Minimum-dimension LSF modelling

L. Lindegren

GAIA–C3–TN–LU–LL–084–01

1 August 2009

Abstract. For the LSF and PSF calibration in the Gaia IDT/IDU we need LSF/PSF models that are both accurate and have as few free parameters as possible. General-purpose functions such as splines are not the best choice since they tend to require more parameters than strictly necessary, and have a tendency to become unstable when fitted to noisy or inhomogeneous data. Using Principal Component Analysis (PCA) of a very large set of theoretically computed, physically plausible, LSF/PSFs, it is possible to derive models with a minimum number of free parameters for given accuracy. A first application to the LSF modelling is described. The results indicate that some 10 free parameters should be sufficient to represent all LSF shapes with an RMS error of 10^{-4}. With the proposed decomposition, the number of free parameters can easily be adapted to the number and quality of available data. The numerically generated basis functions can be represented by splines or similar for easy evaluation. An example is worked out using bi-quartic splines plus analytical wings. The resulting LSF model is fitted to Astrium Campaign #3 test data with excellent results.

1 Introduction

The current LSF and PSF modelling for simulation purposes uses either detailed (oversampled) numerical maps, or some analytical fit to the maps, typically using splines or similar general-purpose functions. A common feature for these models is that they have a high dimensionality, or number of free parameters. In the case of numerical maps, with an oversampling factor of (say) 5, the number of parameters (or bins in this case) ranges from ~ 100 to many thousands. For one-dimensional splines (LSF) the number of parameters is typically of order 35, and the PSF would require at least twice this number (assuming, as a minimum, that the PSF is the product of the AL and AC LSFs).

Using a fairly large number of model parameters is a simple way to ensure that modelling errors are small enough. For example, if the five-fold oversampling in the numerical maps gives too large interpolation errors, one could simply increase the oversampling factor, thus increasing the number of parameters (bins).

This may be an acceptable approach for the simulations, but not for the LSF and PSF calibrations. Here we must strive to minimize the number of free parameters, whilst keeping modelling errors within acceptable bounds. Moreover, the calibration model needs to cope with noisy, irregularly sampled and generally imperfect data. For example, fitting a general spline to noisy data tends to generate unphysical oscillations, especially in the LSF wings, unless the fit is manipulated in various ways by means of ad-hoc constraints, modified knot placements, or similar.

A further difference between simulations and calibrations is that the LSF/PSF used in the simulations can be synthesized in terms of a (potentially large number of) monochromatic or quasi-monochromatic components. A corresponding decomposition will not work for the calibrations, due to the inherent instability of the inverse problem of decomposing the polychromatic function.

This note describes experiments to derive an LSF model of minimal dimension, i.e., with the least possible number of free parameters for a given modelling accuracy. It is based on a statistical analysis, using Principal Component Analysis (PCA), of a large number of theoretical LSFs.

2 Expansion of the LSF in basis functions

Quite generally, the LSF model $\tilde{L}(u)$ can be written as a linear combination of basis functions,

$$\tilde{L}(u) = \sum_{m=0}^{N} c_m B_m(u)\,, \tag{1}$$

where $N+1$ is the number of basis functions $B_m(u)$, and c_m are the coefficients of the basis functions. We use the tilde ($\tilde{}$) to distinguish our model from the true (physical) LSF $L(u)$ discussed in the next section. Our model $\tilde{L}$ is an approximation of L.

N actually stands for the number of degrees of freedom of the model. Since the LSF must be normalized,

$$\int_{-\infty}^{+\infty} \tilde{L}(u)\,\mathrm{d}u = 1\,, \tag{2}$$

it follows that the coefficients must satisfy the constraint

$$\sum_{m=0}^{N} b_m c_m = 1\,, \tag{3}$$

where

$$b_m = \int_{-\infty}^{+\infty} B_m(u)\,\mathrm{d}u\,. \tag{4}$$

Thus there are only N free parameters, although the model uses $N+1$ basis functions.

Although (1) is similar to the spline representation discussed for example in [4], it is much more general and we are presently not restricting ourselves to splines as basis functions. However, an important feature is that the model is *linear* in terms of the unknowns c_m, which simplifies the data processing, e.g., in terms of least-squares fitting to the observed samples.

