Acta Crystallographica Section A
Foundations of Crystallography
ISSN 0108-7673

Received 17 May 1999
Accepted 26 November 1999

The likelihood function in fiber diffraction

Xiang-Qi Mu* and Lee Makowski

Institute of Molecular Biophysics, Florida State University, Tallahassee, FL 32306, USA. Correspondence e-mail: mu@sb.fsu.edu

The likelihood function is an appropriate target function for refinement of molecular structures using fiber diffraction data. However, its practical application to fiber diffraction faces two significant obstacles: (i) the intensities of layer lines in a fiber diffraction pattern usually arise from the superposition of several terms, each equivalent to a crystallographic structure factor, thereby making the calculation significantly more complex than for the crystallographic case; (ii) to describe a molecular structure at the atomic level based on fiber diffraction data, the radial and phase parts of the atomic coordinates must be treated separately owing to the uniaxial symmetry of the structure. These issues are addressed here in order to derive equations of likelihood functions for fiber diffraction. The special case of a single term on a layer line is treated first followed by extension of the method to the multiterm case. A practical difficulty in implementation of likelihood for the multiterm case is that each term has a different variance. An analytical technique is described that allows the conversion of the unequal-variance case to an equal-variance case. This makes it possible to express the likelihood by an explicit formula, allowing a direct implementation of the likelihood calculation. A cylindrically symmetric model is proposed for error distribution of the atomic coordinates in a helical structure. Variances and offset coefficients of the contributing terms in the likelihood functions are expressed in terms of the variance of the atomic coordinates in the cylindrical reference system.

1. Introduction

Crystalline macromolecular structures are usually refined by searching for the global minimum of a target function. This function is comprised of two parts: an energy part and an experimental data part. In X-ray crystallography, the data part of the target function is usually a least-squares residual $\sum w(|F_o| - k|F_c|)^2$, where $|F_o|$ and $|F_c|$ are the observed and calculated structure amplitudes, and k is a scale factor (*e.g.* Hendrickson, 1985; Brünger *et al.*, 1987). Unfortunately, as many authors have pointed out (Bricogne, 1991, 1993; Read, 1990; Pannu & Read, 1996), a least-squares refinement is not appropriate for this problem because of the form of the expected errors.

The errors in the structure factors come from many sources. According to the central limit theorem, when these errors are added together the overall expected error has a Gaussian distribution. The problem is that while the structure factors (usually complex numbers) have errors with Gaussian distribution, the errors in the structure amplitudes (their 'lengths') may have a non-Gaussian distribution. The least-squares refinement is less likely to succeed for a non-Gaussian distribution of errors. Because of this, a better target function, the likelihood function, should be used for refinement of macromolecular structures. (Bricogne, 1991, 1993; Read, 1990; Pannu & Read, 1996).

The situation for fiber diffraction is more complex. The sample for fiber diffraction is typically composed of parallel or nearly parallel diffracting units, randomly rotated about a common axis. The resulting fiber diffraction pattern corresponds to the cylindrical average of the diffraction expected from a single diffracting unit. Usually the intensity in the pattern, $I(R, l/c)$ (where l is the layer-line number, c is the subunit rise along the fiber axis and R the distance from the meridian of the pattern), is the superposition of several components, each corresponding to a cylindrical harmonic of different order. In diffraction from a helix, the number of cylindrical harmonics contributing to a given layer line is limited to those satisfying simple selection rules (Cochran *et al.*, 1952). Because the cylindrical expansion of a function is made as a sum of Bessel functions, the terms in the expansion are refered to as Bessel components. The number of Bessel components contributing to $I(R, l/c)$ increases with R, greatly increasing the difficulty of solving structures at high resolution.

In a diffraction experiment, structural amplitudes are the observed quantities. For an acentric crystal, they are absolute values of complex numbers. For fiber diffraction, however, they are usually the norm of a multidimensional vector. To calculate this norm, an integral must be performed in this multidimensional space. Estimation of errors in structure amplitudes requires explicit analysis of the propagation of

errors during this complex procedure. Some work has been performed on the statistics of fiber diffraction, *i.e.* estimate of the largest likely values for fiber diffraction R factors (Stubbs, 1989; Millane, 1989*a*), and this work provides a basis for some of the analysis presented here. Similar problems arise in the analysis of intensity overlap in X-ray powder diffraction and Bricogne (1991) demonstrated that likelihood is a better target function for structure refinement against overlapped data than a least-squares residual function.

Intensity overlapping also makes the phase problem in fiber diffraction more complicated than that in crystallography (Stubbs & Diamond, 1975; Namba & Stubbs, 1985). The observed intensity has to be decomposed into its components first, followed by determining the phases for each. An approach based on Bayesian statistics has been suggested recently for this problem (Millane & Baskaran, 1997; Baskaran & Millane, 1997, 1998, 1999*a*, 1999*b*). Simulations show that its performance is superior to that of other techniques.

In this paper, we present the formulation of the likelihood functions in fiber diffraction in order to enable their use in the refinement of helical structures. The atomic coordinates in a crystal are usually described in a Cartesian coordinate system. For crystals, it is assumed that the error distribution is spherically symmetric and that all atoms possess the same error distribution. In a helical structure, the atomic coordinates are usually expressed in a cylindrical system, denoted by (r, φ, z). The assumption of identical error distributions in a Cartesian system corresponds to errors in φ, the angular coordinate, that vary with r, the radial coordinate. This greatly complicates the formulation of the likelihood functions. In this paper, a cylindrically symmetric model is proposed for the distribution of atomic coordinate errors and the radial variation of error in φ is taken into consideration. Finally, all formulas are expressed in terms of atomic coordinates and their errors in a cylindrical coordinate system.

2. Preliminaries

The observed intensities of layer lines in a fiber diffraction pattern are cylindrically averaged:

$$\langle I(R, l/c)\rangle = \sum_{k=1}^{K}(A_k^2 + B_k^2), \tag{1}$$

where A_k and B_k may be written as (Stubbs, 1989)

$$A_k = \sum_{j=1}^{N} f_j J_n(2\pi R r_j)\cos\theta_j \tag{2}$$

$$B_k = \sum_{j=1}^{N} f_j J_n(2\pi R r_j)\sin\theta_j, \tag{3}$$

where N is the total number of atoms in a subunit and

$$\theta_j = -n\varphi_j + 2\pi l z_j/c. \tag{4}$$

(r_j, φ_j, z_j) are atomic coordinates in a cylindrical coordinate system and the Bessel orders (n) for layer line l are limited to those satisfying the selection rule (Cochran *et al.*, 1952). At any given resolution, K is finite and determined by the resolution and the maximum radius of the diffracting particle.

