The Zürich Crystallography School
University of Zürich, August 30- September 12, 2009

Optimizing the model of a crystal structure

(1) Model of the diffraction experiment
(2) Linear least squares
(3) Non-linear least squares optimization
(4) Restraints and constraints
(5) Quality of the model
(6) Weighting schemes

Suggested Reading:

- ‘Comparison of silver and molybdenum microfocus X-ray sources for single-crystal structure determination’ L. Krause, R. Herbst-Irmer, G.M. Sheldrick, D. Stalke, J. Appl. Cryst. (2015). 48, 3–10 (good info on data processing and data quality)

- ‘Optimizing the model of a crystal structure’ H.B. Bürgi (Lecture notes from the Zürich school of Crystallography 2013. See http://yaghi.berkeley.edu/links.html#lectures)

- ‘Crystal Structure Refinement - A Crystallographer's Guide to SHELXL’. P. Muller, R. Herbst-Irmer, A. Spek, Th. Schneider, M. Sawaya, OUP 2006

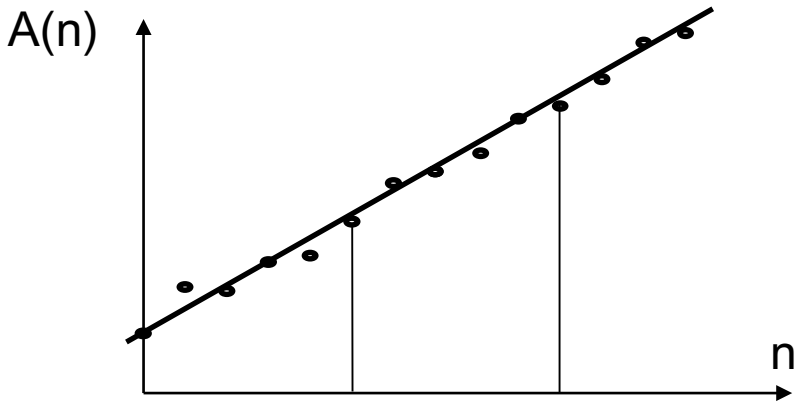
- ‘PLATON SQUEEZE: a tool for the calculation of the disordered solvent contribution to the calculated structure factors’. A. L. Spek, Acta Cryst. C 71 (2015) 9–18.

Model of the diffraction experiment

$$F_{\text{model}}(hkl) = \sum_{i=1}^{q} f_i(\mathbf{S}) \cdot \left[\cos 2\pi(hx_i + ky_i + lz_i) + i \sin 2\pi(hx_i + ky_i + lz_i)\right] \cdot \exp\left[\begin{array}{l} -2\pi^2(h^2 a^{*2} U_i^{11} + k^2 b^{*2} U_i^{22} + l^2 c^{*2} U_i^{33} \\ +2hka^* b^* U_i^{12} + 2hla^* c^* U_i^{13} + 2klb^* c^* U_i^{23}) \end{array}\right] \cdot pop_i$$

$$I_{\text{model}} = |F_{\text{model}}(hkl)|^2 \; g(\text{scale, Flack, extinction, } etc.)$$

Linear least squares



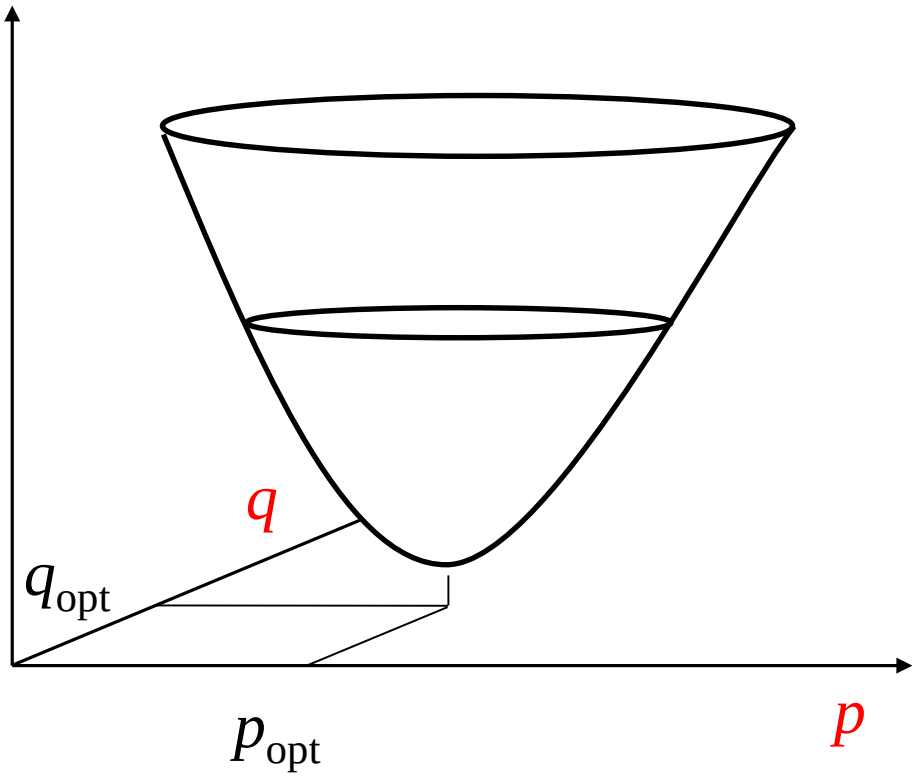
$$A_{\text{model}}(n) = pn + q \qquad (n = 0, \ldots, \mathrm{N})$$

