly because they are affected differently by error propagation. The severity of the differences between the results of the various methods is due to the ‘conditioning’ of the matrix, characterized by its ‘condition number’, CN, which is defined as the ratio of its largest over its smallest eigenvalue. Standard textbooks on numerical analysis (*e.g.* Press *et*

al., 1994) consider that the minimum number of decimal digits to be carried for 'safe' matrix inversion is $\log_{10}(\text{CN})$. Popular routines implementing the above inversion methods in fact considerably outperform this requirement in actual cases because they rearrange the matrix ahead of the decomposition, they process matrix elements in an order dictated by their magnitudes, or they 'pre-condition' the matrix prior to inversion. The effect of those precautions is to minimize error propagation very effectively in actual cases, but there is no guarantee that this will work in all cases. In order to be safe, we therefore discuss the numerical stability of least squares in terms of the conditioning number of the matrix of normal equations and the corresponding minimum number of digits to be carried in calculations.

2.2. Singular value decomposition of the matrix of normal equations

In its general form, SVD expresses any $M \times N$ matrix $\mathbf{A}$, $M \geq N$, as

$$\mathbf{A} = \mathbf{U} \cdot \mathbf{W} \cdot \mathbf{V}^T. \quad (1)$$

The columns of $\mathbf{U}$ are N orthogonal vectors with M components, each vector having length 1. $\mathbf{W}$ is a diagonal matrix of non-negative numbers. $\mathbf{V}^T$ is an orthonormal $N \times N$ matrix, meaning that its columns are orthogonal vectors, each with length 1, while its determinant is 1. It is a theorem (see Strang, 1980) that the decomposition is always possible, even if matrix $\mathbf{A}$ is singular. The decomposition is unique if the diagonal terms of $\mathbf{W}$ can be ranked in decreasing order. Mostly trivial permutations or linear combinations of rows or columns can appear when two or more diagonal terms of $\mathbf{W}$ are exactly equal or zero.

Applied to a symmetric matrix $\mathbf{A}$, SVD performs its decomposition into

$$\mathbf{A} = \mathbf{V} \cdot \mathbf{W} \cdot \mathbf{V}^T, \quad (2)$$

which amounts to the eigen decomposition of $\mathbf{A}$. The practical interest of SVD stems from the fact that it computes, in a numerically stable way, the 'bits and pieces' that are required to perform the inversion of matrix $\mathbf{A}$ in the simplest possible way, but without performing the inversion itself. Cases involving singularity or quasi-singularity are then diagnosed easily in an objective way. Quite often, the diagnostic itself constitutes a straightforward suggestion for a cure.

2.2.1. Interpretation of SVD for normal equations. The least-squares solution of an over-determined linear system of weighted observations $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$ is that of $\mathbf{A}^T \cdot \mathbf{A} \cdot \mathbf{x} = \mathbf{A}^T \cdot \mathbf{b}$. The square and symmetrical matrix $\mathbf{C} = \mathbf{A}^T \cdot \mathbf{A}$ is called the matrix of normal equations. Solutions are obtained as $\mathbf{x} = \mathbf{C}^{-1} \cdot \mathbf{A}^T \cdot \mathbf{b}$. Numerical problems associated with quasi-singularity of $\mathbf{C}$ can be discussed in terms of matrix $\mathbf{C}^{-1}$, called the variance-covariance matrix, as well as the correlation matrix derived from $\mathbf{C}^{-1}$, but this is a very complex topic that can be simplified as follows.

A rotation matrix $\mathbf{R}$ transforms the Euclidian reference system of axes constituted by refinement variables $\mathbf{x}$ into an orthonormal linear combination of them, $\mathbf{x}'$. The elements of $\mathbf{R}$ are the direction cosines of the new axes in the old axes, where an individual element R_{ij} is the projection of unit vector $\mathbf{u}_i'$ onto $\mathbf{u}_j$. Rotation thus transforms the numerical description of a linear system of equations $\mathbf{A} \cdot \mathbf{x}$ while retaining its physical meaning:

$$\mathbf{A}' \cdot \mathbf{x}' = \mathbf{A} \cdot \mathbf{x}.$$

Rotation $\mathbf{R}$ transforms $\mathbf{x}$ into $\mathbf{x}'$ through $\mathbf{x}' = \mathbf{R} \cdot \mathbf{x}$.

Conversely, we can write $\mathbf{x} = \mathbf{R}^T \cdot \mathbf{x}'$; therefore

$$\mathbf{A}' \cdot \mathbf{x}' = \mathbf{A} \cdot \mathbf{R}^T \cdot \mathbf{x}'.$$

Therefore the rotation operates on the transposed system as:

$$\mathbf{A}'^T = (\mathbf{A} \cdot \mathbf{R}^T)^T = \mathbf{R} \cdot \mathbf{A}^T.$$

As a consequence, matrix $\mathbf{C}$ of normal equations transforms as

$$\mathbf{C}' = \mathbf{A}'^T \cdot \mathbf{A}' = \mathbf{R} \cdot \mathbf{A}^T \cdot \mathbf{A} \cdot \mathbf{R}^T = \mathbf{R} \cdot \mathbf{C} \cdot \mathbf{R}^T.$$

As $\mathbf{R}^{-1} = \mathbf{R}^T$ for orthonormal transformations, the inverse of $\mathbf{C}$ transforms as

$$\mathbf{C}'^{-1} = [\mathbf{R} \cdot \mathbf{C} \cdot \mathbf{R}^T]^{-1} = [\mathbf{R}^T]^{-1} \cdot [\mathbf{R} \cdot \mathbf{C}]^{-1} = \mathbf{R} \cdot \mathbf{C}^{-1} \cdot \mathbf{R}^T.$$

In conclusion, rotation of both $\mathbf{C}$ and $\mathbf{C}^{-1}$ is performed through the same expression $\mathbf{M}' = \mathbf{R} \cdot \mathbf{M} \cdot \mathbf{R}^T$. SVD then provides both $\mathbf{D}$ and $\mathbf{R}$ for the diagonal expression $\mathbf{D}$ that $\mathbf{C}$ takes under the very special rotation $\mathbf{R}$ of the reference system made of the variables of the problem. It then becomes straightforward to compute $\mathbf{D}^{-1}$ and rotate it into $\mathbf{C}^{-1}$ through the opposite rotation $\mathbf{R}^T$, meaning $\mathbf{C}^{-1} = \mathbf{R}^T \cdot \mathbf{D}^{-1} \cdot \mathbf{R}$.