The square root of the intensity, $\langle I\rangle^{1/2}$, may be considered as the 'length', or Euclidean norm, of a $2K$-dimensional vector:

$$\boldsymbol{\mathcal{G}} = \begin{pmatrix} A_1 \\ B_1 \\ A_2 \\ B_2 \\ \vdots \\ A_K \\ B_K \end{pmatrix}. \tag{5}$$

The likelihood function for fiber diffraction is proportional to the conditional probability that the observations would be made, given a particular structural model and measurement errors. The proportionality constant may be set equal to one because our main concern is determining the structural parameters corresponding to the maximum likelihood for a model based on fiber diffraction data. The likelihood is defined as:

$$L = P[(|\boldsymbol{\mathcal{G}}_o|)_{\text{all reflections}}; (|\boldsymbol{\mathcal{G}}_c|)_{\text{all reflections}}], \tag{6}$$

where $|\boldsymbol{\mathcal{G}}_o|$ is the square root of the observed intensity with experimental errors, $|\boldsymbol{\mathcal{G}}_c|$ is that calculated based on a structural model and $P[x; y]$ is the conditional probability of x given y. Assuming that all reflections are independent, the complicated probability function, equation (6), reduces to a product of conditional probabilities of individual reflections:

$$L = \prod_{\text{all reflections}} P[|\boldsymbol{\mathcal{G}}_o|; |\boldsymbol{\mathcal{G}}_c|]. \tag{7}$$

If the distribution of the experimental error is $P[|\boldsymbol{\mathcal{G}}_o| - |\boldsymbol{\mathcal{G}}|]$, then

$$P[|\boldsymbol{\mathcal{G}}_o|; |\boldsymbol{\mathcal{G}}_c|] = P[|\boldsymbol{\mathcal{G}}|; |\boldsymbol{\mathcal{G}}_c|] * P[|\boldsymbol{\mathcal{G}}_o| - |\boldsymbol{\mathcal{G}}|], \tag{8}$$

where $*$ denotes convolution and $\boldsymbol{\mathcal{G}}$ is the true structure factor. $P[|\boldsymbol{\mathcal{G}}|; |\boldsymbol{\mathcal{G}}_c|]$ on the right-hand side of the equation is the probability distribution of the true structure amplitudes given the atomic coordinates and the expected errors.

3. Errors of atomic coordinates in a cylindrical system

The atomic coordinates in a crystal are usually expressed in a Cartesian coordinate system. It is assumed in crystallography that the error distribution of these coordinates is spherically symmetric about the center of mass of an atom and that all atoms possess the same error distribution. It is also reasonable to assume a symmetric error distribution for atomic coordinates in a helical structure and assume that the distribution is the same for all atoms in the structure. As shown later, assuming that the error distribution has cylindrical symmetry greatly facilitates the analysis of errors in the atomic coordinates in a fiber. The assumption of identical error distributions is usually made in the Cartesian coordinate system.

The standard coordinate system in the field of fiber diffraction is the cylindrical system. It is convenient to describe the errors of atomic coordinates within the fiber in a

local cylindrical reference system with its origin at the center of mass of the atom and its axis parallel with the fiber axis. (t, α, ω) are used to denote a particular value of the coordinate error. The error distribution is assumed to be cylindrically symmetric, *i.e.* it is independent of α. The variance of the distribution in the plane perpendicular to the fiber axis is denoted by σ_t^2. The error in the third direction (along ω) is independent of that on the plane and its variance is σ_ω^2. For simplicity, we assume that $\sigma^2 = \sigma_t^2 = \sigma_\omega^2$. If the error distribution is Gaussian, an expected value of any function of t, α and ω, $f(t, \alpha, \omega)$, may be calculated as

$$E = [1/(8\pi^3)^{1/2}\sigma^3] \int_0^\infty \int_0^{2\pi} \int_{-\infty}^{\infty} f(t, \alpha, \omega) \times \exp(-t^2/2\sigma^2) \exp(-\omega^2/2\sigma^2) t \, dt \, d\alpha \, d\omega. \quad (9)$$

It is assumed that all atoms in the fiber possess the same error distribution, *i.e.* the variance σ^2 is the same for all atoms. The atomic coordinates in the fiber reference system, denoted by (r, φ, z), is changed, for example, from $(r, 0, 0)$ to $(r + \Delta r, \Delta\varphi, \Delta z)$ owing to the errors, as shown in Fig. 1. The variations of the coordinates are related with the errors (t, α, ω):

$$\begin{cases} \Delta r = (r^2 + t^2 + 2rt\cos\alpha)^{1/2} - r \\ \Delta\varphi = \arctan[t\sin\alpha/(r + t\cos\alpha)] \\ \Delta z = \omega. \end{cases} \quad (10)$$

$\langle\Delta r\rangle$, $\langle\Delta r^2\rangle$, $\langle\Delta\varphi\rangle$ and $\langle\Delta\varphi^2\rangle$ can be calculated using (9). Usually, all of these expected values are r dependent. However, the dependence is significant only when the atom is located near the fiber axis. The results of numerical calculations are shown over a wide range of r in Figs. 2 and 3. For any noncentral atoms (*e.g.* $r \geq 5\sigma$), $\langle\Delta r\rangle$ is nearly zero (*i.e.* $< 0.1\sigma$) and $(\langle\Delta r^2\rangle)^{1/2}$ is very close to σ (within 99.5%). Similarly, we have $\langle\Delta\varphi\rangle = 0$ and $(\langle\Delta\varphi^2\rangle)^{1/2} = \sigma/r$ for noncentral atoms. Also for noncentral atoms, the variance of θ, defined in (4), is

$$\langle\Delta\theta^2\rangle = \sigma^2(n^2/r^2 + 4\pi^2 l^2/c^2). \quad (11)$$

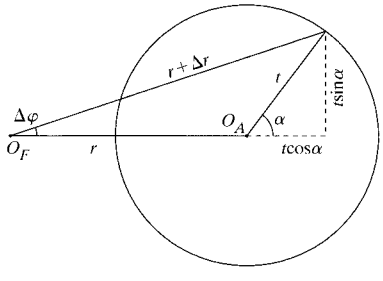

Figure 1
Atomic coordinates in a fiber reference system (origin at O_F) and error distribution in a local cylindrical system (origin at O_A). This is a projection of both systems down the fiber axis. The coordinates of atom A, located at O_A, are $(r, 0, 0)$. The distribution of coordinate errors is cylindrically symmetric around the atom with its symmetric axis parallel to the fiber axis. A specific value of the error in the local system is (t, α, w). The atomic coordinates are changed from $(r, 0, 0)$ to $(r + \Delta r, \Delta\varphi, w)$ owing to this error.

In addition to the variance of coordinate σ^2, the variance $\langle\Delta\theta^2\rangle$ depends also upon n, l and atomic coordinate r. It increases sharply with increasing n and/or l. It can be very large for high-order layer lines and/or for large-order Bessel terms. $\langle\cos\Delta\theta\rangle$ is needed for calculation of the likelihood function, as shown later. For noncentral atoms, $\Delta\varphi = t\sin\alpha/r$ and $\langle\cos\Delta\theta\rangle$ can be derived as follows:

$$\langle\cos\Delta\theta\rangle = [1/(8\pi^3)^{1/2}\sigma^3] \int_0^\infty \int_0^{2\pi} \int_{-\infty}^{\infty} \cos[-(nt/r)\sin\alpha + 2\pi l\omega/c] \times \exp(-t^2/2\sigma^2) \exp(-\omega^2/2\sigma^2) t \, dt \, d\alpha \, d\omega \quad (12)$$

$$= [1/(8\pi^3)^{1/2}\sigma^3] \int_0^\infty t \exp(-t^2/2\sigma^2) \times \left\{ \int_0^{2\pi} \cos[(nt/r)\sin\alpha] \, d\alpha \right\} dt \times \int_{-\infty}^{\infty} \cos(2\pi l\omega/c) \exp(-\omega^2/2\sigma^2) \, d\omega. \quad (13)$$

The second integral in formula (13) is the integral representation of a Bessel function of zero order (Abramowitz & Stegun, 1970, p. 360). The third one is the Fourier cosine transform of the function $\exp(-\omega^2/2\sigma^2)$. So we have

$$\langle\cos\Delta\theta\rangle = [(2\pi)^{1/2}\sigma/(8\pi^3)^{1/2}\sigma^3] \exp(-2\pi^2 l^2\sigma^2/c^2) \times \int_0^\infty t \exp(-t^2/2\sigma^2) 2\pi J_0(nt/r) \, dt \quad (14)$$

$$= \exp[-(\sigma^2/2)(n^2/r^2 + 4\pi^2 l^2/c^2)] \quad (15)$$

$$= \exp(-\tfrac{1}{2}\langle\Delta\theta^2\rangle). \quad (16)$$

The result of integration in (14) was taken from Gradshteyn & Ryzhik (1994, p. 738). As shown in (15) or (16), $\langle\cos\Delta\theta\rangle$ may quickly converge to zero when $\langle\Delta\theta^2\rangle$ is large.