Linear least squares

$$A_{obs}(n) \leftrightarrow A_{\text{model}}(n) = pn + q \qquad (n = 0, \ldots, \mathrm{N})$$

$$S = \sum_{n=0}^{N} w_n \left[A_{\text{obs}}(n) - A_{\text{model}}(n)\right]^2 = \sum_{n=0}^{N} w_n \left[A_{\text{obs}}(n) - pn - q\right]^2$$

$$\frac{\partial S}{\partial p} = \frac{\partial S}{\partial q} = 0$$



Linear least-squares equations

$$p\sum_{n=0}^{N} w_n n^2 + q\sum_{n=0}^{N} w_n n = \sum_{n=0}^{N} w_n A_{\text{obs}}(n) n$$

$$p\sum_{n=0}^{N} w_n n + q\sum_{n=0}^{N} w_n = \sum_{n=0}^{N} w_n A_{\text{obs}}(n)$$

$$\begin{bmatrix} \sum_{n=0}^{N} w_n n^2 & \sum_{n=0}^{N} w_n n \\ \sum_{n=0}^{N} w_n n & \sum_{n=0}^{N} w_n \end{bmatrix} \begin{bmatrix} p \\ q \end{bmatrix} = \begin{bmatrix} \sum_{n=0}^{N} w_n A_{\text{obs}}(n) n \\ \sum_{n=0}^{N} w_n A_{\text{obs}}(n) \end{bmatrix}$$

$$\mathbf{N}\mathbf{p} = \mathbf{\Delta}$$

$$\mathbf{p}_{\text{opt}} = \mathbf{N}^{-1}\mathbf{\Delta}$$

Non-linear least-squares optimization

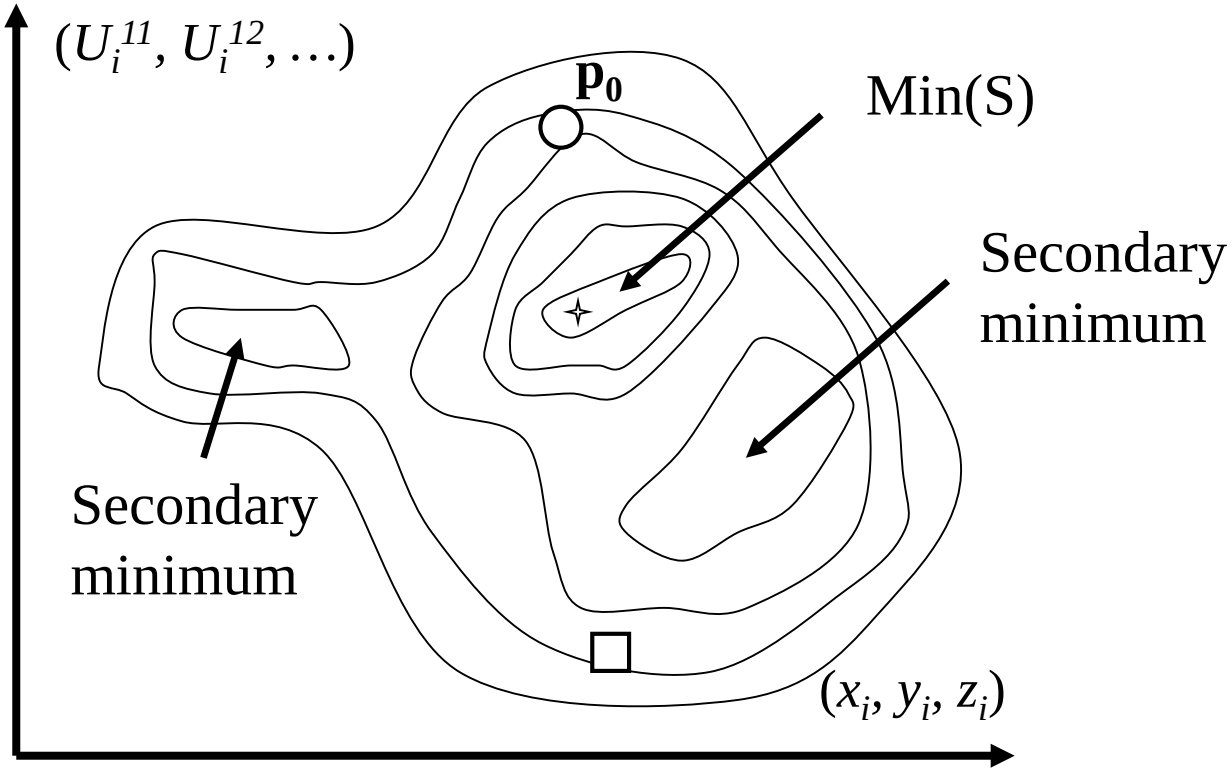
$$I_{obs} \leftrightarrow I_{\text{model}} = |F_{\text{model}}(hkl)|^2 \; g(\text{scale, Flack, extinction, } etc.)$$

$$F_{\text{model}}(hkl) = \sum_{i=1}^{q} f_i(\mathbf{S}) \cdot \left[\cos 2\pi(hx_i + ky_i + lz_i) + i \sin 2\pi(hx_i + ky_i + lz_i)\right] \cdot \exp\left[\begin{array}{l} -2\pi^2(h^2 a^{*2} U_i^{11} + k^2 b^{*2} U_i^{22} + l^2 c^{*2} U_i^{33} \\ +2hka^* b^* U_i^{12} + 2hla^* c^* U_i^{13} + 2klb^* c^* U_i^{23}) \end{array}\right] \cdot pop_i$$

$$S(\mathbf{p}) = \sum_{hkl} w(hkl)\left[I_{\text{obs}}(hkl) - I_{\text{model}}(hkl;\mathbf{p})\right]^2$$

Non-linear least-squares optimization

$$S(\mathbf{p}) = \sum_{hkl} w(hkl)\left[I_{\text{obs}}(hkl) - I_{\text{model}}(hkl;\mathbf{p})\right]^2$$



The shape of S(**p**) is shown with contour lines

Linearization

$\mathbf{p}_0$ is a starting model from e.g. direct methods,
approximate I_{model} and S in the neigborhood of $\mathbf{p}_0$

$$I_{\text{model}}(hkl;\mathbf{p}_0+\Delta\mathbf{p}) = I_{\text{model}}(hkl;\mathbf{p}_0) + \sum_{i=1}^{P}\left(\frac{\partial I_{\text{model}}(hkl;\mathbf{p})}{\partial p_i}\right)_{\mathbf{p}_0}\Delta p_i$$

$$S(\mathbf{p}_0+\Delta\mathbf{p}) = \sum_{hkl} w(hkl)\left[I_{\text{obs}}(hkl) - I_{\text{model}}(hkl;\mathbf{p}_0) - \sum_{i=1}^{P}\left(\frac{\partial I_{\text{model}}(hkl;\mathbf{p})}{\partial p_i}\right)_{\mathbf{p}_0}\Delta p_i \right]^2$$

$$= \sum_{hkl} w(hkl)\left[\Delta I(hkl;\mathbf{p}_0) - \sum_{i=1}^{P}\left(\frac{\partial I_{\text{model}}(hkl;\mathbf{p})}{\partial p_i}\right)_{\mathbf{p}_0}\Delta p_i \right]^2$$

