

# Derivation of likelihood functions and posterior moments used in STARANISO.

by Ian J. Tickle, Global Phasing Ltd.

Version of 11<sup>th</sup> Mar 2018.

This derivation has many features in common with that of French & Wilson's (henceforth 'FW') 'TRUNCATE' algorithm used to obtain the Bayesian posterior half-order moment (expectation of half power) of the intensity (*i.e.* the expected amplitude), given the prior expectation of the X-ray intensity, its experimentally measured value and its standard uncertainty. The expressions for the relevant probability density functions and other relevant equations are repeated here for easy reference.

Popov & Bourenkov (henceforth 'PB') proposed a likelihood function based on the acentric Wilson distribution (centric reflexions were excluded in their algorithm). They used this in an optimisation of the absolute scale and the unconstrained components of the anisotropy tensor of the observed intensities. The derivation of the 'PB' likelihood function is reproduced below (with minor changes in notation). Then a new likelihood function 'FW' based on the FW posterior likelihood is derived. The major difference between the PB and FW functions is that the former is error-free, *i.e.* it takes no account of experimental errors, whereas the latter assumes a Gaussian model for the experimental errors in the observed intensities, exactly as in the FW TRUNCATE treatment. This implies that for observed negative intensities, which can arise only through measurement error, the PB likelihood is undefined and hence all negative intensities must either be excluded from the optimisation of the PB likelihood functions or set to zero. This is likely to have a significant impact on the results in those cases where the true intensity is small relative to the mean intensity, since such intensities do carry information. Hence the PB functions are only of theoretical interest and are not used in practice.

Note that in the case of twinning two implicit assumptions are made, first that all the twin domains have the same anisotropy of diffraction and second that that anisotropy obeys the twin symmetry. In the case of the first assumption, diffraction anisotropy arises from a number of factors, for example the variability in the strengths and the directionality of intermolecular interactions (which is likely to be same for all twin domains), but also from radiation damage (which may well differ between twin domains). In the case of the second assumption, twinning is largely a geometrical effect due to shape symmetry of the twin domain and to symmetry in the interactions at its boundaries (so twinned domains are able to pack together during crystal growth as easily as untwinned ones), but this does not necessarily imply that the diffraction anisotropy obeys that shape symmetry. There is therefore no *a priori* reason why either of these two assumptions should hold in any particular case. In the case that the anisotropy does not obey the twin symmetry knowledge of the twin operators is required.

## Sections:

1. PB likelihood function: untwinned acentric & centric perfect hemihedral twin cases.
2. PB likelihood function: untwinned centric case.
3. PB likelihood function: acentric perfect general and hemihedral twin cases.
4. PB likelihood function: centric perfect general twin case.
5. PB likelihood function: acentric imperfect general and hemihedral twin cases.
6. PB likelihood function: centric imperfect hemihedral twin case.
7. FW likelihood function: untwinned acentric case.
8. FW integral: case of small integer powers of  $J_h$ .
9. FW integral: case of non-integer and large integer powers of  $J_h$ .
10. FW integral: implementation of parabolic cylinder functions in Fortran 90.
11. FW likelihood function & posterior moments: untwinned acentric & centric perfect hemihedral twin cases.
12. FW likelihood function & posterior moments: untwinned centric case.
  
13. FW likelihood function & posterior moments: acentric perfect hemihedral twin case.
14. FW likelihood function & posterior moments: acentric imperfect hemihedral twin case.
15. FW likelihood function & posterior moments: centric imperfect hemihedral twin case.
16. REFERENCE.

## 1. PB likelihood function: untwinned acentric & centric perfect hemihedral twin cases.

For an untwinned acentric or perfect hemihedrally twinned centric reflexion the prior probability density function (PDF) $P_{ah}(J_h   S_h)$ of the true intensity $J_h$ (conditional on its prior expectation $S_h$ ) is assumed to be the acentric Wilson PDF:	
$P_{ah}(J_h   S_h) = S_h^{-1} \exp(-J_h/S_h)$	(1)
The prior expectation $S_h$ of $J_h$ can be expressed in terms of the zone-enhancement factor $\epsilon_h$ , a spherically-symmetric standard 'profile' function $\hat{S}( \mathbf{s}_h )$ , the absolute scale factor $g$ , and the overall anisotropic displacement tensor $\mathbf{U}$ on $F$ :	
$S_h(g, \mathbf{U}) = \epsilon_h \hat{S}( \mathbf{s}_h ) g \exp(-4\pi^2 \tilde{\mathbf{s}}_h \mathbf{U} \mathbf{s}_h)$	(2)
Taking the negative log of (1), the contribution to the negative log-likelihood gain is then given by (omitting terms that are independent of the parameters):	
$-LLG_{ah}(g, \mathbf{U}) = \ln(S_h) + J_h/S_h$	(3)

