Special cases arise when there are combinations of crystallographic and noncrystallographic symmetries, of proper and improper symmetries, or when a noncrystallographic symmetry element maps a cell edge onto itself. In the latter case, the volume of matching density is infinite, and arbitrary limits must be placed upon the mask along one crystal axis.

15.1.2.3.3. *The refinement of noncrystallographic symmetry*

The initial NCS operation obtained from rotation and translation functions or heavy-atom positions can be fine-tuned by a density-space R-factor search in the six-dimensional rotation and translation space. The density-space R factor is defined as

$$R = \sum_{\mathbf{r}} |\rho(\mathbf{r}) - \rho(\mathbf{r}')| \Big/ \sum_{\mathbf{r}} |\rho(\mathbf{r}) + \rho(\mathbf{r}')|, \qquad (15.1.2.24)$$

where $\mathbf{r} = \{xyz\}$ is the set of Cartesian coordinates, $\mathbf{r}' = \Omega\mathbf{r}$ is the NCS-related set of coordinates of $\mathbf{r}$ and Ω represents the NCS operator.

The six-dimensional search is very time-consuming. The search rate can be increased by using only a representative subset of grid points. The NCS operation is systematically altered to find the lowest density-space R factor for the selected subset of grid points.

The solution of the NCS operation from the six-dimensional search can be further refined by the following least-squares procedure. If $\rho(\mathbf{r})$ is related to $\rho(\mathbf{r}')$ by the NCS operation, Ω,

$$\rho(\mathbf{r}') = \rho(\Omega\mathbf{r}). \qquad (15.1.2.25)$$

Here, Ω is a function of $\omega, \Omega = f(\omega)$, where $\omega = \{\alpha, \beta, \gamma, t_x, t_y, t_z\}$ represents the rotation and translation components of the NCS operation. The solution to the NCS parameters, ω, can be obtained by minimizing the density residual between the NCS-related molecules,

$$\varepsilon(\mathbf{r}) = \rho(\mathbf{r}) - \rho(\Omega\mathbf{r}), \qquad (15.1.2.26)$$

using a least-squares formula of the form

$$\left(\frac{\partial\rho}{\partial\omega}\right)^T \left(\frac{\partial\rho}{\partial\omega}\right) \Delta\omega = \left(\frac{\partial\rho}{\partial\omega}\right)^T \varepsilon(\mathbf{r}), \qquad (15.1.2.27)$$

where $\Delta\omega$ is the shift to the NCS parameters. Here,

$$\frac{\partial\rho}{\partial\omega} = \frac{\partial\rho}{\partial\mathbf{r}} \frac{\partial\mathbf{r}}{\partial\omega}. \qquad (15.1.2.28)$$

The partial derivatives, $\partial\rho/\partial\mathbf{r} = \{\partial\rho/\partial x,\ \partial\rho/\partial y,\ \partial\rho/\partial z\}$, can be calculated by Fourier transforms,

$$\begin{aligned}
\frac{\partial\rho}{\partial x} &= -\frac{2\pi i}{V} \sum_{hkl} h F_{hkl} \exp[-2\pi i(hx + ky + lz)] \\
\frac{\partial\rho}{\partial y} &= -\frac{2\pi i}{V} \sum_{hkl} k F_{hkl} \exp[-2\pi i(hx + ky + lz)] \qquad (5.1.2.29) \\
\frac{\partial\rho}{\partial z} &= -\frac{2\pi i}{V} \sum_{hkl} l F_{hkl} \exp[-2\pi i(hx + ky + lz)],
\end{aligned}$$

or more efficiently with a single Fourier transform by the use of spectral B-splines (Cowtan & Main, 1998). $\partial\mathbf{r}/\partial\omega$ is derived analytically based on the relationship between the Cartesian coordinates, r, and the rotational and translational coordinates of the NCS operation, ω,

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \cos\alpha\cos\beta\cos\gamma - \sin\alpha\sin\gamma & -\cos\alpha\cos\beta\sin\gamma - \sin\alpha\sin\gamma & \cos\alpha\sin\beta \\ \sin\alpha\cos\beta\cos\gamma + \cos\alpha\sin\gamma & -\sin\alpha\cos\beta\sin\gamma + \cos\alpha\cos\gamma & \sin\alpha\sin\beta \\ -\sin\beta\cos\gamma & \sin\beta\sin\gamma & \cos\beta \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \begin{pmatrix} t_x \\ t_y \\ t_z \end{pmatrix}. \qquad (15.1.2.30)$$

15.1.2.3.4. *The averaging of NCS-related molecules*

Once the mask and matrices are determined, the electron-density map may be modified by averaging. This can be achieved in one or two stages: The density for each copy of the molecule in the asymmetric unit may be replaced by the averaged density from every copy; however, this becomes slow for high-order NCS (Fig. 15.1.2.4*c*). Alternatively, a single averaged copy of the molecule may be created in an artificial cell [referred to by Rossmann *et al.* (1992) as an *H*-cell], and then each copy of the molecule may be reconstructed in the asymmetric unit from this copy (Fig. 15.1.2.4*d*). This is more efficient for high-order NCS, but additional errors are introduced in the second interpolation.

Interpolation of electron-density values at non-map grid sites is usually required, since the NCS operators will not normally map grid points onto each other. To obtain accurate interpolated values, either a fine grid or a complex interpolation function are required; suitable functions are described in Bricogne (1974) and Cowtan & Main (1998). Solvent flattening and histogram matching are frequently applied after averaging, since histogram matching tends to correct for any smoothing introduced by density interpolation.

In the case of flexible proteins, it may be necessary to average only part of the molecule, in which case the averaging mask will exclude some parts of the unit cell which are indicated as protein by the solvent mask. In other cases, it may be necessary to apply multi-domain averaging; in this case, the protein is divided into rigid domains which can appear in differing orientations. Each domain must then have a separate mask and set of averaging matrices.

Averaging may also be performed across similar molecules in multiple crystal forms (Schuller, 1996); in this case, density modification is performed on each crystal form simultaneously, with averaging of the molecular density across all copies of the molecule in all crystal forms. This is a powerful technique for phase improvement, even when no phasing is available in some crystal forms.

15.1.2.4. *Skeletonization*

The skeletonization method enhances connectivity in the map. This is achieved by locating ridges of density, constructing a graph of linked peaks, and then building a new map using cylinders of density around the graph peaks.

At worse than atomic resolution, the density peaks for bonded atoms are no longer resolved, and so interpretation of the density in terms of atomic positions involves recognition of common motifs in the pattern of ridges in the density. Skeletonization was a tool developed by Greer (1985) to assist model building by tracing high ridges in the electron density to describe the connectivity in the map.

Skeletonization has more recently been adapted to the problem of density modification (Baker, Bystroff *et al.*, 1993; Bystroff *et al.*, 1993; Wilson & Agard, 1993). A skeleton is constructed by tracing the ridges in the map. The resulting ridges form connected 'trees'. These trees may be pruned to remove small unconnected fragments and break circuits to select for protein-like features. A new map may then be built by building density around the links of the skeleton using the profile of a cylindrically averaged atom at the appropriate resolution.

The skeletonization method has been used to add new features to a partial model of a molecule (Baker, Bystroff *et al.*, 1993). An

references