Linearized least-squares equations

$$\frac{\partial S}{\partial \Delta x_i} = \frac{\partial S}{\partial \Delta y_i} = \frac{\partial S}{\partial \Delta z_i} = 0$$

$$\frac{\partial S}{\partial \Delta U_i^{11}} = \frac{\partial S}{\partial \Delta U_i^{22}} = \frac{\partial S}{\partial \Delta U_i^{33}} = \frac{\partial S}{\partial \Delta U_i^{12}} = \frac{\partial S}{\partial \Delta U_i^{13}} = \frac{\partial S}{\partial \Delta U_i^{23}} = 0$$

$$.... = \frac{\partial S}{\partial g} = 0, \quad (i = 1,..., q).$$

$$\sum_{j=1}^{P} \sum_{hkl} w(hkl) \left(\frac{\partial I_{\text{model}}(hkl;\mathbf{p})}{\partial p_i} \right)_{\mathbf{p}_0} \left(\frac{\partial I_{\text{model}}(hkl;\mathbf{p})}{\partial p_j} \right)_{\mathbf{p}_0} \Delta p_j$$

$$= \sum_{hkl} w(hkl) \Delta I(hkl;\mathbf{p}_0) \left(\frac{\partial I_{\text{model}}(hkl;\mathbf{p})}{\partial p_i} \right)_{\mathbf{p}_0} \qquad (i = 1,..., P)$$

$$\mathbf{N}\,\boldsymbol{\Delta p} = \boldsymbol{\Delta}$$

Linearized least-squares equations

$$\mathbf{N}\,\Delta\mathbf{p} = \Delta$$

$$N_{ij} = \sum_{hkl} w(hkl) \left(\frac{\partial I_{\text{model}}(hkl;\mathbf{p})}{\partial p_i} \right)_{\mathbf{p}_0} \left(\frac{\partial I_{\text{model}}(hkl;\mathbf{p})}{\partial p_j} \right)_{\mathbf{p}_0}$$

$$\Delta_i = \sum_{hkl} w(hkl) \Delta I(hkl;\mathbf{p}_0) \left(\frac{\partial I_{\text{model}}(hkl;\mathbf{p})}{\partial p_i} \right)_{\mathbf{p}_0}$$

$$\Delta\mathbf{p} = \mathbf{N}^{-1}\,\Delta$$

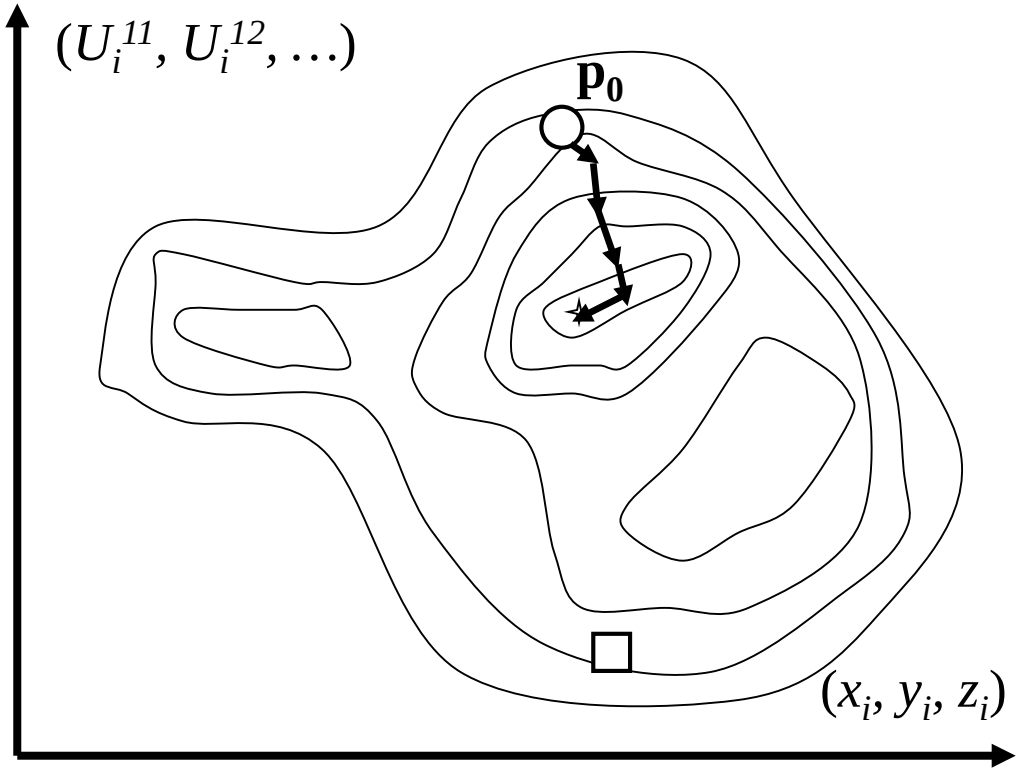
$$\mathbf{p}_1 = \mathbf{p}_0 + \Delta\mathbf{p}$$

$$\ldots$$

$$\mathbf{p}_{k+1} = \mathbf{p}_k + \Delta\mathbf{p}_k \quad \text{to convergence}$$

Non-linear least-squares optimization

$$S(\mathbf{p}) = \sum_{hkl} w(hkl)\left[I_{\text{obs}}(hkl) - I_{\text{model}}(hkl;\mathbf{p})\right]^2$$



The shape of S(**p**) is shown with contour lines

Check convergence

(from SHELXL.lst)