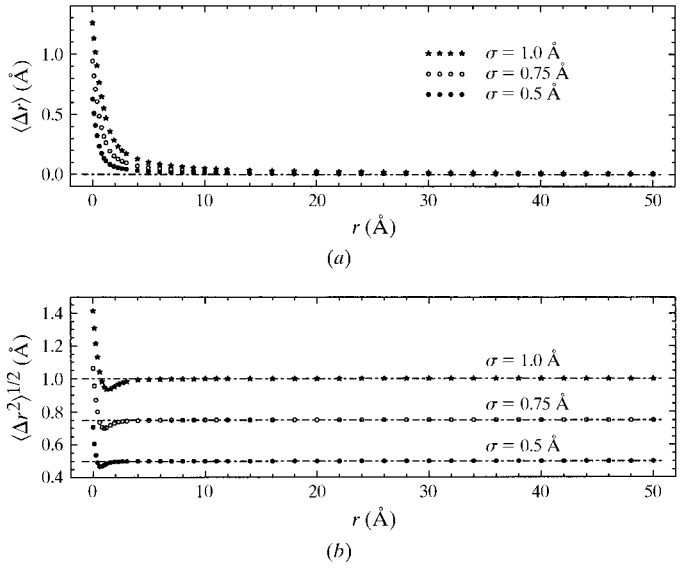

Figure 2
The averaged error of r, $\langle\Delta r\rangle$, and the standard deviation of r, $\langle\Delta r^2\rangle^{1/2}$, for the cylindrically symmetric Gaussian distribution of coordinate errors. The curve of $\langle\Delta r\rangle$ against r and that of $\langle\Delta r^2\rangle^{1/2}$ against r for error distribution of variance $\sigma = 0.5$, $\sigma = 0.75$ and $\sigma = 1.0$ Å are shown. When r is larger than 5σ, $\langle\Delta r\rangle = 0$ and $\langle\Delta r^2\rangle^{1/2} = \sigma$.

A similar derivation shows that $\langle \sin \Delta\theta \rangle = 0$.

4. Likelihood for the case of a single Bessel term

For the special case of a single Bessel-function term on a layer line, the problem reduces to one analogous to an acentric crystal except that special consideration must be made for the atomic coordinates in the fiber structure. In (1), K becomes 1 and A and B (the subscript can be omitted in the single-Bessel-term case) correspond to the real and imaginary parts of crystalline diffraction. In the most general case for a crystal, the corresponding conditional probability is

$$P[\mathbf{F}; \mathbf{F}_c] = (1/2\pi\sigma_c^2) \exp[-|\mathbf{F} - \mathbf{DF}_c|^2/2\sigma_c^2], \qquad (17)$$

where $\mathbf{F}$ and $\mathbf{F}_c$ are structure factors of the true structure and of a related structural model, respectively, $P[\mathbf{F}; \mathbf{F}_c]$ is a two-dimensional Gaussian with centroid of $\mathbf{DF}_c$ and variance of σ_c^2 (*e.g.* Read, 1990; Bricogne, 1991). The conditional probability distribution of structure amplitudes may be developed based on it (*e.g.* Read, 1990).

Equation (17) is also valid for the special case of fiber diffraction in which only a single Bessel-function term makes a contribution to a layer line. Usually, we use $\mathbf{G}$ and $\mathbf{G}_c$ for fiber diffraction in place of $\mathbf{F}$ and $\mathbf{F}_c$. The conditional probability $P[|\mathbf{G}|; |\mathbf{G}_c|]$ may be derived from (17) (*e.g.* Bricogne, 1991) as

$$P[|\mathbf{G}|; |\mathbf{G}_c|] = (|\mathbf{G}|/\sigma_F^2) \exp[-(|\mathbf{G}|^2 + D_F^2|\mathbf{G}_c|^2)/2\sigma_F^2] \times I_0(D_F|\mathbf{G}||\mathbf{G}_c|/\sigma_F^2), \qquad (18)$$

where I_0 is the modified Bessel function of the first kind of zero order. This equation is similar in form to that for an acentric crystal (Read, 1990). To use it for fiber diffraction, we need to know both the relationship between the offset coefficient, D_F, and the atomic coordinates with errors in a cylindrical coordinate system and that between the variance, σ_F^2, and the coordinates. These relationships are shown as follows. Details of the derivation are in Appendix *A*.

$$D_F = S_1 + 4\pi^2 R^2 S_2 \sigma^2 \qquad (19)$$

and

$$\sigma_F^2 = S_3 + 4\pi^2 R^2 S_4 \sigma^2, \qquad (20)$$

where the coefficients S_1, S_2, S_3 and S_4 are

$$S_1 = \frac{\sum_{j=1}^N f_j \langle \cos \Delta\theta_j \rangle J_n(2\pi R r_j) \cos\theta_j}{\sum_{j=1}^N f_j J_n(2\pi R r_j) \cos\theta_j} \qquad (21)$$

$$S_2 = \tfrac{1}{2} \frac{\sum_{j=1}^N f_j \langle \cos \Delta\theta_j \rangle J_n''(2\pi R r_j) \cos\theta_j}{\sum_{j=1}^N f_j J_n(2\pi R r_j) \cos\theta_j} \qquad (22)$$

$$S_3 = \tfrac{1}{2} \sum_{j=1}^N f_j^2 (1 - \langle \cos \Delta\theta_j \rangle^2) [J_n(2\pi R r_j)]^2 \qquad (23)$$

and

$$S_4 = \tfrac{1}{2} \sum_{j=1}^N f_j^2 [J_n'(2\pi R r_j)]^2, \qquad (24)$$

where J_n' and J_n'' are the first and second derivatives of Bessel function J_n with respect to r_j.

It is often more convenient to work with an intensity-based likelihood function. If $|\mathbf{G}| = I^{1/2}$, (18) becomes

$$P[I; I_c] = \left(\frac{1}{2\sigma_F^2}\right) \exp\left(-\frac{I + D_F^2 I_c}{2\sigma_F^2}\right) I_0\left[\frac{D_F (II_c)^{1/2}}{\sigma_F^2}\right]. \qquad (25)$$

In most cases, it is reasonable to assume that the distribution of measurement error is a Gaussian with standard deviation of σ_o. In that case, following the derivation of Pannu & Read (1996) for the case of an acentric crystal, we obtain

$$P[I_o; I_c] = \frac{1}{2(2\pi)^{1/2}\sigma_F^2} \exp\left(-\frac{I_o^2}{2\sigma_o^2} - \frac{D_F^2 I_c}{2\sigma_F^2}\right) \times \sum_{t=0}^{\infty} \frac{1}{t!} \left(\frac{\sigma_o D_F^2 I_c}{4\sigma_F^4}\right)^t \exp\left(\frac{x^2}{4}\right) D_{-t-1}(x), \qquad (26)$$

where $D_{-t-1}(x)$ is a parabolic cylinder function (Abramowitz & Stegun, 1970, p. 687) and

