Acta Crystallographica Section A
Foundations of Crystallography
ISSN 0108-7673

Received 1 April 2011
Accepted 11 May 2011

About the hybrid Fourier syntheses: a probabilistic approach

Maria Cristina Burla,[a,b] Benedetta Carrozzini,[c] Giovanni Luca Cascarano,[c] Carmelo Giacovazzo[b,c]* and Giampiero Polidori[a,b]

[a]Department of Earth Sciences, University of Perugia, 06100 Perugia, Italy, [b]Institute of Crystallography, CNR, Via G. Amendola 122/O, 70126 Bari, Italy, and [c]Dipartimento Geomineralogico, Università di Bari, 70125 Bari, Italy. Correspondence e-mail: carmelo.giacovazzo@ic.cnr.it

The difference electron density has recently been revisited *via* the method of joint probability distribution functions [Burla *et al.* (2010). *Acta Cryst.* A**66**, 347–361]. New Fourier coefficients were devised which were the basis of a new *ab initio* method for the solution of the phase problem (*i.e. VLD, vive la difference*). In this paper we study the joint probability distribution functions $P(F, F_p, F_Q)$, where F_Q is the structure factor corresponding to the ideal hybrid Fourier synthesis $\rho_Q = \tau\rho - \omega\rho_p$ and τ and ω are any pair of real numbers. New Fourier coefficients for the calculations of any hybrid synthesis are obtained, and the properties of the corresponding electron-density maps are discussed. The first applications show the correctness of our theoretical approach and suggest possible applications in phasing procedures.

1. Notation

We will use the same notation employed by Burla, Caliandro *et al.* (2010) (from now on paper I), reported here to make this paper self-consistent. The necessary supplementary notation is added.

ρ, ρ_p: electron densities of the target and of the model structure.

$\rho_q = \rho - \rho_p$: *ideal difference electron density*; summed to ρ_p it exactly provides ρ, no matter the quality of ρ_p.

$\rho_Q = \tau\rho - \omega\rho_p$: *ideal hybrid electron density*; summed to $\omega\rho_p$ it exactly provides $\tau\rho$, no matter the quality of ρ_p.

N: number of atoms in the unit cell for the target structure.

p: number of atoms in the unit cell for the model structure; usually $p \le N$, but it may also be $p > N$.

$f_j, j = 1, \ldots, N$: atomic scattering factors for the target structure (thermal factor included).

$F = \sum_{j=1}^{N} f_j \exp(2\pi i \mathbf{h}\mathbf{r}_j) = |F| \exp(i\varphi)$: structure factor of the target structure.

$F_p = \sum_{j=1}^{p} f_j \exp(2\pi i \mathbf{h}\mathbf{r}'_j) = |F| \exp(i\varphi_p)$, where $\mathbf{r}'_j = \mathbf{r}_j + \Delta\mathbf{r}_j$: structure factor of the model structure. $\Delta\mathbf{r}_j$ is the misfit between the model and target atomic positions.

$F_q = F - F_p = |F_q| \exp(i\varphi_q)$: structure factor of the *ideal difference structure*.

$F_Q = \tau F - \omega F_p = |F_Q| \exp(i\varphi_Q)$: structure factor of the *ideal hybrid electron density*.

$E = A + iB = R\exp(i\varphi)$, $E_p = A_p + iB_p = R\exp(i\varphi_p)$, $E_q = A_q + iB_q = R_q \exp(i\varphi_q)$, $E_Q = A_Q + iB_Q = R_Q \exp(i\varphi_Q)$: normalized structure factors of F, F_p, F_q and F_Q, respectively.

$\Sigma_N = \sum_{j=1}^{N} f_j^2$, $\Sigma_p = \sum_{j=1}^{p} f_j^2$.

R'_p, R'_q, R'_Q: structure factors pseudonormalized with respect to the target structure (*i.e.* $R'_p = |F_p|/\Sigma_N^{1/2}$, $R'_q = |F_q|/\Sigma_N^{1/2}$, $R'_Q = |F_Q|/\Sigma_N^{1/2}$).

$D = \langle \cos(2\pi \mathbf{h}\Delta\mathbf{r}) \rangle$: the average is performed per resolution shell.

$\sigma_A = D(\Sigma_p/\Sigma_N)^{1/2}$.

$\sigma_R^2 = \langle |\mu|^2 \rangle / \Sigma_N$, where $\langle |\mu|^2 \rangle$ is the measurement error.

$e = 1 + \sigma_R^2$.

$I_i(x)$: modified Bessel function of order i.

$m = \langle \cos(\varphi - \varphi_p) \rangle = I_1(X)/I_0(X) = D_1(X)$, where $X = 2\sigma_A R R_p/(e - \sigma_A^2)$. m is calculated from the (von Mises type) conditional probability $P(\varphi | R, R_p, \varphi_p)$.

EDM: electron-density modification.

DEDM: difference electron-density modification.

RESID: the sum is over the $\mathbf{h}$ reflections.

CORR: correlation factor between the electron-density map calculated by using observed moduli and phases φ_p, and the map computed *via* observed moduli and phases φ calculated from deposited coordinates.

2. Introduction

Calculating an observed Fourier synthesis (say $\rho_{\rm obs}$) is routine work today: observed moduli are combined with model phases *via* a weight m taking into account phase uncertainty,

$$m = D_1[2\sigma_A R R_p/(1 - \sigma_A^2)]. \tag{1}$$

As mentioned in §1, the value of m arises from the joint probability distribution $P(R, R_p, \varphi, \varphi_p)$ obtained by Srinivasan & Ramachandran (1965): the weight is able to take into account errors in the model coordinates and generalizes a

doi:10.1107/S0108767311017843

previous weighting scheme suggested by Sim (1959), based on structure models without error. More recently a paper by Caliandro *et al.* (2005) generalized further on such weight, to also take into account measurement errors,

$$m = D_1[2\sigma_A R R_p/(e - \sigma_A^2)]. \tag{2}$$

Today the residual electron density (see, among others, Cochran, 1951; Henderson & Moffat, 1971; Nixon & North, 1976; Ursby & Bourgeois, 1997) is usually calculated according to Read (1986): he proposed to reduce the model bias component by using the coefficient

$$(m|F| - D|F_p|)\exp(i\varphi_p). \tag{3}$$

In paper I new coefficients for a difference Fourier synthesis were obtained *via* the study of the joint probability distribution

$$P(R, R_p, R_q, \varphi, \varphi_p, \varphi_q). \tag{4}$$

Equation (4) was derived by taking into account errors in both the model and measurements, and suggested the following coefficient,

$$\left[(mR - \sigma_A R_p) - R'_p(1 - D)\left(\frac{e - \sigma_A^2}{1 - \sigma_A^2}\right)\right]\exp(i\varphi_p), \tag{5}$$

containing the classical Read *difference term* $(mR - \sigma_A R_p)\exp(i\varphi_p)$ and the *flipping term* $-R'_p(1 - D) \times (e - \sigma_A^2/1 - \sigma_A^2)\exp(i\varphi_p)$. The applications clearly showed that the difference electron density calculated *via* the coefficients in equation (5) is well correlated with the ideal ρ_q map even if the model structure is random, and coincides with the synthesis calculated *via* the coefficients in equation (3) when the model is well correlated with the target structure. The mean features of the corresponding electron-density map are not traditional: the map shows very strong negative peaks where model atoms do not overlap with target atoms, medium-intensity negative peaks where model and target atoms overlap, and medium-intensity positive peaks where target atoms do not overlap with model atoms.

