Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Error estimation and bias correction in phase-improvement calculations

Kevin Cowtan

University of York, Heslington, York YO1 5DD, England

Correspondence e-mail: cowtan@yorvic.york.ac.uk

Received 27 January 1999
Accepted 27 May 1999

With the rise of Bayesian methods in crystallography, the error estimates attached to estimated phases are becoming as important as the phase estimates themselves. Phase improvement by density modification can cause problems in this environment because the quality of the resulting phases is usually overestimated. This problem is addressed by an extension of the γ correction [Abrahams (1997). *Acta Cryst.* D**53**, 371–376] to arbitrary density-modification techniques. The degree to which the improved phases are biased by the features of the initial map is investigated in order to determine the limits of the resulting procedure and the quality of the phase-error estimates.

1. Introduction

Phase improvement by density modification is widely used for the improvement of phases from MIR/MAD data. Improved phases, along with weights representing the estimated phase error, may then be used to calculate a weighted map for interpretation in a graphics program. Phase improvement can often lead to an interpretable map when the experimental phasing alone is insufficient.

The aim of density-modification calculations has frequently been to obtain the best (most interpretable) map and some success has been achieved in this respect. However, advances in other areas of crystallographic computing, in particular the increased use of statistical techniques, create another imperative: that the phase probability distributions, or the weights attached to the phases, should be as realistic as possible. For some purposes, such as model refinement, it may be better to have a poorer phase set with reasonable weights than a better phase set with unrealistic weights. In particular, the following situations are common.

(i) The use of maximum-likelihood refinement with phase constraints (Pannu *et al.*, 1998) provides a powerful method for improving the restraint-to-parameter ratio for protein refinement, improving the radius of convergence and the quality of the final model. Initial phase improvement can further improve refinement results, but only if the error estimates on the phases are reasonable.

(ii) The sequential use of several density-modification methods (possibly from different software packages) to obtain further phase improvement depends on realistic phase-error estimates at each stage, since most phase-improvement procedures have been optimized to be used after a maximum-likelihood phasing procedure (*e.g. SHARP*; La Fortelle & Bricogne, 1997), which already give good phase-error estimates. For best results in subsequent phase-improvement stages, the resulting phase-error estimates should be equally reliable to those from modern phasing software.

In practice, almost all density-modification methods lead to badly overestimated weights (or badly underestimated phase errors; see Cowtan & Main, 1996). This leads to significant problems in both phased refinement and use of multiple phase-improvement methods. In the case of phased refinement, errors will be introduced into the model to fit the estimated phases to within the (underestimated) error bounds. Phases from density modification have been used with good results in phased refinement calculations, but only after the application of a 'blur factor' to reduce the associated weights (Murshudov, 1997).

In the case of multiple density-modification procedures, subsequent procedures will be unable to overcome the phase biases introduced by the first procedure in the chain. Since most phase-improvement calculations involve multiple cycles of density modification, the same effect also limits the ultimate effectiveness of individual density-modification calculations.

These problems have been addressed in the past by various techniques with the aim of obtaining better maps. Some of these techniques will be examined in more detail in order to gauge the suitability of the output phases in weights for use in phased refinement and other critical applications and to determine protocols under which reliable error estimates may be obtained.

2. Overview of phase improvement by density modification

Conventional density-modification calculations combine information in both real and reciprocal space, and therefore elements of the calculation are performed in both spaces. Some initial phase information is required: for example, from a MIR/MAD experiment. This is usually represented as a phase probability distribution $P(\varphi_{\rm obs})$ described in terms of Hendrickson–Lattmann coefficients (Hendrickson & Lattmann, 1970).

The calculation then proceeds as follows.

(i) Centroid map coefficients, representing best estimate of the map coefficients given the observed magnitudes and phase probabilities, are calculated. These map coefficients can be described by a best phase and a weight to be applied to the observed magnitudes.

(ii) An initial map is calculated from the initial map coefficients by FFT.

(iii) Density constraints are applied to produce a modified map from the initial map.

(iv) Modified map coefficients are determined from the modified map by inverse FFT.

(v) The agreement between the modified map coefficients and the observed structure-factor magnitudes is used to estimate the phase errors in the modified phases. The estimated phase error is used to form a phase probability distribution for the modified phase.

(vi) The experimental phase probability distribution and the phase information from the modified map are combined to produce an updated phase probability distribution. This combined distribution is obtained by multiplying the two probability distributions.

The improved phase probabilities are then used to start another cycle and the process is iterated until no further improvement appears to be occurring.

Clearly, the multiplication of phase probability distributions in step (vi) is only valid when those distributions contain independent information. This is not the case, since the experimental phases were used in calculating the initial map for modification. This problem forms the core concern of this work.

The mathematical symbols used in this paper will be defined by restating the calculation in mathematical terms as follows.

The initial data consist of the observed structure-factor magnitudes $|F_{\rm obs}(\mathbf{h})|$ and some initial phase probability distributions $P[\varphi_{\rm obs}(\mathbf{h})]$. At each cycle of density modification, an updated phase probability distribution $P[\varphi_{(i)}(\mathbf{h})]$ will be calculated, where the subscript i represents the cycle number. $P[\varphi_{\rm obs}(\mathbf{h})]$ will become $P[\varphi_{(0)}(\mathbf{h})]$. The density-modification cycle then proceeds as follows.

Step 1. Calculate the initial (centroid) map coefficients, $F_{\rm init}$.

$$F_{\rm init}(\mathbf{h}) = w|F_{\rm obs}(\mathbf{h})| \exp(i\varphi_{\rm best}), \tag{1}$$

where

$$w^2 = \left\{\int_0^{2\pi} P[\varphi_{i-1}(\mathbf{h})]\cos(\varphi)\mathrm{d}\varphi\right\}^2 + \left\{\int_0^{2\pi} P[\varphi_{i-1}(\mathbf{h})]\sin(\varphi)\mathrm{d}\varphi\right\}^2,$$
$$\varphi_{\rm best} = \tan^{-1}\left\{\int_0^{2\pi} P[\varphi_{i-1}(\mathbf{h})]\sin(\varphi)\mathrm{d}\varphi \Big/ \int_0^{2\pi} P[\varphi_{i-1}(\mathbf{h})]\cos(\varphi)\mathrm{d}\varphi\right\}. \tag{2}$$

Optionally, difference map coefficients could be used on subsequent cycles. (F is used in this paper to represent a general map coefficient, not necessary a structure factor or estimate. $\mathbf{h}$, $\mathbf{k}$, $\mathbf{x}$ and $\mathbf{y}$ are assumed to be vectors throughout.)

Step 2. The initial map $\rho_{\rm init}$ is calculated by FFT,

$$\rho_{\rm init}(\mathbf{x}) = (1/V)\sum_{\forall\mathbf{h}} F_{\rm init}(\mathbf{h})\exp(-2\pi i\mathbf{h}.\mathbf{x}). \tag{3}$$

Step 3. The modified map $\rho_{\rm mod}$ is calculated from the initial map by application of some set of density constraints

$$\rho_{\rm mod}(\mathbf{x}) = f[\rho_{\rm init}(\mathbf{y})\ \forall\ \mathbf{y}]. \tag{4}$$

Step 4. The modified map coefficients are calculated by inverse FFT,

$$F_{\rm mod}(\mathbf{h}) = |F_{\rm mod}(\mathbf{h})|\exp(i\varphi_{\rm mod}) = \int \rho_{\rm mod}(\mathbf{x})\exp(2\pi i\mathbf{h}.\mathbf{x})\mathrm{d}\mathbf{x}. \tag{5}$$

A phase probability distribution for the modified phases is estimated using

$$P[\varphi_{\rm mod}(\mathbf{h})] \propto \exp[X\cos(\varphi - \varphi_{\rm mod})], \tag{6}$$

where

$$X = [2\sigma_A/(1-\sigma_A^2)]|E_{\rm obs}||E_{\rm mod}|. \quad (7)$$

$|E_{\rm obs}|$ and $|E_{\rm mod}|$ are the normalized magnitudes of the observed structure factor and the modified map coefficient, and σ_A is calculated by the method of Read (1986).

The updated phase probability distribution is calculated as

$$P[\varphi_{(i)}(\mathbf{h})] = P[\varphi_{\rm mod}(\mathbf{h})]P[\varphi_{\rm obs}(\mathbf{h})]. \quad (8)$$

The following additional symbols and terminology are used in the paper. $F_{\rm adj}$, map coefficients adjusted by some means to reduce the dependence on $F_{\rm init}$, thus allowing meaningful phase combination as described in steps 5 and 6. $F_{\rm omit}$, map coefficients obtained from a full reflection-omit calculation (see §4.1 and Cowtan & Main, 1996).

3. Correlation in phase combination

In the phase-combination calculation, the updated phase probability distribution is obtained by multiplying the initial phase probability by the distribution obtained from the modified map,

$$P[\varphi_{(i)}] = P[\varphi_{\rm mod}]P[\varphi_{\rm obs}]. \quad (9)$$

Implicit in this step is the assumption that the probability distributions for the initial and modified phases are independent. Since the centroid of the initial phase probability distribution is used in calculating the initial map for modification, this assumption is clearly wrong.