Actually, the general LSF model should be written in the slightly modified form

$$\tilde{L}(u) = \sum_{m=0}^{N} c_m B_m(u - \delta u)\,. \tag{5}$$

The shift δu is related to the definition of the origin of the LSF [6], which needs to be decoupled from the rest of the model. The full model therefore has $N+1$ degrees of freedom, when the origin parameter δu is included. The determination of δu is a separate problem which will not be further discussed here. Thus the subsequent analysis may be restricted to the unshifted model with $\delta u = 0$, that is Eq. (1).

3 Physical model of the LSF

The elements of the physical model that go into calculating the detailed PSF and LSF have been reviewed in several previous papers. Briefly, they are as follows:

- The optical wavefront error (WFE) $W(x, y)$, which depends on the FOV and field angles, but is independent of wavelength and supposedly very stable over time. The WFE uniquely determines the optical PSF as function of wavelength.
- The pixel binning and four-phase TDI charge transfer, that can be represented by a convolution with a known, constant and wavelength-independent function.
- The charge diffusion, which depends (at least) on wavelength.
- The spectral energy distribution (SED) of the object. The polychromatic LSF/PSF is the photon-weighted average of the monochromatic LSF/PSFs.
- AL and AC smearing due to image motion relative to the mean charge motion during exposure. This can be modelled as a convolution with a 2D box function with known AL and AC widths. The width could vary from 0 to some maximum value ($\sim$0.5 pixel AL, few pixels AC).
- Finally, the LSF is obtained by integrating the PSF over a certain number of AC pixels and re-normalizing to unit area.

The AL and AC smearing due to image motion is special because (i) it is variable between observations on the same CCD, (ii) the amount of smearing is known for each observation, and (iii) it can be applied to un-smeared basis functions in the form of lookup tables. Therefore, this effect need not be considered here.

Using the steps outlined above (but neglecting the smearing) we can compute the physically expected LSF, denoted $L(u)$, for arbitrary WFE and SED. This calculation can in principle be done at arbitrary resolution (sub-pixel sampling) and out to arbitrarily large values of u.

4 Deriving the optimal basis functions

For any given number of free parameters, N, we now want to derive the most accurate approximation $\tilde{L}$ of the form (1) to the calculated L. We also want to find out how good the approximation is as function of N. It turns out that Principal Component Analysis (PCA) provides the solution to both problems. The following is basically a description of PCA in the context of the present problem.

First, a large number of polychromatic LSFs must be generated, using many different WFE maps and SEDs. The procedure for choosing the WFEs and SEDs is described in Sect. 5. Let $L_k(u)$ be the LSF for the kth WFE-SED combination, where $k = 1 \ldots K$.[1] Each LSF is represented by M equidistantly sampled points, constituting the vector (column matrix) $\boldsymbol{L}_k = [\ L_k(u_1)\ L_k(u_2)\ \ldots\ L_k(u_M)\ ]'$, where $u_m = u_1 + (m-1)\Delta u$. The sampling interval Δu should be a fraction of a pixel (at most $\Delta u = 1/8$), and $[u_1,\ u_M]$ must span the relevant central part of the LSF (say $\pm$20 pixels).

[1] For consistency with the MATLAB implementation in Sect. 5, indices are here counted from 1.

The ensemble of K LSFs can be contained in the $M \times K$ matrix $\boldsymbol{L} = [\ \boldsymbol{L}_1\ \boldsymbol{L}_2\ \ldots\ \boldsymbol{L}_K\ ]$. We compute first the mean LSF

$$\boldsymbol{B}_0 = \frac{1}{K}\sum_{k=1}^{K} \boldsymbol{L}_k\,, \tag{6}$$

which is the 0th basis vector. Next we compute the deviations from the mean profile, $\boldsymbol{X}_k = \boldsymbol{L}_k - \boldsymbol{B}_0$, to obtain the $M \times K$ matrix $\boldsymbol{X} = [\ \boldsymbol{X}_1\ \boldsymbol{X}_2\ \ldots\ \boldsymbol{X}_K\ ]$. The covariance matrix is