$$x = (\sigma_o^2 - 2\sigma_F^2 I_o)/2\sigma_o\sigma_F^2. \qquad (27)$$

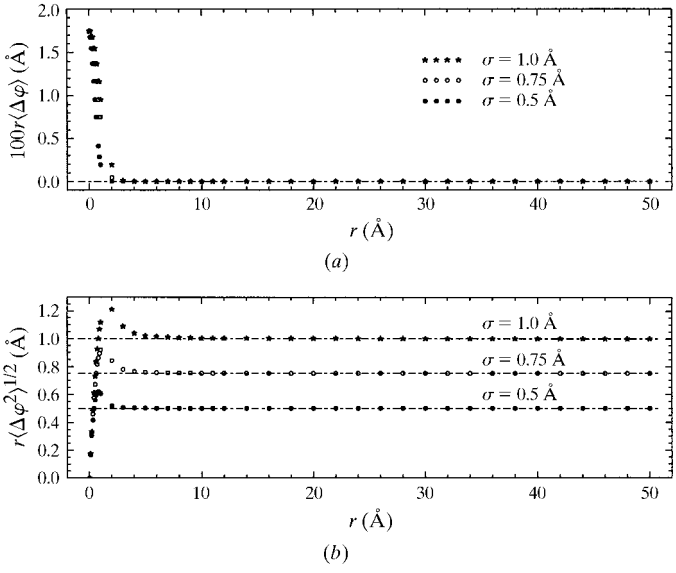

Figure 3
The averaged error of φ, $\langle\Delta\varphi\rangle$, and the standard deviation of φ, $\langle\Delta\varphi^2\rangle^{1/2}$, for the cylindrically symmetric Gaussian distribution of coordinate errors. Curves of $\langle\Delta\varphi\rangle$ against r are in (*a*). Those of $r\langle\Delta\varphi^2\rangle^{1/2}$ against r are in (*b*). When r is not too small, $\langle\Delta\varphi\rangle = 0$ and $r\langle\Delta\varphi^2\rangle^{1/2} = \sigma$ or $\langle\Delta\varphi^2\rangle^{1/2} = \sigma/r$.

5. Likelihood for the case of multiple Bessel terms

Usually more than one Bessel-function terms make contributions to the layer-line intensities in a fiber diffraction pattern. The conditional probability in this case is

$$P[\mathcal{G}; \mathcal{G}_c] = \prod_{k=1}^K (1/2\pi\sigma_k^2) \exp\{-[(A_k - D_k A_{ck})^2 + (B_k - D_k B_{ck})^2]/2\sigma_k^2\}, \qquad (28)$$

where A_k and B_k are components of the $2K$-dimensional vector $\mathcal{G}$, A_{ck} and B_{ck} are those of $\mathcal{G}_c$, σ_k^2 and D_k are the variance and the offset coefficient of the kth Bessel term,

respectively. These parameters can be dissimilar for different Bessel terms, making it difficult to derive an explicit formula for the probability. Bricogne (1991) gave a generalized result for this unequal-variance case. Expressed in terms of generalized hypergeometric functions in several variables, his formula converges impossibly slowly. The general unequal-variance case is also considered by Baskaran & Millane (1998).

We suggest here an alternate procedure for the calculation of the conditional probability. Equation (28) may be rewritten as

$$P[\boldsymbol{\mathcal{G}};\boldsymbol{\mathcal{G}}_c]=\prod_{k=1}^{K}(\lambda_k^2/2\pi\sigma_F^2)\exp\{-[(\lambda_k A_k-\lambda_k D_k A_{ck})^2+(\lambda_k B_k-\lambda_k D_k B_{ck})^2]/2\sigma_F^2\}, \tag{29}$$

where the λ_k are chosen so that

$$\lambda_k=\sigma_F/\sigma_k, \tag{30}$$

where σ_F is the variance of the Bessel term with the smallest order on the layer line l.

Three new $2K$-dimensional vectors are defined as

$$\boldsymbol{\mathcal{H}}=\begin{pmatrix}\lambda_1A_1\\ \lambda_1B_1\\ \lambda_2A_2\\ \lambda_2B_2\\ \vdots\\ \lambda_KA_K\\ \lambda_KB_K\end{pmatrix} \tag{31}$$

$$\boldsymbol{\mathcal{H}}_c=\begin{pmatrix}\lambda_1A_{c1}\\ \lambda_1B_{c1}\\ \lambda_2A_{c2}\\ \lambda_2B_{c2}\\ \vdots\\ \lambda_KA_{cK}\\ \lambda_KB_{cK}\end{pmatrix} \tag{32}$$

$$\boldsymbol{\mathcal{Q}}=\begin{pmatrix}\lambda_1D_1A_{c1}\\ \lambda_1D_1B_{c1}\\ \lambda_2D_2A_{c2}\\ \lambda_2D_2B_{c2}\\ \vdots\\ \lambda_KD_KA_{cK}\\ \lambda_KD_KB_{cK}\end{pmatrix}. \tag{33}$$

If the factors λ_k^2 before the exponentials in (29) are replaced by 1, the resultant can be represented by $P[\boldsymbol{\mathcal{H}};\boldsymbol{\mathcal{H}}_c]$. We have:

$$P[\boldsymbol{\mathcal{G}};\boldsymbol{\mathcal{G}}_c]=WP[\boldsymbol{\mathcal{H}};\boldsymbol{\mathcal{H}}_c], \tag{34}$$

where

$$W=\prod_{k=1}^{K}\lambda_k^2. \tag{35}$$

The new function $P[\boldsymbol{\mathcal{H}};\boldsymbol{\mathcal{H}}_c]$ differs from $P[\boldsymbol{\mathcal{G}};\boldsymbol{\mathcal{G}}_c]$ in that all the variances are identical for the former while they can be dissimilar for the latter. The vector $\boldsymbol{\mathcal{H}}$ is produced by changing the components of $\boldsymbol{\mathcal{G}}$ as shown in (31). The ratio of lengths of the two vectors is defined as

$$\Lambda=|\boldsymbol{\mathcal{H}}|/|\boldsymbol{\mathcal{G}}|. \tag{36}$$

Similarly, $\boldsymbol{\mathcal{H}}_c$ comes from $\boldsymbol{\mathcal{G}}_c$ as shown in (32). A similar ratio may be defined as:

$$\Lambda_c=|\boldsymbol{\mathcal{H}}_c|/|\boldsymbol{\mathcal{G}}_c|. \tag{37}$$

The procedure of changing vectors $\boldsymbol{\mathcal{G}}$ and $\boldsymbol{\mathcal{G}}_c$ into $\boldsymbol{\mathcal{H}}$ and $\boldsymbol{\mathcal{H}}_c$ allows an unequal-variance case to be converted into an equal-variance one. To obtain an equation for the conditional probability for intensities, or its square root, (29) has to be integrated over the surface of a hypersphere in the $2K$-dimensional space. Integration of this kind was completed by Stubbs (1989) when he calculated the largest likely R factor for fiber diffraction; and the corresponding calculation for the conditional probability may be found in Bricogne (1991). After the integration of $P[\boldsymbol{\mathcal{H}};\boldsymbol{\mathcal{H}}_c]$, we obtain

$$P[|\boldsymbol{\mathcal{G}}|;|\boldsymbol{\mathcal{G}}_c|]=\frac{\Omega_{2K}W(\Lambda|\boldsymbol{\mathcal{G}}|)^{2K-1}}{(2\pi)^K\sigma_F^{2K}}\exp\left(-\frac{\Lambda^2|\boldsymbol{\mathcal{G}}|^2+|\boldsymbol{\mathcal{Q}}|^2}{2\sigma_F^2}\right)\times{}_0F_1\left[K;\left(\frac{\Lambda|\boldsymbol{\mathcal{G}}||\boldsymbol{\mathcal{Q}}|}{2\sigma_F^2}\right)^2\right], \tag{38}$$

where Ω_{2K} is the surface area of a unit hypersphere of the $2K$-dimensional space, ${}_0F_1$ is one of the simplest generalized hypergeometric functions and is related to the Bessel function (Abramowitz & Stegun, 1970, p. 377). If one rewrites (38) in terms of the modified Bessel function, it becomes

$$P[|\boldsymbol{\mathcal{G}}|;|\boldsymbol{\mathcal{G}}_c|]=\frac{W(\Lambda|\boldsymbol{\mathcal{G}}|)^{2K-1}}{\sigma_F^{2K}}\left(\frac{\Lambda|\boldsymbol{\mathcal{G}}||\boldsymbol{\mathcal{Q}}|}{\sigma_F^2}\right)^{-(K-1)/2}\times\exp\left(-\frac{\Lambda^2|\boldsymbol{\mathcal{G}}|^2+|\boldsymbol{\mathcal{Q}}|^2}{2\sigma_F^2}\right)I_{K-1}\left(\frac{\Lambda|\boldsymbol{\mathcal{G}}||\boldsymbol{\mathcal{Q}}|}{\sigma_F^2}\right), \tag{39}$$

where I_{K-1} is the modified Bessel function of order $(K-1)$. This equation reduces to (18) when $K=1$.