The approaches considered here try to improve the combined phase estimates by isolating that portion of the modified map coefficients which is independent of the initial Fourier coefficients used in the map calculation. These modified map coefficients may then be used in the σ_A calculation to form independent probability estimates. For the resultant modified phase probability distribution to be independent of the initial phase distribution, each modified map coefficient must be independent of the corresponding initial map coefficient both in magnitude and phase.

In order to obtain independent phase estimates, a new set of adjusted map coefficients, $F_{\rm adj}(\mathbf{h})$, will be calculated, which aim to be independent of the initial map coefficients $F_{\rm init}(\mathbf{h})$.

One way of ensuring that the adjusted map coefficients are independent of the initial map coefficients is to obtain them from a completely independent source, so that the initial map coefficients play no part in the determination of the new map coefficients. Clearly this is impractical for density-modification calculations, which can only improve existing maps. However, since phase combination is performed one reflection at a time, it is possible to produce a scheme in which the new map coefficient for a particular reflection depends only on the initial map coefficients for other reflections. This guarantees that the modified map coefficient is independent of the initial map coefficient as long as initial map coefficients with differing Miller indices are independent from each other. Since conventional MIR/MAD phasing does not exploit phase relationships between reflections, this assumption is valid for experimental phasing.

An important limitation of this approach is that it is only valid for the first cycle of density modification. Density modification exploits phase relationships between reflections and so the modified phases for reflections of differing Miller index are no longer independent. The breakdown of the approach is shown in §4.7 and gives rise to the same problems which were encountered in traditional density-modification procedures, as documented by Cowtan & Main (1996). A full theoretical model of a multicycle density-modification calculation might suggest an answer to this problem, but such a model is not yet available. Nonetheless, in practical tests the approach described here provides a significant improvement over existing methods.

The resulting adjusted map coefficient $F_{\rm adj}(\mathbf{h})$ will vary as the initial map coefficients for other reflections $F_{\rm init}(\mathbf{k})$, $\mathbf{k} \neq \mathbf{h}$ vary, but will be unchanged as $F_{\rm init}(\mathbf{h})$ is varied. This may equivalently be stated in terms of the derivative of $F_{\rm adj}(\mathbf{h})$ with respect to $F_{\rm init}(\mathbf{h})$,

$$\frac{\partial F_{\rm adj}(\mathbf{h})}{\partial F_{\rm init}(\mathbf{h})} = 0 \;\forall\; F_{\rm init}(\mathbf{h}). \quad (10)$$

No restriction is placed on the general derivative term $\partial F_{\rm mod}(\mathbf{h})/\partial F_{\rm init}(\mathbf{k})$ for $\mathbf{k} \neq \mathbf{h}$.

4. Approaches to correlation removal

4.1. Reflection omit

The reflection-omit scheme, described by Cowtan & Main (1996), removes the dependence from the modified map coefficients by omitting $F_{\rm init}(\mathbf{h})$ from the calculation of the map which is used to generate $F_{\rm mod}(\mathbf{h})$. It is not practical to calculate a map with each reflection omitted in turn, so it is normal to divide the reflections into 10–20 sets and calculate a map with each set omitted in turn. Each map is modified and back-transformed to produce estimates for the omitted reflections.

An arbitrary density modification is a set of functions $M_{\mathbf{h}}$ which return a value for each modified map coefficient dependent on the values of all the initial map coefficients,

$$F_{\rm mod}(\mathbf{h}) = M_{\mathbf{h}}[(F_{\rm init}(\mathbf{k}_1), F_{\rm init}(\mathbf{k}_2), \ldots, F_{\rm init}(\mathbf{k}_i = \mathbf{h}), \ldots, F_{\rm init}(\mathbf{k}_N)]. \quad (11)$$

In a full omit calculation, the initial map coefficient $F_{\rm init}(\mathbf{h})$ is omitted from the calculation of $F_{\rm mod}(\mathbf{h})$,

$$F_{\rm omit}(\mathbf{h}) = M_h[F_{\rm init}(\mathbf{k}_1), F_{\rm init}(\mathbf{k}_2), \ldots, 0, \ldots, F_{\rm init}(\mathbf{k}_N)]; \quad (12)$$

therefore, $F_{\rm omit}(\mathbf{h})$ takes the same value for any value of $F_{\rm init}(\mathbf{h})$. The gradient term $\partial F_{\rm omit}(\mathbf{h})/\partial F_{\rm init}(\mathbf{h})$ is identically zero for any density-modification method.

The difficulty of this approach is that it is time-consuming to use many sets of reflections. However, noise is introduced into the modified map in proportion to the number of reflections in each set, therefore fewer sets lead to noisier phases.

4.2. The γ correction

The γ correction was introduced by Abrahams (1997) as a means to improve the results of solvent-flattening calculations. It represents a correction to the diagonal elements of the gradient matrix, achieved by subtracting the initial map coefficients, scaled by some real factor, from the modified map coefficients,

$$F_{\rm adj}(\mathbf{h}) = F_{\rm mod}(\mathbf{h}) - \gamma F_{\rm init}(\mathbf{h}); \tag{13}$$

therefore,

$$\frac{\partial F_{\rm adj}(\mathbf{h})}{\partial F_{\rm init}(\mathbf{k})} = \frac{\partial F_{\rm mod}(\mathbf{h})}{\partial F_{\rm init}(\mathbf{k})} - \gamma\delta_{\mathbf{h}=\mathbf{k}}, \tag{14}$$

where in Abrahams' formulation the correction γ is assumed to be constant and equal for all reflections and is obtained from theoretical consideration of the density-modification technique. Comparing (10) and (14), it is clear that γ should take the value of the diagonal element of the derivative matrix. The value of γ is constant only if the functions $M_{\mathbf{h}}$ are linear, *i.e.* the $F_{\rm mod}(\mathbf{h})$ are linear combinations of the $F_{\rm init}(\mathbf{h})$. In the general case, γ may also be a function of $\mathbf{h}$.

Abrahams' work concentrates on the calculation of γ for solvent flattening, although he suggests expressions for histogram matching and some other methods. Techniques are examined here for estimating γ for an arbitrary unknown density modification, including cases where γ may take on different values for different groups of reflections.

4.2.1. The theoretical γ. Abrahams (1997) calculates a theoretical value of γ for the case when the map modification can be expressed in the following form,

$$\rho_{\rm mod}(\mathbf{x}) = g(\mathbf{x})\rho_{\rm init}(\mathbf{x}). \tag{15}$$

In the case of solvent flattening, $g(\mathbf{x})$ is a function which is 1 in the protein region and 0 in the solvent region (given appropriate adjustment of the origin term to bring the mean of the solvent to zero). In reciprocal space, the product becomes a convolution,

$$F_{\rm mod}(\mathbf{h}) = (1/V)\sum_{\mathbf{k}} G(\mathbf{h}-\mathbf{k})F_{\rm init}(\mathbf{k}), \tag{16}$$

where G is the Fourier transform of g. In this case,

$$\gamma = \frac{\partial F_{\rm mod}(\mathbf{h})}{\partial F_{\rm init}(\mathbf{h})} = (1/V)G(000); \tag{17}$$

thus, γ is constant and equal for all reflections. $(1/V)G(000) = \bar{g}$; therefore, for solvent flattening γ is equal to the fraction of the unit cell occupied by protein.

The extension of this reasoning to averaging is straightforward; however, other techniques such as histogram matching present some difficulties which will be examined in more detail in §4.4. The prediction of γ is further complicated when multiple techniques are applied simultaneously.

Application of the γ correction to solvent flattening has the effect of inverting or 'flipping' the solvent region of the map; therefore, the modified density in the solvent region is anti-correlated with the initial density. However, the weights for the modified phases are normally significantly smaller than the weights for the initial phases, so the improved map after phase combination generally shows positive, if reduced, features in the solvent region.

4.2.2. The empirical γ. It is more convenient for many purposes to have a single algorithm which is applicable to bias reduction for any form of density modification. It would therefore be useful to be able to estimate a value for γ from the values of the structure factors alone without reference to which density modifications have been applied.

One approach is to assume that the modified map coefficient is made up from the initial map coefficient multiplied by γ, plus some new component $F_{\rm indep}(\mathbf{h})$ which is independent in magnitude and phase. As has been shown, the first part of this assumption holds, at least for solvent flattening:

$$F_{\rm mod}(\mathbf{h}) = F_{\rm indep}(\mathbf{h}) + \gamma F_{\rm init}(\mathbf{h}). \tag{18}$$

To calculate an estimate for γ in this expression, assuming that γ is constant for all reflections, we multiply this equation by $F_{\rm init}(-\mathbf{h})$ and sum over all reflections,

$$\sum_{\mathbf{h}\neq 000} F_{\rm mod}(\mathbf{h})F_{\rm init}(-\mathbf{h}) = \sum_{\mathbf{h}\neq 000} F_{\rm indep}(\mathbf{h})F_{\rm init}(-\mathbf{h}) + \gamma \sum_{\mathbf{h}\neq 000} |F_{\rm init}(\mathbf{h})|^2. \tag{19}$$

If $F_{\rm indep}(\mathbf{h})$ is independent of $F_{\rm init}(\mathbf{h})$ in phase, the first summation on the right-hand side of this equation will be small. Then,

$$\gamma = \sum_{\mathbf{h}\neq 000} F_{\rm mod}(\mathbf{h})F_{\rm init}(-\mathbf{h}) \Big/ \sum_{\mathbf{h}\neq 000} [F_{\rm init}(\mathbf{h})]^2. \tag{20}$$

Summation over all reflections means that γ is real.