## 2. PB likelihood function: untwinned centric case.

For untwinned centric reflexions we use the centric Wilson PDF as the prior $P_{ch}(J_h   S_h)$ :	
$P_{ch}(J_h   S_h) = (2\pi S_h J_h)^{-1/2} \exp(-J_h/2 S_h)$	(4)
Again, taking the negative log of (4), the contribution to the negative log-likelihood gain is given by:	
$-LLG_{ch}(g, \mathbf{U}) = 1/2(\ln(2\pi S_h J_h) + J_h/S_h)$	(5)
PB appear not to have included centric reflexions in their treatment, though it is not clear why, given that it's such a simple modification.	

### 3. PB likelihood function: acentric perfect general and hemihedral twin cases.

In the perfect general twin case ( $N$ equal twin domains; usually $N = 2$ : hemihedry, 4: tetartohedry or 8: ogdohedry), the PDF for acentric reflexions is:	
$P_{\text{at } h}(J_h   S_h) = \frac{N^N J_h^{N-1}}{(N-1)! S_h^N} \exp(-N J_h / S_h)$	(6)
Again, taking the log of (6), the contribution to the negative log-likelihood gain is:	
$-\text{LLG}_{\text{at } h}(g, \mathbf{U}) = N(\ln(S_h) + J_h / S_h)$	(7)
In the perfect hemihedral twin case ( $N = 2$ ) the PDF and -LLG for acentric reflexions are therefore:	
$P_{\text{at } h}(J_h   S_h) = 4 S_h^{-2} J_h \exp(-2 J_h / S_h)$	(8)
$-\text{LLG}_{\text{at } h}(g, \mathbf{U}) = \ln(S_h^2 / 4 J_h) + 2 J_h / S_h$	(9)

### 4. PB likelihood function: centric perfect general twin case.

In the perfect general twin case ( $N$ equal twin domains) the PDF for centric reflexions is:	
$P_{\text{ct } h}(J_h   S_h) = \frac{(N/2)^{N/2} J_h^{N/2-1}}{\Gamma(N/2) S_h^{N/2}} \exp(-N J_h / 2 S_h)$	(10)
$-\text{LLG}_{\text{ct } h}(g, \mathbf{U}, \alpha) = (N/2) (\ln(S_h) + J_h / S_h)$	(11)
For $N = 2$ this reduces to the acentric untwinned case (3).	

## 5. PB likelihood function: acentric imperfect general and hemihedral twin cases.

In the imperfect general twin case ( $N$ unequal twin domains) the PDF for acentric reflexions is:	
$P_{\text{at}h}(J_h   S_h) = S_h^{-1} \sum_i^N \alpha_i^{-1} \left( \prod_{j=1, j \neq i}^N \frac{\alpha_i}{\alpha_i - \alpha_j} \right) \exp(-J_h / \alpha_i S_h)$	(12)
In the imperfect hemihedral twin case ( $N = 2$ ) this reduces to:	
$P_{\text{at}h}(J_h   S_h) = ((1-2\alpha)S_h)^{-1} (\exp(-J_h/(1-\alpha)S_h) - \exp(-J_h/\alpha S_h))$	(13)
$-\text{LLG}_{\text{at}h}(J_h   S_h) = -\ln(P_{\text{at}h}(J_h   S_h))$	(14)

## 6. PB likelihood function: centric imperfect hemihedral twin case.

In the imperfect hemihedral twin case the PDF for centric reflexions is:	
$P_{\text{ct}h}(J_h   S_h) = (2\sqrt{\alpha(1-\alpha)}S_h)^{-1} \exp(-J_h/2\alpha S_h) {}_1F_1(1/2; 1; (1-2\alpha)J_h/2\alpha(1-\alpha)S_h)$ $= (2\sqrt{\alpha(1-\alpha)}S_h)^{-1} \exp(-J_h/4\alpha(1-\alpha)S_h) I_0((1-2\alpha)J_h/4\alpha(1-\alpha)S_h)$ $= (2\sqrt{\alpha(1-\alpha)}S_h)^{-1} \exp(-J_h/2(1-\alpha)S_h) I'_0((1-2\alpha)J_h/4\alpha(1-\alpha)S_h)$	(15)
$-\text{LLG}_{\text{ct}h}(J_h   S_h) = -\ln(P_{\text{ct}h}(J_h   S_h))$	(16)

where  ${}_1F_1(a; b; z)$  is [Kummer's confluent hypergeometric function](#),  $I_0(z)$  is the zero-order [modified Bessel function](#) and  $I'_0(x)$  is the scaled zero-order modified Bessel function:  $I'_0(x) = \exp(-x) I_0(x)$ . This avoids the division by  $\alpha$  for  $\alpha$  near to zero. In (15) it is not necessary to deal specifically with the edge cases where  $\alpha$  is near 0 or 0.5 since unlike (13) there is no cancellation of large terms.