The unusual properties of the new difference electron density were employed by Burla, Giacovazzo & Polidori (2010) to design the *VLD* (*vive la difference*) algorithm, aimed at passing from an initial random model to the target structure. The method was able to solve small and medium-sized structures, and proteins (Burla *et al.*, 2011).

Numerous papers may be found in the literature dedicated to hybrid Fourier syntheses $\rho_Q = \tau\rho - \omega\rho_p$ (among others, Ramachandran & Srinivasan, 1970; Dodson & Vijayan, 1971; Main, 1979; Vijayan, 1980), where τ and ω are any pair of real numbers. According to Read (1986) they may be estimated *via* the coefficients

$$(\tau m|F| - \omega D|F_p|)\exp(i\varphi_p). \tag{6}$$

From equation (6) coefficients for the observed synthesis are obtained by fixing $\tau = 1, \omega = 0$, coefficients for the difference synthesis by choosing $\tau = 1, \omega = 1$.

On the basis of the above considerations it seems advisable to apply the method of joint probability distribution functions, used in paper I for the study of the difference Fourier synthesis, to derive the best coefficients to calculate hybrid Fourier syntheses. The first aim of this paper is therefore the study of the distribution $P(E, E_p, E_Q)$: we will also discuss the main features of the new syntheses and suggest their potential applications in phasing procedures.

3. The distribution $P(E, E_p, E_Q)$

Let

$$\rho_Q(r) = \tau\rho(r) - \omega\rho_p(r) \tag{7}$$

be the ideal hybrid synthesis: τ and ω are any pair of rational numbers. By Fourier transform of equation (7) the following relation is obtained:

$$F_{Qh} = \tau F_h - \omega F_{ph} = |F_{Qh}|\exp(i\varphi_{Qh}).$$

To obtain an estimate of the synthesis [equation (7)] we will derive the joint probability distribution function $P(E, E_p, E_Q)$ under the following conditions:

(*a*) The coordinates of the vectors $\mathbf{r}_j$, $j = 1, \ldots, N$ are the primitive random variables, assumed to be uniformly distributed in the unit cell.

(*b*) The variables $\Delta\mathbf{r}_j$, $j = 1, \ldots, p$, are local variables randomly distributed around zero.

(*c*) Two supplementary primitive random variables, μ and ϑ, are introduced, arising from the experimental uncertainty of the observed structure-factor moduli. Accordingly

$$F = \textstyle\sum f_j \exp(2\pi i\mathbf{h}\mathbf{r}_j) + |\mu|\exp(i\vartheta):$$

$|\mu|\exp(i\vartheta)$ may be considered to be the complex error over the structure factor.

(*d*) All the primitive random variables are assumed to be statistically independent of each other.

Such a statistical model leads to the following mathematical definitions,

$$\begin{aligned}
A &= \left\{\sum_{j=1}^{N} f_j \cos(2\pi\mathbf{h}\mathbf{r}_j) + |\mu|\cos\vartheta\right\}\Big/(\varepsilon\Sigma_N)^{1/2}\\
B &= \left\{\sum_{j=1}^{N} f_j \sin(2\pi\mathbf{h}\mathbf{r}_j) + |\mu|\sin\vartheta\right\}\Big/(\varepsilon\Sigma_N)^{1/2}\\
A_p &= \sum_{j=1}^{p} f_j \cos[2\pi\mathbf{h}(\mathbf{r}_j + \Delta\mathbf{r}_j)]/(\varepsilon\Sigma_p)^{1/2}\\
B_p &= \sum_{j=1}^{p} f_j \sin[2\pi\mathbf{h}(\mathbf{r}_j + \Delta\mathbf{r}_j)]/(\varepsilon\Sigma_p)^{1/2}\\
A_Q &= \left\{\tau\sum_{j=1}^{N} f_j \cos(2\pi\mathbf{h}\mathbf{r}_j) - \omega\sum_{j=1}^{p} f_j \cos[2\pi\mathbf{h}(\mathbf{r}_j + \Delta\mathbf{r}_j)]\right\}\Big/(\varepsilon\Sigma_Q)^{1/2}\\
B_Q &= \left\{\tau\sum_{j=1}^{N} f_j \sin(2\pi\mathbf{h}\mathbf{r}_j) - \omega\sum_{j=1}^{p} f_j \sin[2\pi\mathbf{h}(\mathbf{r}_j + \Delta\mathbf{r}_j)]\right\}\Big/(\varepsilon\Sigma_Q)^{1/2},
\end{aligned} \tag{8}$$

where ε is the statistical Wilson coefficient, correcting for the expected intensities in reciprocal-lattice zones.

We obtained for Σ_Q, the scattering power corresponding to the Q structure, the following expression:

$$\Sigma_Q = \tau^2\Sigma_N + \omega^2\Sigma_p - 2\omega\tau D\Sigma_p. \tag{9}$$

The value of Σ_Q depends on the quality of the model: it tends to $\Sigma_N(\tau-\omega)^2$ when $\sigma_A = 1$ (in this case D approaches unity and Σ_p approaches Σ_N), and to $\tau^2\Sigma_N + \omega^2\Sigma_p$ when ρ_p progressively loses (up to $D = 0$) its isomorphism with ρ. In the latter case Σ_Q may be much larger than Σ_N, and $\rho_Q(\mathbf{r})$ will show τN positive peaks and ωp negative peaks.

Let us now consider how the normalized structure factors of the three derivative structures are correlated: their correlation or anticorrelation is basic for the next calculations and defines the algebraic expression of the joint probability distribution $P(E, E_p, E_Q)$. We obtain

$$\langle EE_p\rangle = \sigma_A; \tag{10}$$

$$\langle E_pE_Q\rangle = \frac{(\tau D-\omega)\Sigma_p}{\Sigma_p^{1/2}\Sigma_Q^{1/2}} = \frac{\tau\sigma_A\Sigma_N^{1/2} - \omega\Sigma_p^{1/2}}{\Sigma_Q^{1/2}}; \tag{11}$$

$$\langle E\,E_Q\rangle = \frac{\tau\Sigma_N - \omega D\Sigma_p}{\Sigma_N^{1/2}\Sigma_Q^{1/2}}\frac{\tau\Sigma_N^{1/2} - \omega\sigma_A\Sigma_p^{1/2}}{\Sigma_Q^{1/2}}. \tag{12}$$

We note:

(*a*) Since $0 \le \sigma_A \le 1$, $\langle RR_p\cos(\varphi-\varphi_p)\rangle$ is not expected to be negative: its value should increase (up to 1) when the model becomes closer to the target structure.

(*b*) $\langle R_pR_Q\cos(\varphi_p-\varphi_Q)\rangle$ attains its minimum value [say $-\omega\Sigma_p^{1/2}/(\tau^2\Sigma_N+\omega^2\Sigma_p)^{1/2}$] in the case of a complete lack of isomorphism: then E_p and E_Q are strongly anticorrelated. The anticorrelation attains a maximum when $\omega \gg \tau$: this conclusion has to be kept in mind because a strong anticorrelation between E_p and E_Q allows φ_Q to be estimated given φ_p, even when φ_p is random. If $\tau > \omega$ and σ_A is sufficiently large, positive values of $\langle E_pE_Q\rangle$ are allowed: then ρ_Q becomes closer to ρ, and consequently E_p and E_Q are positively correlated.