Applying this correction will remove the correlation between the initial and modified map coefficients. This has the disadvantage that only features which were not present in the initial map coefficients can appear in the adjusted coefficients; it is impossible for the empirical γ-corrected map to confirm features which are already present. Since there will usually be some indication of the correct features in the initial map (*e.g.* flatness in the solvent, agreement between NCS-related densities), the empirical γ will generally be overestimated.

The method as outlined also assumes that γ is constant for all reflections, although it would be possible to calculate γ separately for different groups of reflections: for example, by grouping reflections in resolution shells.

4.2.3. The perturbation γ. A better approach would be to obtain a direct estimate of diagonal terms of the gradient matrix, allowing current features of the data to be reinforced if they are genuinely indicated by the density modification.

This may be achieved by applying a perturbation to the initial map and measuring the size of the corresponding perturbation in the modified map. The density modification must be performed twice, once for the unperturbed and once for the perturbed data.

Let the perturbation to the initial map coefficients be $\Delta F_{\rm init}(\mathbf{h})$ and the corresponding perturbation in the modified map be $\Delta F_{\rm mod}(\mathbf{h})$. Then, by the chain rule,

$$\Delta F_{\rm mod}(\mathbf{h}) \simeq \sum_{\forall \mathbf{k}} \frac{\partial F_{\rm mod}(\mathbf{h})}{\partial F_{\rm init}(\mathbf{k})} \Delta F_{\rm init}(\mathbf{k}). \quad (21)$$

Multiplying both sides by $\Delta F_{\rm init}(-\mathbf{h})$ and summing over some subset H of the reflections,

$$\begin{aligned}\sum_{\mathbf{h}\in H} \Delta F_{\rm mod}(\mathbf{h})\Delta F_{\rm init}(-\mathbf{h}) &\simeq \sum_{\mathbf{h}\in H}\sum_{\forall \mathbf{k}} \frac{\partial F_{\rm mod}(\mathbf{h})}{\partial F_{\rm init}(\mathbf{k})} \Delta F_{\rm init}(\mathbf{k})\Delta F_{\rm init}(-\mathbf{h}) \\ &\simeq \sum_{\mathbf{h}\in H}\left[\frac{\partial F_{\rm mod}(\mathbf{h})}{\partial F_{\rm mod}(\mathbf{h})}|\Delta F_{\rm init}(\mathbf{h})|^2 \right. \\ &\quad \left. + \sum_{\mathbf{k}\neq\mathbf{h}} \frac{\partial F_{\rm mod}(\mathbf{h})}{\partial F_{\rm init}(\mathbf{k})} \Delta F_{\rm init}(\mathbf{k})\Delta F_{\rm init}(-\mathbf{h})\right]. \end{aligned} \quad (22)$$

Rearranging,

$$\begin{aligned}\sum_{\mathbf{h}\in H} \frac{\partial F_{\rm mod}(\mathbf{h})}{\partial F_{\rm mod}(\mathbf{h})}|\Delta F_{\rm init}(\mathbf{h})|^2 &\simeq \sum_{\mathbf{h}\in H}\left[\Delta F_{\rm mod}(\mathbf{h})\Delta F_{\rm init}(-\mathbf{h}) \right. \\ &\quad \left. - \Delta F_{\rm init}(-\mathbf{h})\sum_{\mathbf{k}\neq\mathbf{h}} \frac{\partial F_{\rm mod}(\mathbf{h})}{\partial F_{\rm init}(\mathbf{k})} \Delta F_{\rm init}(\mathbf{k})\right]. \end{aligned} \quad (23)$$

The final summation in (23) (over $\mathbf{k} \neq \mathbf{h}$) is the derivative of the reflection omit equation (12), which could be called $\Delta F_{\rm omit}(\mathbf{h})$. This may be expected to be independent of $\Delta F_{\rm init}(\mathbf{h})$ for the same reasons as for the omit calculation and this term will therefore be small. As a result, if $\partial F_{\rm mod}(\mathbf{h})/\partial F_{\rm init}(\mathbf{h})$ is constant for all reflections in the set H, then

$$\gamma_H = \frac{\partial F_{\rm mod}(\mathbf{h})}{\partial F_{\rm init}(\mathbf{h})}_{\mathbf{h}\in H} \simeq \frac{\sum_{\mathbf{h}\in H} \Delta F_{\rm mod}(\mathbf{h})\Delta F_{\rm init}(-\mathbf{h})}{\sum_{\mathbf{h}\in H} |\Delta F_{\rm init}(\mathbf{h})|^2}. \quad (24)$$

The perturbation γ provides an estimate for any subset of the diagonal terms of the gradient matrix, under the assumption that these terms are equal for the chosen subset of reflections. The resultant adjusted map coefficients will be independent of the initial map coefficients for a map modification for which these elements are equal and constant, *i.e.* the modification must be linear.

4.3. Correlation removal and solvent flattening

The theoretical γ for solvent flattening has already been described in §4.2.1, where it was shown that $\gamma = f_p$, where f_p is the fraction of the unit cell occupied by protein.

The perturbation γ, since it estimates the diagonal terms of the derivative matrix, will agree with the theoretical γ to within the limits of the approximation in (24).

The reflection-omit calculation for solvent flattening can be described as follows. Consider the calculation of $F_{\rm omit}(\mathbf{h})$, in which only the map coefficient $F_{\rm init}(\mathbf{h})$ has been omitted from the map. The initial map coefficients may then be written $F_{\rm init}(\mathbf{k}) - \delta_{\mathbf{h}=\mathbf{k}}F_{\rm init}(\mathbf{h})$. Substituting this expression into the equation for solvent flattening (16) in §4.2.1, we obtain the equation

$$F_{\rm omit}(\mathbf{h}) = (1/V)\sum_{\mathbf{k}} G(\mathbf{h}-\mathbf{k})F_{\rm init}(\mathbf{k}) - (1/V)G(000)F_{\rm init}(\mathbf{h}). \quad (25)$$

Since $(1/V)G(000) = f_p$, this is clearly identical to the result for the theoretical γ. In practice, there will be some noise introduced through simultaneously omitting multiple reflections.

The empirical γ is estimated by rewriting equation (25) as $F_{\rm mod}(\mathbf{h}) = f_p\, F_{\rm init}(\mathbf{h}) + F_{\rm omit}(\mathbf{h})$ and substituting in (20):

$$\gamma = \frac{\sum_{\mathbf{h}\neq 000}\left[f_p F_{\rm init}(\mathbf{h}) + F_{\rm omit}(\mathbf{h})\right]F_{\rm init}(-\mathbf{h})}{\sum_{\mathbf{h}\neq 000}|F_{\rm init}(\mathbf{h})|^2} \quad (26)$$

$$= f_p + \frac{\sum_{\mathbf{h}\neq 000} F_{\rm init}(-\mathbf{h})F_{\rm omit}(\mathbf{h})}{\sum_{\mathbf{h}\neq 000}|F_{\rm init}(\mathbf{h})|^2}. \quad (27)$$

If $F_{\rm omit}(\mathbf{h})$ and $F_{\rm init}(\mathbf{h})$ are independent in phase, the first term in (27) will be dominant and then the empirical γ will agree with the other approaches. This may hold approximately for a noisy initial map, where phase correlation amongst reflection is insignificant. However, if the solvent is already flat in the initial map, then $F_{\rm mod}(\mathbf{h}) = F_{\rm init}(\mathbf{h})$ and $F_{\rm omit}(\mathbf{h}) = (1 - f_p)\, F_{\rm init}(\mathbf{h})$ and the empirical γ is therefore equal to 1. In general, the solvent will be flatter than the protein; therefore, the empirical estimate for γ is generally overestimated.

This is an important distinction between the empirical γ and the other approaches. The other approaches all return a scaled version of the current structure factors when applied to an already flattened map. This reflects the fact that solvent flattening is confirming a feature already present in the map; therefore, the current phasing should be reinforced. By contrast, since the empirical γ only allows new uncorrelated features to appear in the adjusted data, when flattening an already flat map no new phasing is introduced.

As a result, the empirical γ may discard some useful information. However, the other approaches may give rise to problems when flattening is applied over multiple cycles to the same data. This problem is considered further in §4.7.

4.4. Correlation removal and histogram matching

Histogram matching (Zhang & Main, 1990) is an effective complement to solvent flattening, since while solvent flattening provides updated values for the density in the solvent region of the map, histogram matching provides values for the protein region. It is also particularly effective for increasing the resolution of a map (Zhang *et al.*, 1997).

Histogram matching involves the rescaling of density values in the map to obtain some desired density histogram, which is a known function of resolution and arises from the shape and spacing of atoms in protein structures. The rescaling is applied by giving each point in the protein a value based only on the initial value of the density at that point,

$$\rho_{\rm mod}(\mathbf{x}) = f[\rho_{\rm init}(\mathbf{x})], \quad (28)$$

where f is a monotonically increasing function.