2.2.2. What difference is there with e.g. LU decomposition? It is well known that SVD computes $\mathbf{D}$ in a numerically stable way. Each diagonal element of $\mathbf{D}^{-1}$ is the inverse of the corresponding element of $\mathbf{D}$. Inversion of a diagonal matrix $\mathbf{D}$ with non-zero elements, each in the range 10^{1000} to 10^{-1000} , then presents no special numerical problem for 64-bit real arithmetic. Numerical problems analogous to those that plague LU decomposition of quasi-singular systems in fact arise during rotation by $\mathbf{R}^T$ of $\mathbf{D}^{-1}$ to give $\mathbf{C}^{-1} = \mathbf{R}^T \cdot \mathbf{D}^{-1} \cdot \mathbf{R}$. If implemented in this way, inversion through SVD is just as prone to numerical problems as if it were performed by LU decomposition, for example. The main difference arises from the fact that SVD provides a diagnostic for potential problems. This diagnostic often hints at a cure.

A somewhat better way to proceed would be to rotate the normal vector $\mathbf{v} = \mathbf{A}^T \cdot \mathbf{b}$ to the diagonal system, giving $\mathbf{v}' = \mathbf{A}'^T \cdot \mathbf{b} = \mathbf{R} \cdot \mathbf{A}^T \cdot \mathbf{b}$. Solutions of $\mathbf{D} \cdot \mathbf{x}' = \mathbf{v}'$ are $x'_i = v'_i / D_{ii}$, which is a numerically stable calculation in the diagonal reference system as long as D_{ii} is in the range $10^{\pm 1000}$. Back-rotation of $\mathbf{x}'$ into $\mathbf{x}$ with $\mathbf{x} = \mathbf{R}^T \cdot \mathbf{x}'$ is not quite as prone to error propagation as calculation of $\mathbf{C}^{-1}$. This approach may give numerically better solutions $\mathbf{x}$, but determining the variance/covariance of $\mathbf{x}$ requires knowledge of $\mathbf{C}^{-1}$ anyway.

2.2.3. SVD diagnostic with the condition number. A diagnostic is always possible. Error propagation occurs during the reverse rotation to obtain $\mathbf{C}^{-1}$ from $\mathbf{D}^{-1}$ when $\log_{10}$ of ratios of diagonal values in $\mathbf{D}^{-1}$ approach the number of digits carried by the computer. The condition number of the system is in wide use to discuss the numerical stability of solutions to specific least-squares problems, but to our knowledge it does not seem to have been applied to the discussion of the numerical stability of Rietveld refinements.

2.2.4. Rank-reduction cure. Singular vectors (eigenvectors), *i.e.* the unit vectors along the new axes that diagonalize matrix $\mathbf{C}$ in the Euclidian space defined by the variables used for stating the least-squares problem, are the columns of matrix $\mathbf{R}$. In the case of a zero singular value (zero eigenvalue), the corresponding column of $\mathbf{R}$ then spells out the relationship that is identically zero in terms of the variables for the problem. The easy cure for this is called 'rank reduction'. The inverse of the zero singular value, which should then be infinite at face value, is instead set to zero and the user then proceeds with rotation of $\mathbf{D}^{-1}$ back into $\mathbf{C}^{-1}$, calculated in the regular way. The effect of rank reduction is to withdraw the corresponding linear combination of variables from the least-squares process. Values and estimated standard deviations (e.s.d.'s) will nevertheless be obtained for all variables in the formulation of the problem because back-rotation of $\mathbf{D}$ with $\mathbf{R}^T$ populates all elements in matrix $\mathbf{C}^{-1}$, unless the corresponding singular vector had only one non-zero

component. In that case, rank reduction corresponds to elimination of a variable in the formulation of the problem. Rank reduction applied to SVD results then allows solution without reformulation of systems with zero determinants, *i.e.* singular systems. The extracted values and e.s.d.'s are correlated according to the linear relationship defined by the components of the singular vector with zero singular value.

The same treatment is applied to singular values with condition numbers that exceed the number of digits carried by the computer. Their corresponding values extracted by least squares would be meaningless anyway. Borderline cases where the conditioning ratio is only just poor require judgment, for instance, assignment of a sensible but somewhat arbitrary value to a variable whose refinement is then subsequently halted. In all cases, SVD is invaluable as a tool to point out objectively to the user what exact problems the least squares is facing, as shown in §4.1 below with examples.

Deciding whether a variable is superfluous or not requires a degree of judgment. The 'red flag' is that the ratio of the value for that variable divided by its e.s.d. is less than say 2 or 3, possibly indicating an insignificant numerical value, but that is not the whole story. The real criterion is that the 'physically sensible range' for the value divided by that e.s.d. must be quite small, say less than 0.5. The implication of such a fact is that, whatever the value might be within its physically meaningful range, its contribution to observations would be swamped by the noise in measurements anyway. It is then advisable to eliminate that variable from the model because its inclusion decreases the degree of over-determination of the system without a possibility to clarify any issue. Such tests of value divided by the e.s.d. are straightforward to implement, either with SVD or with the inverse of the normal matrix, but require knowledge of the normal vector. Unfortunately, not all Rietveld programs currently allow printing of the normal vector.

2.2.5. Byte-flooding cure. The conditioning number obtained from SVD tells us the minimum number of decimal digits required to solve the least-squares system without spoiling numerical results through error propagation. The brute-force byte-flooding cure then consists of implementing matrix inversion and matrix multiplication by a vector using arbitrary-precision arithmetic, carrying far more than that minimum number of digits.

2.2.6. Matrix pre-conditioning. Pre-conditioning normal matrices is a very simple but effective manipulation used by most inversion algorithms, those in *TOPAS* (Bruker AXS, 2005) and *GSAS* (Larson & Von Dreele, 2004) included. It can take several forms (see *e.g.* Hendrickson & Konnert, 1980; Guillot *et al.*, 2001), one of which is as follows.

Suppose that we can write $\mathbf{C} = \mathbf{P} \cdot \mathbf{B} \cdot \mathbf{P}$, where $\mathbf{C} = \mathbf{A}^T \cdot \mathbf{A}$ is the matrix of normal equations and $\mathbf{P}$ is an arbitrary matrix. We can then write $\mathbf{C}^{-1} = [\mathbf{B} \cdot \mathbf{P}]^{-1} \cdot \mathbf{P}^{-1} = \mathbf{P}^{-1} \cdot \mathbf{B}^{-1} \cdot \mathbf{P}^{-1}$. If we choose $\mathbf{P}$ to be the diagonal matrix with element $P_{ii} = (C_{ii})^{1/2}$, it follows that $\mathbf{P}^{-1}$ is the diagonal matrix with element $P_{ii}^{-1} = (C_{ii})^{-1/2}$. We accordingly have $B_{ij} = C_{ij}(C_{ii}C_{jj})^{-1/2}$, while $C_{ij}^{-1} = B_{ij}^{-1}(C_{ii}C_{jj})^{-1/2}$.