N	value	esd	shift/esd	parameter
1	0.11374	0.00031	-0.026	OSF
2	0.50684	0.02871	0.027	FVAR 2
3	0.42375	9.99999	0.000	FVAR 3
4	0.12506	0.00529	0.089	FVAR 4
5	0.00023	0.00005	-0.025	EXTI

Mean shift/esd = 0.032 Maximum = 0.209 for U11 O1B_b

Max. shift = 0.003 A for O1B_b Max. dU = 0.001 for O1B_b

Convergence for final model

(Acta Cryst standard)

N	value	esd	shift/esd	parameter
1	0.11374	0.00031	0.000	OSF
2	0.50729	0.02870	0.001	FVAR 2
3	0.42375	9.99999	0.000	FVAR 3
4	0.12520	0.00528	0.001	FVAR 4
5	0.00023	0.00005	0.000	EXTI

Mean shift/esd = 0.001 Maximum = 0.006 for U12 O1B_b

Max. shift = 0.000 A for O1B_b Max. dU = 0.000 for O1B_b

Constraints

x, *y*, *z* *on general position*
x, *y*, 0 *in mirror plane*
x, *x*, *z* *in mirror plane*

Special position constraints
(due to symmetry, set automatically)

U^{11}, U^{22}, U^{33}, U^{12}, U^{23}, U^{13}
U^{11}, U^{22}, U^{33}, 0, 0, U^{13}
U^{11}, U^{11}, U^{33}, 0, 0, 0

Optional constraints

10+q.rst	add 10 to an x y z U^{ij} or pop will fix its value to q.rst
FVAR osf p ..	for making any x, y, z, U^{ij} or pop parameter the same for different atoms (10k+q.rst)
EXYZ, EADP	constraining atoms to the same x y z and U^{ij}
AFIX mn	for fixing geometries of H-X bonds and rings

Complementary disordered atoms

at a distance close to or below the resolution limit

EADP O2A O2B

FVAR 0.11362 0.89498

		x	y	z	pop	U
O1A	3	0.170717	0.093893	0.000000	20.50000	0.00671
O1B	3	0.194774	0.076868	0.000000	-20.50000	0.00671

Note:

D(O1A..O1B) = 0.61 Å, Data Resolution ~0.7 Å

Population for SHELX of O1A: 0.5 * FVAR(2)

Population for SHELX of O1B: 0.5 * [1 - FVAR(2)]

Both atoms on mirror plane, i.e. z = 0 and multiplicity mult = 2

Population at site of O1A: 0.5 * FVAR(2) * mult = 0.89498

Population at site of O1B: 0.5 * [1 - FVAR(2)] * mult = 0.10502

Geometric Restraints

$$S' = S(\mathbf{p}) + \sum_j w_i [d_{j,1}^{\text{model}}(\mathbf{p}) - d_{j,2}^{\text{model}}(\mathbf{p})]^2$$

SAME, SADI

SADI 0.03 C20 C16 C16 C18 C18 C20

w = 1/(0.03*0.03)

$$S' = S(\mathbf{p}) + \sum_j w_j [d_j^{\text{model}}(\mathbf{p}) - d_j^{\text{restraint}}]^2$$

DFIX, DANG, CHIV, FLAT

DFIX 1.25 0.02 O8 N10 O7 N10 O6 N9 O5 N9

$d^{\text{restraint}}$ = 1.25, w = 1/(0.02*0.02)

http://shelx.uni-ac.gwdg.de/SHELX/shelx97.pdf (SHELX MANUAL)

Other Restraints

$$S' = S(\mathbf{p}) + \sum_j w_j [q_j^{\text{model}}(\mathbf{p}) - q_j^{\text{restraint}}]^2$$

For ADPs: ISOR, SIMU, DELU

For occupations: SUMP (to include analytical data)

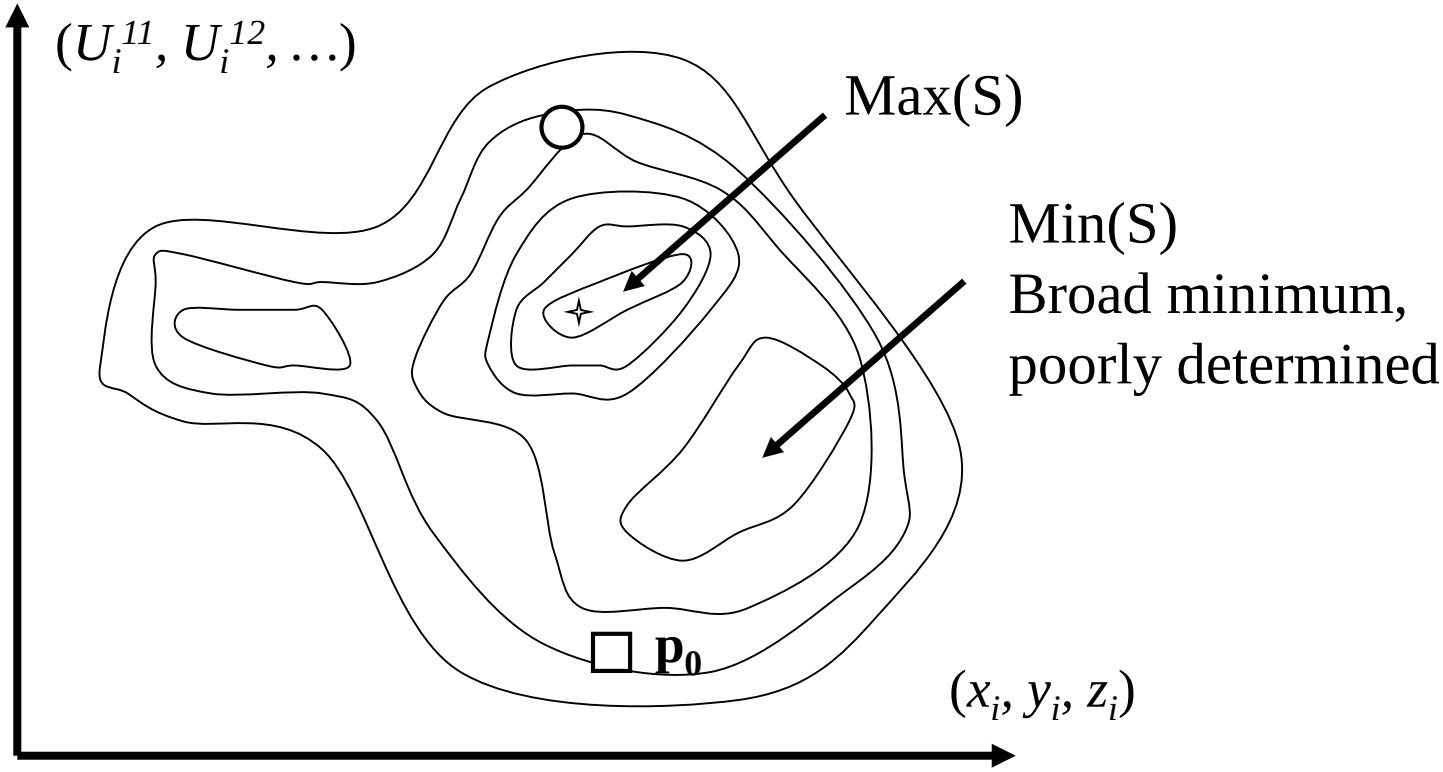
$$S' = S(\mathbf{p}) + \sum_j w_j [p_{j,k+1} - p_{j,k}]^2 = S + \sum_j w_j \Delta p_{j,k}^2$$