## 7. FW likelihood function: untwinned acentric case.

Since the prior distribution  $P_h(J_h | S_h)$  of the true intensity  $J_h$  (conditional on the prior expectation  $S_h$  of  $J_h$ ) is independent of the error distribution  $P_h(I_h | J_h, \sigma_h)$  of the measured intensity  $I_h$ , then the joint distribution of  $I_h$  and  $J_h$  is simply the product of their respective densities:

$$P_h(I_h, J_h | S_h, \sigma_h) = P_h(I_h | J_h, \sigma_h) P_h(J_h | S_h) \quad (17)$$

The PDF of  $I_h$  conditional on  $S_h$  and the standard deviation  $\sigma_h$  of  $I_h$  is therefore given by marginalising out the unknown  $J_h$  from the joint distribution:

$$\begin{aligned} P_h(I_h | S_h, \sigma_h) &= \int_0^{\infty} P_h(I_h, J_h | S_h, \sigma_h) dJ_h \\ &= \int_0^{\infty} P_h(I_h | J_h, \sigma_h) P_h(J_h | S_h) dJ_h \end{aligned} \quad (18)$$

We assume that  $P_h(I_h | J_h, \sigma_h)$  is the normal error distribution of  $I_h$  :

$$P_h(I_h | J_h, \sigma_h) = (\sigma_h \sqrt{2\pi})^{-1} \exp(-(I_h - J_h)^2 / 2\sigma_h^2) \quad (20)$$

In the untwinned acentric case:

$$P_h(J_h | S_h) = S_h^{-1} \exp(-J_h/S_h)$$

so in this case:

$$P_h(I_h | S_h, \sigma_h) = (\sigma_h \sqrt{2\pi} S_h)^{-1} \int_0^{\infty} \exp(-(I_h - J_h)^2 / 2\sigma_h^2) \exp(-J_h/S_h) dJ_h \quad (21)$$

## 8. FW integral: case of small integer powers of $J_h$ .

The prior distribution  $P_h(J_h | S_h)$  of  $J_h$  takes different forms depending on the centrality of  $h$  and the twin fraction(s). Also, very similar integrals will be required later for the evaluation in the various specific cases of the posterior moments  $\langle J_h^m \rangle$  of  $J_h$ , so it is convenient to generalise the form of the required integral  $Q_h(\mu, \nu, \xi)$  here, incorporating the order  $m$  of the moment:

$$\begin{aligned}
 Q_h(\mu, \nu, \xi) &= (\sigma_h \sqrt{2\pi} S_h^\mu)^{-1} \int_0^\infty J_h^\nu \exp(-(I_h - J_h)^2 / 2\sigma_h^2) \exp(-\xi J_h / S_h) dJ_h \\
 &= (\sigma_h \sqrt{2\pi} S_h^\mu)^{-1} \int_0^\infty J_h^\nu \exp(-(I_h - J_h)^2 / 2\sigma_h^2 - \xi J_h / S_h) dJ_h \\
 &= (\sigma_h \sqrt{2\pi} S_h^\mu)^{-1} \int_0^\infty J_h^\nu \exp(-I_h^2 / 2\sigma_h^2 + I_h J_h / \sigma_h^2 - J_h^2 / 2\sigma_h^2 - \xi J_h / S_h) dJ_h \\
 &= (\sigma_h \sqrt{2\pi} S_h^\mu)^{-1} \exp((\xi \sigma_h)^2 / 2 S_h^2 - \xi I_h / S_h) \\
 &\quad \int_0^\infty J_h^\nu \exp(-(J_h / \sigma_h - (I_h / \sigma_h - \xi \sigma_h / S_h))^2 / 2) dJ_h \\
 &= \sigma_h^\nu (\sqrt{2\pi} S_h^\mu)^{-1} \exp((\xi \sigma_h)^2 / 2 S_h^2 - \xi I_h / S_h) \int_0^\infty u_h^\nu \exp(-(u_h - t_h(\xi))^2 / 2) du_h
 \end{aligned} \tag{22}$$

where  $u_h = J_h / \sigma_h$  and  $t_h(\xi) = I_h / \sigma_h - \xi \sigma_h / S_h$ .