(*c*) if $\tau \ge \omega$ then E and E_Q are positively correlated particularly when D = 0: in this case $RR_q\cos(\varphi-\varphi_q) \ge \tau[\Sigma_N(\tau^2\Sigma_N+\omega^2\Sigma_p)]^{1/2}$.

The characteristic function of the distribution $P\big(E, E_p, E_Q\big)$ is

$$\begin{aligned} C(u, u_p, u_Q, v, v_p, v_Q) = \exp\Big\{ &-(1/4)[e(u^2+v^2) + (u_p^2+v_p^2) \\ &+ (u_Q^2+v_Q^2) + 2\sigma_A(uu_p+vv_p) \\ &+ 2\sigma_{AQ}(uu_Q+vv_Q) \\ &+ 2\sigma_{ApQ}(u_pu_Q+v_pv_Q)]\Big\}, \end{aligned} \tag{13}$$

where u, u_p, u_Q, v, v_p, v_Q are carrying variables associated with A, A_p, A_Q, B, B_p, B_Q, respectively,

$$\sigma_{AQ} = \frac{\tau\Sigma_N^{1/2} - \omega\sigma_A\Sigma_p^{1/2}}{\Sigma_Q^{1/2}}$$

$$\sigma_{ApQ} = \frac{\tau\sigma_A\Sigma_N^{1/2} - \omega\Sigma_p^{1/2}}{\Sigma_Q^{1/2}}.$$

The distribution $P(A, A_p, A_Q, B, B_p, B_Q)$ is the Fourier transform of equation (13). After some calculations, not quoted for brevity, we obtain in polar coordinates

$$\begin{aligned} &P(R, R_p, R_Q, \varphi, \varphi_p, \varphi_Q) \\ &\quad\simeq \pi^{-3}e^{-1}(\det\mathbf{L})^{-1}RR_pR_Q\exp\{-[\lambda_{11}R^2 + \lambda_{22}R_p^2 \\ &\qquad + \lambda_{33}R_Q^2 + 2\lambda_{12}RR_p\cos(\varphi-\varphi_p) + 2\lambda_{13}RR_Q\cos(\varphi-\varphi_Q) \\ &\qquad + 2\lambda_{23}R_pR_Q\cos(\varphi_p-\varphi_Q)]\}. \end{aligned} \tag{14}$$

The distribution in equation (14) is a six-dimensional Gaussian distribution, the coefficients of which are stated below:

$$(\det\mathbf{L}) = \frac{(e-1)(1-\sigma_A^2)\tau^2\Sigma_N}{e\Sigma_Q},$$

$$\lambda_{11} = \frac{1}{(e-1)},\ \ \lambda_{22} = \frac{\Sigma_Q}{\tau^2\Sigma_N}\frac{1}{1-\Sigma_A^2} + \frac{\Sigma_p}{\tau^2\Sigma_N}\frac{\omega^2}{e-1},$$

$$\lambda_{33} = \frac{\Sigma_Q}{\tau^2\Sigma_N}\left[\frac{1}{e-1} + \frac{1}{1-\sigma_A^2}\right],\ \ \lambda_{12} = -\frac{\omega}{\tau}\left(\frac{\Sigma_p}{\Sigma_N}\right)^{1/2}\frac{1}{e-1},$$

$$\lambda_{13} = -\frac{1}{\tau}\left(\frac{\Sigma_Q}{\Sigma_N}\right)^{1/2}\frac{1}{e-1},$$

$$\lambda_{23} = \frac{\big(\Sigma_p\Sigma_Q\big)^{1/2}}{\Sigma_N}\left[\frac{\omega(e-\sigma_A^2)}{\tau^2(e-1)(1-\sigma_A^2)}\right] - \left(\frac{\Sigma_Q}{\Sigma_N}\right)^{1/2}\frac{\sigma_A}{\tau(1-\sigma_A^2)}.$$

The distribution given by equation (14) is the basic result of this paper.

4. The conditional distribution $P(\varphi_Q\,|\,R, R_p, R_Q, \varphi_p)$

If we have a model, φ_p is known and only two conditional distributions will be of interest: $P\big(\varphi_Q|R, R_p, R_Q, \varphi_p\big)$ and $P(\varphi|R, R_p, R_Q, \varphi_p, \varphi_Q)$. In both cases R_Q is assumed to be a known parameter, but, as will be emphasized in the following sections, only estimates of R_Q are available in practice. Indeed prior knowledge of R_p, R and φ_p does not geometrically fix the value of R_Q. Accordingly, the distributions $P\big(\varphi_Q|R, R_p, R_Q, \varphi_p\big)$ and $P(\varphi|R, R_p, R_Q, \varphi_p, \varphi_Q)$ should be considered as asymptotic, accurate in the case in which the accuracy of the R_Q estimate is large. To provide the user with a simple example, the general practice for the unweighted difference Fourier synthesis was to assume $R_Q = |R - R_p|$ and derive for φ_Q the following estimate: $\varphi = \varphi_p$ if $R - R_p > 0$, $\varphi = \varphi_p + \pi$ if $R - R_p < 0$. Our distribution $P\big(\varphi_Q|R, R_p, R_Q, \varphi_p\big)$ allows the estimation of the accuracy of φ_Q given as R, R_p, R_Q, φ_p, or, more realistically, given as R, R_p, φ_p, and an estimate of R_Q.

In this section we will derive the conditional distribution $P\big(\varphi_Q|R, R_p, R_Q, \varphi_p\big)$; the study of $P(\varphi|R, R_p, R_Q, \varphi_p, \varphi_Q)$ will be discussed in §5. To simplify the calculations we will first use the assumption $\varphi_p \simeq \varphi$ (this holds when both R and R_p are sufficiently large), to obtain

$$P(\varphi_Q|R, R_p, R_Q, \varphi_p) \simeq [2\pi I_0(G_Q)]^{-1} \exp\{G_Q \cos(\varphi_Q - \varphi_p)\}, \tag{15}$$

where

$$G_Q = -2R_Q(\lambda_{13}R + \lambda_{23}R_p).$$

In accordance with §3,

$$G_Q = \frac{2R'_Q}{\tau^2(e-1)}\left\{(\tau R - \omega R'_p) - (\omega - \tau D)\frac{(e-1)}{1-\sigma_A^2}R'_p\right\}, \tag{16}$$

and, in terms of structure factors,

$$G_Q = \frac{2}{\tau^2(e-1)}\frac{|F_Q|}{\Sigma_N}\left\{(\tau|F| - \omega|F_p|) - (\omega - \tau D)\frac{(e-1)}{1-\sigma_A^2}|F_p|\right\}. \tag{17}$$

If we replace the condition $\varphi_p \simeq \varphi$ by the better approximation (Read, 1986) $|F|\exp(i\varphi) = m|F|\exp(i\varphi_p)$, we obtain

$$G_Q = \frac{2R'_Q}{\tau^2(e-1)}\left\{(\tau m R - \omega\sigma_A R_p) - R'_p\left[\omega(1-D) + (\omega - \tau D)\frac{(e-1)}{1-\sigma_A^2}\right]\right\} \tag{18}$$

and, in terms of structure factors,

$$G_Q = \frac{2}{\tau^2(e-1)}\frac{|F_Q|}{\Sigma_N}\left\{(\tau m|F| - \omega D|F_p|) - |F_p|\left[\omega(1-D) + (\omega - \tau D)\frac{(e-1)}{1-\sigma_A^2}\right]\right\}. \tag{19}$$

We observe:

(*a*) G_Q is the sum of two terms: the first includes the classical *difference term* $(\tau m|F| - \omega D|F_p|)$, the second in some conditions flips the contribution of the model electron density. It really flips if ω is sufficiently large (*e.g.* if $\omega > \tau$) and D is sufficiently small (model badly correlated with the target).