$$\boldsymbol{C} = \frac{1}{K}\,\boldsymbol{X}\boldsymbol{X}'\,. \tag{7}$$

Singular value decomposition (SVD) then gives[2]

$$\boldsymbol{C} = \boldsymbol{B}\boldsymbol{D}\boldsymbol{B}'\,, \tag{8}$$

where $\boldsymbol{B}$ is an orthonormal $M \times M$ matrix of basis vectors and $\boldsymbol{D}$ a diagonal matrix of singular values. Standard SVD routines (e.g., in MATLAB) returns the SVD decomposition in order of non-increasing size of the singular values, i.e., $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_M$ where $\sigma_m = D_{mm}$ is the mth singular value.

The columns $\boldsymbol{B}_m$ in $\boldsymbol{B} = [\ \boldsymbol{B}_1\ \boldsymbol{B}_2\ \ldots\ \boldsymbol{B}_M\ ]$ form an orthonormal set of basis vectors for the expansion of arbitrary vectors of length M whose mean value is zero. That is, any LSF can be exactly written as the linear combination of the M basis vectors,

$$\boldsymbol{L} = \boldsymbol{B}_0 + \sum_{m=1}^{M} c_m \boldsymbol{B}_m = \boldsymbol{B}_0 + \boldsymbol{B}\boldsymbol{c} \tag{9}$$

for some coefficient vector $\boldsymbol{c} = [\ c_1\ c_2\ \ldots\ c_M\ ]'$. From the orthonormality of $\boldsymbol{B}$ we have

$$\boldsymbol{c} = \boldsymbol{B}'(\boldsymbol{L} - \boldsymbol{B}_0) \tag{10}$$

or $c_m = \boldsymbol{B}'_m(\boldsymbol{L} - \boldsymbol{B}_0)$. The really nice thing about (9) is that the truncated expansion,

$$\tilde{\boldsymbol{L}} = \boldsymbol{B}_0 + \sum_{m=1}^{N} c_m \boldsymbol{B}_m\,, \qquad N < M \tag{11}$$

has minimum expected RMS error among all linear models with N free parameters, if expectations are calculated over a population of LSFs with mean value $\boldsymbol{B}_0$ and covariance $\boldsymbol{C}$. This immediately follows from the readily derived formula for the sum of squared residuals,

$$\mathrm{E}\left[(\tilde{\boldsymbol{L}} - \boldsymbol{L})'(\tilde{\boldsymbol{L}} - \boldsymbol{L})\right] = \sum_{m=N+1}^{M} \sigma_m\,, \tag{12}$$

and the non-increasing order of the singular values. This therefore completely solves the problem of finding the minimum-order linear model for any specified upper limit on the expected RMS error.

In Eq. (11) the value $c_0 = 1$ is implied for the coefficient of the basis vector $\boldsymbol{B}_0$. This would be fine if $b_m = 0$ for all $m > 0$. In practice this cannot be guaranteed, and it is then better to retain c_0 as an adjustable parameter and re-normalize the coefficient vector after fitting, so that Eq. (3) is satisfied.

[2] $\boldsymbol{C}$ being real and symmetric, the general SVD form $\boldsymbol{V}\boldsymbol{D}\boldsymbol{U}^*$ simplifies to $\boldsymbol{V}\boldsymbol{D}\boldsymbol{V}'$.

5 A first application to the LSF modelling

The proposed methodology has been tested, using a total of 20,000 theoretical LSFs computed as outlined below and analysed according to Sect. 4.

5.1 Generation of WFE maps and monochromatic LSFs

First, 200 random WFE maps were computed by using a 6th order polynomial decomposition

$$W(x,y) = \sum_{\substack{i+j>1 \\ i+j<6}} Q_{ij}\, N_i(2x/D)\, N_j(2y/H) \tag{13}$$

where $N_n(x) = (2n+1)^{1/2} P_n(x)$ are normalized Legendre polynomials and $D = 1.45$ m, $H = 0.5$ m the pupil dimensions. For each map a target RMS WFE, W_0, was selected as a random number with uniform density between 40 and 60 nm. Then, each of the 18 coefficients $Q_{02},\ Q_{11},\ \ldots,\ Q_{50}$ were taken from a normal distribution with zero mean and standard deviation $W_0/\sqrt{18}$. The actual WFE, $W = (\sum_{ij} Q_{ij}^2)^{1/2}$, therefore had a somewhat wider distribution from 22 to 77 nm, which is quite realistic for Gaia.