For small positive integer values (0, 1 or 2) of  $\nu$  the integral in (22) can be expressed in terms of the normal probability function  $\phi$  and its integral  $\Phi$ :

$$\begin{aligned}
 Q_h(\mu, 0, \xi) &= (\sqrt{2\pi} S_h^\mu)^{-1} \exp((\xi \sigma_h)^2 / 2 S_h^2 - \xi I_h / S_h) \int_0^\infty \exp(-(u_h - t_h(\xi))^2 / 2) du_h \\
 &= S_h^{-\mu} \exp((\xi \sigma_h)^2 / 2 S_h^2 - \xi I_h / S_h) \Phi(t_h(\xi))
 \end{aligned} \tag{23}$$

$$\begin{aligned}
 Q_h(\mu, 1, \xi) &= \sigma_h (\sqrt{2\pi} S_h^\mu)^{-1} \exp((\xi \sigma_h)^2 / 2 S_h^2 - \xi I_h / S_h) \int_0^\infty u_h \exp(-(u_h - t_h(\xi))^2 / 2) du_h \\
 &= \sigma_h S_h^{-\mu} \exp((\xi \sigma_h)^2 / 2 S_h^2 - \xi I_h / S_h) (t_h(\xi) \Phi(t_h(\xi)) + \phi(t_h(\xi)))
 \end{aligned} \tag{24}$$

$$\begin{aligned}
 Q_h(\mu, 2, \xi) &= \sigma_h^2 (\sqrt{2\pi} S_h^\mu)^{-1} \exp((\xi \sigma_h)^2 / 2 S_h^2 - \xi I_h / S_h) \int_0^\infty u_h^2 \exp(-(u_h - t_h(\xi))^2 / 2) du_h \\
 &= \sigma_h^2 S_h^{-\mu} \exp((\xi \sigma_h)^2 / 2 S_h^2 - \xi I_h / S_h) ((t_h^2(\xi) + 1) \Phi(t_h(\xi)) + t_h(\xi) \phi(t_h(\xi)))
 \end{aligned} \tag{25}$$

## 9. FW integral: case of non-integer and large integer powers of $J_h$ .

<p>For other values of <math>\nu</math> the integral in (22) can be expressed in terms of the parabolic cylinder functions <math>D</math>, or alternatively <math>U</math> and <math>V</math> [<a href="http://mathworld.wolfram.com/ParabolicCylinderFunction.html">http://mathworld.wolfram.com/ParabolicCylinderFunction.html</a>]. Factoring out the constant term in the integral:</p>	
$\int_0^{\infty} u^{\nu} \exp(-(u-t)^2/2) du = \exp(-t^2/2) \int_0^{\infty} u^{\nu} \exp(-u^2/2+tu) du$	(26)
<p>Using the formula in Gradshteyn &amp; Ryzhik, 7th Ed., p.365, eqn. 3.462(1) for <math>\nu &gt; 0</math> and <math>\beta &gt; 0</math>:</p>	
$\int_0^{\infty} x^{\nu-1} \exp(-\beta x^2 - \gamma x) dx = (2\beta)^{-\nu/2} \Gamma(\nu) \exp(\gamma^2/8\beta) D_{-\nu}(\gamma/\sqrt{2\beta})$	(27)
<p>Making the necessary substitutions:</p>	
$\exp(-t^2/2) \int_0^{\infty} u^{\nu} \exp(-u^2/2+tu) du = \Gamma(\nu+1) \exp(-t^2/4) D_{-\nu-1}(-t)$	(28)
<p>An alternative definition of the PCFs (see above MathWorld site) is:</p>	
$U(a, x) = D_{-a-1/2}(x)$	(29)
<p>Therefore for <math>a = \nu + 1/2</math>:</p>	
$\int_0^{\infty} u^{\nu} \exp(-(u-t)^2/2) du = \Gamma(\nu+1) \exp(-t^2/4) U(\nu+1/2, -t)$	(30)
<p>Therefore from (22) the required integral in the general case is:</p>	
$Q_h(\mu, \nu, \xi) = \sigma_h^{\nu} (\sqrt{2\pi} S_h^{\mu})^{-1} \exp((\xi \sigma_h)^2 / 2 S_h^2 - \xi I_h / S_h) \Gamma(\nu+1) \exp(-t_h^2(\xi)/4) U(\nu+1/2, -t_h(\xi))$	(31)



## 10. FW integral: implementation of parabolic cylinder functions in Fortran 90.

A Fortran 90 subroutine is available to calculate the PCFs  $U(a,x)$  and  $V(a,x)$  for  $x \geq 0$  by numerical methods [Netlib software repository: algorithm toms/850; see also <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.100.6920&rep=rep1&type=pdf>].