The intensity-based probability function may be obtained using a procedure analogous to that for the single-Bessel-term case leading to the result

$$P[I;I_c]=\frac{W\Lambda^{2K-1}I^{K-1}}{2(K-1)!\sigma_F^{2K}}\exp\left(-\frac{\Lambda^2I+|\boldsymbol{\mathcal{Q}}|^2}{2\sigma_F^2}\right)\times{}_0F_1\left[K;\left(\frac{\Lambda|\boldsymbol{\mathcal{Q}}|I^{1/2}}{2\sigma_F^2}\right)^2\right]. \tag{40}$$

If the distribution of measurement errors is a Gaussian with variance σ_o^2, then we have

$$P[I_o; I_c] = \int_0^\infty P(I_o; I)P(I; I_c)\,\mathrm{d}I \quad (41)$$

$$= \int_0^\infty [1/(2\pi)^{1/2}\sigma_o]\exp[-(I - I_o)^2/2\sigma_o^2]P(I; I_c)\,\mathrm{d}I. \quad (42)$$

The generalized hypergeometric function in (40) can be expanded into a series (Lebedev, 1972, p. 275), allowing the integral to be completed by termwise integration. The conditional probability, $P[I; I_c]$, takes the form

$$P[I; I_c] = \frac{W\Lambda^{2K-1}I^{K-1}}{2\sigma_F^{2K}}\exp\left(-\frac{\Lambda^2 I + |\boldsymbol{Q}|^2}{2\sigma_F^2}\right) \times \sum_{t=0}^{\infty}\frac{1}{\Gamma(K+t)t!}\left(\frac{\Lambda^2|\boldsymbol{Q}|^2 I}{4\sigma_F^4}\right)^t. \quad (43)$$

Substitution of this formula into (42) gives

$$P(I_o; I_c) = \frac{W\Lambda^{2K-1}}{2(2\pi)^{1/2}\sigma_o\sigma_F^{2K}}\exp\left(-\frac{I_o^2}{2\sigma_o^2} - \frac{|\boldsymbol{Q}|^2}{2\sigma_F^2}\right) \times \sum_{t=0}^{\infty}\frac{1}{\Gamma(K+t)t!}\left(\frac{\Lambda^2|\boldsymbol{Q}|^2}{4\sigma_F^4}\right)^t \times \int_0^\infty I^{K+t-1}\exp\left[-\frac{I^2}{2\sigma_o^2} - \frac{I(\sigma_o^2\Lambda - 2\sigma_F^2 I_o)}{2\sigma_o^2\sigma_F^2}\right]\mathrm{d}I. \quad (44)$$

The integral in (44) has an analytical solution (Gradshteyn & Ryzhik, 1994, p. 384):

$$\int_0^\infty x^{\nu-1}\exp(-\beta x^2 - \gamma x)\,\mathrm{d}x = [\Gamma(\nu)/(2\beta)^{\nu/2}]\exp(\gamma^2/8\beta)D_{-\nu}[\gamma(2\beta)^{1/2}], \quad (45)$$

where $D_{-\nu}[\gamma/(2\beta)^{1/2}]$ is a parabolic cylinder function (Abramowitz & Stegun, 1970, p. 687). It follows that

$$\int_0^\infty I^{K+t-1}\exp\left[-\frac{I^2}{2\sigma_o^2} - \frac{I(\sigma_o^2\Lambda - 2\sigma_F^2 I_o)}{2\sigma_o^2\sigma_F^2}\right]\mathrm{d}I = \sigma_o^{K+t}\Gamma(K+t)\exp\left[\frac{(\sigma_o^2\Lambda - 2\sigma_F^2 I_o)^2}{16\sigma_o^2\sigma_F^4}\right] \times D_{-K-t}\left(\frac{\sigma_o^2\Lambda - 2\sigma_F^2 I_o}{2\sigma_o\sigma_F^2}\right). \quad (46)$$

The final form of $P[I_o; I_c]$ is then

$$P[I_o; I_c] = \frac{\sigma_o^{K-1}W\Lambda^{2K-1}}{2(2\pi)^{1/2}\sigma_F^{2K}}\exp\left[-\frac{I_o^2}{2\sigma_o^2} - \frac{|\boldsymbol{Q}|^2}{2\sigma_F^2}\right] \times \sum_{t=0}^{\infty}\frac{1}{t!}\left(\frac{\sigma_o\Lambda^2|\boldsymbol{Q}|^2}{4\sigma_F^4}\right)^t\exp\left(\frac{x^2}{4}\right)D_{-K-t}(x), \quad (47)$$

where

$$x = (\sigma_o^2\Lambda - 2\sigma_F^2 I_o)/2\sigma_o\sigma_F^2. \quad (48)$$

Equation (47) reduces to (26) for the single-Bessel-term case when $K = 1$.

Equations (39) and (47) are the final forms used to calculate the conditional probability and the likelihood for the case of multiple Bessel terms. Theoretically speaking, no assumption was made in the derivation. An unknown parameter Λ, however, was introduced when converting the unequal-variance case into the equal-variance one. Actually, this parameter is neither observable nor computable. Assumption or approximation has to be made to estimate its value in implementation of likelihood calculation.

6. Parameters of likelihood

In addition to the experimental data with errors and atomic coordinates with errors, we need some special parameters to calculate the likelihood function. They are the offset coefficient D_F and the variance σ_F for both the single-Bessel-term case and the multiple-term case, and the metric ratio Λ only for the second case. The first two parameters are related to the variances of the atomic coordinates as shown in (19) and (20). These equations make it possible to adjust the coordinate variances at the same time as the whole structure is under refinement.

Here we show example calculations of D_F and σ_F on the structure of tobacco mosaic virus (TMV). This structure was refined to 2.9 Å resolution by X-ray fiber diffraction (Namba *et al.*, 1989). The atomic coordinates of TMV and specific values of σ are used with (15) to obtain $\langle\cos\Delta\theta\rangle$. S_1, S_2, S_3 and S_4 are then obtained with (21)–(24). Finally, D_F and σ_F are calculated using (19) and (20), respectively.

6.1. Offset coefficient D_F

The offset coefficients of TMV with $\sigma = 0.5$, 0.75 and 1 Å are shown in Fig. 4. D_F decreases with increasing R and with increasing σ for any Bessel term on any layer line. Each curve of D_F against R on the graphics may be fitted to a polynomial:

$$D_F = a - bR^2 + cR^4, \quad (49)$$

where a, b and c are constants for particular values of l, n and σ.