It is not immediately clear that anything has been gained by the above manipulation. Instead of inverting matrix $\mathbf{C}$, we calculate matrix $\mathbf{B}$ from $\mathbf{C}$, invert $\mathbf{B}$ and calculate $\mathbf{C}^{-1}$ from $\mathbf{B}^{-1}$. For reasons for which no rigorous general justification seems to exist, the condition number for matrix $\mathbf{B}$ is often many orders of magnitude lower than that for matrix $\mathbf{C}$ for actual least-squares problems, structure refinement in particular. Calculation of $\mathbf{B}^{-1}$ is then numerically much more stable than that of $\mathbf{C}^{-1}$. Calculation of $\mathbf{P}^{-1}$ gives rise to no error propagation. Calculation of $\mathbf{C}^{-1}$ in this indirect way is then numerically far more stable than if performed directly on $\mathbf{C}$, but again there is no firm proof that this will always be the case.

3. Implementation of tools for SVD analysis and accurate inversion

We have developed a program called *SVDdiagnostic* (distributed freely at <http://www.tothcanada.com>), under Microsoft Windows, that can read the matrix of normal equations from pristine *TOPAS* (V3.0; Bruker AXS, 2005) or *GSAS* (Larson & Von Dreele, 2004) output files. From this input, *SVDdiagnostic* computes the SVD of the matrix using an adaptation of the routine *SVDcmp* from Press *et al.* (1994), and produces two diagnostics from it. The first diagnostic is the condition number accompanied by its interpretation as good/fair/poor/borderline/ill conditioning of the matrix for 'safe' double-precision processing of the matrix. The second diagnostic is an output of the m largest and n smallest singular values for the matrix accompanied by the direction cosines of the corresponding singular vectors, each component labelled with the name of the corresponding refinement variable. The values of m and n are user-supplied *via* a dialog, together with input and output file names. With minimal practice, users quickly become adept at reading those direction cosines that indicate mutually perpendicular unit vectors in the Euclidian space of the refinement variables and that are best and worst determined by the least-squares processing of the observations supplied. In the case of zero or very small singular values, those direction cosines are the coefficients of the linear relation between the refinement variables that is exactly or approximately obeyed. If desired, *SVDdiagnostic* can also pre-condition the normal matrix according to the scheme in §2.2.6 and process it in the above way.

In a different program, we performed LU decomposition and solution of the linear system made from the matrix of normal equations and the vector using arbitrary-precision arithmetic for comparison with printed refinement results. Actual Rietveld problems are not singular, but they are routinely so ill-conditioned that the reliability of double-precision matrix inversion is in doubt. Use of standard inversion algorithms with 16-byte arithmetic would reliably solve most Rietveld systems, but most Fortran Windows compilers allow double, but not quadruple precision. We then adapted arbitrary-precision arithmetic routines from ch. 20 in *Numerical Recipes in Fortran* (Press *et al.*, 1994) and assembled a library to perform addition, subtraction, multiplication, division and square-root calculation for arbitrary-precision real numbers. We then modified the routines *ludcmp* and *lubksb* from ch. 2 in *Numerical Recipes in Fortran* (Press *et al.*, 1994) to link with that library, thus performing LU decomposition and matrix inversion while carrying any desired number of digits. Using 36-byte (equivalent to 86 decimal digits) representation for the mantissa and 4-byte integers for the exponent, allowing exponents up to a billion if needed, this gave printouts of reliable values for the elements of the inverse matrix $\mathbf{C}^{-1}$, as well as for the solution vector $\mathbf{x}$ derived from the vector printout by the Rietveld package. In its current implementation, that library can perform the five basic operations on real numbers with up to 255 bytes in their mantissa, *i.e.* more than 600 significant decimal digits. Today's PCs are so fast that unoptimized 36-byte inversion of a 33×33 matrix takes about 4 s on a 3 GHz AMD Athlon64 processor operated in 32-bit mode, including the ASCII input, as well as the conversion and output of the results.

4. Results and discussion

The reason why we undertook the present study is the comparison of the numerical stability of the standard crystallographic refinement from Mercier *et al.* (2006) with its crystal-chemical counterpart. We

Table 1Examples of the output of *SVDdiagnostic* for problematic cases of Rietveld refinements of $P6_3/m$ fluorapatite.(a) Standard crystallographic refinement with *TOPAS*; $R_{wp} = 8.673\%$, GOF = 1.550.

UNPROCESSED normal matrix														
Condition number for matrix of normal equations = 0.785E+13														
Error propagation is likely to spoil 13 trailing decimal digits out of probably 14.														
Problem is ill-conditioned for double-precision matrix inversion unless error propagation is well taken care of.														
Use eigenvectors for small eigenvalues below to diagnose quasi singularity.														
Eigenvectors for cr11.out ranked according to eigenvalues are printed as columns below														
Eigenvector	#:	1..	21	22	23	24	25	26	27	28	29	30	31	32
zero_error	1:	0.000..	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
divslit	2:	0.000..	-0.001	0.001	-0.002	0.000	-0.001	0.001	-0.001	0.001	-0.001	0.000	0.000	-0.002
bkg1040163200	3:	0.000..	0.028	-0.018	0.655	0.152	0.475	0.098	0.374	-0.008	0.334	0.079	0.228	-0.031
bkg1040163201	4:	0.000..	0.010	-0.008	-0.032	0.063	0.054	0.342	0.163	0.483	-0.150	0.618	-0.459	-0.007
bkg1040163202	5:	0.000..	-0.014	0.019	-0.555	-0.167	-0.075	0.093	0.328	0.123	0.526	0.223	0.448	0.021
bkg1040163203	6:	0.000..	-0.012	0.033	0.034	-0.360	0.026	-0.634	0.235	-0.320	0.187	0.313	-0.417	-0.020
bkg1040163204	7:	0.000..	0.001	0.023	0.401	-0.027	-0.487	-0.290	-0.355	0.275	0.053	0.405	0.385	0.016
bkg1040163205	8:	0.000..	0.006	0.000	-0.002	0.598	-0.337	0.151	-0.148	-0.323	0.523	0.067	-0.317	-0.008
bkg1040163206	9:	0.000..	0.004	-0.007	-0.276	0.391	0.482	-0.201	-0.289	-0.285	-0.235	0.455	0.266	0.000
bkg1040163207	10:	0.000..	-0.008	0.008	-0.003	-0.372	0.391	0.135	-0.662	0.150	0.446	-0.032	-0.170	-0.003
bkg1040163208	11:	0.000..	-0.003	0.008	0.148	-0.404	-0.163	0.544	-0.042	-0.604	-0.151	0.298	0.119	0.001
p1SCALE	12:	1.000..	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
CrystalliteSize13:	0.000..	0.000	-0.001	-0.026	-0.006	-0.024	0.003	-0.014	0.012	-0.007	-0.002	0.023	-0.999	
alat	14:	0.000..	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
clat	15:	0.000..	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
A1z	16:	0.000..	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
beqA1	17:	0.000..	-0.088	0.011	0.007	0.007	0.011	0.005	0.007	0.007	0.002	0.005	0.000	-0.001
A2x	18:	0.000..	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
A2y	19:	0.000..	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
beqA2	20:	0.000..	-0.061	0.005	0.007	0.007	0.009	0.013	0.003	0.006	-0.001	0.003	0.000	0.000
beqB	21:	0.000..	0.061	0.008	0.017	0.007	0.007	0.017	0.007	0.011	0.001	0.001	0.000	-0.001
Bx	22:	0.000..	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
By	23:	0.000..	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
beqO	24:	0.000..	-0.992	0.016	0.024	0.021	0.011	0.010	0.008	0.004	-0.003	0.002	0.001	-0.001
O1x	25:	0.000..	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
O1y	26:	0.000..	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
O2x	27:	0.000..	-0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
O2y	28:	0.000..	-0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
O3x	29:	0.000..	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
O3y	30:	0.000..	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
O3z	31:	0.001..	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
beqF	32:	0.000..	0.018	0.999	0.008	0.027	0.022	0.024	0.005	0.007	-0.016	-0.017	0.000	-0.001
Eigenvalues	:	1.78E+12..	1.97E+03	2.68E+02	4.16E+01	3.56E+01	3.01E+01	2.30E+01	2.08E+01	1.44E+01	6.69E+00	5.11E+00	1.15E+00	2.26E-01