Shift limiting restraints: DAMP

http://shelx.uni-ac.gwdg.de/SHELX/shelx97.pdf (SHELX MANUAL)

DAMP restraints

$$S(\mathbf{p}) = \sum_{hkl} w(hkl)\left[I_{\text{obs}}(hkl) - I_{\text{model}}(hkl;\mathbf{p})\right]^2$$



The shape of S(**p**) is shown with contour lines

Model of the diffraction experiment

(hopefully complete and correct)

$$F_{\text{model}}(hkl) = \sum_{i=1}^{q} f_i(\text{S}) \cdot \left[\cos 2\pi(hx_i + ky_i + lz_i) + i \sin 2\pi(hx_i + ky_i + lz_i)\right] \cdot \exp\left[\begin{array}{l} -2\pi^2(h^2 a^{*2} U_i^{11} + k^2 b^{*2} U_i^{22} + l^2 c^{*2} U_i^{33} \\ +2hka^*b^*U_i^{12} + 2hla^*c^*U_i^{13} + 2klb^*c^*U_i^{23}) \end{array}\right] \cdot pop_i$$

$$I_{\text{model}} = |F_{\text{model}}(hkl)|^2 \; g(\text{scale, Flack, extinction, } etc.)$$

including all constraints and restraints

Global quality indicators

$$R1 = \frac{\sum_{hkl} \left| \left| F_{\text{obs}}(hkl) \right| - \left| F_{\text{model}}(hkl;\mathbf{p}) \right| \right|}{\sum_{hkl} \left| F_{\text{obs}}(hkl) \right|}$$

$$wR2 = \left\{ \frac{\sum_{hkl} w(hkl) \left[I_{\text{obs}}(hkl) - I_{\text{model}}(hkl;\mathbf{p}) \right]^2}{\sum_{hkl} w(hkl) \left[I_{\text{obs}}(hkl) \right]^2} \right\}^{1/2} = \left\{ \frac{S(\mathbf{p})}{\sum_{hkl} w(hkl) \left[I_{\text{obs}}(hkl) \right]^2} \right\}^{1/2}$$

$$\text{GOF} = \left\{ \frac{\sum_{hkl} w(hkl) \left[I_{\text{obs}}(hkl) - I_{\text{model}}(hkl;\mathbf{p}) \right]^2}{n_{\text{obs}} - n_{\text{par}}} \right\}^{1/2} = \left(\frac{S(\mathbf{p})}{n_{\text{obs}} - n_{\text{par}}} \right)^{1/2}$$

$\sigma(F_o^2)$ and weighting schemes

- The reliability of an $F_o^2(hkl)$ is indicated by ist weight w(hkl).
  Ideally $w(hkl) = 1/\sigma^2[F_o^2(hkl)]$
  in practice $\sigma^2[F_o^2(hkl)]$ are fudged at least twice.

- $\sigma_{raw}^2[F_o^2\,(hkl)] = [I(hkl) + n*B(hkl)]$ Poisson statistics

- Sheldrick's error model (SADABS):
  $\sigma^2[F_o^2\,(hkl)] = (K * \sigma_{raw}\,[F_o^2\,(hkl)])^2 + (g * [F_o^2(hkl)])^2$

Ideal values: $K = 1, \quad g = 0$

Acceptable values $0.7 < K < 1.3,$

$0 < g < 0.05$

For K and g outside acceptable ranges check your sample quality and data integration

$\sigma(F_o^2)$ and weighting schemes

(from SHELXL.lst)

- How does SHELXL modify $\sigma(F_o^2)$'s for least-squares refinement?

- Inconsistent equivalents etc.

h	k	l	Fo^2	Sigma(Fo^2)	N	Esd of mean(Fo^2)
-9	18	2	2.02	0.59	9	3.07

1 Inconsistent equivalents

Sigma(Fo^2) = $(\Sigma\ \sigma^2[F_o^2(hkl)] / N)^{1/2}$

Σ: sum over equivalents

Esd of mean(Fo^2) = $\{\Sigma\ [F_o^2(hkl) - \langle F_o^2(hkl)\rangle]/N(N-1)\}^{1/2}$

$\sigma(F_o^2)$ and weighting schemes

How does SHELXL MODIFY sigma^2(Fo^2) to obtain weights?

Weight = 1 / [sigma^2(Fo^2) + (a * P)^2 + b * P]

where P = (Max (Fo^2, 0) + 2 * Fc^2) / 3

a and b are chosen such that GOF (~$S^{1/2}$) is as close to 1 as possible for each of 10 groups in Fc/Fcmax

NOTE: FOR GOOD DATA AND A GOOD MODEL
a ~ 0.00 – 0.05, b ~ 0 – 5,
THE SMALLER, THE BETTER!