6.2. Variance σ_F^2

Curves of σ_F *versus* R are shown in Fig. 5 for various values of σ. As expected, σ_F increases with increasing σ. The variances for any Bessel term are zero when R is less than a critical value and quickly reach a maximum for R larger than this critical value. In the long tail part of the curve, the variance decreases slowly with increasing R. Curves of σ_F *versus* R of various Bessel terms on some layer lines are shown in Fig. 6. Both the starting point and the maximum of the curve move to a higher value of R for a larger absolute value of n. However, the tail parts of curves on the same layer line are similar.

6.3. Metric ratio Λ

The differences in σ_F's, shown in Fig. 6, produce the principal obstacle for calculation of the likelihood function for the multiple-Bessel-term case. This problem may be solved by converting to an equal-variance case as suggested above. Components of the vector $\mathcal{G}$ experience changes during the

conversion, as do components of vector $\mathcal{G}_c$, changed at the same time, but possibly with different factors. Theoretically speaking, the lengths of two vectors can be changed by different ratios after the conversion from an unequal-variance case to an equal-variance case. In other words, Λ and Λ_c, defined in (36) and (37), may be different. Λ is needed for calculation of the likelihood function. It is, however, unobservable and uncomputable. On the other hand, Λ_c may be calculated for any structural model. The assumption we make here is that

$$\Lambda = \Lambda_c, \tag{50}$$

i.e. the metric ratio of the 'true' intensity is similar to that of the calculated.

7. Discussion

From the point of view of statistics theory, the likelihood function is an appropriate target for refinement of structures based on fiber diffraction and is demonstrably superior to minimization of a least-squares residual. An explicit relationship between the likelihood function and errors of atomic coordinates in the helical structural models is necessary in such a refinement process.

In previous work on statistics of fiber diffraction (Stubbs, 1989; Millane, 1989*a*,*b*, 1990*a*,*b*), only error in θ was taken into account; that of r was omitted. Furthermore, these works assumed that the variances of Bessel functions contributing to a point in a diffraction pattern were identical. The principal result of these studies was estimation of the largest likely R factors for fiber diffraction and this result is unlikely to be affected by these assumptions. However, for calculation of a likelihood function, these simplifying assumptions cannot be made.

The likelihood functions have been explicitly expressed in terms of atomic coordinate errors in the cylindrical reference system in this work. A cylindrical symmetry has been assumed for the distribution of coordinate errors in a helical structure.

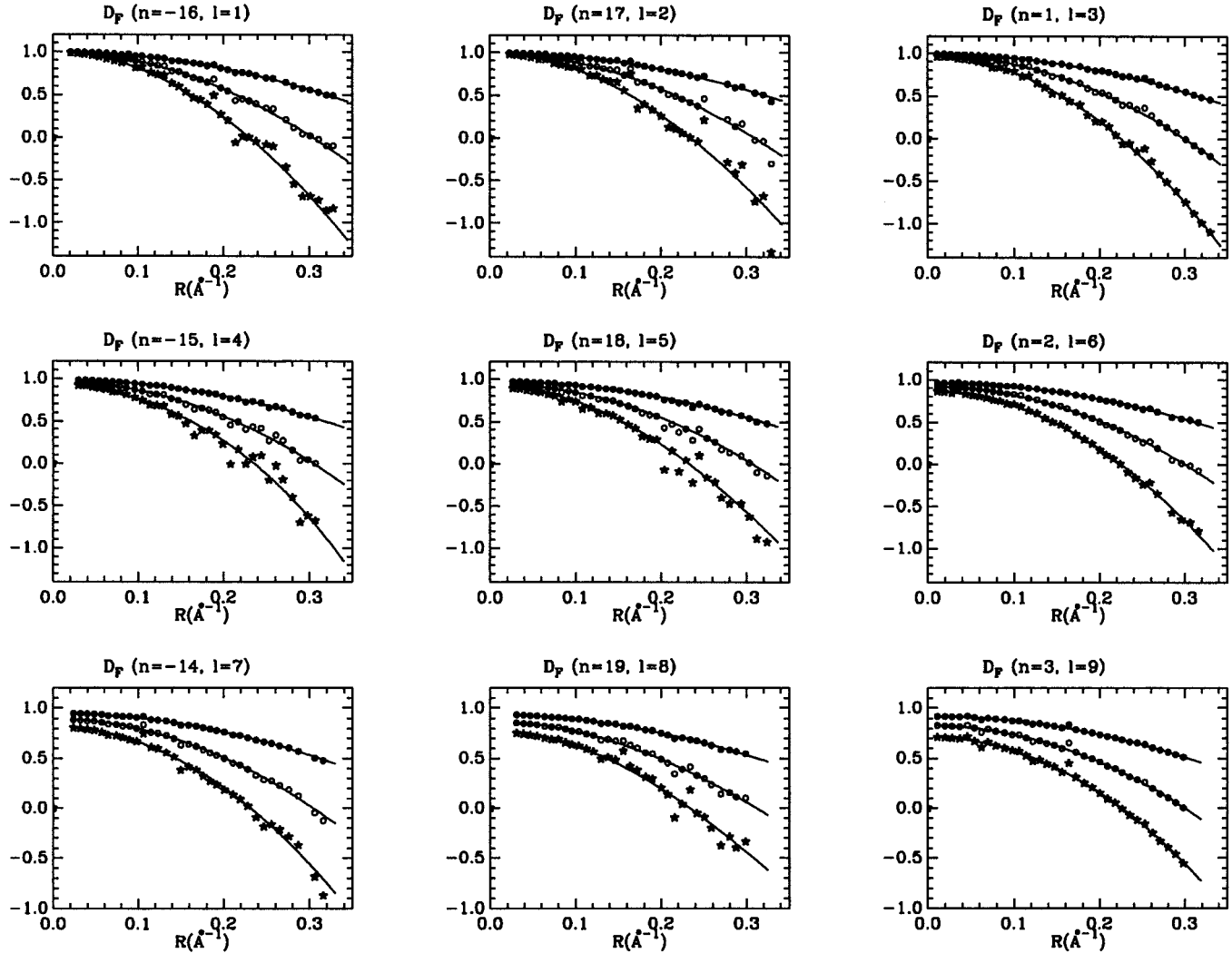

Figure 4
Curves of D_F against R for error distribution of variance $\sigma = 0.5$ Å (●●●), $\sigma = 0.75$ Å (○○○) and $\sigma = 1.0$ Å (★★★). They are calculated for the structure of TMV (Namba *et al.*, 1989). Each of the solid lines is obtained by a polynomial: $a - bR^2 + cR^4$.

In terms of the variance of this distribution, the formulas of the offset coefficient D_F and the variance σ_F, the parameters in the conditional probability $P[\mathcal{G};\mathcal{G}_c]$, have been derived. The integral of $P[\mathcal{G};\mathcal{G}_c]$ has been performed to obtain the final formula of the likelihood by converting the unequal-variance problem to an equal-variance one.