In order to obtain an equivalent formula in reciprocal space, the function f can be replaced by a power-series expansion,

$$\rho_{\rm mod}(\mathbf{x}) = c_0 + c_1\rho_{\rm init}(\mathbf{x}) + c_2\rho_{\rm init}(\mathbf{x})^2 + c_3\rho_{\rm init}(\mathbf{x})^3 + \ \ldots \quad (29)$$

The Fourier transform of this equation is

$$\begin{aligned} F_{\rm mod}(\mathbf{h}) = {} & Vc_0\delta_{\mathbf{h}=000} + c_1F_{\rm init}(\mathbf{h}) \\ & + c_2(1/V)\sum_k F_{\rm init}(\mathbf{k})F_{\rm init}(\mathbf{h}-\mathbf{k}) \\ & + c_3(1/V^2)\sum_{\mathbf{k}}\sum_{\mathbf{l}}F_{\rm init}(\mathbf{k})F_{\rm init}(l)F_{\rm init}(\mathbf{h}-\mathbf{k}-\mathbf{l}) + \ \ldots \end{aligned} \quad (30)$$

Ignoring the origin term, each modified map coefficient is primarily dependent on the corresponding initial map coefficient, followed by a Sayre term, a quartet term and so on.

Abrahams suggests that a theoretical γ correction may be calculated for histogram matching (Abrahams, 1997; §5) from the formula $\gamma = \bar{g}$, where $g(\mathbf{x})$ is calculated by dividing the modified map by the initial map according to the formula

$$g(\mathbf{x}) = [\rho_{\rm mod}(\mathbf{x}) + c(\mathbf{x})]\big/[\rho_{\rm init}(\mathbf{x}) + c(\mathbf{x})], \quad (31)$$

where $c(\mathbf{x})$ is chosen to ensure that the numerator is zero at any map point where the denominator is zero. Unfortunately, $c(\mathbf{x})$ is under-determined; for example, $c(\mathbf{x}) = -\rho_{\rm mod}(\mathbf{x})$ satisfies this condition but gives $g(\mathbf{x}) = 0$ for all $\mathbf{x}$. From (29), when $\rho_{\rm init}(\mathbf{x}) = 0$, $\rho_{\rm mod}(\mathbf{x}) = c_0$, so in practice it is sufficient to add a constant in the numerator to avoid singularities. An additional feature of histogram matching is that the modified map is independent of the mean and variance of the initial map and the positions of the zeros in the initial map are arbitrarily dependent on its origin term.

A slight modification of this approach which avoids these difficulties is to calculate the least-squares straight line fitting $\rho_{\rm mod}(\mathbf{x})$ to $\rho_{\rm init}(\mathbf{x})$. The gradient of this line will give an estimate for γ and the intercept is ignored. m and c are calculated to minimize R in the expression

$$R = \sum_{\mathbf{x}}\{[m\rho_{\rm init}(\mathbf{x}) + c] - \rho_{\rm mod}(\mathbf{x})\}^2, \quad (32)$$

giving

$$\gamma = m = \frac{\overline{\rho_{\rm init}\rho_{\rm mod}} - \overline{\rho_{\rm init}}\ \overline{\rho_{\rm mod}}}{\overline{\rho_{\rm init}^2} - \overline{\rho_{\rm init}}^2}. \quad (33)$$

The Fourier transform of this equation is identical to the empirical estimate for γ described in §4.2.2.

Some problems with this approach appear when we consider the derivatives of the modified map with respect to the initial map coefficients. From (30),

$$\begin{aligned} \frac{\partial F_{\rm mod}(\mathbf{h})}{\partial F_{\rm init}(\mathbf{h})} = {} & c_1 + c_2(2/V)F_{\rm init}(0) \\ & + c_3(3/V^2)\sum_{\mathbf{k}}F_{\rm init}(\mathbf{k})F_{\rm init}(\mathbf{h}-\mathbf{k}). \end{aligned} \quad (34)$$

The density modification is non-linear; therefore, no γ correction will completely remove the effect of $F_{\rm init}(\mathbf{h})$ on $F_{\rm mod}(\mathbf{h})$. Using $\gamma = c_1$ will give zero gradient near $F_{\rm init}(\mathbf{h}) = 0$; the perturbation γ will give zero gradient near the current values of $F_{\rm init}$ and the empirical γ will force the integral of the gradient between zero and the current value to zero.

In this case, the reflection-omit method behaves differently from the various γ corrections because the dependence of $F_{\rm mod}(\mathbf{h})$ on $F_{\rm init}(\mathbf{h})$ is explicitly removed, and therefore the gradient term is constant and zero even for non-linear density modifications.

4.5. Correlation removal and averaging

When performing averaging alone (*i.e.* without solvent flattening), the region of the cell outside the averaging mask is unmodified. The density inside the averaging mask is scaled down by a factor of $1/N_{\rm ncs}$ and combined with the reoriented density from elsewhere in the cell (where $N_{\rm ncs}$ is the order of the non-crystallographic symmetry). Each map coefficient of the rotated density will be largely independent of the corresponding initial map coefficient; the value of γ, reflecting the fraction of the original signal in the modified map, is expected to be $f_s + f_p/N_{\rm ncs}$, where f_s and f_p are the solvent and protein fractions, respectively.

When solvent flattening is applied, the solvent density is no longer conserved and so the expected value of γ will be $f_p/N_{\rm ncs}$, as given by Abrahams (1997).

Note that averaging may be described as multiplication of the original signal by some mask function, which may now contain the values 1 and $1/N_{\rm ncs}$ (and 0 if flattening is performed), followed by addition of a signal which is somewhat independent of the original density. The strong parallel with the solvent-flattening calculation suggests that the behaviour of the various approaches to correlation removal will be similar to the case of solvent flattening.

4.6. Correlation removal and other density modifications

Multi-resolution modification is a technique first employed in the *dm* package, version 1.8 (Cowtan, 1998) and exploits the fact that electron-density histograms have been predicted over a wide range of resolutions. Solvent flattening and histogram matching are therefore applied at several resolutions as follows. A low-resolution map is initially calculated from a set of reflections truncated to the lower resolution. This map is modified by solvent flattening and histogram matching using the electron-density histogram at that resolution. The resulting map coefficients (which extend to higher resolution) are averaged with the initial map coefficients. The new map coefficients are then used to calculate a higher resolution map. The process is then repeated using higher resolution cutoffs until all the data has been included. Most of the improvement is obtained by performing density modification at two resolutions, with comparatively small gains from further stages.

When the perturbation method was used to estimate the γ correction in this case, it was found that different values of γ are required in the different resolution shells. In this case, it is necessary to calculate the γ correction as a function of resolution. This approach may also be necessary for other complex density modifications; for example, atomization calculations such as the *ARP*/*wARP* procedure (Lamzin & Wilson, 1997).

4.7. Density modification over multiple cycles

The problem of correlation has further implications when density modification is applied over multiple cycles. As was shown in §4.3, application of solvent flattening to an already flattened map will reinforce the existing phasing, since the solvent flattening is assumed to be new information. Over many cycles, this would lead to the solvent constraint overwhelming the contribution from the experimental phasing.

This problem is addressed in practical density-modification implementations by performing phase combination from the weighted modified phases back to the initial experimental phases at each cycle, rather than to the phases obtained from the previous density-modification cycle as might be expected. In practice, if phase combination is performed using the phase probability distributions from the previous cycle, the figures of merit converge rapidly towards 1.0 with little improvement in the phases after the first cycle.

A further difficulty arises because the density constraints imply phase relationships throughout reciprocal space. The bias-reduction techniques described in this paper depend on assuring that $F_{\rm adj}(\mathbf{h})$ is primarily dependent on $F_{\rm init}(\mathbf{k})$, $\mathbf{k} \neq \mathbf{h}$, and is minimally dependent on $F_{\rm init}(\mathbf{h})$. However, after a single cycle of density modification, all the other reflections will contain contributions from $F_{\rm init}(\mathbf{h})$, which in turn will influence the value of $F_{\rm adj}(\mathbf{h})$ on subsequent cycles.

A simple illustration can be seen in the case of applying solvent flattening twice to the same map. In the case of 50% solvent, Abrahams (1997) showed that the application of γ-corrected solvent flattening was identical to flipping (inverting) the density in the solvent region (introduced by Abrahams & Leslie, 1996). The resulting map coefficient $F_{\rm adj}(\mathbf{h})$ is only dependent on $F_{\rm init}(\mathbf{k})$, $\mathbf{k} \neq \mathbf{h}$, and not $F_{\rm init}(\mathbf{h})$. However, applying the same density modification again flips the solvent back to its initial value; therefore, the initial data is restored through phase relationships with the rest of the reflections.

Let $F_0(\mathbf{h})$ be the coefficients of the initial map, $F_1(\mathbf{h})$ the coefficients after the first modification of the map and $F_2(\mathbf{h})$ the coefficients after the second modification of the map. For a γ-corrected flattening calculation, the coefficients are related as follows

$$F_1(\mathbf{h}) = (1/V) \sum_{\mathbf{k}\neq\mathbf{h}} G(\mathbf{h}-\mathbf{k}) F_0(\mathbf{k}), \tag{35}$$

$$F_2(\mathbf{h}) = (1/V) \sum_{\mathbf{k}\neq\mathbf{h}} G(\mathbf{h}-\mathbf{k}) F_1(\mathbf{k}). \tag{36}$$

The derivative matrix of $F_2(\mathbf{h})$ with respect to $F_0(\mathbf{l})$ is then

$$\begin{aligned}\frac{\partial F_2(\mathbf{h})}{\partial F_0(\mathbf{l})} &= \sum_{\mathbf{k}} \frac{\partial F_2(\mathbf{h})}{\partial F_1(\mathbf{k})} \frac{\partial F_1(\mathbf{k})}{\partial F_0(\mathbf{l})} \\ &= (1/V^2) \sum_{\mathbf{k}\neq\mathbf{h},\mathbf{l}} G(\mathbf{h}-\mathbf{k}) G(\mathbf{k}-\mathbf{l}).\end{aligned} \tag{37}$$

The diagonal elements of this matrix are given by

$$\frac{\partial F_2(\mathbf{h})}{\partial F_0(\mathbf{h})} = (1/V^2) \sum_{\mathbf{k}\neq\mathbf{h}} |G(\mathbf{h}-\mathbf{k})|^2. \tag{38}$$

Even if the diagonal elements of both derivative matrices are zero, the diagonal elements of the product will be non-zero and will increase with the strength of the phase relationships in reciprocal space.