(b) Crystal-chemical refinement with *TOPAS*; $R_{wp} = 9.231$, GOF = 1.650.

UNPROCESSED normal matrix														
Condition number for matrix of normal equations = 0.790E+13														
Error propagation is likely to spoil 13 trailing decimal digits out of probably 14.														
Problem is ill-conditioned for double-precision matrix inversion unless error propagation is well taken care of.														
Use eigenvectors for small eigenvalues below to diagnose quasi singularity.														
Eigenvectors for ch13.out ranked according to eigenvalues are printed as columns below														
Eigenvector #	#:	1..	17	18	19	20	21	22	23	24	25	26	27	28
dA101	1:	0.000...	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
DeltaA10	2:	0.000...	-0.001	-0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
deltaA1	3:	0.000...	0.027	-0.143	0.000	0.003	-0.006	0.005	0.008	-0.003	-0.004	-0.001	0.000	0.000
alphaA1	4:	0.000...	0.009	-0.089	0.001	0.001	-0.001	0.004	0.003	-0.001	0.000	0.000	0.000	0.000
SB0	5:	0.000...	0.000	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
tauOB0	6:	0.000...	-0.005	0.276	-0.001	-0.007	0.006	-0.017	-0.009	0.001	0.003	-0.001	0.001	0.000
rhoA2	7:	0.000...	0.001	0.005	0.000	0.000	0.000	-0.001	0.000	0.000	0.000	0.000	0.000	0.000
alphaA2	8:	0.000...	-0.007	-0.012	-0.002	-0.008	-0.001	-0.001	0.002	0.000	0.002	0.000	0.000	0.000
dA203	9:	0.000...	0.000	-0.006	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
phi03A203	10:	0.000...	-0.018	0.945	-0.012	0.001	0.005	-0.022	-0.017	0.009	0.000	-0.003	0.001	-0.002
zero_error	11:	0.000...	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
divslit	12:	0.000...	0.001	0.000	-0.002	0.000	0.001	-0.001	0.001	0.001	0.001	0.000	0.000	-0.002
bkg805787520	13:	0.000...	-0.017	0.002	0.650	0.170	-0.488	-0.123	-0.344	-0.011	-0.339	0.082	0.228	-0.032
bkg805787521	14:	0.000...	-0.006	-0.014	-0.033	0.063	-0.058	-0.347	-0.143	0.481	0.151	0.619	-0.460	-0.006
bkg805787522	15:	0.000...	0.020	-0.016	-0.544	-0.186	0.076	-0.108	-0.328	0.125	-0.526	0.225	0.448	0.020
bkg805787523	16:	0.000...	0.030	0.017	0.059	-0.363	-0.018	0.617	-0.265	-0.321	-0.194	0.311	-0.417	-0.019
bkg805787524	17:	0.000...	0.020	0.015	0.401	0.000	0.495	0.305	0.335	0.276	-0.049	0.401	0.385	0.017
bkg805787525	18:	0.000...	0.003	0.002	-0.032	0.607	0.328	-0.140	0.155	-0.319	-0.522	0.064	-0.317	-0.009
bkg805787526	19:	0.000...	-0.006	0.015	-0.297	0.367	-0.481	0.226	0.281	-0.286	0.237	0.455	0.265	0.000
bkg805787527	20:	0.000...	0.006	0.013	-0.002	-0.377	-0.371	-0.095	0.679	0.158	-0.442	-0.034	-0.169	-0.003
bkg805787528	21:	0.000...	0.007	-0.006	0.159	-0.392	0.171	-0.545	0.080	-0.603	0.146	0.298	0.117	0.001
p1SCALE	22:	1.000...	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
p1CS	23:	0.000...	-0.001	-0.002	-0.025	-0.007	0.024	-0.002	0.013	0.013	0.008	-0.002	0.022	-0.999
beqA1	24:	0.000...	0.014	-0.015	0.009	0.006	-0.011	-0.005	-0.007	0.008	-0.002	0.005	-0.001	-0.001
beqA2	25:	0.000...	0.003	0.002	0.007	0.006	-0.009	-0.013	-0.002	0.006	0.001	0.003	0.000	0.000
beqO	26:	0.000...	0.006	-0.004	0.022	0.020	-0.013	-0.009	-0.005	0.004	0.003	0.002	0.001	-0.001
beqB	27:	0.000...	0.004	0.010	0.016	0.008	-0.007	-0.017	-0.004	0.011	0.000	0.001	0.000	-0.001
beqX	28:	0.000...	0.998	0.022	0.009	0.023	-0.022	-0.021	-0.003	0.008	0.017	-0.016	0.000	-0.002
Eigenvalues	:	1.84E+12..	2.86E+02	1.26E+02	4.22E+01	3.58E+01	3.02E+01	2.33E+01	2.12E+01	1.45E+01	6.76E+00	5.15E+00	1.16E+00	2.33E-01

Table 1 (continued)

(c) Standard crystallographic refinement with *GSAS*; $R_{wp} = 9.44\%$, GOF = 2.853.