(*b*) The reliability parameter G_Q increases with R'_Q, and is inversely proportional to τ^2. Values $\tau \gg \omega$ do not allow reasonable φ_Q estimates when the model is poorly correlated with the target: indeed, for such values of τ, ρ_Q tends to coincide with $\rho_{\rm obs}$, which by hypothesis is a poor representation of the target. Values $\tau \gg \omega$ permit good φ_Q estimates if the model is well correlated with the target: in this case ρ_Q again tends to coincide with $\rho_{\rm obs}$ and φ_Q is expected to be close to φ.

(*c*) The reliability of the φ_Q estimates increases when $\tau \gg \omega$. If the model is poorly correlated with the target, the Read term vanishes and the flipping term becomes quite large. As a consequence $|G_Q|$ becomes very large, so satisfying the condition for good φ_Q estimates. Their quality is also expected to be high when the model is strongly correlated with the target. This result suggests particular attention towards such mixed Fourier syntheses.

The above conclusion agrees well with the result recently obtained by Giacovazzo & Mazzone (2011), according to which the variance in a generic point **r** of $\rho_Q(\mathbf{r}) = \tau\rho - \omega\rho_p$ is proportional to τ^2: *i.e.*

$$\mathrm{var}\,\rho_Q(\mathbf{r}) = \tau^2\,\mathrm{var}\,\rho(\mathbf{r}),$$

where var $\rho(\mathbf{r})$ is the variance at the point **r** of the electron density $\rho(\mathbf{r})$. Evidently, the variance of$\rho_Q(\mathbf{r})$, and therefore the expected error for the φ_Q estimate, is expected to be smaller when $\tau < \omega$ than when $\tau > \omega$.

5. The conditional distribution $P(\varphi\,|\,R, R_p, R_Q, \varphi_p, \varphi_Q)$

From equation (14) the conditional distribution

$$P(\varphi|R, R_p, R_Q, \varphi_p, \varphi_Q) = [2\pi I_0(\xi_Q)]^{-1}\exp[\xi_Q\cos(\varphi - \psi)] \tag{20}$$

is obtained, where ψ is the most probable value of φ, given by

$$\tan\psi = \frac{\omega R'_p \sin\varphi_p + R'_Q\sin\varphi_Q}{\omega R'_p\cos\varphi_p + R'_Q\cos\varphi_Q}, \tag{21}$$

and

$$\xi_Q = 2R[\tau(e-1)]^{-1}[\omega^2 R'^2_p + R'^2_Q + 2\omega R'_p R'_Q\cos(\varphi_p - \varphi_Q)]^{1/2} \tag{22}$$

is its reliability factor.

Equations (21) and (22) provide the best estimate of φ given $R, R_p, R_Q, \varphi_p, \varphi_Q$ *via* the sum of two contributions, the first arising from ρ_p and the second from ρ_Q. Again the reliability factor is large when the ratio ω/τ is large, thus confirming that the most interesting Fourier syntheses are those for which $\omega \geq \tau$.

In equation (21) the pseudonormalized structure factor R'_Q may be expressed in terms of the normalized structure factor R_Q and of the ratio (frequently known *a priori*) Σ_p/Σ_N using equation (9),

$$R'_Q = R_Q(\Sigma_Q/\Sigma_N)^{1/2},$$

where

$$(\Sigma_Q/\Sigma_N)^{1/2} = \left[\tau^2 + \omega^2\frac{\Sigma_p}{\Sigma_N} - 2\omega\tau D\frac{\Sigma_p}{\Sigma_N}\right]^{1/2}.$$

Equation (21) suggests that the most probable value of φ may be derived without any weighting scheme, owing to the fact that $\tau\rho$ is just the sum of $\omega\rho_p$ and ρ_Q. However, in practice, values of E_Q remain unknown and only their estimates become available: this suggests that a more general tangent expression is

$$\tan\psi = \frac{w_p\omega R'_p\sin\varphi_p + w_Q R'_Q\sin\varphi_Q}{w_p\omega R'_p\cos\varphi_p + w_Q R'_Q\cos\varphi_Q} \tag{23}$$

and

$$\xi_Q = 2R(e-1)^{-1}[w_p^2\omega^2 R'^2_p + w_Q^2 R'^2_Q + 2\omega w_p w_Q R'_p R'_Q\cos(\varphi_p - \varphi_Q)]^{1/2}. \tag{24}$$

6. Coefficients for ($\tau F_o - \omega F_c$) Fourier syntheses

The ideal ρ_Q map is essentially a sum of Fourier syntheses and may be decomposed in a pair of components in different ways. For example, the first component may be τF_o and the second component may be ωF_c, but the decomposition may also be made according to the following rule:

$$\rho_Q = (\tau - \omega)\rho + \omega(\rho - \rho_p) \quad \text{if} \quad \tau > \omega,$$
$$\rho_Q = (\tau - \omega)\rho_p + \tau(\rho - \rho_p) \quad \text{if} \quad \tau < \omega.$$

In this case, in accordance with Read (2001), ρ_Q may be calculated:

(i) if $\tau > \omega$, *via* coefficients

$$[(\tau - \omega)m|F| + \omega(m|F| - D|F_p|)]\exp(i\varphi_p);$$

(ii) if $\tau < \omega$, *via* coefficients

$$[(\tau - \omega)D|F_p| + \tau(m|F| - D|F_p|)]\exp(i\varphi_p).$$

In both cases the coefficients reduce to $(\tau m|F| - \omega D|F_p|)\exp(i\varphi_p)$, or equivalently, in the case of E maps, $(\tau mR - \omega\sigma_A R_p)\exp(i\varphi_p)$.

The theoretical results obtained in §4 suggest the following Fourier coefficients:

$$\left\{(\tau mR - \omega\sigma_A R_p) - R'_p\left[\omega(1-D) + (\omega - \tau D)\frac{(e-1)}{1-\sigma_A^2}\right]\right\}\exp(i\varphi_p), \tag{25}$$

or, in terms of structure factors,

$$\left\{(\tau m|F| - \omega D|F_p|) - |F_p|\left[\omega(1-D) + (\omega - \tau D)\frac{(e-1)}{1-\sigma_A^2}\right]\right\}\exp(i\varphi_p). \tag{26}$$

If $\omega = \tau = 1$, equations (25) and (26) reduce to the difference Fourier coefficients derived in paper I.