To counter this problem, it might be possible to perform bias reduction with respect to several maps from different stages of the calculation; however, in a real calculation the effect of phase weighting and combination at each stage make such an approach extremely complex.

5. Test calculations

To test some of the ideas described here, data were used from the structure of RNAase from *Streptomyces aureofaciens* (Ševčík *et al.*, 1991). The structure consists of two molecules of 96 amino acids in the asymmetric unit, including one α-helix and a twisted three-strand antiparallel β-sheet. The structure was solved using multiple isomorphous derivatives and refined to 1.8 Å.

This data set was chosen because the derivative data were all available and the structure was suitable for testing of both averaging and non-averaging calculations. The phasing was calculated using an earlier data set to 2.4 Å and the two weaker (mercury and iodine) derivatives of the three collected at that resolution. MIR phases were calculated using the *SHARP* program (La Fortelle & Bricogne, 1997); *SHARP* was chosen because it not only provides good phase estimates, but has less tendency to overestimate the accuracy of the estimated phases (and thus the FOMs) than other programs. The mean figure-of-merit of the initial phases to 2.5 Å resolution was 0.35, although beyond 3.1 Å there was only one derivative and the phasing was very weak. The mean phase error was $\sim$73° and the map correlation to the final map was 0.40. The resulting map shows broken connectivity along the main chain and many links between chains; this map would be difficult to interpret and thus provides an effective trial for density modification.

The various correlation-removal techniques described in this paper were implemented in an experimental version of the *dm* density-modification software (Cowtan, 1998). For the purposes of understanding the propagation of bias in phase-improvement calculations, some simplifications were made in the density-modification algorithm.

(i) All reflections are used at every cycle of the calculation. It has often been the practice to introduce reflections during the course of a density-modification calculation; for example, to perform gradual phase extension for lower to higher resolution shells and reduce overestimation of FOMs for weakly extrapolated reflections. (However, use of all reflections is the recommended approach for the *dm* software in reflection-omit mode).

(ii) The calculations are run for a small number of cycles. Performing many cycles of density modification may lead to maps which are more easily interpreted but more noisy when judged in terms of weighted mean phase error or map correlation to the final model. (This is a direct result of the weights being overestimated.) The problem of which map is 'best'

depends on the purpose for which it is to be used, so the behaviour of the calculations is only examined over enough cycles to demonstrate the points raised in this paper.

Thus, the results presented here are intended less as a direct comparison of the various approaches, but are an attempt to elucidate the mechanism of signal propagation and error estimation in phase-improvement calculations.

The perturbation γ was implemented by introducing a perturbation to each acentric reflection whose magnitude was of 10% of the mean intensity in each resolution shell and whose phase was random. The perturbations are used in two ways: firstly, to estimate the mean of the diagonal elements of the derivative matrix in order to test the independence of the initial and modified map coefficients and, secondly, as an estimate for the γ correction to remove that dependence.

5.1. Sources of error in the modified phase probabilties

The initial phases will contain a certain level of error, which ideally should be reflected in the probability distributions attached to those phases. In order to compare different density-modification schemes, additional errors introduced during phase improvement must be compared. Types of error introduced during phase improvement include the following.

(i) Bias from the initial map. This occurs if the density-modification calculation increases the weight of the features already present in the initial map without the support of additional information from the density-modification constraints. This source of error can be examined by applying a noise signal to the initial map and measuring whether that noise signal has been reinforced in the final map. The results are presented in §5.2 and §5.3.

(ii) Over-weighting of the modified phases. This occurs when the estimated phases errors attached to the modified phases are systematically smaller than the actual errors in the phases. This can cripple any subsequent calculation which depends on the phase-error estimates. Over-weighting is tested in §5.4. Furthermore, over-weighting of the phases from an early cycle of a density-modification calculation will also contribute to bias in the following cycles (§5.3).

(iii) Noise introduced during the calculation. Noise is introduced by sources including errors in masks, flattening of ordered solvent, incorrect density histograms and omission of missing or free-set reflections. The relative effect of the combination of bias, over-weighting and noise between different density-modification schemes may be examined by comparing the final maps from phase improvement. These results are presented in §5.5.

5.2. Correlation tests

5.2.1. Estimation of γ. The first tests were to establish the level of dependence between initial and modified map coefficients under a variety of density modifications and bias-reduction schemes. This was achieved by using the perturbation γ to estimate the mean of the diagonal elements of the derivative matrix.

Table 1
Estimated γ corrections for solvent flattening, histogram matching, averaging and averaging plus flattening.

The theoretical γ is calculated for the formulae in §4.3 and §4.5.

Density modification	Estimated γ		
	Theoretical	Perturbation	Empirical
Solvent flattening	0.55	0.56	0.63
Histogram matching	†	1.79	1.97
Averaging	0.725	0.70	0.75
Averaging + solvent flattening	0.275	0.32	0.46

† For histogram matching, the only theoretical estimate for γ is the empirical γ; see §4.4.

Note that the mean of the diagonal elements is not an absolute indicator of independence, since the elements of the diagonal matrix may be scaled by simply scaling the modified map; however, the density-modification techniques in *dm* are all implemented in such a way as to preserve the data on an absolute scale and the values therefore contain significant information.

The values for γ calculated using theoretical, perturbation and empirical estimates with solvent flattening, histogram matching, averaging and flattening plus averaging are shown in Table 1.

For solvent flattening, averaging, and flattening plus averaging, the theoretical values match those predicted by the perturbation γ very well. However, in all cases the empirical γ is overestimated. This suggests that the initial map already agrees to some extent with the density constraints: the solvent is already flatter than the protein, the protein already has a histogram skewed away from Gaussian and the NCS-related domains already agree somewhat. These features are present in the initial map, and the empirical γ tries to ensure that they are not present in the adjusted map.

Note that the estimated γ for histogram matching is significantly greater than 1, revealing that the histogram-matching process is scaling up the initial map coefficients. Whether it is also adding useful information is not apparent at this stage. Further tests reveal that the estimated γ for histogram matching varies with resolution, map quality and solvent content, and prediction of a theoretical value is therefore probably impractical. The empirical γ is overestimated and therefore unsuitable for use as a theoretical estimate.

The theoretical estimates for γ for the other density modifications are confirmed by the perturbation method. The theoretical γ can be calculated for solvent flattening and averaging calculations and only requires a single density-modification step; therefore, this approach is quicker when only these methods are employed.

5.2.2. Application of the γ correction. The estimated values for $\langle \partial F_{\rm mod}(\mathbf{h})/\partial F_{\rm init}(\mathbf{h}) \rangle$ for the adjusted map coefficients arising from the reflection-omit method and the various γ corrections are shown in Table 2. The derivative with no bias correction agrees well with the previous estimate for the perturbation γ (the small differences arise from the use of a different random seed).

Table 2
Estimated $\partial F_{\rm mod}(\mathbf{h})/\partial F_{\rm init}(\mathbf{h})$ after bias correction by reflection-omit and γ corrections for solvent flattening, histogram matching, averaging and averaging plus flattening.

The theoretical γ is calculated using the formulae in §4.3 and §4.5.

Density modification	$\langle\partial F_{\rm mod(h)}/\partial F_{\rm init(h)}\rangle$ after correction				
	None	Omit	Theoretical γ	Perturbation γ	Empirical γ
Solvent flattening	0.54	−0.01	0.01	−0.02	−0.11
Histogram matching	1.77	−0.03	†	−0.01	−0.19
Averaging	0.70	0.01	−0.03	−0.01	−0.06
Averaging + solvent flattening	0.30	0.00	0.02	−0.02	−0.16

† For histogram matching, the only theoretical estimate for γ is the empirical γ; see §4.4.

The reflection-omit method, the theoretical and perturbation γ all give map coefficients which are independent of the initial estimates to within the range of error introduced by the perturbation (again excepting the theoretical γ for histogram matching).

The perturbation γ overestimates the dependence between the initial and modified map coefficients, and as a result the corrected coefficients have an inverse dependence on the initial values. Thus, phase combination with these coefficients will suppress genuine features in the initial map.

The value of $\langle\partial F_{\rm mod}(\mathbf{h})/\partial F_{\rm init}(\mathbf{h})\rangle$ estimated by the perturbation methods was also examined as a function of resolution. Under all the density modifications described above, $\partial F_{\rm mod}(\mathbf{h})/\partial F_{\rm init}(\mathbf{h})$ does not appear to vary significantly with resolution. This result is expected from the theory for solvent flattening and averaging; however, it is interesting that it also holds for histogram matching.