UNPROCESSED normal matrix														
Condition number for matrix of normal equations = 0.223E+18														
Error propagation is likely to spoil ALL digits in some elements of the normal matrix.														
System is singular for double-precision matrix inversion unless error propagation is well taken care of.														
Use eigenvectors below to diagnose singularity.														
Eigenvectors for best-fit-so-far.txt ranked according to eigenvalues are printed as columns below														
Eigenvector #:	#:	1 ..	27	28	29	30	31	32	33	34	35	36	37	38
1 1X	1:	0.000 ..	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
1 1Y	2:	0.000 ..	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
1 1UI50	3:	0.000 ..	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
1 2X	4:	0.000 ..	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
1 2Y	5:	0.000 ..	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
1 2UI50	6:	0.000 ..	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
1 3X	7:	0.000 ..	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
1 3Y	8:	0.000 ..	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
1 3Z	9:	0.000 ..	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
1 3UI50	10:	0.000 ..	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
1 4UI50	11:	0.000 ..	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
1 5X	12:	0.000 ..	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
1 5Y	13:	0.000 ..	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
1 5UI50	14:	0.000 ..	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
1 6Z	15:	0.000 ..	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
1 6UI50	16:	0.000 ..	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
1 7X	17:	0.000 ..	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
1 7Y	18:	0.000 ..	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
1 7UI50	19:	0.000 ..	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
1 RM11	20:	0.978 ..	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
1 RM33	21:	0.208 ..	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
1 HSC1	22:	0.000 ..	0.034	-0.160	-0.121	0.320	0.383	-0.132	0.229	-0.015	0.035	0.004	-0.003	0.000
ZERO 1	23:	0.000 ..	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.000	-0.001	0.000	0.000
1 1PF 1	24:	0.000 ..	0.055	-0.142	0.144	-0.130	-0.240	-0.536	0.349	-0.013	-0.513	0.006	0.002	0.000
1 1PF 2	25:	0.000 ..	0.032	-0.055	0.079	-0.094	-0.169	-0.241	0.119	0.020	0.820	0.000	-0.008	0.001
1 1PF 3	26:	0.000 ..	-0.049	0.102	-0.096	0.094	0.072	0.185	-0.123	-0.005	-0.236	-0.002	0.003	0.000
1 1PF 4	27:	0.000 ..	0.018	-0.040	0.036	-0.030	-0.021	-0.136	0.113	-0.001	0.022	0.002	-0.001	0.000
1 1PF 5	28:	0.000 ..	-0.016	0.039	-0.052	0.045	0.121	0.390	-0.290	0.002	-0.057	-0.005	0.003	0.000
1BK 1	29:	0.000 ..	-0.305	0.473	-0.361	0.441	-0.171	-0.269	-0.013	-0.225	0.025	0.183	-0.226	0.319
1BK 2	30:	0.000 ..	0.027	0.160	-0.108	-0.016	-0.316	-0.181	-0.364	0.123	-0.023	0.458	0.218	-0.599
1BK 3	31:	0.000 ..	0.047	-0.436	0.197	-0.103	-0.174	-0.035	-0.431	-0.145	-0.027	0.419	-0.178	0.523
1BK 4	32:	0.000 ..	-0.039	-0.382	0.047	0.306	0.402	-0.103	0.015	-0.322	0.037	0.313	-0.218	-0.397
1BK 5	33:	0.000 ..	0.413	0.334	0.038	-0.105	0.468	-0.131	0.083	0.237	0.012	0.465	0.285	0.278
1BK 6	34:	0.000 ..	-0.089	0.404	0.330	-0.415	0.098	0.160	0.196	-0.248	-0.010	0.213	-0.542	-0.158
1BK 7	35:	0.000 ..	-0.637	-0.164	-0.128	-0.235	0.007	0.233	0.354	-0.057	0.016	0.355	0.424	0.080
1BK 8	36:	0.000 ..	0.228	-0.223	-0.608	-0.148	-0.145	0.194	0.255	-0.008	0.375	0.205	-0.444	-0.023
1BK 9	37:	0.000 ..	0.469	0.029	0.092	0.294	-0.392	0.392	0.349	-0.427	0.009	0.160	0.213	0.005
1BK10	38:	0.000 ..	-0.196	0.009	0.507	0.463	-0.146	0.162	0.153	0.604	-0.007	0.167	-0.156	0.006
Eigenvalues	:	5.53E+14 ..	4.01E+01	3.89E+01	3.42E+01	3.06E+01	2.24E+01	1.68E+01	1.12E+01	6.77E+00	2.20E+00	6.22E-01	5.07E-02	2.48E-03

also practiced on a variety of recent Rietveld cases as tests, but we only report on the above item.

4.1. Evaluation of numerical stability using SVD

Examples of the output of *SVDdiagnostic* for problematic cases of crystallographic or crystal-chemical refinements are given in Table 1. The global diagnostic is given numerically in the form of the condition number (which dictates the number of decimal digits required for safe inversion of the least-squares matrix) and in a brief sentence. The smallest eigenvalues at the bottom right of the printouts in Tables 1(a)–1(c) are the main culprit for the poor conditioning. The column of numbers above the 21st largest eigenvalue in Table 1(a) is the corresponding eigenvector whose modulus must be 1 by definition. The value of essentially 1 in front of variable beqF in that column tells us that this column is mostly about beqF. The $\sim 10^{10}$ ratio of the first eigenvalue to the 21st eigenvalue tells us that this is a weak, but by no means redundant refinement variable. Similar remarks also apply to the column above the 17th eigenvalue in Table 1(b), which shows variable phi03A203 (*i.e.* $\varphi_{03-AII-O3}$; see Table 1 of Mercier *et al.*, 2006) to be the main component of this eigenvector.

The significant components for eigenvectors giving rise to low eigenvalues in *TOPAS*, from the 22nd to 32nd eigenvalues in Table 1(a) and from the 18th to 28th eigenvalues in Table 1(b), tell us that these eigenvalues are all concerned with linear combinations of variables having labels starting with bkg. These are the coefficients of a Chebyshev polynomial expansion of the background with a user-defined number of terms refined by *TOPAS*, all together or not at all. Being unable to sort out those parameters, most of them redundant

or insignificant, we then accepted a given set of their numerical values and did not perform further refinement on them for both kinds of refinement. Variables describing profiles were also found to give rise to small eigenvalues (Table 2) and were thus also fixed. In the same way, occupancy factors and corresponding isotropic thermal parameters were read to be redundant on eigenvectors. We accordingly fixed occupancies to full occupancy in both refinements, while refining only the isotropic thermal factors. This did not alter the values of the agreement factors (R_{wp}) and goodness-of-fit (GOF), or the refined values, but it brought the condition number for both kinds of refinement close to the limit between good and fair conditioning of the matrix (Table 3). This shows that both approaches are similarly conditioned, with an apparent slight advantage with the standard refinement. As argued by Mercier *et al.* (2006), the latter advantage is clearly reversed by the much lower e.s.d.'s of the crystal-chemical refinement.

As the least squares is ill-conditioned but not singular, another way to proceed would have been to use more digits than the minimum of 13–17 digits required (Tables 2 and 3) in the derivation of the solution for these particular cases (as implemented and discussed in §§4.2–4.3), but we could not do that because there is currently no provision to print the normal equations vector in *TOPAS*.

4.2. Actual numerical stability of the Rietveld procedure

This item should be examined from two perspectives: with and without pre-conditioning. The information derived is different and useful in each case.