Simplified expressions may be obtained by neglecting the contribution of $(\omega - \tau D)[(e-1)/(1-\sigma_A^2)] = (\omega - \tau D)[\sigma_R^2/(1-\sigma_A^2)]$ (this is particularly straightforward when the model is far away from the target structure). Then the Fourier coefficients (see Main, 1979)

$$(\tau mR - \omega R'_p)\exp(i\varphi_p) \tag{27}$$

or

$$(\tau m|F| - \omega|F_p|)\exp(i\varphi_p) \tag{28}$$

arise. The coefficients (27) and (28), when used as coefficients of a Fourier synthesis, give rise to an electron-density map which is the difference between a weighted observed Fourier synthesis, scaled by the factor τ, and a calculated synthesis scaled by ω. This schematization however does not help to understand the main features of the map. Its properties can be better understood if the coefficients (27) or (28) are rewritten as

$$[\tau(mR - \sigma_A R_p) - (\omega - \tau D)R'_p]\exp(i\varphi_p) \tag{29}$$

or

$$[\tau(m|F| - D|F_p|) - (\omega - \tau D)|F_p|]\exp(i\varphi_p). \tag{30}$$

We note:

(*a*) The first term in equation (29) coincides with the classical *difference term*. Roughly speaking, it is expected to generate electron-density maps with positive peaks where atoms of the target do not overlap with atoms of the model structure, and negative peaks where atoms of the model do not overlap with atoms of the target structure. The peak intensities are magnified by the factor τ. It is well known that maps calculated using the *difference term* as a coefficient do not provide useful information when the model is poor.

(*b*) The second term of equation (29) flips the model electron density if $\omega > \tau D$: the flipping is stronger when the model is poor. If $\tau = \omega$ we are in the case described in paper I: if the model is very poor, the flipping term is dominant with respect to the difference term and the corresponding difference Fourier synthesis is well correlated with the ideal ρ_q map. If $\tau > \omega$ the corresponding map is constituted by a dominant observed Fourier synthesis plus a difference Fourier synthesis: very high τ/ω ratios makes the latter negligible no matter whether the model is poor or not. The flipping term becomes more and more dominant with increasing values of the ω/τ ratio (this feature is particularly dominant when the model is poor): in this situation the negative model peak intensities are emphasized with respect to the peak intensities of the map calculated *via* Read coefficients. If $\omega \gg \tau$ the Fourier coefficients reduce to those of the flipped model electron density: this extreme case is not very useful for phasing because it only depends on the model.

(*c*) Hybrid Fourier syntheses, as well as observed Fourier syntheses, are often involved in cyclic EDM procedures (Cowtan, 1994, 1999; Abrahams, 1997; Abrahams & Leslie, 1996; Giacovazzo & Siliqi, 1997) where electron-density maps are first modified according to suitable criteria, and then Fourier inverted to generate new structure-factor estimates. It is expected that EDM techniques lead to models more correlated with the target structure. To better describe the role of the hybrid Fourier syntheses calculated with coefficients (27) or (28) in EDM procedures we adopt the following notation: $\varphi_{\rm est}$ and $\varphi_{Q\rm est}$ are the phases used in the calculation of the observed and of the hybrid Fourier synthesis, respectively, $\varphi_{\rm inv}$ and $\varphi_{Q\rm inv}$ are the corresponding phase values obtained by Fourier inversion of the modified maps. We observe:

(*d*) In the observed Fourier syntheses $\varphi_{\rm est} = \varphi_p$ always, and $m|F|$ or $m|E|$ is the coefficient modulus. In a hybrid Fourier synthesis there is always a subset of reflections for which $\varphi_{Q\rm est} = \varphi_p + \pi$ (see Table 1). Therefore, hybrid and observed Fourier syntheses should never be confused (see below).

(*e*) With increasing values of the τ/ω ratio, hybrid Fourier syntheses become more and more similar to the observed synthesis (indeed the percentage of reflections for which $\varphi_{Q\rm est} = \varphi_p + \pi$ diminishes with the ratio τ/ω, see Table 1). If the model is badly correlated with the target such hybrid maps will also be badly correlated with the corresponding ideal hybrid maps.

Table 1
$\varphi_{\rm est}$ values for each type of hybrid Fourier syntheses.

Type	$\varphi_{\rm est}$
$m\|F\|\exp(i\varphi_p)$	$\varphi = \varphi_p$ always
$m\|F\| - D\|F_p\|\exp(i\varphi_p)$	$\varphi = \varphi_p + \pi$ if $\|F\| < D\|F_p\|/m$, otherwise $\varphi_{\rm est} = \varphi_p$
$m\|F\| - \|F_p\|\exp(i\varphi_p)$	$\varphi = \varphi_p + \pi$ if $\|F\| < \|F_p\|/m$, otherwise $\varphi_{\rm est} = \varphi_p$
$2m\|F\| - D\|F_p\|\exp(i\varphi_p)$	$\varphi = \varphi_p + \pi$ if $\|F\| < D\|F_p\|/(2m)$, otherwise $\varphi_{\rm est} = \varphi_p$
$2m\|F\| - \|F_p\|\exp(i\varphi_p)$	$\varphi = \varphi_p + \pi$ if $\|F\| < \|F_p\|/(2m)$, otherwise $\varphi_{\rm est} = \varphi_p$
$\tau m\|F\| - \omega D\|F_p\|\exp(i\varphi_p)$	$\varphi = \varphi_p + \pi$ if $\|F\| < (\omega/\tau m)D\|F_p\|$, otherwise $\varphi_{\rm est} = \varphi_p$
$\tau m\|F\| - \omega\|F_p\|\exp(i\varphi_p)$	$\varphi = \varphi_p + \pi$ if $\|F\| < (\omega/\tau)D\|F_p\|$, otherwise $\varphi_{\rm est} = \varphi_p$

Table 2
Protein 1e8a, two available models of different accuracy.

See text for details. For each model we show, for the lowest-order hybrid Fourier syntheses, the percentage of reflections (perc1 and perc2) for which $\varphi_{Q{\rm est}} = \varphi_p + \pi$.

Synthesis type	perc1 (%)	perc2 (%)
$m\|F\|\exp(i\varphi_p)$	0	0
$(m\|F\| - D\|F_p\|)\exp(i\varphi_p)$	62	53
$(m\|F\| - \|F_p\|)\exp(i\varphi_p)$	95	65
$(2m\|F\| - D\|F_p\|)\exp(i\varphi_p)$	37	22
$(2m\|F\| - \|F_p\|)\exp(i\varphi_p)$	77	29
$(m\|F\| - 2D\|F_p\|)\exp(i\varphi_p)$	86	91
$(m\|F\| - 2\|F_p\|)\exp(i\varphi_p)$	99	95

(*f*) For hybrid syntheses calculated *via* $\tau m|F| - \omega|F_p|$ coefficients, the percentage of reflections for which $\varphi_{Q{\rm est}} = \varphi_p + \pi$ is larger than for a $\tau m|F| - \omega D|F_p|$ synthesis (see Table 1). In other words, $\tau m|F| - \omega|F_p|$ syntheses emphasize the weight of the difference Fourier synthesis. That is the reason why the convergence from a random model to the target needs the combination of the $\tau m|F| - \omega|F_p|$ synthesis with the tangent step described in §5.