To avoid floating-point underflows and overflows in the exponential terms and to improve the relative accuracy of the result, the algorithm returns scaled PCFs:

$$\tilde{U}(a, x) = F(a, x) U(a, x) \quad (32)$$

$$\tilde{V}(a, x) = V(a, x) / F(a, x) \quad (33)$$

$$V(a, x) = (\Gamma(1/2+a)/\pi) (\sin(\pi a) D_{-a-1/2}(x) + D_{-a-1/2}(-x)) \quad (34)$$

where the scaling factor  $F(a,x)$  is:

$$F(a, x) = (x/2 + \sqrt{x^2/4+a})^a \exp((x/2) \sqrt{x^2/4+a} - a/2) \quad (35)$$

Therefore now writing (30) in terms of the scaled PCFs:

$$\begin{aligned} \int_0^{\infty} u^{\nu} \exp(-(u-t)^2/2) du &= \Gamma(1/2+a) \exp(-t^2/4) U(a, -t) \\ &= \Gamma(1/2+a) \exp(-t^2/4) \tilde{U}(a, -t) / F(a, -t) \\ &= \Gamma(1/2+a) (-t/2 + \sqrt{t^2/4+a})^{-a} \\ &\quad \exp(-t^2/4 + (t/2) \sqrt{t^2/4+a} + a/2) \tilde{U}(a, -t) \end{aligned} \quad (36)$$

Since the PCF algorithm used cannot return function values for  $x < 0$ , the following transformation must be used:

$$\begin{aligned} U(a, x) &= \pi V(a, -x) / \Gamma(1/2+a) - \sin(\pi a) U(a, -x) \\ &= F(a, -x) (\pi \tilde{V}(a, -x) / \Gamma(1/2+a) - \sin(\pi a) \tilde{U}(a, -x) / F^2(a, -x)) \end{aligned} \quad (37)$$

or in terms of  $t$  for  $t > 0$ :

$$U(a, -t) = F(a, t) (\pi \tilde{V}(a, t) / \Gamma(1/2+a) - \sin(\pi a) \tilde{U}(a, t) / F^2(a, t)) \quad (38)$$

Substitution of (38) in (36) and again combining exponential terms as before to avoid floating-point overflows then gives the required value of the integral for  $t > 0$ :

$$\begin{aligned} \int_0^{\infty} u^{\nu} \exp(-(u-t)^2/2) du &= \Gamma(1/2+a) (t/2 + \sqrt{t^2/4+a})^a \exp(-t^2/4 + (t/2) \sqrt{t^2/4+a} - a/2) \\ &\quad (\pi \tilde{V}(a, t) / \Gamma(1/2+a) - \sin(\pi a) \tilde{U}(a, t) / F^2(a, t)) \end{aligned} \quad (39)$$

# 11. FW likelihood function & posterior moments: untwinned acentric & centric perfect hemihedral twin cases.

<p>In the untwinned acentric &amp; centric perfect hemihedral twin cases, the required form of the Q integral is:</p>	
$  \begin{aligned}  P_h(I_h   S_h, \sigma_h) &= (\sigma_h \sqrt{2\pi} S_h)^{-1} \int_0^\infty \exp(-(I_h - J_h)^2 / 2\sigma_h^2) \exp(-J_h / S_h) dJ_h \\  &= Q_h(1, 0, 1) \\  &= S_h^{-1} \exp(\sigma_h^2 / 2 S_h^2 - I_h / S_h) \Phi(t_h(1))  \end{aligned}  $	(40)
<p>where <math>t_h(1) = I_h / \sigma_h - \sigma_h / S_h</math>. Hence the contribution to the negative log-likelihood gain is:</p>	
$-LLG_{ah}(g, \mathbf{U}) = \ln(S_h / \Phi(t_h(1))) - \sigma_h^2 / 2 S_h^2 + I_h / S_h$	(41)
<p>In the general untwinned case the posterior density of <math>J_h</math> from Bayes' theorem is the normalised joint probability density (14) <math>P_h(I_h, J_h   S_h, \sigma_h)</math> :</p>	
$  \begin{aligned}  P_h(J_h   I_h, S_h, \sigma_h) &= P_h(I_h, J_h   S_h, \sigma_h) / \int_0^\infty P_h(I_h, J_h   S_h, \sigma_h) dJ_h \\  &= P_h(I_h   J_h, \sigma_h) P_h(J_h   S_h) / \int_0^\infty P_h(I_h   J_h, \sigma_h) P_h(J_h   S_h) dJ_h  \end{aligned}  $	(42)
<p>The <math>m</math>'th posterior moment (expectation of <math>m</math>'th power) of <math>J_h</math> is then:</p>	
$\langle J_h^m \rangle = \int_0^\infty J_h^m P_h(J_h   I_h, S_h, \sigma_h) dJ_h$	(43)
<p>In the untwinned acentric and centric perfect hemihedral twin cases this is:</p>	
$  \begin{aligned}  \langle J_h^m \rangle &= \int_0^\infty J_h^m P_h(I_h   J_h, \sigma_h) P_h(J_h   S_h) dJ_h / \int_0^\infty P_h(I_h   J_h, \sigma_h) P_h(J_h   S_h) dJ_h \\  &= Q_h(0, m, 1) / Q_h(0, 0, 1)  \end{aligned}  $	(44)
<p>Note that for all the moments the value of <math>\mu</math> is irrelevant since it cancels out. In this case the posterior half- (<math>m = 1/2</math>) and first-order (<math>m = 1</math>) moments (<i>i.e.</i> the estimates of the true amplitude and intensity) are:</p>	
$\langle F_h \rangle = \langle J_h^{1/2} \rangle = (\sigma_h / 2\pi)^{1/2} \Gamma(\frac{3}{2}) \exp(-t_h^2(1) / 4) U(1, -t_h(1)) / \Phi(t_h(1))$	(45)
$  \begin{aligned}  \langle J_h \rangle &= (t_h(1) \Phi(t_h(1)) + \phi(t_h(1))) / \Phi(t_h(1)) \\  &= t_h(1) + \phi(t_h(1)) / \Phi(t_h(1))  \end{aligned}  $	(46)
<p>The standard deviation of <math>F</math> is the square root of the second moment of the deviation <math>F</math> from the mean:</p>	
$  \begin{aligned}  \sigma(F_h) &= \langle (F_h - \langle F_h \rangle)^2 \rangle^{1/2} \\  &= \langle F_h^2 \rangle - \langle F_h \rangle^2 \\  &= \langle J_h \rangle - \langle J_h^{1/2} \rangle^2  \end{aligned}  $	(47)