Note that the absence of tip-tilt terms in (13) means that the origin $u = v = 0$ corresponds to the PSF centroid in the limit of infinite wavelength [3]. This seems to be a physically well-chosen reference point, but it should be realized that it does not represent the origin actually used in the data processing, which is fixed by convention [6]. This is why the δu is needed in the complete model (5).

For each WFE map, monochromatic PSFs were computed using the algorithms described in [2] and [4]. The main parameters for the calculation were $N_x = 1024$, $N_y = 512$, $s_u = s_v = 0.125$ pixel for the sampling in u and v, and $\sigma_u = \sigma_v = 4$ μm for the charge diffusion. The PSFs were normalized and collapsed to one-dimensional LSFs. Only the 321 central points, corresponding to $u = \pm 20$ pixels at 0.125 pixel resolution, were stored for the further analysis.

Monochromatic LSFs were thus computed for 39 wavelengths, geometrically spaced from 330 to 1015 nm (i.e., with a constant factor 1.03 between successive wavelengths). Associated with each wavelength point was a weight proportional to the AF response function (product of transmittance and quantum efficiency) and the wavelength spacing of the points.

5.2 Generation of SEDs and polychromatic LSFs

Next, 50 random SEDs were generated, independently for each WFE map. For simplicity, this consisted of a Planck function, with temperature parameter $\theta = (5040\ \mathrm{K})/T$ randomly selected in the interval 0.2 to 2 (i.e., 2520 K to 25200 K). Using a uniform distribution in θ rather than T puts more emphasis on cool stars, as in the real sky. To simulate spectral variations, the Planck function in each of the 39 wavelength points was multiplied by a log-normal random variable with standard deviation 0.3 neper. Figure 1 shows a few representative SEDs generated in this way.

Folding the monochromatic LSFs with the SEDs (and the instrument response through the weight factors mentioned in Sect. 5.1) resulted in 50 polychromatic LSFs for each WFE map. The 200 WFE maps and 50 SEDs thus gave 10,000 different LSFs.

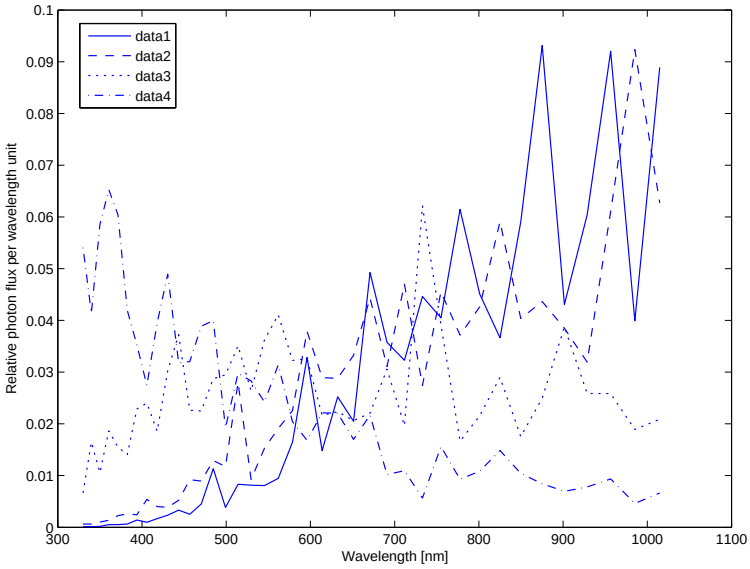

FIGURE 1: A representative set of random spectra.

This set was doubled in size by augmenting it with the reversed LSFs (mirrored in $u = 0$). This trick ensures that the PCA provides basis functions that are either even or odd, which is desirable for reasons of symmetry. The total number of LSFs was therefore $K = 20000$, each with $M = 321$ data points. The subsequent analysis followed the PCA outlined in Sect. 4.