(*g*) For a hybrid Fourier synthesis with $\tau < \omega$, the number of reflections for which $\varphi = \varphi_p + \pi$ is larger than for a hybrid Fourier synthesis with $\tau > \omega$ (indeed larger is the number of reflections for which $|F| < (\omega/\tau m)D|F_p|$ or $|F| < (\omega/\tau)D|F_p|$, see Table 1). As a consequence, a hybrid density map calculated with $\tau < \omega$ will differ from an observed synthesis more than a map with $\tau > \omega$.

(*h*) To allow the reader to appreciate some numerical examples of subsets of reflections for which $\varphi_{Q{\rm est}} = \varphi_p + \pi$, Table 2 shows the results for the protein 1e8a (experimental data resolution 1.95 Å, 17 174 measured unique reflections) in two different situations: in the first a poor model is available (mean phase error = 70°), in the second a good model has already been obtained (mean phase error = 49°). The reader will easily verify that:

(i) a hybrid synthesis should never be confused with an observed synthesis: indeed in the hybrid there is always a non-negligible percentage of reflections to which $\varphi_{Q{\rm est}} = \varphi_p + \pi$ is assigned;

(ii) for the $\tau m|F| - \omega|F_p|$ synthesis the percentage of reflections for which $\varphi_{Q{\rm est}} = \varphi_p + \pi$ is always larger than for the $\tau m|F| - \omega D|F_p|$ synthesis;

(iii) the percentages tend to diminish when the model improves, particularly for $\tau m|F| - \omega|F_p|$ syntheses;

(iv) the percentages for $\omega > \tau$ are larger than for $\omega < \tau$.

Both $\tau m|F| - \omega D|F_p|$ and $\tau m|F| - \omega|F_p|$ Fourier syntheses may converge to the ideal $(\tau - \omega)m|F|$ synthesis if a virtuous EDM technique is applied (then F_p tends to F and both m and D tend to unity). Particularly useful cases are $2m|F| - D|F_p|$ and $2m|F| - |F_p|$ maps. The first was originally devised to reduce the model bias. Perfect convergence is not always guaranteed: it depends on the initial error (large initial errors usually do not allow convergence), on data resolution *etc.* Some classes of reflections are more resistant to convergence, owing to the intrinsic nature of the Fourier synthesis. For example, during an EDM procedure using $\tau m|F| - \omega D|F_p|$ syntheses, the reflections for which $|F| < (\omega/\tau m)D|F_p|$ tend to show (after the map Fourier inversion) the value $\varphi_{Q{\rm inv}} = \varphi_{Q{\rm est}} = \varphi_p + \pi$, so disturbing the convergence process. This is mainly due to the intrinsic difference between ρ_Q and ρ, and therefore between φ_Q and φ. If the *VLD* technique is used (*i.e.* the phase indications provided by the modified hybrid Fourier map are combined with the tangent formula as described in §5), this tendency may be diminished.

From the above considerations the following conclusions arise: if one is interested in phasing a target structure starting from a random or very poor model, the most interesting Fourier syntheses are those with $\omega > \tau$: these syntheses explore the phase space faster because they allow rapid changes of the model, and imply accurate estimates of the φ_Q phases. The passage from φ_Q to φ requires an additional step, *i.e.* the use of the tangent procedure described in §5. If one has to refine a model correlated with the target structure (*e.g.* obtained by molecular replacement or by anomalous-dispersion techniques) hybrid Fourier syntheses with $\omega \le \tau$ are also advisable.

7. Applications

The following preliminary tests aimed to check the validity of the probabilistic theory described above: in particular, we checked whether the quality of the hybrid maps complies with theoretical expectations. We used the same 18 proteins employed as test cases in paper I (see Table 3). The models were found *via* molecular replacement using the program *REMO09* (Caliandro *et al.*, 2009): accordingly, φ_p are the phase values available at the end of the molecular-replacement process (without any attempt at phase refinement, to have a wide range of CORR). The test structures were arranged in decreasing order of CORR, so that the high quality models are at the top of Table 3. The columns Target and Model indicate the PDB code of the target and model structure, and RES is the data resolution. The corresponding $\langle\Delta\varphi\rangle$ values (say the average phase error between model and target structure) are also quoted, as reference values for the phase errors reported in Table 4. The reader should note that the CORR values in Table 3 of this paper do not coincide with

Table 3
Details of the test structures.

For each test structure: Target and Model are the PDB codes of the target and of the model, respectively, Res is the target data resolution limit (in Å); NresT and NresM are the number of residues for the target and model, respectively; CORR is the correlation factor between the electron-density map calculated using observed normalized moduli and phases φ_p, and the map calculated *via* observed normalized moduli and phases φ calculated from deposited coordinates; $\langle|\Delta\varphi|\rangle$ is the corresponding average phase error. PDB codes followed by primes indicate the same target structure but different model structures.

Target	Res	NresT	Model	NresM	$\langle\|\Delta\varphi\|\rangle$	CORR
1kf3	1.0	124	7rsa	124	26	0.93
6rhn	2.2	115	4rhn	115	29	0.91
2sar	1.8	192	1ucl, chain *A*	96	41	0.83
1zs0	1.6	163	1i76	163	43	0.79
1a6m	1.0	151	1mbc	153	43	0.77
1na7	2.4	329	1m2r	327	45	0.76
2p0g	2.3	318	2oka	336	50	0.75
1lys	1.7	258	2ihl	129	57	0.63
1kqw	1.8	134	1opa	133	59	0.60
6ebx′	1.7	124	3ebx, 2 copies	124	59	0.59
6ebx	1.7	124	3ebx	62	59	0.59
9pti′	1.2	58	1lri	98	87	0.01
2pby	2.1	1155	1mki	598	89	0.01
2iff	2.6	556	1hem	129	88	0.00
9pti	1.2	58	3ebx	62	88	0.00
1yxa	2.1	740	1qlp	372	89	0.00
1s31	2.7	273	1c8z	265	89	−0.01
1cgn	2.2	127	2ccy	127	90	−0.01

those reported in Table 1 of paper I: the values reported here were obtained using the latest release of *REMO09*. To increase the CORR range and to study the accuracy of our theory for very poor models, the last seven test cases deliberately correspond to false molecular-replacement solutions. Of particular interest is the test case 2iff, for which we used a partial model (129 residues against 556 of the model). In all our calculations maps were calculated *via* normalized moduli.

To compare the various Fourier syntheses with ideal ones [*i.e.* those calculated *via* coefficients $E_{Q\tau\omega} = \tau E - \omega E_p = |E_{Q\tau\omega}|\exp(i\varphi_{Q\tau\omega})$] we used the following notation in Table 3:

$$(\tau m R - \omega\sigma_A R_p) \tag{31}$$

is the best Fourier coefficient according to Read (1986). It will be denoted by $(E_{Q\tau\omega})_{31}$ from now on;

$(\varphi_{Q\tau\omega})_{31} = \varphi_p$ or $\varphi_p + \pi$ according to whether $(E_{Q\tau\omega})_{31}$ is positive or negative;

$(\rho_{Q\tau\omega})_{31}$ will denote the corresponding hybrid map.

$$(\tau m R - \omega\sigma_A R_p) - R'_p\left[\omega(1-D) + (\omega - \tau D)\frac{(e-1)}{1-\sigma_A^2}\right] \tag{32}$$

is the best Fourier coefficient according to this paper. It will be denoted by $(E_{Q\tau\omega})_{32}$ from now on;

$(\varphi_{Q\tau\omega})_{32} = \varphi_p$ or $\varphi_p + \pi$ according to whether $(E_{Q\tau\omega})_{32}$ is positive or negative;

$(\rho_{Q\tau\omega})_{32}$ will denote the corresponding hybrid map.