## 12. FW likelihood function & posterior moments: untwinned centric case.

For untwinned centric reflexions we have similarly:	
$  \begin{aligned}  P_{ch}(I_h   S_h, \sigma_h) &= (2\pi \sigma_h S_h)^{-1} \int_0^\infty J_h^{-1/2} \exp(-(I_h - J_h)^2 / 2\sigma_h^2) \exp(-J_h / 2S_h) dJ_h \\  &= (2\pi)^{-1/2} Q_h(1/2, -1/2, 1/2) \\  &= (2\pi(\sigma_h S_h)^{1/2})^{-1} \exp(\sigma_h^2 / 8 S_h^2 - I_h / 2S_h) \Gamma(1/2) \exp(-t_h^2(1/2) / 4) U(0, -t_h(1/2))  \end{aligned}  $	(48)
where $t_h(1/2) = I_h / \sigma_h - \sigma_h / 2S_h$ . Hence the contribution to the negative log-likelihood gain in the centric case is:	
$-LLG_{ch}(g, \mathbf{U}) = \ln(2\pi(\sigma_h S_h)^{1/2} / \Gamma(1/2) \exp(-t_h^2(1/2) / 4)) U(0, -t_h(1/2)) - \sigma_h^2 / 8 S_h^2 + I_h / 2S_h$	(49)
Note that the centric contribution to the FW function is not simply related to the acentric one, as it was for the PB function.	
The $m$ 'th posterior moment of $J_h$ in this case is then:	
$  \begin{aligned}  \langle J_h^m \rangle &= Q_h(0, m-1/2, 1/2) / Q_h(0, -1/2, 1/2) \\  &= U(m, -t_h(1/2)) / U(0, -t_h(1/2))  \end{aligned}  $	(50)
The half- and first-order posterior moments of $J_h$ are then:	
$  \begin{aligned}  \langle J_h^{1/2} \rangle &= Q_h(0, 0, 1/2) / Q_h(0, -1/2, 1/2) \\  &= (2\pi \sigma_h)^{1/2} \Phi(t_h(1/2)) / \Gamma(1/2) \exp(-t_h^2(1/2) / 4) U(0, -t_h(1/2)) \\  \langle J_h \rangle &= Q_h(0, 1/2, 1/2) / Q_h(0, -1/2, 1/2) \\  &= \sigma_h \Gamma(3/2) U(1, -t_h(1/2)) / \Gamma(1/2) U(0, -t_h(1/2))  \end{aligned}  $	(51)