4.2.1. Diagnostic with preconditioning. The important output of *SVDdiagnostic* in this case is the condition number of matrix **B** from §2.2.6. Its value is usually around 10^2 – 10^5 , which is reassuring. It tells us that double-precision arithmetic is sufficient for numerically stable inversion of **B**, and therefore of **C** in this indirect way. A condition number with value greater than 10^{10} – 10^{12} would be abnormally high, indicating inversion trouble ahead. It could indicate that pre-conditioning did not perform well on this problem, but most likely it would

point to a suspicious model with *e.g.* pseudo-translations in it, making it singular or quasi-singular. In this case where the condition number of matrix **B** would be greater than 10^{10} , the singular vectors defining the linear combination of variables that is identically zero should be read from the SVD analysis of the normal matrix **C**. This is because the variables of **C** are the variables of the problem, while those of the pre-conditioned matrix **B** are rescaled variables, each with its own scaling factor.

Table 2

Examples of ill-conditioned Rietveld refinements.

All background parameters: fixed. Profile parameters: divslit fixed, CrystalliteSize refined.

(a) Standard crystallographic refinement with *TOPAS*; $R_{wp} = 8.696\%$, GOF = 1.553.

UNPROCESSED normal matrix												
Condition number for matrix of normal equations = 0.342E+13												
Error propagation is likely to spoil 13 trailing decimal digits out of probably 14.												
Problem is ill-conditioned for double-precision matrix inversion unless error propagation is well taken care of.												
Use eigenvectors for small eigenvalues below to diagnose quasi singularity.												
Eigenvectors for cr08.out ranked according to eigenvalues are printed as columns below												
Eigenvector	#:	1	2	3	4	5 ..	17	18	19	20	21	22
zero_error	1:	0.000	0.159	-0.015	0.004	-0.010 ..	0.000	0.000	0.000	0.000	0.000	0.000
p1SCALE	2:	1.000	0.000	0.000	0.000	0.000 ..	0.000	0.000	0.000	0.000	0.000	0.000
CrystalliteSize	3:	0.000	0.000	0.000	0.000	0.000 ..	0.000	0.000	0.000	0.000	0.001	-1.000
alat	4:	0.000	-0.834	-0.538	-0.001	-0.001 ..	0.000	0.000	0.000	0.000	0.000	0.000
clat	5:	0.000	-0.529	0.843	-0.001	0.000 ..	0.000	0.000	0.000	0.000	0.000	0.000
Alz	6:	0.000	0.000	0.000	-0.008	-0.075 ..	0.000	0.000	-0.001	0.000	0.000	0.000
begA1	7:	0.000	0.000	0.000	0.000	-0.001 ..	-0.313	0.745	-0.582	0.091	-0.012	0.000
A2x	8:	0.000	0.000	-0.001	0.583	-0.001 ..	0.001	0.001	0.001	0.000	0.000	0.000
A2y	9:	0.000	0.003	-0.001	-0.808	-0.036 ..	0.000	0.000	0.000	0.000	0.000	0.000
begA2	10:	0.000	0.000	0.000	0.000	0.000 ..	-0.927	-0.369	0.035	0.061	-0.005	0.000
begB	11:	0.000	0.000	0.000	0.000	0.001 ..	-0.195	0.555	0.806	-0.063	-0.009	0.000
Bx	12:	0.000	0.001	0.004	-0.022	0.693 ..	0.000	0.001	-0.001	0.000	0.000	0.000
By	13:	0.000	-0.001	0.005	0.030	-0.701 ..	0.001	0.000	0.001	0.000	0.000	0.000
begO	14:	0.000	0.000	0.000	0.000	0.000 ..	0.073	-0.010	0.102	0.992	-0.018	0.000
Olx	15:	0.000	0.000	0.000	0.004	-0.002 ..	0.000	0.001	-0.001	0.000	0.000	0.000
Oly	16:	0.000	0.000	-0.001	0.012	0.049 ..	-0.001	0.000	0.000	0.000	0.000	0.000
O2x	17:	0.000	0.000	0.001	0.020	0.016 ..	-0.001	0.001	-0.002	0.002	0.000	0.000
O2y	18:	0.000	0.000	0.000	-0.070	-0.071 ..	-0.001	0.000	-0.001	0.001	0.000	0.000
O3x	19:	0.000	-0.001	0.001	-0.007	0.068 ..	0.000	0.000	0.000	-0.001	0.000	0.000
O3y	20:	0.000	0.001	-0.001	-0.017	-0.085 ..	0.000	0.000	-0.001	-0.001	0.000	0.000
O3z	21:	0.001	0.000	0.001	-0.005	0.045 ..	0.001	-0.001	0.000	0.000	0.000	0.000
begF	22:	0.000	0.000	0.000	0.000	0.000 ..	0.009	-0.011	-0.002	-0.019	-1.000	-0.001
Eigenvalues	:	1.80E+12	7.04E+09	2.62E+09	3.69E+08	2.17E+08 ..	9.59E+03	5.13E+03	3.58E+03	2.03E+03	2.73E+02	5.27E-01

(b) Crystal-chemical refinement with *TOPAS*; $R_{wp} = 9.251\%$, GOF = 1.652.

UNPROCESSED normal matrix												
Condition number for matrix of normal equations = 0.347E+13												
Error propagation is likely to spoil 13 trailing decimal digits out of probably 14.												
Problem is ill-conditioned for double-precision matrix inversion unless error propagation is well taken care of.												
Use eigenvectors for small eigenvalues below to diagnose quasi singularity.												
Eigenvectors for ch11.out ranked according to eigenvalues are printed as columns below												
Eigenvector #	#:	1	2	3	4	5 ..	13	14	15	16	17	18
dA1O1	1:	0.000	0.718	-0.212	0.009	-0.613 ..	-0.001	0.000	0.000	0.000	0.000	0.000
DeltaA1O	2:	0.000	0.434	-0.129	0.014	0.233 ..	0.002	0.001	0.000	0.001	-0.002	0.000
deltaA1	3:	0.000	0.003	-0.001	0.000	0.008 ..	-0.157	-0.071	0.502	-0.036	-0.077	0.000
alphaA1	4:	0.000	0.015	-0.005	0.001	0.023 ..	0.005	0.020	0.203	-0.005	-0.040	0.000
dBO	5:	0.000	0.536	0.522	0.037	0.552 ..	0.001	-0.001	0.000	0.000	0.001	0.000
tauOBO	6:	0.000	0.006	-0.009	0.000	0.016 ..	-0.049	0.007	-0.813	0.032	0.172	0.000
rhoA2	7:	0.000	0.000	0.000	0.000	-0.195 ..	-0.001	-0.003	-0.003	-0.001	0.004	0.000
alphaA2	8:	0.000	0.000	0.000	0.000	0.000 ..	0.813	0.477	0.047	0.006	-0.011	0.000
dA2O3	9:	0.000	-0.089	0.815	0.018	-0.476 ..	-0.001	0.001	-0.009	0.000	-0.007	0.000
phiO3A2O3	10:	0.000	-0.001	0.006	0.000	-0.004 ..	0.010	-0.006	0.194	0.035	0.980	0.000
zero_error	11:	0.000	-0.031	-0.030	0.999	-0.010 ..	0.000	0.000	0.000	0.000	0.000	0.000
p1SCALE	12:	1.000	0.000	0.000	0.000	0.000 ..	0.000	0.000	0.000	0.000	0.000	0.000
CrystalliteSize	13:	0.000	0.000	0.000	0.000	0.000 ..	0.000	0.000	0.000	0.000	0.000	1.000
begA1	14:	0.000	0.000	0.000	0.000	0.001 ..	0.016	-0.070	0.062	-0.017	-0.014	0.000
begA2	15:	0.000	0.000	0.000	0.000	0.000 ..	0.151	-0.003	-0.006	-0.004	0.002	0.000
begO	16:	0.000	0.000	0.000	0.000	-0.001 ..	0.456	-0.873	-0.024	-0.009	-0.003	0.000
begB	17:	0.000	0.000	0.000	0.000	0.000 ..	-0.284	-0.004	-0.034	-0.002	0.011	0.000
begX	18:	0.000	0.000	0.000	0.000	0.000 ..	0.005	0.014	-0.039	-0.998	0.043	0.000
Eigenvalues	:	1.87E+12	1.57E+11	1.95E+10	6.60E+07	8.96E+06 ..	2.65E+03	2.11E+03	5.97E+02	2.93E+02	7.34E+01	5.39E-01