$\langle\Delta\varphi_{Q\tau\omega}\rangle_{31} = \langle|(\varphi_{Q\tau\omega})_{31} - \varphi_{Q\tau\omega}|\rangle$. It is the average phase error of the hybrid Fourier synthesis, calculated *via* coefficients (31), with respect to the *ideal hybrid Fourier synthesis*. $(\mathrm{CORR}_{Q\tau\omega})_{31}$ is the correlation between the corresponding hybrid electron-density maps.

$\langle\Delta\varphi_{Q\tau\omega}\rangle_{32} = \langle|(\varphi_{Q\tau\omega})_{32} - \varphi_{Q\tau\omega}|\rangle$. It is the average phase error of the hybrid Fourier synthesis, according to coefficients (32), with respect to the *ideal hybrid Fourier synthesis*. $(\mathrm{CORR}_{Q\tau\omega})_{32}$ is the correlation between the corresponding hybrid electron-density maps.

To check how well $(\rho_{Q\tau\omega})_{31}$ and $(\rho_{Q\tau\omega})_{32}$ approximate the ideal $\rho_{Q\tau\omega}$ map we will only check the pairs $(\tau, \omega) = (1, 1), (2, 1), (1, 2)$: the trend for different (τ, ω) pairs may be easily derived from our results. In Table 4 for each test structure the values of $(\mathrm{CORR}_{Q\tau\omega})_{31}$, $\langle\Delta\varphi_{Q\tau\omega}\rangle_{31}$, $(\mathrm{CORR}_{Q\tau\omega})_{32}$ and $\langle\Delta\varphi_{Q\tau\omega}\rangle_{32}$ are shown for $(\tau, \omega) = (1, 1), (2, 1), (1, 2)$. We observe:

(i) As a general trend, $(\mathrm{CORR}_{Q11})_{31}$ decreases and $\langle\Delta\varphi_{Q11}\rangle_{31}$ increases with decreasing values of CORR. The results confirm those obtained in paper I and the common belief that the classic difference Fourier synthesis provides useful information only when the model structure is sufficiently accurate.

(ii) In accordance with theoretical expectations, $\langle\Delta\varphi_{Q11}\rangle_{32}$ is nearly equivalent to $\langle\Delta\varphi_{Q11}\rangle_{31}$ when good models are available, and is by far superior in bad or random models, for which it attains the minimum values. $(\mathrm{CORR}_{Q11})_{32}$ is always larger than $(\mathrm{CORR}_{Q11})_{31}$ even for good models: that suggests a better correlation between moduli of the ideal difference synthesis and moduli of the coefficients (32).

(iii) To understand the behavior of the ρ_{Q21} estimates, we notice that

$$\rho_{Q21} = \rho_{Q10} + \rho_{Q11} \equiv \rho + \rho_{Q11}$$

and that any estimate of ρ_{Q21} will be the sum of estimates: *e.g.*

$$(\rho_{Q21})_{31} = \rho_{\mathrm{obs}} + (\rho_{Q11})_{31} \quad \text{and} \quad (\rho_{Q21})_{32} = \rho_{\mathrm{obs}} + (\rho_{Q11})_{32}.$$

If CORR is sufficiently large, ρ_{obs} is expected to be very close to ρ and ρ_{Q11} is expected to be negligible with respect to ρ_{obs}: as a consequence $(\rho_{Q21})_{31}$ and $(\rho_{Q21})_{32}$ will be good approximations of ρ_{Q21} (see the corresponding columns in Table 4). Evidently the good qualities of $(\rho_{Q21})_{31}$ and $(\rho_{Q21})_{32}$ rely on the good quality of ρ_{obs}: that is the reason why hybrid syntheses with $\tau = 2$ and $\omega = 1$ are successfully used to reduce the model bias in protein crystallography.

Let us now consider the extreme cases in which CORR is very small (last lines in Table 4). Then ρ is badly estimated by ρ_{obs} and ρ_{Q11} is no longer negligible with respect to ρ_{obs}. Since ρ_{Q11} is badly estimated *via* coefficients (31), $(\rho_{Q21})_{31}$ will be a poor approximation of ρ_{Q21}. The reader will notice that negative values of CORR are obtained when the model is poor and coefficients (31) are used. That is probably due to the fact that the model has been deliberately translated in the solvent region of the target.

Different behavior is expected for $(\rho_{Q21})_{32}$ when CORR is small: indeed, even if ρ_{obs} is a poor approximation of ρ, $(\rho_{Q11})_{32}$ is a relatively good estimate of ρ_{Q11}: accordingly, $(\rho_{Q21})_{32}$ is a relatively good approximation of ρ_{Q21}, no matter whether the model structure is poor or accurate.

(iv) In accordance with (iii)

$$\rho_{Q12} = \rho_p + \rho_{Q11}$$

Table 4
Average phase errors and correlation values for the test structures.

$\langle\Delta\varphi_{Q\tau\omega}\rangle_{31} = \langle|(\varphi_{Q\tau\omega})_{31} - \varphi_{Q\tau\omega}|\rangle$ and $\langle\Delta\varphi_{Q\tau\omega}\rangle_{32} = \langle|(\varphi_{Q\tau\omega})_{32} - \varphi_{Q\tau\omega}|\rangle$ are the average phase errors of the hybrid Fourier synthesis, calculated *via* the coefficients (31) and (32) respectively, with respect to the *ideal hybrid Fourier synthesis*. $(\mathrm{CORR}_{Q\tau\omega})_{31}$ and $(\mathrm{CORR}_{Q\tau\omega})_{32}$ are the correlation values between the corresponding maps. The calculations are limited to $(\tau, \omega) = (1, 1), (2, 1), (1, 2)$.