5.3 Resulting basis functions

Figure 2 shows all 321 singular values σ_m plotted on a logarithmic scale. The curve suggests that the true dimensionality of the data set is around 100. More interesting is the RMS residual as function of the number of retained components, N, shown in Fig. 3. This shows that it is possible to represent this set of LSFs with an RMS error of 10^{-3} with only 5 components (in addition to the mean profile), and to 10^{-4} with about 10 components.

The mean LSF profile, $\boldsymbol{B}_0$ or $B_0(u)$, is shown in Fig. 4. By construction it is completely symmetric.

Figures 5–6 show the first 12 basis functions $\boldsymbol{B}_m$ or $B_m(u)$, $m = 1 \ldots 12$. They have a remarkably regular appearance, suggesting that the random set is large enough for stable evaluation of at least the low-order basis functions. It is interesting to note that the first component represents a variation in LSF width, relative to the mean LSF, and that the basis functions thereafter alternate (almost) regularly between odd and even functions of increasing complexity. For example, the second component is odd, representing a certain asymmetry in the profile. However, the asymmetry is not just a shift of the profile – that would have a basis function proportional to the derivative of the mean profile, while $B_2(u)$ is more complex. A simple shift is actually not permitted, since the WFE has no tilt term ($Q_{10} = 0$). It is also noteworthy that the outer parts of the basis functions (beyond ± 10 pixels) remain smooth up to high orders.

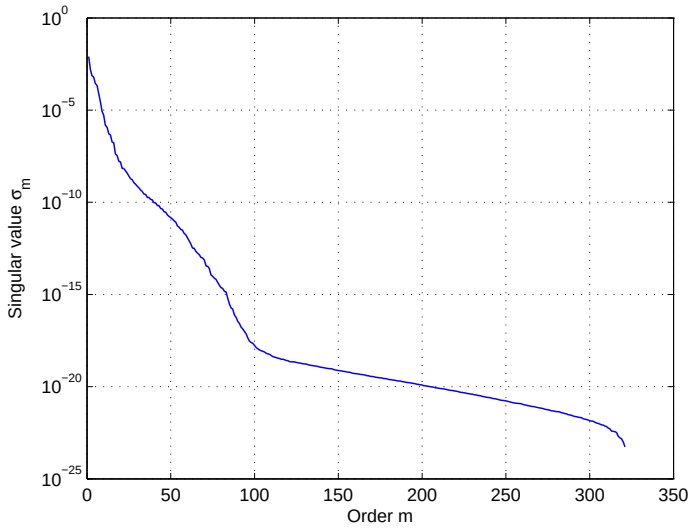

FIGURE 2: The singular values (σ_m) of the LSF decomposition.

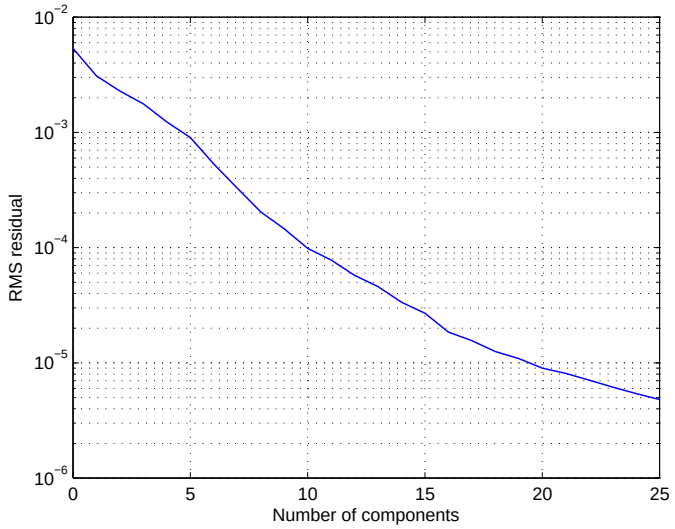

FIGURE 3: The overall RMS residual of the approximation, as function of the number of components, N.

Figure 7, finally, shows a few examples of the fit obtained by using the mean LSF plus the first 12 basis functions. The left diagrams show the accurately computed LSFs (dots), with the least-squares fits superposed (lines); the right diagrams show the residuals. With $N = 12$ basis functions, the overall RMS residual according to Fig. 3 is about 7×10^{-5}. The plots in Fig. 7 show that the largest residuals are usually obtained in the core of the LSF, where it may reach about 4 times the overall RMS value.