computer programs

Table 3

Final refinements obtained after diagnostic by SVD.

All background and profile parameters needed to be fixed.

(a) Standard crystallographic refinement with *TOPAS*; $R_{wp} = 8.695\%$, GOF = 1.553.

UNPROCESSED normal matrix												
Condition number for matrix of normal equations = 0.659E+10												
Error propagation is likely to spoil 10 trailing decimal digits out of probably 14.												
Problem poorly conditioned for double-precision matrix inversion unless error propagation is well taken care of.												
Use eigenvectors for small eigenvalues below to diagnose quasi singularity.												
Eigenvectors for cr10.out ranked according to eigenvalues are printed as columns below												
Eigenvector	#:	1	2	3	4	5	6 ..	17	18	19	20	21
zero_error	1:	0.000	0.159	-0.015	0.004	-0.010	0.007 ..	0.000	0.000	0.000	0.000	0.000
pISCALE	2:	1.000	0.000	0.000	0.000	0.000	0.000 ..	0.000	0.000	0.000	0.000	0.000
alat	3:	0.000	-0.834	-0.538	-0.001	-0.001	0.001 ..	0.000	0.000	0.000	0.000	0.000
clat	4:	0.000	-0.529	0.843	-0.001	0.000	0.005 ..	0.000	0.000	0.000	0.000	0.000
A1z	5:	0.000	0.000	0.000	-0.008	-0.075	-0.009 ..	0.000	0.000	-0.001	0.000	0.000
begA1	6:	0.000	0.000	0.000	0.000	-0.001	0.000 ..	-0.313	0.745	-0.582	-0.091	-0.012
A2x	7:	0.000	0.000	-0.001	0.583	-0.001	0.663 ..	0.001	0.001	0.001	0.000	0.000
A2y	8:	0.000	0.003	-0.001	0.000	-0.036	0.467 ..	0.000	0.000	0.000	0.000	0.000
begA2	9:	0.000	0.000	0.000	0.000	0.000	0.001 ..	-0.927	-0.369	0.035	-0.061	-0.005
begB	10:	0.000	0.000	0.000	0.000	0.001	0.000 ..	-0.195	0.556	0.806	0.064	-0.009
Bx	11:	0.000	0.001	0.004	-0.022	0.693	-0.180 ..	0.000	0.001	-0.001	0.000	0.000
By	12:	0.000	-0.001	0.005	0.030	-0.701	-0.282 ..	0.001	0.000	0.001	0.000	0.000
begO	13:	0.000	0.000	0.000	0.000	0.000	0.000 ..	0.073	-0.010	0.103	-0.992	-0.018
O1x	14:	0.000	0.000	0.000	0.004	-0.002	0.057 ..	0.000	0.001	-0.001	0.000	0.000
O1y	15:	0.000	0.000	-0.001	0.012	0.049	-0.074 ..	-0.001	0.000	0.000	0.000	0.000
O2x	16:	0.000	0.000	0.001	0.020	0.016	0.018 ..	-0.001	0.001	-0.002	-0.002	0.000
O2y	17:	0.000	0.000	0.000	-0.070	-0.071	0.019 ..	-0.001	0.000	-0.001	-0.001	0.000
O3x	18:	0.000	-0.001	0.001	-0.007	0.068	-0.342 ..	0.000	0.000	0.000	0.001	0.000
O3y	19:	0.000	0.001	-0.001	-0.017	-0.085	0.321 ..	0.000	0.000	-0.001	0.001	0.000
O3z	20:	0.001	0.000	0.001	-0.005	0.045	-0.024 ..	0.001	-0.001	0.000	0.000	0.000
begF	21:	0.000	0.000	0.000	0.000	0.000	0.000 ..	0.009	-0.011	-0.002	0.019	-1.000
Eigenvalues	:	1.80E+12	7.04E+09	2.62E+09	3.69E+08	2.17E+08	9.32E+07 ..	9.59E+03	5.13E+03	3.58E+03	2.03E+03	2.73E+02

(b) Crystal-chemical refinement with *TOPAS*; $R_{wp} = 9.231\%$, GOF = 1.648.