### 13. FW likelihood function & posterior moments: acentric perfect hemihedral twin case.

For perfect hemihedrally twinned acentric reflexions, substituting (8) and (20) in (18) and again discarding factors that are independent of the parameters $g$ and $\mathbf{U}$ :	
$  \begin{aligned}  P_{ath}(I_h   S_h, \sigma_h) &= 4 (\sigma_h \sqrt{2\pi} S_h^2)^{-1} \int_0^\infty J_h \exp(-(I_h - J_h)^2 / 2\sigma_h^2) \exp(-2J_h/S_h) dJ_h \\  &= 4 Q_h(2, 1, 2) \\  &= 4 \sigma_h S_h^{-2} \exp(2\sigma_h^2/S_h^2 - 2I_h/S_h) (t_h(2) \Phi(t_h(2)) + \phi(t_h(2)))  \end{aligned}  $	(52)
where $t_h(2) = I_h/\sigma_h - 2\sigma_h/S_h$ . Hence the contribution to the negative log-likelihood gain in the acentric perfect hemihedral twin case is:	
$-\text{LLG}_{ath}(g, \mathbf{U}) = \ln(S_h^2 / 4 \sigma_h (t_h(2) \Phi(t_h(2)) + \phi(t_h(2)))) - 2\sigma_h^2/S_h^2 + 2I_h/S_h$	(53)
The $m$ 'th posterior moment of $J_h$ in this case is then:	
$\langle J_h^m \rangle = Q_h(0, m+1, 2) / Q_h(0, 1, 2)$	(54)
The half- and first-order posterior moments of $J_h$ are:	
$  \begin{aligned}  \langle J_h^{1/2} \rangle &= Q_h(0, \frac{3}{2}, 2) / Q_h(0, 1, 2) \\  &= (\sigma_h/2\pi)^{1/2} \Gamma(\frac{5}{2}) \exp(-t_h^2(2)/4) U(2, -t_h(2)) / (t_h(2) \Phi(t_h(2)) + \phi(t_h(2))) \\  \langle J_h \rangle &= Q_h(0, 2, 2) / Q_h(0, 1, 2) \\  &= \sigma_h ((t_h^2(2)+1) \Phi(t_h(2)) + t_h(2) \phi(t_h(2))) / (t_h(2) \Phi(t_h(2)) + \phi(t_h(2))) \\  &= \sigma_h (t_h^2(2) + 1 + t_h(2) \phi(t_h(2)) / \Phi(t_h(2))) / (t_h(2) + \phi(t_h(2)) / \Phi(t_h(2)))  \end{aligned}  $	(55)

#### 14. FW likelihood function & posterior moments: acentric imperfect hemihedral twin case.

The PDF and -LLG in the acentric imperfect hemihedral twin case is:	
$P_{ath}(I_h   S_h, \sigma_h, \alpha) = f(1-\alpha) + f(\alpha)$	
where:	
$\begin{aligned} f(x) &= (\sigma_h \sqrt{2\pi} (2x-1) S_h)^{-1} \int_0^{\infty} \exp(-(I_h - J_h)^2 / 2\sigma_h^2) \exp(-J_h/x S_h) dJ_h \\ &= (2x-1)^{-1} Q_h(1, 0, x^{-1}) \end{aligned}$	
$P_{ath}(I_h   S_h, \sigma_h, \alpha) = (1-2\alpha)^{-1} (Q_h(1, 0, (1-\alpha)^{-1}) - Q_h(1, 0, \alpha^{-1}))$	
$-LLG_{ath}(g, \mathbf{U}, \alpha) = -\ln(P_{ath}(I_h   S_h, \sigma_h, \alpha))$	
<p>In (58) &amp; (59) care must be taken to deal with values of the twin fraction <math>\alpha</math> near zero causing ill-conditioning which can occur during optimisation of <math>\alpha</math>. We use the scaled cumulative normal density, factoring out the exponentially increasing component of the function:</p> <p><math>\Phi'(t_h(\alpha^{-1})) = \exp(t_h^2(\alpha^{-1})/2) \Phi(t_h(\alpha^{-1}))</math> whenever <math>t_h(\alpha^{-1}) \leq 0</math> in the calculation of <math>Q_h(1, 0, \alpha^{-1})</math> using (23). It is not necessary to deal similarly with values of <math>\alpha</math> near 1 in the first term of (58) since we can always replace <math>1-\alpha</math> by <math>\alpha</math> (<i>i.e.</i> by definition <math>\alpha</math> can always be moved to the range <math>0 \leq \alpha \leq 0.5</math>). By using a first-order Taylor expansion in <math>1-2\alpha</math> we can deal with ill-conditioning due to <math>\alpha</math> near the perfect twin case 0.5 (since then both numerator and denominator in (58) are near zero).</p> <p>The <math>m</math>'th posterior moment of <math>J_h</math> in this case is then:</p>	
$\langle J_h^m \rangle = \frac{(Q_h(0, m, (1-\alpha)^{-1}) - Q_h(0, m, \alpha^{-1}))}{(Q_h(0, 0, (1-\alpha)^{-1}) - Q_h(0, 0, \alpha^{-1}))}$	
The half- and first-order posterior moments of $J_h$ are:	
$\begin{aligned} \langle J_h^{1/2} \rangle &= \frac{(Q_h(0, 1/2, (1-\alpha)^{-1}) - Q_h(0, 1/2, \alpha^{-1}))}{(Q_h(0, 0, (1-\alpha)^{-1}) - Q_h(0, 0, \alpha^{-1}))} \\ &= \frac{(\sigma_h/2\pi)^{1/2} \Gamma(\frac{3}{2}) (\exp(-t_h^2((1-\alpha)^{-1})/4) \mathbf{U}(1, -t_h((1-\alpha)^{-1})) - \exp(-t_h^2(\alpha^{-1})/4) \mathbf{U}(1, -t_h(\alpha^{-1})))}{(\Phi(t_h((1-\alpha)^{-1})) - \Phi(t_h(\alpha^{-1})))} \end{aligned}$	
$\begin{aligned} \langle J_h \rangle &= \frac{(Q_h(0, 1, (1-\alpha)^{-1}) - Q_h(0, 1, \alpha^{-1}))}{(Q_h(0, 0, (1-\alpha)^{-1}) - Q_h(0, 0, \alpha^{-1}))} \\ &= \frac{\sigma_h (t_h((1-\alpha)^{-1}) \Phi(t_h((1-\alpha)^{-1})) + \phi(t_h((1-\alpha)^{-1})) - t_h(\alpha^{-1}) \Phi(t_h(\alpha^{-1})) - \phi(t_h(\alpha^{-1})))}{(\Phi(t_h((1-\alpha)^{-1})) - \Phi(t_h(\alpha^{-1})))} \end{aligned}$	