NOTE: IF a > 0.05, B > 5,
QUESTION THE SAMPLE, DATA, INTEGRATION AND/OR MODEL

Semi-global quality indicators

(in reciprocal space, from SHELXL.lst)

R1 = 0.0529, wR2 = 0.1151,

Analysis of variance for reflections employed in refinement

K = Mean[Fo^2] / Mean[Fc^2] for group

Fc/Fc(max)	0.000	0.014	0.028	0.048	0.063	0.083	0.109	0.134	0.172	0.231
Number in group	58.	52.	56.	56.	51.	55.	55.	55.	53.	
GooF	1.149	1.438	1.576	1.719	1.654	0.867	0.849	0.773	1.139	
K	-1.165	0.034	0.577	0.626	0.750	0.989	1.037	1.029	1.037	

Resolution(A)	0.78	0.81	0.85	0.90	0.94	1.01	1.10	1.23	1.43	1.88
Number in group	56.	56.	55.	52.	56.	53.	55.	54.	55.	
GooF	1.451	1.396	1.438	1.388	1.490	1.371	1.150	0.896	0.862	
K	0.797	1.002	0.995	1.009	1.011	1.031	1.022	1.020	1.016	
R1	0.296	0.182	0.149	0.101	0.116	0.069	0.060	0.037	0.021	

Recommended weighting scheme: WGHT 0.0110 162.5899

Local quality indicators

(in reciprocal space, from SHELXL.lst)

Most Disagreeable Reflections

Error/esd is sqrt(wD^2/<wD^2>), w is given by the weight formula, D = Fo^2-Fc^2 and <> refers to the average over all reflections.

h	k	l	Fo^2	Fc^2	Error/esd	Fc/Fc(max)	Resolution(A)
0	2	2	63080.22	3488.57	16.96	0.064	7.33
2	2	2	124011.70	241363.84	8.07	0.530	5.99
0	6	6	33451.04	19797.35	4.64	0.152	2.44
0	2	8	23532.70	13514.00	4.35	0.125	2.51
2	2	6	15895.94	9150.49	3.68	0.103	3.13
0	4	8	1529.09	60.39	3.66	0.008	2.32
2	2	8	54521.12	74668.03	3.48	0.295	2.44
2	4	4	200053.92	158813.53	3.34	0.430	3.46
0	2	6	32.98	1893.38	3.28	0.047	3.28
1	5	5	97056.38	75962.81	3.12	0.297	2.90
1	3	3	49886.82	63090.50	2.55	0.271	4.76

Local quality indicators

(reciprocal space)

Deviations from expectation values for Gaussian distribution of errors (example with 745 observations)

		present example
➢ 3 sigma	0.3 %	10/745 => 1.3 %
➢ 4 sigma	0.01 %	4/745 => 0.5 %
➢ 5 sigma	0.0001 %	2/745 => 0.25 %

The present example shows two outliers (0 2 2) and (2 2 2) which have to be accounted for somehow, e.g. experimental error or incomplete model

Local quality indicators

(direct space, from SHELXL.lst)

Standard uncertainties (s.u.)

ATOM	x	y	z	sof	U11	U22	…..
Zr1	0.11941	0.00000	0.00000	0.12500	0.00141	0.00286	
0.0007	0.00003	0.00000	0.00000	0.00000	0.00036	0.00027	
O1A_a	0.17088	0.09373	0.00000	0.45000	0.00671	0.00601	
0.0045	0.00017	0.00016	0.00000	0.00000	0.00163	0.00156	
C1_a	0.15149	0.15149	0.00000	0.22500	0.00896	0.00896	
0.0090	0.00022	0.00022	0.00000	0.00000	0.00180	0.00180	

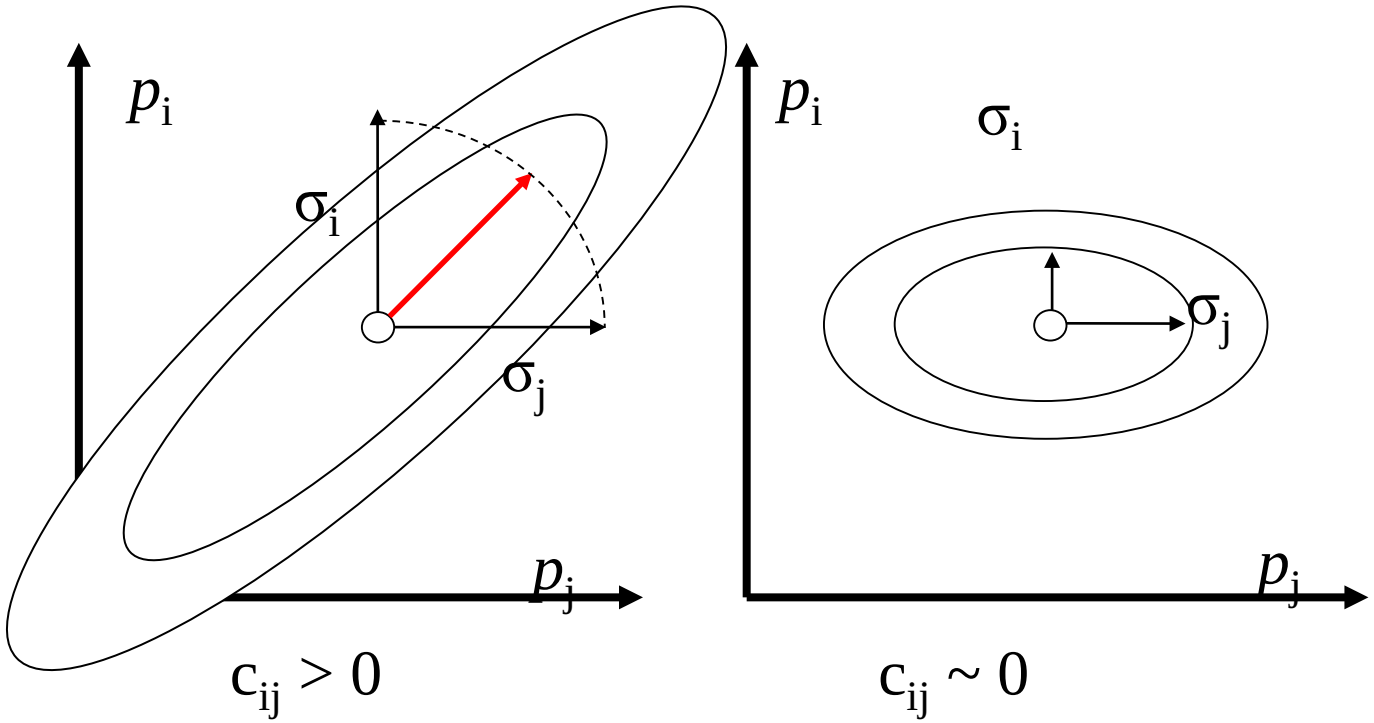
Red numbers: isotropic positional uncertainties σ_i (in Å) !

