

Blocked Matrices Pandora's Box

For numerically stable refinements, it is not necessary to compute the full normal matrix at every intermediate stage of an analysis. Smaller sub-blocks of the full matrix may be adequate.

However, if there is instability, especially as a consequence of data shortage or pseudo symmetry, it is wisest to use the full matrix

Using the full matrix is almost always the safest procedure, though it may not be the most efficient.

46/62

Solvent of Crystallisation

One really important use for blocking the matrix is when the bulk of the structure is well behaved, but there is a small part (solvent, small ligand) which is proving difficult to model.

In this situation, the analyst usually proposes a model, performs some refinement, and then assesses the outcome.

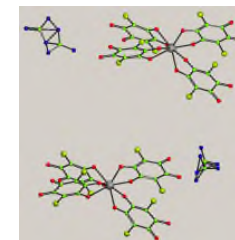
It is not cost-effective to re-refine the whole of the well-behaved part during this experimental stage, though everything must be refined together to finish the analysis

47/62

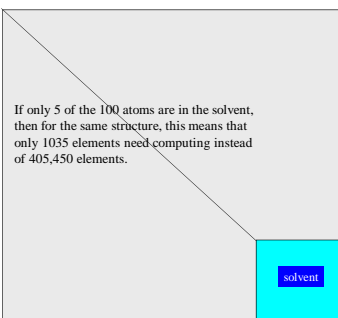
Example Structure

Consider a structure composed of 100 atoms in the main molecule, and some disordered solvent.

If the main molecule is well behaved, it can be left partly refined until the disorder is sorted out.



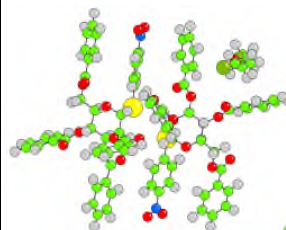
48/62



If only 5 of the 100 atoms are in the solvent, then for the same structure, this means that only 1035 elements need computing instead of 405,450 elements.

49/62

Solvent of Crystallisation

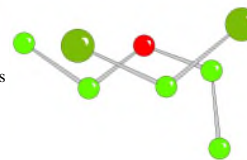


The structure contains 106 well-resolved atoms, plus 8 atoms in the solvent.

Refining the main part aniso plus solvent iso gives 986 parameters (486,098 LSQ elements)

Refining just the solvent while it is being sorted out gives 32 parameters (512 LSQ elements)

A cycle of refinement is 1000 times faster.



Blocked Matrices

For mesomolecular structures (100-1000) atoms, conventional full matrix least squares becomes increasingly inefficient, but conjugate gradient methods show no advantage.

Dividing up the normal matrix into sub matrices (either within a single cycle, or in successive cycles) is a cost-effective way of handling medium sized problems.

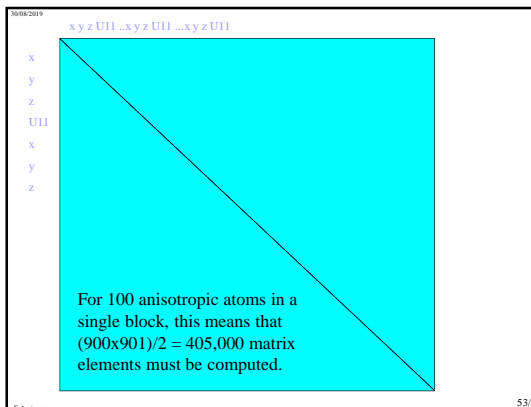
51/62

Matrix Schemes

For 100 anisotropic atoms in a single block, this means that $(900 \times 901)/2 = 405,000$ matrix elements must be computed.

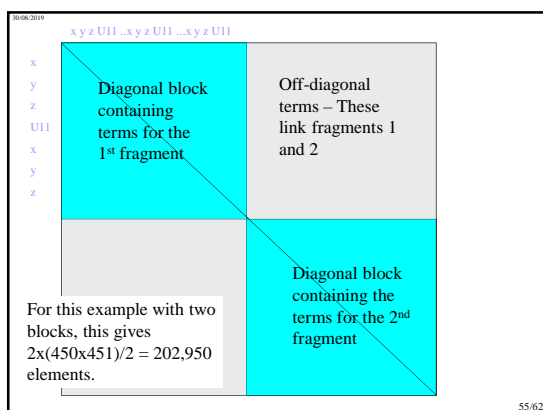
If there are 10 reflections per parameter, this gives **4 million sums-of-products**.

Various schemes are available to reduce the size of the computation.



Large-Block Approximation

If the structure consists of non-bonded molecular fragments, **and there is no pseudo-symmetry between them**, then each fragment can be refined in a separate matrix block. For this example with two blocks, this gives $2 \times (450 \times 451)/2 = 202,950$ elements.



Refining Fragments in Alternate Cycles

This is similar to the large-block approximation.

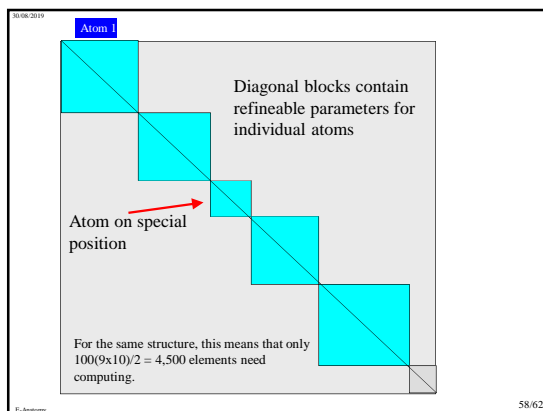
If the full matrix refinement is unstable because of high correlations, this method will appear to fix it, but in fact only conceals the problem.

Atom-Block Diagonal

This is the most extreme scheme, and is used in SIR92 etc. It puts all the refinable parameters for each atom into an individual block.

For the same structure, this means that only $100(9 \times 10)/2 = 4,500$ elements need computing.

However, the rate of convergence is often poor, (or non-existent). Remember that structure factors and derivatives also have to be computed for each cycle.



10/08/2019

Matrix Elements

Full Matrix	405,000
2 Large Blocks	202,950
Atom Blocks	4,500

The atom-block refinement is almost 100 times faster per cycle than full matrix.

However, more cycles must be computed and convergence is less certain.

59/62

- 10/08/2019
- ### Some Useful Recipes
- Some suggested blocking schemes are:
1. A block for x, y, z separately, plus one for Uiso and scale.
 2. A block for x's, and a block for U's and scale
 3. As in 2, but with adps divided into two blocks, plus dummy Uiso and scale.
- Parameters in directions related by a very oblique angle must be in the same block.**
- 60/62

10/08/2019

Conclusion

For an unknown material, the safest strategy is almost always to use a full matrix refinement.

Dividing a structure to avoid singularities is *always* dangerous, and may lead to the structure being 'Marshed'.

However, for very large structures, full matrix is a wasteful strategy (since many off-diagonal elements will sum to zero), and may actually be slow to converge if rounding errors are allowed to accumulate.

61/62