5.3. Correlation over multiple cycles

To examine the effect of correlation over multiple cycles of density modification, the perturbation procedure used to estimate $\langle\partial F_{\rm mod}(\mathbf{h})/\partial F_{\rm init}(\mathbf{h})\rangle$ in the previous section was modified to apply a full phase-improvement calculation.

The adjusted map coefficients were incorporated in a full iterated density-modification calculation as follows. Weights were calculated for the phases associated with the adjusted map coefficients, using the σ_A approach of Read (1986). Phase combination was performed between the newly weighted phases and the MIR phase probabilities (or, in a later test, the phase probability distributions from the previous cycle). The combined phases and weights were used in the calculation of a new weighted map for the next cycle of density modification.

To test for bias over multiple cycles, a perturbation was introduced into the initial map coefficients (but is not present in the phase probability distributions employed in phase combination). The corresponding perturbation in the final map after multiple cycles of density modification was then measured for the various bias corrections and density modifications. Note that these bias estimates are lower bounds, since each phase combination will introduce further bias with respect to the (unperturbed) MIR phases.

Estimated values of $\langle\partial F_{\rm dm}(\mathbf{h})/\partial F_{\rm init}(\mathbf{h})\rangle$, where $F_{\rm dm}(\mathbf{h})$ is the map coefficient after multiple cycles of density modification and phase combination, are shown in Fig. 1. The estimated bias is shown for density-modification schemes using the uncorrected σ_A phase weighting, the perturbation γ correction and for the empirical γ correction. (The reflection-omit and theoretical γ results are very similar to the perturbation γ and so have been omitted). Figs. 1(*a*) and 1(*b*) show the results for solvent flattening and flattening plus averaging, respectively.

In Fig. 1(*a*), the modified phases with no bias correction remain biased with respect to the initial map for the whole of the calculation. After a single cycle, the perturbation γ gives a map which is unbiased with respect to the initial map. However, after a second cycle the map is again strongly biased by the initial map, as was predicted in §4.7. Subsequently, even cycles show bias with respect to the initial map and odd cycles do not. The behaviour of the empirical γ is similar to the perturbation γ with respect to bias. The inclusion of the phase weighting and combination steps means that the level of bias is no longer subject to simple estimation.

Similar tests on the level of bias in the maps after multiple cycles of density modification were conducted using histogram matching alone and NCS averaging alone. In both cases, the results were very similar to those obtained using solvent flattening alone.

Once averaging and solvent flattening are combined (Fig. 1*b*) the problem of bias becomes far less serious in the uncorrected density-modification scheme, suggesting that even with twofold averaging and solvent flattening the phases are reasonably well determined. The perturbation γ shows less oscillation than in previous cases. However, the empirical γ is still overcorrecting, leading to alternate maps which are positively and negatively correlated with the initial map.

Note that these results explain the oscillations in map quality from cycle to cycle observed in Cowtan & Main (1996). Maps on even cycles are systematically worse owing to stronger bias with respect to the initial map.

5.4. Figure-of-merit estimation

The critical test of the bias-reduction schemes described here is the accuracy of the estimated phase errors (figures of merit) which are obtained. Ideally, the figure-of-merit $w_{\rm est}(\mathbf{h})$ should be an estimate of $\cos[\delta\varphi(\mathbf{h})]$, where $\delta\varphi(\mathbf{h})$ is the error in the estimated phase for that reflection.

The figure-of-merit (FOM) overestimation can be measured by calculating the FOM overestimation factor m which gives the best fit between $(1/m)w_{\rm est}(\mathbf{h})$ and $\cos[\delta\varphi(\mathbf{h})]$ for the entire data set. [Thus, for reflections with $w_{\rm est}(\mathbf{h}) = 1$, $\delta\varphi(\mathbf{h})$ should equal 0 and for reflections with $w_{\rm est}(\mathbf{h}) = 0$, $\delta\varphi(\mathbf{h})$ should be evenly distributed over 0–2π.]

The FOM overestimation factor m is shown in Fig. 2(*a*) for the MIR data after 1–6 cycles of solvent flattening. The initial FOMs for the MIR data are slightly overestimated; the error

treatment in the *SHARP* program is comprehensive, so this initial error overestimation may well be the result of errors in the final model phases. Results are shown for no bias correction, perturbation and empirical γ correction. (The reflection omit and theoretical γ results are indistinguishable from the perturbation γ result.)

After a single cycle of solvent flattening with no bias correction, the overestimation of FOMs is considerably worse and worsens further over subsequent cycles. Examining the data itself reveals that the mean estimated FOM $\langle w_{\rm est}(\mathbf{h})\rangle$ rapidly approaches 1.0, even though the phases stop improving and the mean phase error is still $\sim$70°.

With the perturbation γ, the FOM overestimation factor is unchanged after the first cycle, although overestimation occurs on subsequent cycles. Although bias with respect to the initial map only appears in even-numbered cycles, overall bias worsens as the cycles proceed because bias is being introduced with respect to every previous map in the calculation. The mean FOM increases every cycle, but much more slowly than without bias correction.

The empirical γ causes the least overestimation of FOMs, suggesting that the figures of merit are most reliable; however, examining the data shows that the phases are not improving either.

Similar results are obtained for all the bias-removal techniques using histogram matching or averaging.

Fig. 2(*b*) shows the FOM overestimation *m* for the MIR and density-modified data using both solvent flattening and averaging. Note that in this case even with no bias removal the error estimation is much better than for solvent flattening

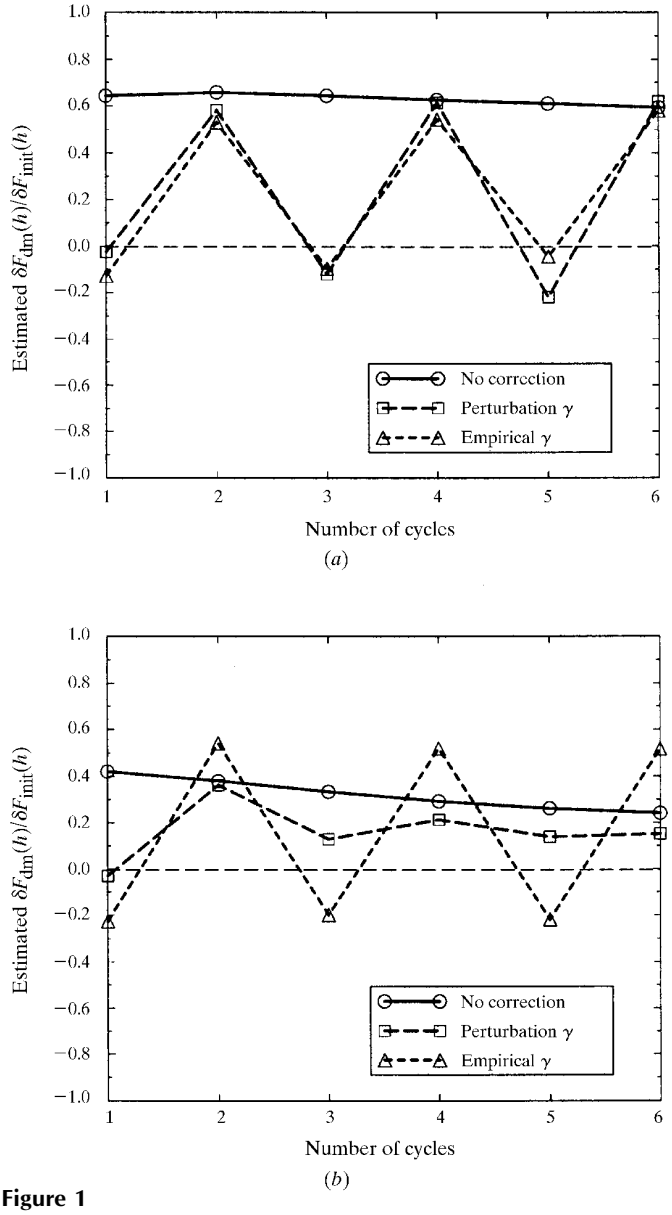

Figure 1
Estimated values of $\langle \partial F_{\rm dm(h)}/\partial F_{\rm init(h)}\rangle$ for (*a*) six cycles of solvent flattening and (*b*) six cycles of flattening plus averaging, as a function of number of density-modification cycles using no bias correction, perturbation γ and empirical γ.