Approximate uncertainty in distance d between atoms i and j

$$\sigma(d) \sim (\sigma_i^2 + \sigma_j^2)^{1/2}$$

Parameter correlation

$$S(\mathbf{p}) = \sum_{hkl} w(hkl)\left[I_{\text{obs}}(hkl) - I_{\text{model}}(hkl;\mathbf{p})\right]^2$$



$c_{ij} >$ 0.98: increasing both p_i and p_j has little effect on R
$p_i + p_j$ is poorly determined, $p_i - p_j$ is well determined

$c_{ij} <$ -0.98: increasing pi and decreasing pj has little effect on R
$p_i - p_j$ is poorly determined, $p_i + p_j$ is well determined

Local quality indicators

(direct space, from SHELXL.lst)

Does the model obey the restraints?

DFIX 1.25 O8 N10 O7 N10 O6 N9 O5 N9
… 26 more DFIX restraints
SADI C20 C16 C16 C18 C18 C20 C15 C17 C17 C19 C19 C15
… 7 more SADI restraints, 4 FLAT restraints
… 8 DELU restraints, 8 SIMU restraints, 7 EADP constraints

R1 = 0.2100 for 4848 data, 640 parameters and 977 restraints

Summary of restraints applied in FINAL cycle

	DFIX	SAME/SADI	FLAT	DELU	SIMU
Number	42	120	45	212	558
rms sigma	0.0200	0.0200	0.1000	0.0100	0.0528
rms deviation	0.0711	0.0219	0.0637	0.0053	0.0300

Local quality indicators

(direct space, from SHELXL.lst)

Does the model obey the restraints?

DFIX 1.25 O8 N10 O7 N10 O6 N9 O5 N9

really means

DFIX 1.25 0.02 O8 N10 O7 N10 O6 N9 O5 N9

Disagreeable restraints before cycle xy

Observed	Target	Error	Sigma	Restraint
1.3390	1.2500	0.0890	0.0200	DFIX O8 N10

Bond lengths and angles

O8 - N10 1.3390 (0.0172)

NOTE: This standard uncertainty is meaningless as it is determined by the DFIX instruction!

Local quality indicators

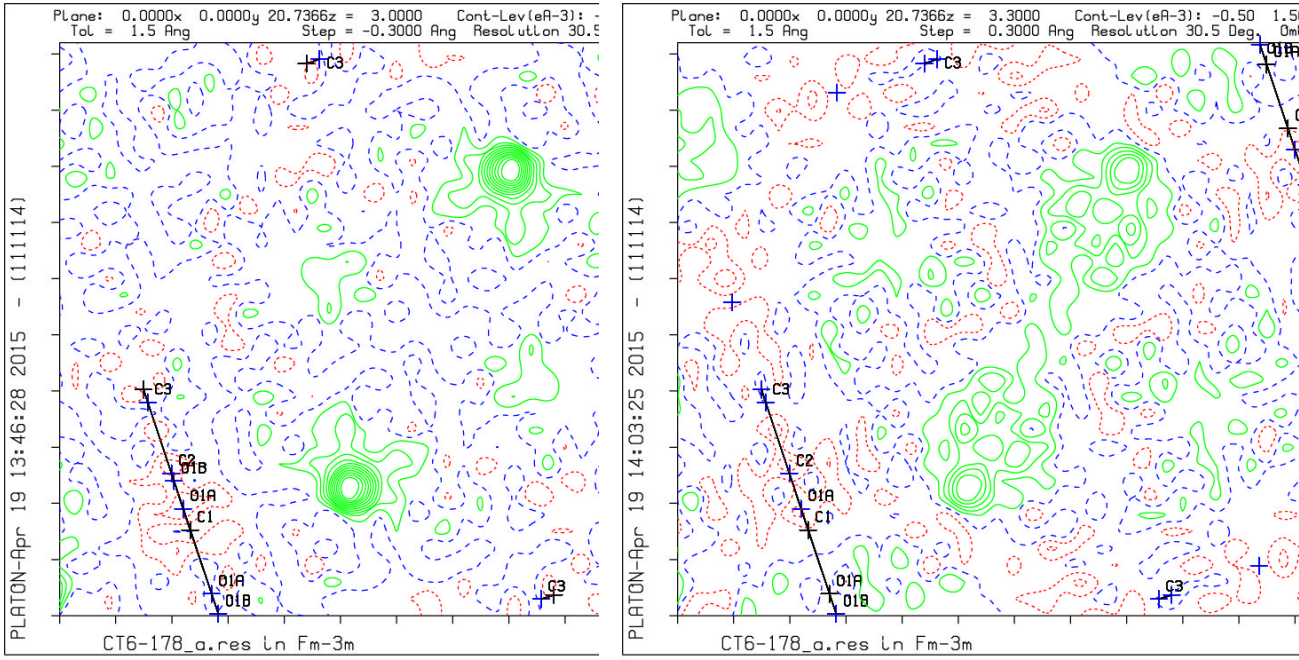
(direct space, from PLATON)

Difference Fourier synthesis showing residual electron density

Q-peaks or Contour map

Q-peaks are generally informative as long as significantly occupied sites have not been included in the structural model

Contour maps are advised when disordered solvents need to be modeled (most easily produced with PLATON)