APPENDIX *A*

The real part of a structure factor calculated from a structural model is

$$A_c = \sum_{j=1}^{N} f_j J_n(2\pi R r_j^c)\cos\theta_j^c, \tag{51}$$

where r_j^c and θ_j^c are atomic coordinates of the model with errors

$$r_j^c = r_j + \Delta r_j \tag{52}$$

$$\theta_j^c = \theta_j + \Delta\theta_j. \tag{53}$$

The difference in the real parts of the calculated and true structure factors is

$$\Delta A = A_c - A = \sum_{j=1}^{N} \delta_j, \tag{54}$$

where

$$\delta_j = f_j J_n[2\pi R(r_j + \Delta r_j)]\cos(\theta_j + \Delta\theta_j) - f_j J_n(2\pi R r_j)\cos\theta_j. \tag{55}$$

The function $J_n[2\pi R(r_j + \Delta r_j)]$ may be expanded in a Taylor series. If Δr_j is small, it can be approximated as

$$J_n[2\pi R(r_j + \Delta r_j)] \cong J_n(2\pi R r_j) + 2\pi R\Delta r_j J_n'(2\pi R r_j) + 2\pi^2 R^2(\Delta r_j)^2 J_n''(2\pi R r_j), \tag{56}$$

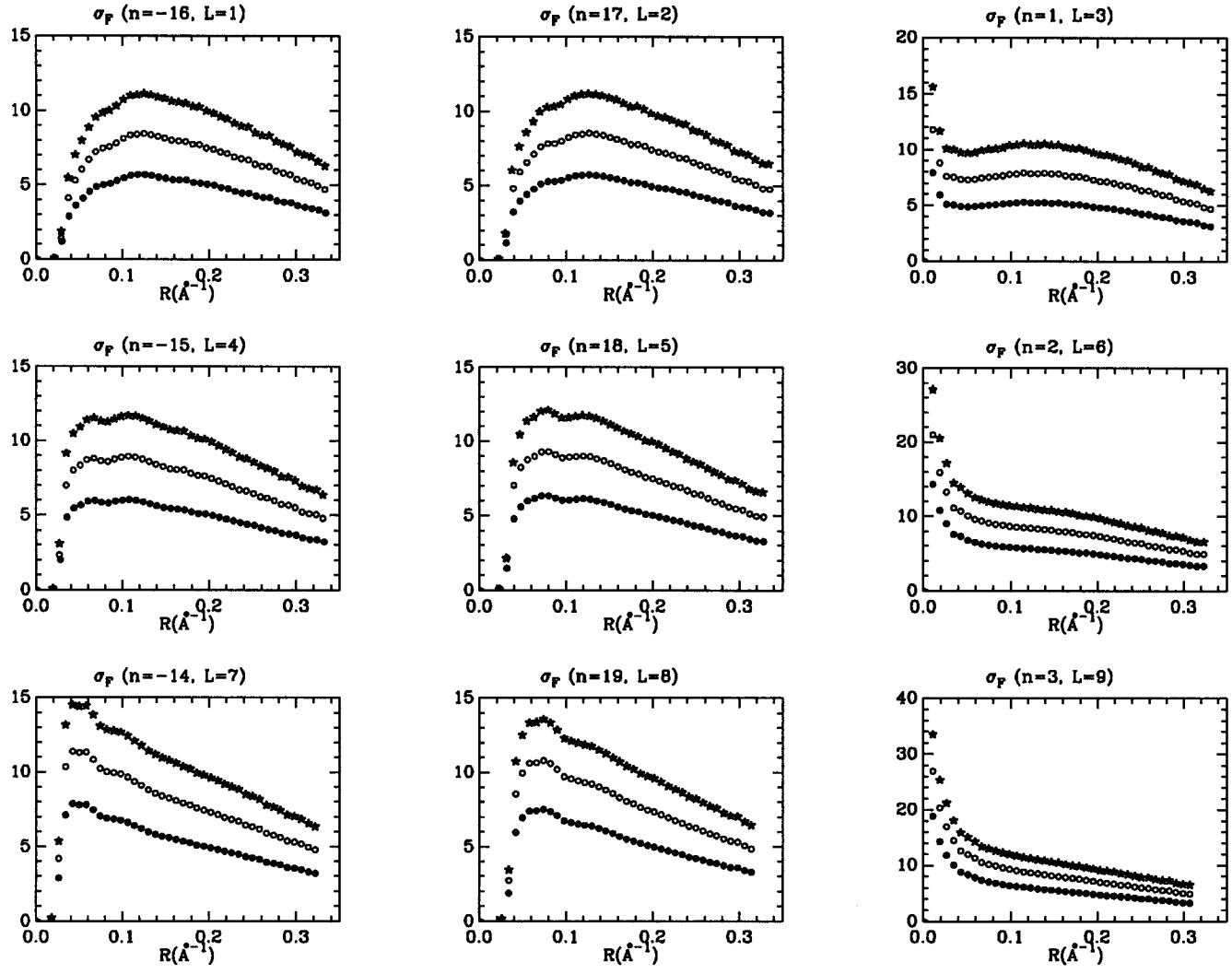

Figure 5
Curves of σ_F against R for error distribution of variance $\sigma = 0.5$ Å (• • •), $\sigma = 0.75$ Å (○ ○ ○) and $\sigma = 1.0$ Å (⋆ ⋆ ⋆). They are calculated for the structure of TMV (Namba *et al.*, 1989). σ_F increases with increasing atomic coordinate error. σ_F is zero when R is smaller than a critical value and quickly reaches its maximum after the starting point of the curve. No maximum is observed on curves of low-order Bessel terms. In the tail part of the curve, σ_F decreases slowly with increasing R.

where $J'_n(2\pi Rr_j)$ and $J''_n(2\pi Rr_j)$ are the first and second derivatives of $J_n(2\pi Rr_j)$ with respect to r_j. They are obtained by the recurrence relations of Bessel functions (Abramowitz & Stegun, 1970, p. 361)

$$J'_n(2\pi Rr_j) = \tfrac{1}{2}[J_{n-1}(2\pi Rr_j) - J_{n+1}(2\pi Rr_j)] \quad (57)$$

$$J''_n(2\pi Rr_j) = \tfrac{1}{4}[J_{n+2}(2\pi Rr_j) + J_{n-2}(2\pi Rr_j) - 2J_n(2\pi Rr_j)]. \quad (58)$$

It was known that $\langle \Delta r\rangle = 0$, $\langle \Delta r^2\rangle = \sigma^2$ and $\langle \sin \Delta\theta\rangle = 0$ for noncentral atoms. Assuming that Δr_j and $\Delta\theta_j$ vary independently and that all atoms have the same variance σ^2, we have the expected value of δ_j:

$$\langle \delta_j\rangle = f_j[J_n(2\pi Rr_j) + 2\pi^2R^2\sigma^2J''_n(2\pi Rr_j)]\cos\theta_j\langle\cos\Delta\theta_j\rangle - f_jJ_n(2\pi Rr_j)\cos\theta_j. \quad (59)$$

This may be rewritten as

$$\langle \delta_j\rangle = (\langle\cos\Delta\theta_j\rangle - 1)f_jJ_n(2\pi Rr_j)\cos\theta_j + 2\pi^2R^2\sigma^2\langle\cos\Delta\theta_j\rangle f_jJ''_n(2\pi Rr_j)\cos\theta_j. \quad (60)$$

The averaged difference in A, $\langle A\rangle$, should be a summation of $\langle\delta_j\rangle$ over j. If the offset coefficient D_A is defined by

$$\langle \Delta A\rangle = (D_A - 1)A, \quad (61)$$

we have

$$D_A = (1/A)\left[\sum_{j=1}^{N} f_j\langle\cos\Delta\theta_j\rangle J_n(2\pi Rr_j)\cos\theta_j + 2\pi^2R^2\sigma^2\sum_{j=1}^{N} f_j\langle\cos\Delta\theta_j\rangle J''_j(2\pi Rr_j)\cos\theta_j\right], \quad (62)$$