## 15. FW likelihood function & posterior moments: centric imperfect hemihedral twin case.

The PDF in the centric imperfect hemihedral twin case is:

$$\begin{aligned}
 P_{cth}(I_h | S_h, \sigma_h, \alpha) &= (2\sigma_h \sqrt{2\pi\alpha(1-\alpha)} S_h)^{-1} \int_0^\infty \exp(-(I_h - J_h)^2 / 2\sigma_h^2 - J_h / 2\alpha S_h) \\
 &\quad {}_1F_1(1/2; 1; (1-2\alpha) J_h / 2\alpha(1-\alpha) S_h) dJ_h \\
 &= (2\sigma_h \sqrt{2\pi\alpha(1-\alpha)} S_h)^{-1} \int_0^\infty \exp(-(I_h - J_h)^2 / 2\sigma_h^2 - J_h / 4\alpha(1-\alpha) S_h) \\
 &\quad I_0((1-2\alpha) J_h / 4\alpha(1-\alpha) S_h) dJ_h \\
 &= (2\sigma_h \sqrt{2\pi\alpha(1-\alpha)} S_h)^{-1} \int_0^\infty \exp(-(I_h - J_h)^2 / 2\sigma_h^2 - J_h / 2(1-\alpha) S_h) \\
 &\quad I'_0((1-2\alpha) J_h / 4\alpha(1-\alpha) S_h) dJ_h
 \end{aligned} \tag{63}$$

$$-\text{LLG}_{cth}(g, \mathbf{U}, \alpha) = -\ln(P_{cth}(I_h | S_h, \sigma_h, \alpha)) \tag{64}$$

In (63) we again use a scaled function, in this case the scaled modified Bessel function  $I'_0$  in order to factor out the exponentially increasing component of the function and thus avoid floating-point overflow. Note that in this case there are no ill-conditioning issues from the division by  $\alpha$  for values of  $\alpha$  near to zero.

The general  $m$ 'th order posterior moment of  $J_h$  in this case is then:

$$\begin{aligned}
 \langle J_h^m \rangle &= \frac{\int_0^\infty J_h^m \exp(-(I_h - J_h)^2 / 2\sigma_h^2 - J_h / 2(1-\alpha) S_h) I'_0((1-2\alpha) J_h / 4\alpha(1-\alpha) S_h) dJ_h}{\int_0^\infty \exp(-(I_h - J_h)^2 / 2\sigma_h^2 - J_h / 2(1-\alpha) S_h) I'_0((1-2\alpha) J_h / 4\alpha(1-\alpha) S_h) dJ_h}
 \end{aligned} \tag{65}$$

This expression cannot be further simplified, so the integrals must be evaluated numerically.

## 16. REFERENCE

Rees, D.C. (1982) "A General Theory of X-ray Intensity Statistics for Twins by Merohedry". *Acta Cryst.* **A38**, 201-7.