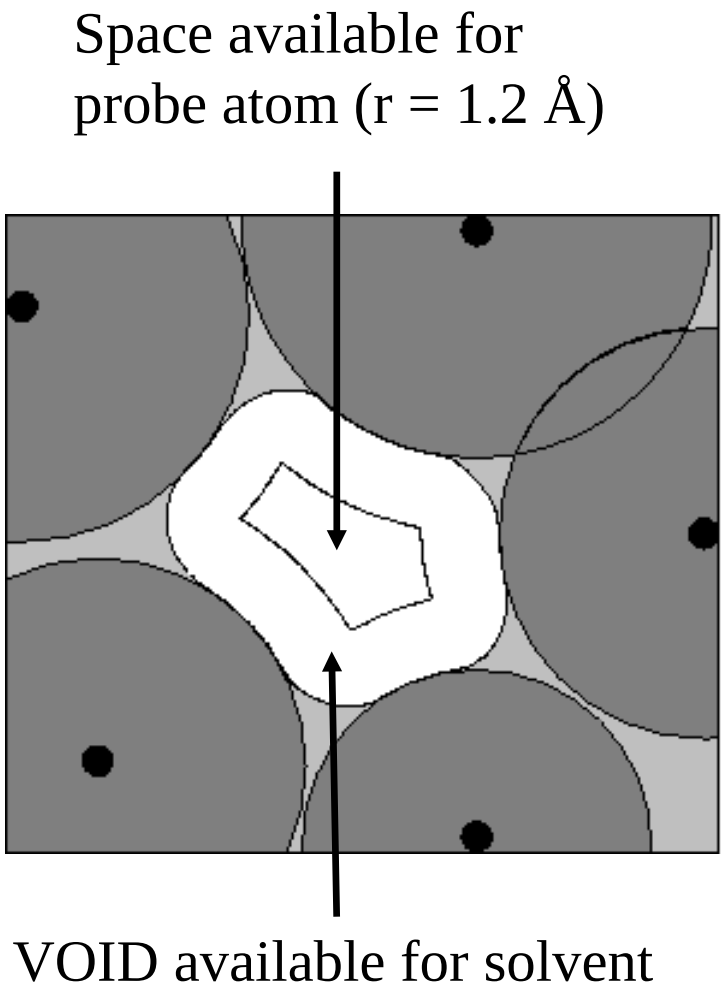
The Disordered Solvent Problem

- Molecules of interest often co-crystallize (only) with the inclusion of a suitable solvent molecule.
- Solvent molecules often fill voids in a structure with little interaction and are often located on symmetry sites and with population less than 1.0
- Sometimes even the nature of the (mixture) of included solvent(s) is unclear.
- Inclusion of the scattering contribution of the solvent to the structure factors can be done either with an (elaborate) disorder model or with the SQUEEZE approach.

From: Ton Spek, PLATON Workshop, Chicago, 24-July-2010 and
A. L. Spek, Acta Cryst. C 71 (2015) 9–18

How does SQUEEZE work?

- Determine VOID region
- Calculate residual density (difference Fourier)
- Count e- in VOID region
- Calculate their contribution to the model structure factors and model phases
- Calculate residual density with improved model phases
- Iterate to convergence



How does SQUEEZE work?

$$\Delta\rho = V^{-1}\sum_{h}[k\,|F_h^o|\exp(i\varphi_h^m) - |F_h^m|\exp(i\varphi_h^m)]\exp[-2\pi i(\mathbf{h}\cdot\mathbf{r})]$$

$$F_{000}^s = \Delta V_g \sum_{k}^{VOID}\Delta\rho(\mathbf{r}_k)$$

- Count e- in VOID

$$F_h^s = \Delta V_g \sum_{k}^{VOID}\Delta\rho(\mathbf{r}_k)\exp[2\pi i(\mathbf{h}\cdot\mathbf{r})]$$

- Calculate contribution of e- in VOID to structure factor

$$F_h^M = F_h^m + F_h^s = |F_h^M|\exp(i\varphi_h^M)$$

- Calculate improved phases

$$\Delta\rho = V^{-1}\sum_{h}[k\,|F_h^o|\exp(i\varphi_h^M) - |F_h^m|\exp(i\varphi_h^m)]\exp[-2\pi i(\mathbf{h}\cdot\mathbf{r})] + V^{-1}F_{000}^s$$

Results from SQUEEZE

- Number of Missing Reflections = 12

- Number of Missing Low Order Reflections = 2
  (0 0 2) Fc(solv) = - 112.4, (1 1 1) Fc(solv) = -156.7

- Volume of VOID 4554 $Å^3$ containing 603 electrons
  (corresponds to ~35 water molecules whose volume would be approximately $35 \cdot 15\ Å^3 \sim 500\ Å^3$)

SinT/L	<FcMod>	<FcSolv>	<FcTot>	<Fo>	N	R(Mod)	R(Tot)
0.05	63.65	180.47	243.97	244.96	1.	0.740	0.004
0.10	324.94	39.76	286.62	287.24	6.	0.137	0.025
0.15	307.25	24.84	314.25	317.72	13.	0.096	0.027

Results from SQUEEZE,

EXAMPLE FROM A MOF-STRUCTURE:

- Without modeling pore content:

R1 = 0.0436 WGHT 0.0551 191.2836

- SQUEEZE only (~600 e^-)

R1 = 0.0251 WGHT 0.0132 36.5623

NOTE: the phase improved difference Fourier after SQUEEZE is often clearer and a starting point for atomic modeling of solvent density

- Modeling most density in pores, no SQUEEZE:

R1 = 0.0271 WGHT 0.0301 27.8971

Disordered solvent: discrete modeling or SQUEEZE?

- SEE: ‘PLATON SQUEEZE: a tool for the calculation of the disordered solvent contribution to the calculated structure factors’. A. L. Spek, Acta Cryst. C 71 (2015) 9–18.

- QUOTE: ‘The development of an atomistic model of the disordered solvent is generally to be preferred wherever possible, in particular when the disorder can be described easily with constraint and restraint tools, such as those available in SHELXL.’

- QUOTE: ‘The time invested in devising an unsatisfactorily parametrized disordered solvent model is not always considered to be worth the effort. This applies in particular in the context of a routine (service) structure determination intended to characterize the chemistry of the main component in the crystal.’