where

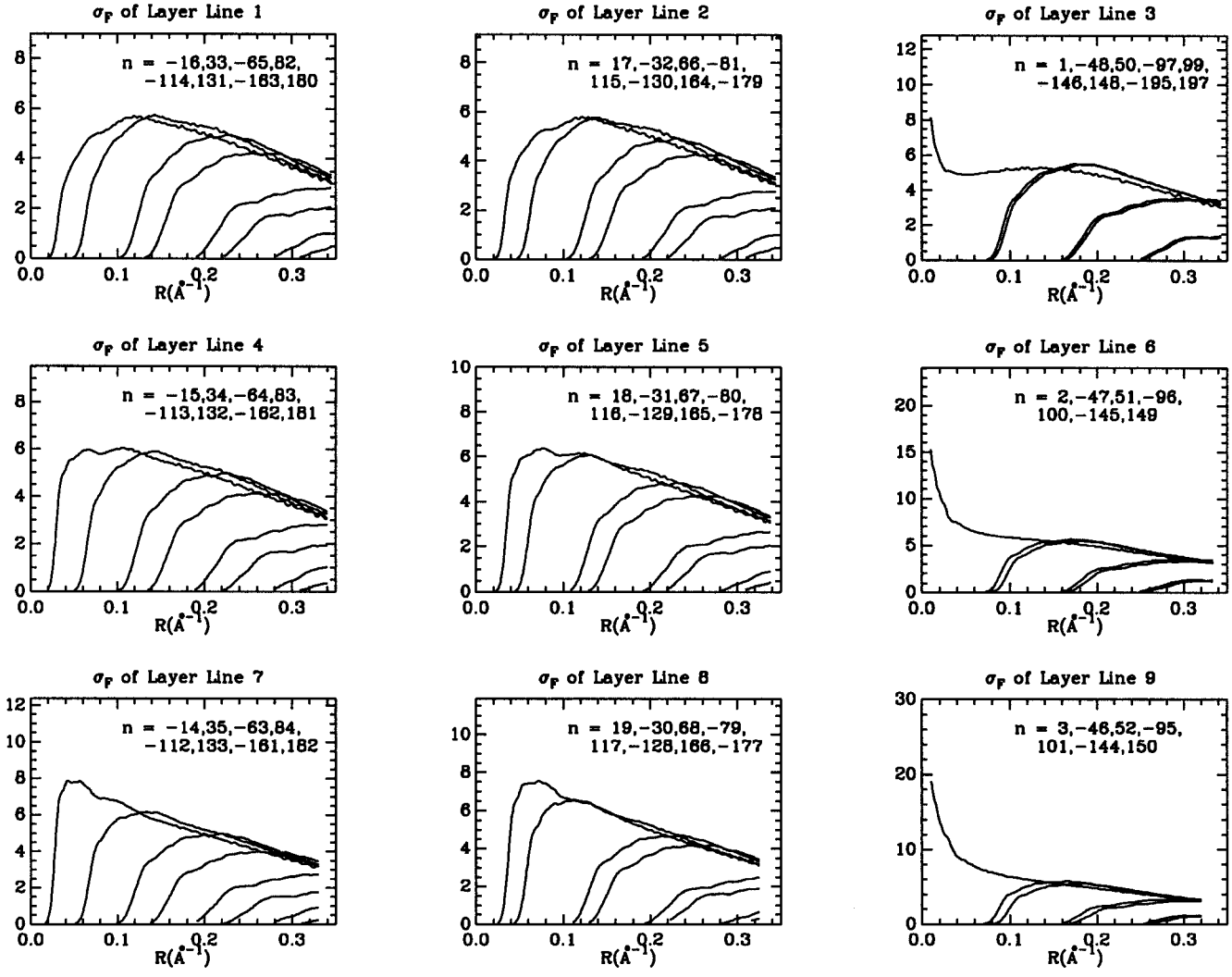

Figure 6
Curves of σ_F against R for error distribution of variance $\sigma = 0.5$ Å. For comparison, all curves of various Bessel terms on the same layer line are put together in a panel. The starting point of the curve depends on the order of the Bessel term. It is at higher R for high order of Bessel term. Tail parts of the curves on the same layer line are similar. When the difference in $|n|$ is small, two neighboring curves on a layer line can be almost identical, *e.g.* the curve for $n = -47$, $l = 6$ and that for $n = 51$, $l = 6$.

$$A = \sum_{j=1}^{N} f_j J_n(2\pi R r_j)\cos\theta_j. \quad (63)$$

Substitution of (55) into the definition of the variance, σ_j^2,

$$\sigma_j^2 = \langle\delta_j^2\rangle - \langle\delta_j\rangle^2, \quad (64)$$

followed by a summation over j leads to the desired result:

$$\sigma_A^2 = \sum_{j=1}^{N} f_j^2(1 - \langle\cos\Delta\theta_j\rangle^2)[J_n(2\pi R r_j)]^2(\sin\theta_j)^2 + 4\pi^2 R^2\sigma^2 \sum_{j=1}^{N} f_j^2[J_n'(2\pi R r_j)]^2(\cos\theta_j)^2. \quad (65)$$

For random values of θ_j, this equation becomes (Stubbs, 1989)

$$\sigma_A^2 = \tfrac{1}{2}\sum_{j=1}^{N} f_j^2(1 - \langle\cos\Delta\theta_j\rangle^2)[J_n(2\pi R r_j)]^2 + 2\pi^2 R^2\sigma^2 \sum_{j=1}^{N} f_j^2[J_n'(2\pi R r_j)]^2. \quad (66)$$

Similar formulas can be obtained also for D_B and σ_B^2. They are equal to D_A and σ_A^2, respectively. Notation of D_F ($= D_A = D_B$) and σ_F ($= \sigma_A = \sigma_B$) is used and (62), (65) are rewritten as (19) and (20) in the text.

The authors thank the referees for their valuable comments. This work was supported by a grant from the National Science Foundation.

References

Abramowitz, M. & Stegun, I. A. (1970). *Handbook of Mathematical Functions.* New York: Dover.
Baskaran, S. & Millane, R. P. (1997). *Proc. SPIE*, **3170**, 227–237.
Baskaran, S. & Millane, R. P. (1998). *Proc. SPIE*, **3459**, 29–38.
Baskaran, S. & Millane, R. P. (1999*a*). *J. Opt. Soc. Am.* A**16**, 236–245.
Baskaran, S. & Millane, R. P. (1999*b*). *IEEE Trans. Image Process*. **8**, 1420–1434.
Bricogne, G. (1991). *Acta Cryst.* A**47**, 803–829.
Bricogne, G. (1993). *Acta Cryst.* D**49**, 37–60.
Brünger, A. T., Kuriyan, J. & Karplus, M. (1987). *Science*, **235**, 458–460.
Cochran, W., Crick, F. H. C. & Vand, V. (1952). *Acta Cryst.* **5**, 581–586.
Gradshteyn, I. S. & Ryzhik, I. M. (1994). *Tables of Integrals, Series and Products*, 5th ed. New York: Academic Press.
Hendrickson, W. A. (1985). *Methods Enzymol.* **115**, 252–270.
Lebedev, N. N. (1972). *Special Functions and their Applications.* New York: Dover.
Millane, R. P. (1989*a*). *Acta Cryst.* A**45**, 258–260.
Millane, R. P. (1989*b*). *Acta Cryst.* A**45**, 573–576.
Millane, R. P. (1990*a*). *Acta Cryst.* A**46**, 68–72.
Millane, R. P. (1990*b*). *Acta Cryst.* A**46**, 552–559.
Millane, R. P. & Baskaran, S. (1997). *Fiber Diffr. Rev.* **6**, 14–18.
Namba, K., Pattanayek, R. & Stubbs, G. (1989). *J. Mol. Biol.* **208**, 307–325.
Namba, K. & Stubbs, G. (1985). *Acta Cryst.* A**41**, 252–262.
Pannu, N. S. & Read, R. J. (1996). *Acta Cryst.* A**52**, 659–668.
Read, R. J. (1990). *Acta Cryst.* A**46**, 900–912.
Stubbs, G. (1989). *Acta Cryst.* A**45**, 254–258.
Stubbs, G. & Diamond, R. (1975). *Acta Cryst.* A**31**, 709–718.