UNPROCESSED normal matrix												
Condition number for matrix of normal equations = 0.221E+11												
Error propagation is likely to spoil 10 trailing decimal digits out of probably 14.												
Problem is ill-conditioned for double-precision matrix inversion unless error propagation is well taken care of.												
Use eigenvectors for small eigenvalues below to diagnose quasi singularity.												
Eigenvectors for chl2.out ranked according to eigenvalues are printed as columns below												
Eigenvector	#:	1	2	3	4	5	6 ..	13	14	15	16	17
dA1O1	1:	0.000	0.718	-0.213	0.009	-0.610	0.255 ..	-0.001	0.000	0.000	0.000	0.000
DeltaA1O	2:	0.000	0.433	-0.129	0.014	0.221	-0.840 ..	0.002	0.001	0.000	-0.001	0.002
deltaA1	3:	0.000	0.003	-0.001	0.000	0.008	-0.005 ..	-0.161	-0.074	0.500	0.028	0.086
alphaA1	4:	0.000	0.016	-0.005	0.001	0.024	0.004 ..	0.007	0.017	0.195	0.000	0.051
dB0	5:	0.000	0.537	0.521	0.037	0.557	0.295 ..	0.001	-0.001	0.000	0.000	-0.001
tauOB0	6:	0.000	0.006	-0.009	0.000	0.016	0.001 ..	-0.051	0.003	-0.811	-0.020	-0.185
rhoA2	7:	0.000	0.000	0.000	0.000	-0.194	-0.279 ..	-0.001	-0.003	-0.003	0.001	-0.004
alphaA2	8:	0.000	0.000	0.000	0.000	0.000	-0.001 ..	0.814	0.472	0.046	-0.006	0.010
dA2O3	9:	0.000	-0.088	0.816	0.018	-0.480	-0.255 ..	-0.001	0.001	-0.009	0.000	0.007
phlO3A2O3	10:	0.000	-0.001	0.006	0.000	-0.004	-0.003 ..	0.008	-0.003	0.211	-0.017	-0.977
zero_error	11:	0.000	-0.031	-0.030	0.999	-0.010	0.003 ..	0.000	0.000	0.000	0.000	0.000
pISCALE	12:	1.000	0.000	0.000	0.000	0.000	0.000 ..	0.000	0.000	0.000	0.000	0.000
begA1	13:	0.000	0.000	0.000	0.000	0.001	0.005 ..	0.021	-0.075	0.066	0.016	0.015
begA2	14:	0.000	0.000	0.000	0.000	0.000	0.002 ..	0.150	-0.003	-0.008	0.005	-0.002
begO	15:	0.000	0.000	0.000	0.000	-0.001	-0.001 ..	0.451	-0.875	-0.023	0.009	0.001
begB	16:	0.000	0.000	0.000	0.000	0.000	-0.003 ..	-0.287	-0.001	-0.032	0.002	-0.013
begX	17:	0.000	0.000	0.000	0.000	0.000	-0.001 ..	0.004	0.014	-0.027	0.999	-0.023
Eigenvalues	:	1.87E+12	1.57E+11	1.95E+10	6.59E+07	8.98E+06	3.09E+06 ..	2.66E+03	2.12E+03	5.97E+02	2.91E+02	8.47E+01

4.2.2. Diagnostic on normal matrix. From printouts of SVD decompositions of Rietveld refinements for a variety of problems, we have observed that they are rarely well conditioned over the range of crystallographic variables. The cell data variables and especially the scale factor contribute overwhelmingly to the normal equations, creating huge eigenvalues with *TOPAS* (see Table 1b). Structural variables (*i.e.* atomic positions and thermal factors) affect integrated intensities. They then affect profile points through combination with profile-fitting variables. At SVD time, the eigenvalues for structure variables then appear as midgets in front of the heavy-weight variables such as scale factor and cell data. Variables describing profiles and backgrounds can give rise to even smaller eigenvalues (Tables 1–

3). In itself, this does not constitute a problem because the systems are not singular. The problem is that this large range of eigenvalues is not matched by a sufficient number of digits in the arithmetic. Some inversion algorithms are remarkably stable and are more tolerant of large condition numbers. The first 8 digits of large and small elements of the inverse matrix for a 33-variable problem with 2×10^{17} condition number processed with the double-precision *ludcmp/lubksb* pair of routines from *Numerical Recipes in Fortran* (Press *et al.*, 1994) were correct when compared with results using 36-byte arithmetic. Surprisingly, the worst offender was not among the smallest elements and its symmetrical equivalent computed with 10 correct digits. This numerical feat is probably due to the line swapping

feature combined with a form of preconditioning ahead of the LU decomposition by Crout's method in this well tuned algorithm. The astonishing stability of this and similar inversion algorithms is the likely reason why most Rietveld refinements do not misbehave in spite of their poor condition numbers, but trusting that they will never misbehave would be a dangerous assumption to make, hence the usefulness of *SVDdiagnostic*.

4.3. Suggestion to developers of Rietveld packages

Our *SVDdiagnostic* program processes output files from two widespread Rietveld packages, *GSAS* and *TOPAS* (V3.0). With minimal work, it could also process output from any other powder or single-crystal package, in fact any least-squares program that prints the matrix of normal equations. Provided that the matrix and the vector could be given with sufficient precision, it would become possible to solve externally the system of equations, and either confirm independently the convergence of the solution, or derive the input for the next cycle using an 'overkill' number of digits in calculating the solution. All that is needed is the optional ASCII output of the matrix and vector with a format giving access to all significant digits, like D22.14 or E21.13 in Fortran. As that moderately voluminous output is not intended for printing but for re-reading, its actual size does not matter much. Such an output would give users the possibility to assess the convergence, diagnose convergence problems or produce a truly converged solution even in ill-conditioned cases.

5. Conclusions

The present analysis contains important general considerations for objective diagnostics of ill-conditioned least squares, allowing their recovery by rank reduction or numerical overkill. The program *SVDdiagnostic* that resulted from those considerations is available for free distribution at <http://www.tothcanada.com>. In this analysis, singular value decomposition was found to be of invaluable help.

SVD analyses of the initial matrix and its pre-conditioned equivalent each contain important hints about numerical stability of the inversion itself and about redundant variables which somehow spoil the e.s.d.'s without possibly contributing any useful information about the model.

We thank Ooi Siew Ting (Nanyang Technological University, Singapore) for performing Rietveld refinements using *TOPAS* (V3.0). We thank Alan A. Coelho and Robert B. Von Dreele for helpful hints. We also thank both anonymous referees for wise suggestions about the presentation and the substance of this study. This work was supported through the NRC/A*STAR Joint Research Programme on 'Advanced Ceramic Methods for the Co-stabilization and Recycling of Incinerator Fly Ash with Industrial Wastes'.

References

- Bruker AXS (2005). *TOPAS*, Versions 2.1 and 3.0. Bruker AXS, Karlsruhe, Germany.
- Guillot, B., Viry, L., Guillot, R., Lecomte, C. & Jelsch, C. (2001). *J. Appl. Cryst.* **34**, 214–223.
- Hendrickson, W. A. & Konnert, J. H. (1980). *Computing in Crystallography*, edited by R. Diamond, S. Ramaseshan & K. Venkatesan, pp. 13.01–13.25. IUCr/Indian Academy of Sciences, Bangalore.
- Larson, A. C. & Von Dreele, R. B. (2004). *General Structure Analysis System (GSAS)*, Report LAUR 86-748. Los Alamos National Laboratory, New Mexico, USA.
- Mercier, P. H. J., Le Page, Y., Whitfield, P. S. & Mitchell, L. D. (2006). *J. Appl. Cryst.* **39**, 369–375.
- Mercier, P. H. J., Le Page, Y., Whitfield, P. S., Mitchell, L. D., Davidson, I. D. & White, T. J. (2005). *Acta Cryst.* **B61**, 635–655.
- Press, W. H., Flannery, B. P., Teukolsky, S. A. & Vetterling, V. T. (1994). *Numerical Recipes in Fortran: the Art of Scientific Computing*, 2nd ed., reprinted with corrections. New York: Cambridge University Press.
- Strang, G. (1980). *Linear Algebra and its Applications*, 2nd ed. New York: Academic Press.
- Watkin, D. (1994). *Acta Cryst.* **A50**, 411–437.