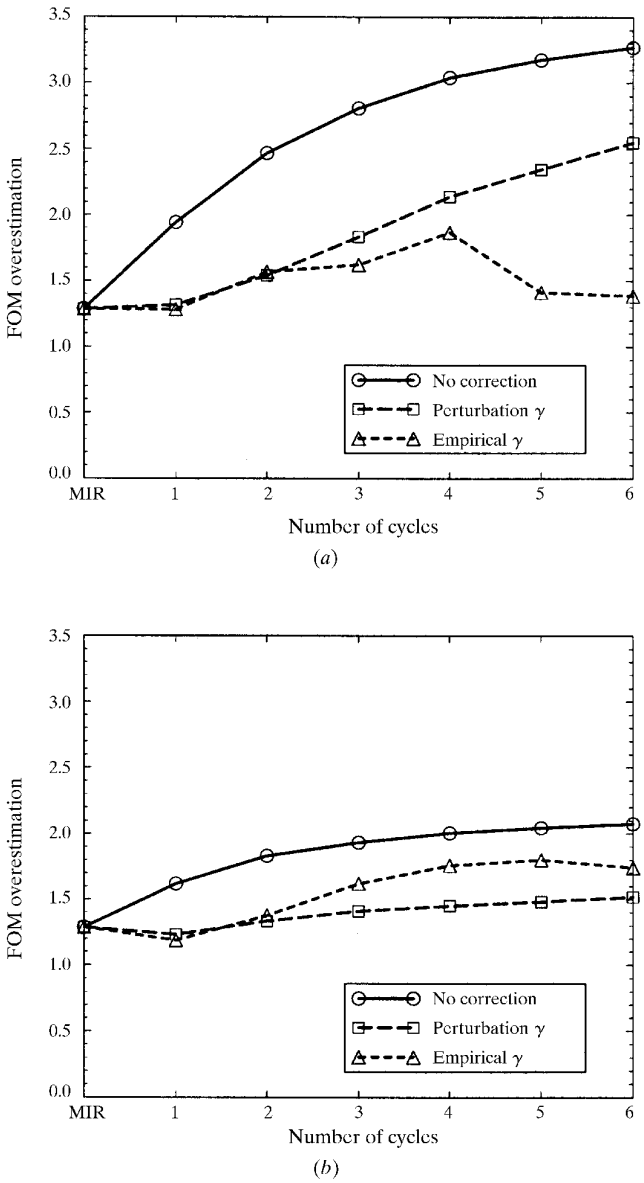

Figure 2
FOM overestimation for MIR data and after 1–6 cycles of (*a*) solvent flattening and (*b*) flattening plus averaging.

alone (Fig. 2). Bias removal is effective in reducing the remaining FOM overestimation. This again confirms that when averaging and another constraint are available, the phases are fairly well determined.

5.5. Phase-improvement results

The quality of the improved map will depend on both the quality of the phases and the quality of the associated weights which are used in calculation of a weighted map. The overall quality of all of this information may be judged by calculating the correlation coefficient between the weighted map using the density-modified phases and weights with the final map obtained from the refined model at the same resolution,

$$\text{correl} = \frac{\overline{\rho_{\rm dm}\rho_{\rm cal}} - \overline{\rho_{\rm dm}}\,\overline{\rho_{\rm cal}}}{[(\overline{\rho_{\rm dm}^2} - \overline{\rho_{\rm dm}}^2)(\overline{\rho_{\rm cal}^2} - \overline{\rho_{\rm cal}}^2)]^{1/2}}. \qquad (39)$$

The map correlation coefficient is highly sensitive to the overall B factor, but for these comparative trials using an identical initial data set it provides an adequate comparison. The map correlations presented here are calculated over the whole electron-density map; however, similar results were obtained when the correlation was calculated over the protein region alone.

Fig. 3 shows the map correlation coefficient for the MIR map and 1–6 cycles of solvent flattening with various bias-reduction techniques. With no bias reduction, the map improves for a single cycle, but beyond that the map deteriorates again. The poor quality of the initial map in this case makes the final map very sensitive to bias from FOM overestimation, hence the poor result in this case.

Reflection-omit and the theoretical and perturbation γ corrections give similar results, with a more substantial improvement on the first cycle and roughly constant correlation coefficient for the next two cycles before the map again deteriorates owing to FOM overestimation. Note that the reflection-omit calculation gives a very slightly poorer result than the other two methods; this is a consequence of the noise introduced by omitting large batches of reflections. As the number of omit sets is increased to 100 or 1000, the results for the reflection-omit calculation approach those for the other methods.

The empirical γ performs well for one cycle, but on the second cycle the map is worse than the initial map. The good features introduced in the first cycle are being removed by the overcorrection applied in the second cycle. To avoid this problem, an additional test was performed: the phase probability distributions obtained after empirical γ correction at each cycle were multiplied by the combined phase probability distribution from the previous cycle (as opposed to the experimental distribution), as was suggested in §4.2.2. The results are better, but more cycles are required to achieve the best map, which is still poorer than for the perturbation γ.

Fig. 4 shows the map correlation coefficient for the MIR map and 1–6 cycles of histogram matching. With no bias correction, there is a similar improvement as for solvent flattening over a single cycle. However, with reflection omit or the perturbation γ there is a much greater improvement over a single cycle, and the improvement continues for two further cycles before oscillating as described in §5.3. The empirical γ behaves in a similar manner as for solvent flattening, although combination back to the previous cycle helps less in this case.

Fig. 5 shows the map correlation coefficient for the MIR map and 1–6 cycles of both solvent flattening and histogram matching. The results are fairly similar to those for histogram matching alone, again suggesting that this is the more powerful constraint. With no bias reduction, the map is somewhat improved over histogram matching alone, suggesting that the problem is somewhat better determined. Interestingly, when

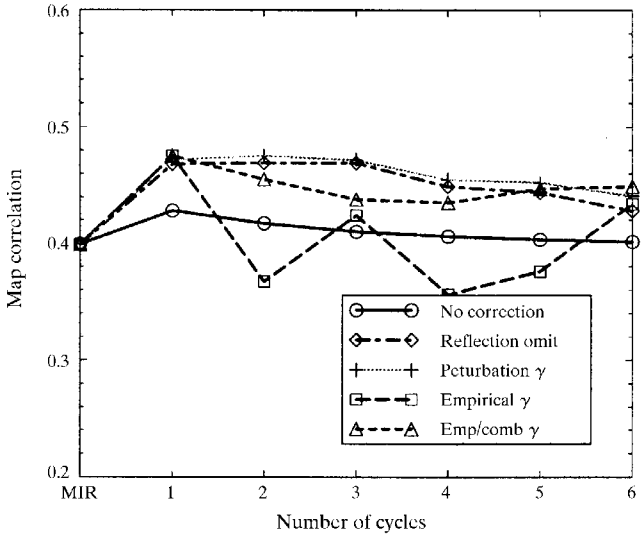

Figure 3
Map improvement measured by map correlation coefficient for maps from MIR and 1–6 cycles of solvent flattening. The results of the theoretical γ correction are omitted because they are indistinguishable from the perturbation γ results. Emp/comb γ refers to the empirical γ correction with phase combination back to the previous cycle.

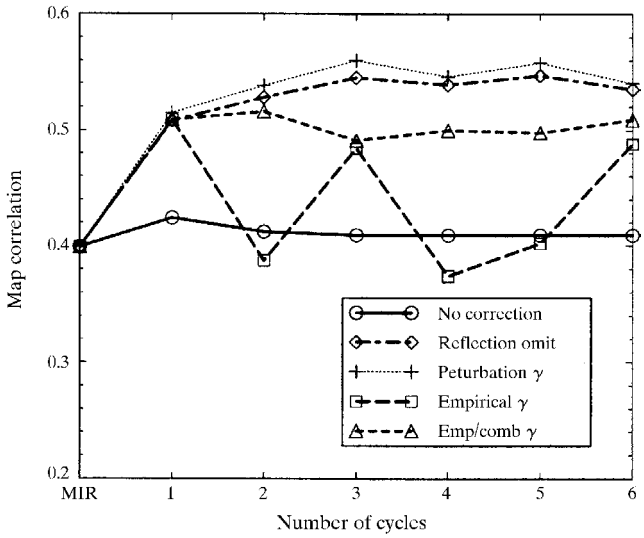

Figure 4
Map improvement measured by map correlation coefficient for maps from MIR and 1–6 cycles of histogram matching.

bias correction is applied the results are very slightly worse than for histogram matching alone. It is suspected that one important effect of both constraints is to increase the contrast between solvent and protein, and actually forcing solvent flatness as well is detrimental at this resolution. As resolution drops, histogram matching becomes less effective, so the combination may still be useful in many cases.

Fig. 6 shows the map correlation coefficient for the MIR map and 1–6 cycles of solvent flattening, histogram matching and averaging. In this case, the results continue to improve as the calculation progresses with or without bias reduction; however, the results from the bias-reduction calculation are better. This again suggests that the calculation is now quite well determined.

The optimum number of cycles is determined by the cumulative effect of the various sources of error in the calculation. In the absence of these sources of error, the map should continue to improve over many cycles. However, overweighting increases quickly with the number of cycles in all cases except the averaging calculation; as a result, phase improvement stops beyond 2–3 cycles in these cases. The bias with respect to the initial map makes a small additional contribution, favouring calculation over an odd number of cycles.

To illustrate the effect of bias removal on the resulting map, a test was performed using using histogram matching alone. The solvent boundary was determined from the final model rather than from the initial map in order to isolate the effects of errors in the solvent boundary. Density modification was performed with and without bias correction and the resulting maps were compared.

The region of density around Phe89 in the structure is shown in Fig. 7 for the MIR map and for density-modified maps with no bias correction and with the perturbation γ correction. The MIR map shows poor connectivity, with several links between adjacent chains and some small breaks in the main-chain density. Density modification with no bias correction breaks some of the cross-links, but does not restore the chain breaks. With the perturbation γ, the connectivity is almost completely restored and much of the side-chain density is present. The phase errors in the three maps are 73, 69 and 61°, respectively.

Table 3
Comparison of MIR, solvent flattened (SF), histogram matched (HM), and flattened + matched (SF + HM) maps at various resolutions.

The improved phases are after a single cycle of density modification and perturbation γ correction. The correlations are all against the 2.5 Å map.