Target	$\langle\Delta\varphi_{Q11}\rangle_{31}$ (°)/ $(\mathrm{CORR}_{Q11})_{31}$	$\langle\Delta\varphi_{Q11}\rangle_{32}$ (°)/ $(\mathrm{CORR}_{Q11})_{32}$	$\langle\Delta\varphi_{Q21}\rangle_{31}$ (°)/ $(\mathrm{CORR}_{Q21})_{31}$	$\langle\Delta\varphi_{Q21}\rangle_{32}$ (°)/ $(\mathrm{CORR}_{Q21})_{32}$	$\langle\Delta\varphi_{Q12}\rangle_{31}$ (°)/ $(\mathrm{CORR}_{Q12})_{31}$	$\langle\Delta\varphi_{Q12}\rangle_{32}$ (°)/ $(\mathrm{CORR}_{Q12})_{32}$
1kf3	49/0.68	46/0.71	40/0.81	39/0.82	22/0.96	21/0.96
6rhn	52/0.60	57/0.62	46/0.76	55/0.74	25/0.94	25/0.95
2sar	54/0.56	51/0.64	54/0.64	53/0.65	29/0.91	27/0.92
1zs0	56/0.51	54/0.60	56/0.60	58/0.61	30/0.89	29/0.91
1a6m	48/0.73	46/0.80	52/0.70	51/0.72	40/0.91	38/0.94
1na7	58/0.47	55/0.60	61/.54	64/0.56	32/0.86	29/0.91
2p0g	61/0.48	51/0.58	63/0.52	56/0.51	37/0.88	32/0.91
1lys	63/0.45	64/0.49	67/0.49	68/0.50	42/0.78	42/0.79
1kqw	58/0.38	54/0.62	65/0.37	66/0.43	34/0.82	30/0.90
6ebx′	59/0.38	55/0.61	64/0.36	65/0.42	35/0.83	32/0.89
6ebx	60/0.42	62/0.48	64/0.45	66/0.46	41/0.77	41/0.79
9pti′	59/0.02	24/0.92	88/−0.29	42/0.77	26/0.61	12/0.98
2pby	65/0.03	46/0.71	82/−0.17	65/0.43	38/0.58	28/0.90
2iff	81/−0.01	72/0.39	89/−0.11	83/0.10	56/0.49	52/0.69
9pti	60/0.02	28/0.89	88/−0.27	49/0.70	28/0.63	15/0.97
1yxa	65/0.02	46/0.71	81/−0.16	63/0.43	42/0.53	28/0.90
1s31	66/0.02	48/0.69	83/−0.25	70/0.28	38/0.74	27/0.90
1cgn	66/0.02	50/0.69	81/−0.18	68/0.42	39/0.56	30/0.89

and, for their estimates,

$$(\rho_{Q12})_{31} = \rho_p + (\rho_{Q11})_{31} \quad \text{and} \quad (\rho_{Q12})_{32} = \rho_p + (\rho_{Q11})_{32}.$$

Since ρ_p is fixed by prior information, the only uncertain source for the phases arises from the quality of the ρ_{Q11} estimate.

If CORR is high, ρ_{Q11} is negligible with respect to ρ_p: in these cases ρ_{Q11} is well estimated using both coefficients (31) and (32), and, as a result, both $(\rho_{Q12})_{31}$ and $(\rho_{Q12})_{32}$ are good approximations of ρ_{Q12}.

If CORR is small $(\rho_{Q11})_{31}$ is a poor estimate of ρ_{Q11}, but its combination with ρ_p makes $(\rho_{Q12})_{31}$ a sufficiently good estimate of ρ_{Q12}. Obviously the best ρ_{Q12} estimates are obtained by using coefficients (32) (see the last column in Table 4).

(v) For 2iff the efficiency of coefficients (32) is smaller than in other cases. The reason was already theoretically foreseen in paper I for the difference electron density when $\Sigma_p \gg \Sigma_N$: its correctness was checked by Burla, Caliandro *et al.* (2010), Burla, Giacovazzo & Polidori (2010) and Burla *et al.* (2011) in the first applications of the *VLD* algorithm. To better understand the results for other types of hybrid syntheses, we rewrite coefficient (32) as

$$(\tau mR - \omega\sigma_A R_p) - R_p\Big\{\omega\big[(\Sigma_p/\Sigma_N)^{1/2} - \sigma_A\big] + \big[\omega(\Sigma_p/\Sigma_N)^{1/2} - \tau\sigma_A\big]\frac{(e-1)}{1-\sigma_A^2}\Big\}.$$

When $\Sigma_p/\Sigma_N \gg 1$, σ_A is small (by definition) and the flipping term is vanishing unless $\omega > \tau$. In this last case the flipping term is strengthened and maps calculated *via* coefficients (32) provide a good approximation of the ideal hybrid synthesis. In simpler words, coefficients (32) work better when the random model shows a scattering power equivalent to (rather than smaller than) that of the target. According to this result, in the *VLD* algorithm complete random models are always assumed to be starting points for the phasing process.

(vi) The properties of the hybrid Fourier syntheses with $\tau = 3$, $\omega = 1$ or $\tau = 1$, $\omega = 3$ can be easily estimated from (i)–(iv). It may be stated that syntheses for which $\tau \gg \omega$ would be less able to reduce the model bias during EDM procedures and syntheses with $\omega \gg \tau$ would be too biased toward the model to be useful.

8. Conclusions

We have described a probabilistic method for the study of the joint probability distribution $P(E, E_p, E_Q)$, where the three normalized structure factors are the Fourier transform of the target, of a model and of a hybrid electron density, respectively.

The theory allows the best coefficients for any type of hybrid Fourier synthesis and the estimates of their reliability to be derived. The new coefficients are sums of the Read coefficient and of a flipping component, which is dominant when the model is poor. Emphasis is given to hybrid syntheses of the type $\tau\rho - \omega\rho_p$ with $\omega \geq \tau$: the coefficients suggested in this paper are expected to allow their accurate calculation even when model and target structures are uncorrelated. The applications to practical cases show full agreement between theoretical expectations and experimental results.

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.

Burla, M. C., Caliandro, R., Giacovazzo, C. & Polidori, G. (2010). *Acta Cryst.* A**66**, 347–361.

Burla, M. C., Giacovazzo, C. & Polidori, G. (2010). *J. Appl. Cryst.* **43**, 825–836.

Burla, M. C., Giacovazzo, C. & Polidori, G. (2011). *J. Appl. Cryst.* **44**, 193–199.

Caliandro, R., Carrozzini, B., Cascarano, G. L., De Caro, L., Giacovazzo, C., Moustiakimov, M. & Siliqi, D. (2005). *Acta Cryst.* A**61**, 343–349.

Caliandro, R., Carrozzini, B., Cascarano, G. L., Giacovazzo, C., Mazzone, A. & Siliqi, D. (2009). *Acta Cryst.* A**65**, 512–527.

Cochran, W. (1951). *Acta Cryst.* **4**, 408–411.

Cowtan, K. (1994). *Jnt CCP4/ESF–EACBM Newsl. Protein Crystallogr.* **31**, 34–38.

Cowtan, K. (1999). *Acta Cryst.* D**55**, 1555–1567.

Dodson, E. & Vijayan, M. (1971). *Acta Cryst.* B**27**, 2402–2411.

Giacovazzo, C. & Mazzone, A. (2011). *Acta Cryst.* A**67**, 210–218.

Giacovazzo, C. & Siliqi, D. (1997). *Acta Cryst.* A**53**, 789–798.

Henderson, R. & Moffat, J. K. (1971). *Acta Cryst.* B**27**, 1414–1420.

Main, P. (1979). *Acta Cryst.* A**35**, 779–785.

Nixon, P. E. & North, A. C. T. (1976). *Acta Cryst.* A**32**, 325–333.

Ramachandran, G. N. & Srinivasan, R. (1970). *Fourier Methods in Crystallography.* New York: John Wiley.

Read, R. J. (1986). *Acta Cryst.* A**42**, 140–149.

Read, R. J. (2001). *International Tables for Crystallography*, Vol. F, edited by M. G. Rossmann & E. Arnold, pp. 325–331. Dordrecht: Kluwer Academic Publishers.

Sim, G. A. (1959). *Acta Cryst.* **12**, 813–815.

Srinivasan, R. & Ramachandran, G. N. (1965). *Acta Cryst.* **19**, 1008–1014.

Ursby, T. & Bourgeois, D. (1997). *Acta Cryst.* A**53**, 564–575.

Vijayan, M. (1980). *Acta Cryst.* A**36**, 295–298.