Resolution (Å)	Map correlation			
	MIR	SF	HM	SF + HM
4.5	0.293	0.323	0.332	0.328
4.0	0.335	0.387	0.395	0.393
3.5	0.370	0.425	0.452	0.442
3.0	0.394	0.460	0.501	0.489
2.5	0.399	0.472	0.515	0.506

5.6. Solvent flattening and histogram matching

Early results (Zhang & Main, 1990) suggested that solvent flattening and histogram matching contributed roughly equally to map improvement, with solvent flattening more powerful for refining existing phases and histogram matching more effective for phase extension.

In the light of the results obtained here, those conclusions are seen to be inaccurate. The phase relationships implicit in histogram matching are more powerful than those implicit in solvent flattening (for similar volumes of protein and solvent); however, histogram matching suffers more from bias by the initial map. Of course for extrapolated phases this is not a problem, and so histogram matching appeared particularly effective in this case. Once a bias-reduction method is intro-

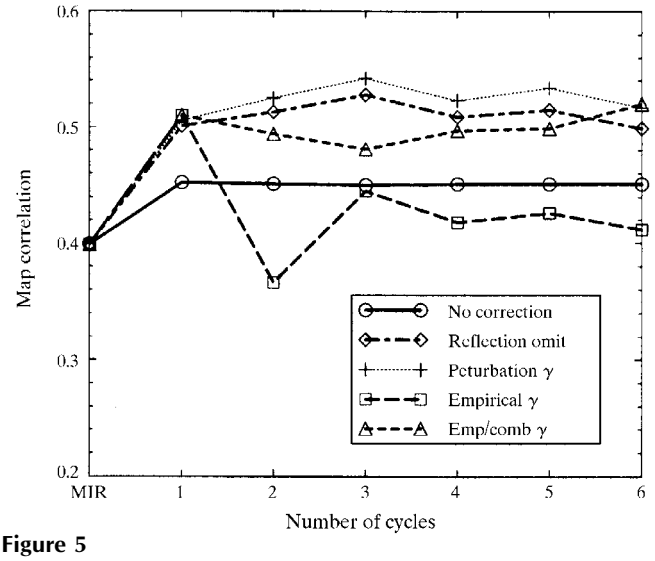

Figure 5
Map improvement measured by map correlation coefficient for maps from MIR and 1–6 cycles of solvent flattening and histogram matching.

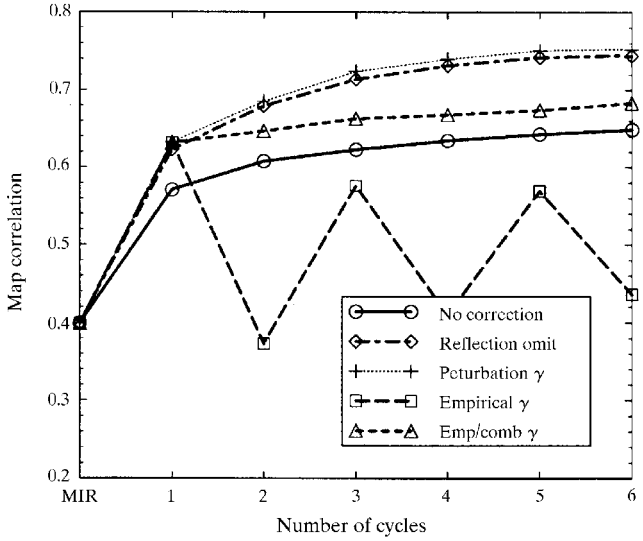

Figure 6
Map improvement measured by map correlation coefficient for maps from MIR and 1–6 cycles of solvent flattening, histogram matching and averaging. Note the altered scale of the *y* axis.

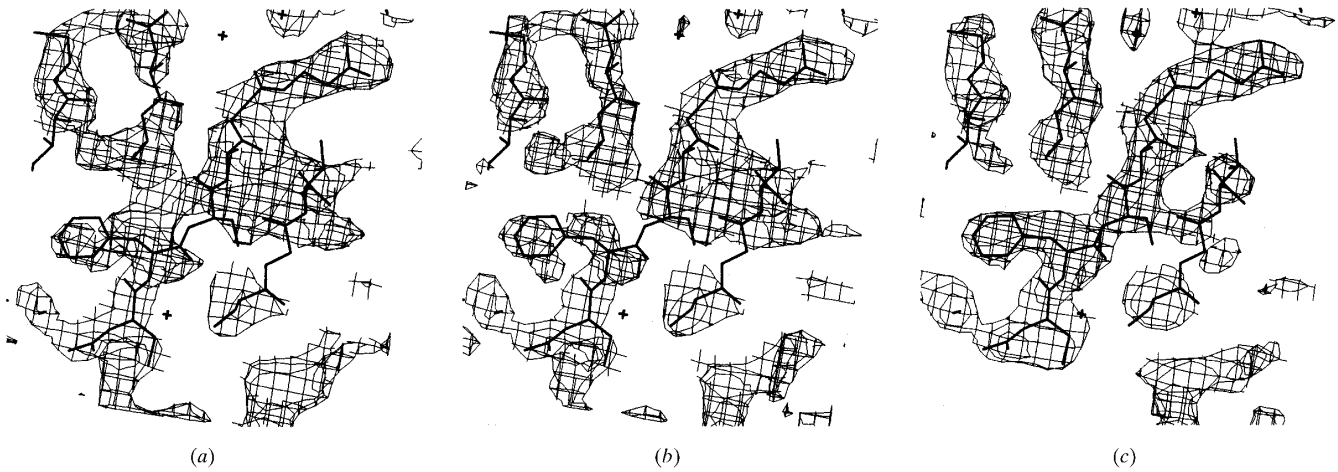

Figure 7
Electron density for RNAase around residue Phe98 for (*a*) MIR map, (*b*) histogram matching with no bias correction, (*c*) histogram matching with perturbation γ correction.

duced, histogram matching is seen to be significantly more effective.

Table 3 shows the results of applying solvent flattening, histogram matching and both techniques at once to the MIR data after truncation to various resolutions (the data is therefore slightly better than a true data set at that resolution). At lower resolutions, all the techniques give limited improvement (which is unsurprising, as the truncated maps are very poor indeed) and the difference between solvent flattening, histogram matching and their combination is small. As resolution improves, the phase improvement becomes more effective, as does the difference between the solvent-flattened map and the histogram-matched map.

6. Conclusions

The treatment of bias in density-modification calculations presented here has contributed to a further understanding of how bias arises and how it may be reduced in order to achieve good maps and realistic estimates of phase error.

Bias reduction by Abrahams' γ correction, although theoretically only applicable to density modifications based on a linear function of the initial map, has been shown to be sufficient for most solvent flattening, averaging, histogram matching and multi-resolution modification. The theoretical estimates of γ given by Abrahams are ideal for combinations of solvent flattening and averaging; however, for the other density modifications the perturbation method provides a good prediction of γ. In the case of extremely non-linear density modifications, it may still be necessary to resort to the slower reflection-omit calculation.

Over multiple cycles of density modification, it has been shown that the final phase-error estimates are underestimated even when bias-reduction methods are used, since after the first cycle of density modification the phase estimates for different reflections are no longer independent. For reliable error estimates, density modification should only be performed over a single cycle. Best maps are obtained when density modification is performed over an odd number of cycles; therefore, with weak density constraints the best map is often obtained after three or five cycles. Error estimates over multiple cycles improve as the density constraints become stronger, so for averaging calculations it is usually possible to run many more cycles of phase improvement and the resulting phase-error estimates may be used with care.

In addition, more light has been shed on the effect of histogram matching. The reciprocal-space phase constraints implied by the histogram constraint are more powerful than the phase constraints implied by solvent flattening; however, this has been masked in the past by the fact that histogram-matched maps were far more strongly biased by the initial map. After bias removal, histogram matching is significantly more powerful than solvent flattening for comparable volumes of protein and solvent.

Dr Cowtan received funding from United Kingdom BBSRC (grant number 87/B03785) for this work.

References

Abrahams, J. P. (1997). *Acta Cryst.* D**53**, 371–376.
Abrahams, J. P. & Leslie, A. G. W. (1996). *Acta Cryst.* D**52**, 30–42.
Cowtan, K. D. (1998). *dm. Software for Density Modification, Versions* 1.8–2.0. Department of Chemistry, University of York, Heslington, York, England.
Cowtan, K. D. & Main, P. (1996). *Acta Cryst.* D**52**, 43–48.
Hendrickson, W. A. & Lattmann, E. E. (1970). *Acta Cryst.* B**26**, 136–143.
La Fortelle, E. de & Bricogne, G. (1997). *Methods Enzymol.* **276**, 472–494.
Lamzin, V. S. & Wilson, K. S. (1997). *Methods Enzymol.* **277**, 269–305.
Murshudov, G. (1997). Personal communication.
Pannu, N. S., Murshudov, G. N., Dodson, E. J. & Read R. J. (1998). *Acta Cryst.* D**54**, 1285–1294.
Read, R. J. (1986). *Acta Cryst.* A**42**, 140–149.
Ševčík, J., Dodson, E. & Dodson, G. G. (1991). *Acta Cryst.* B**47**, 240–253.
Zhang, K. Y. J., Cowtan, K. & Main, P. (1997). *Methods Enzymol.* **277**, 53–64.
Zhang, K. Y. J. & Main, P. (1990). *Acta Cryst.* A**46**, 41–46.