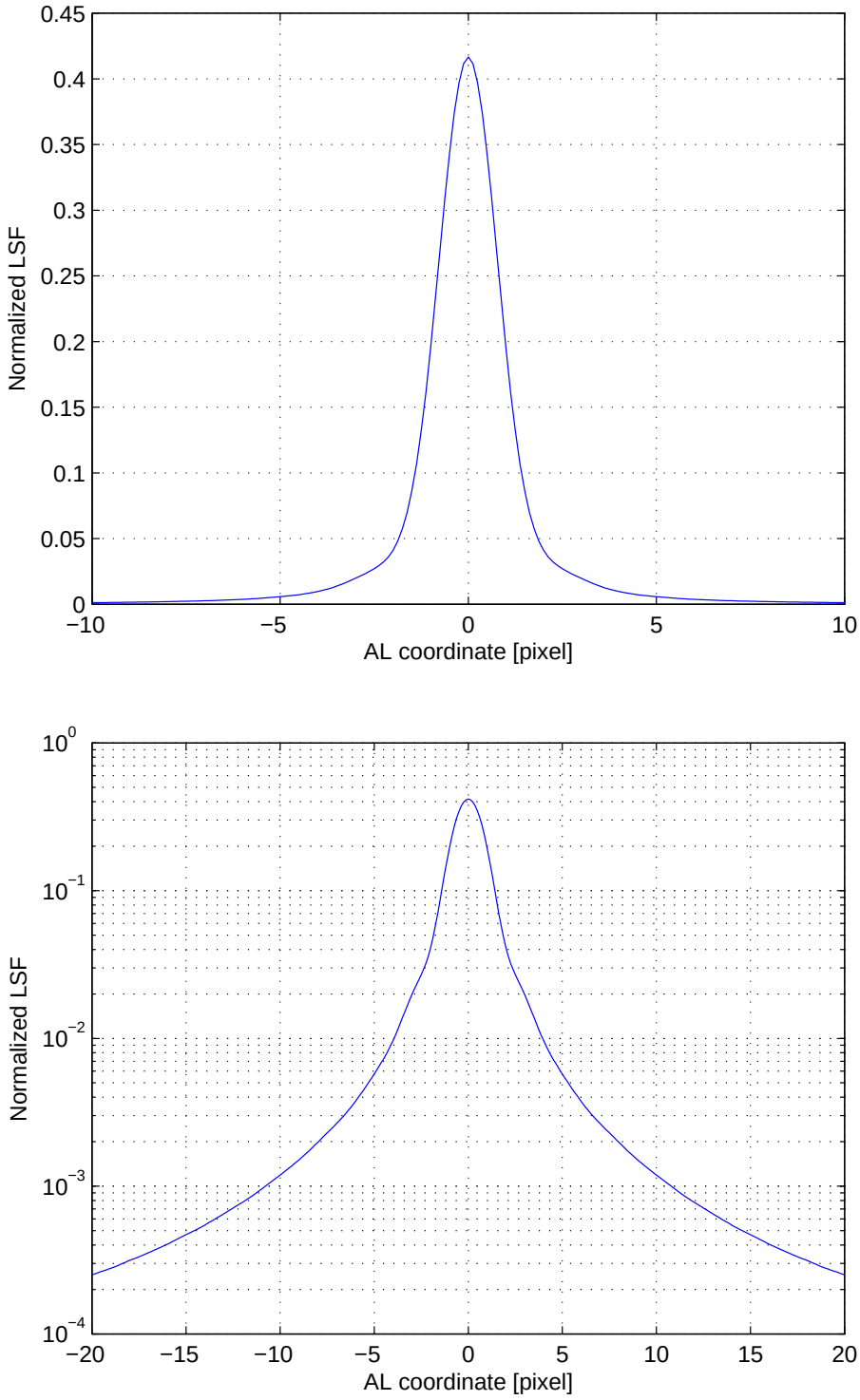

FIGURE 4: The mean LSF, or 0th basis function, $B_0(u)$.

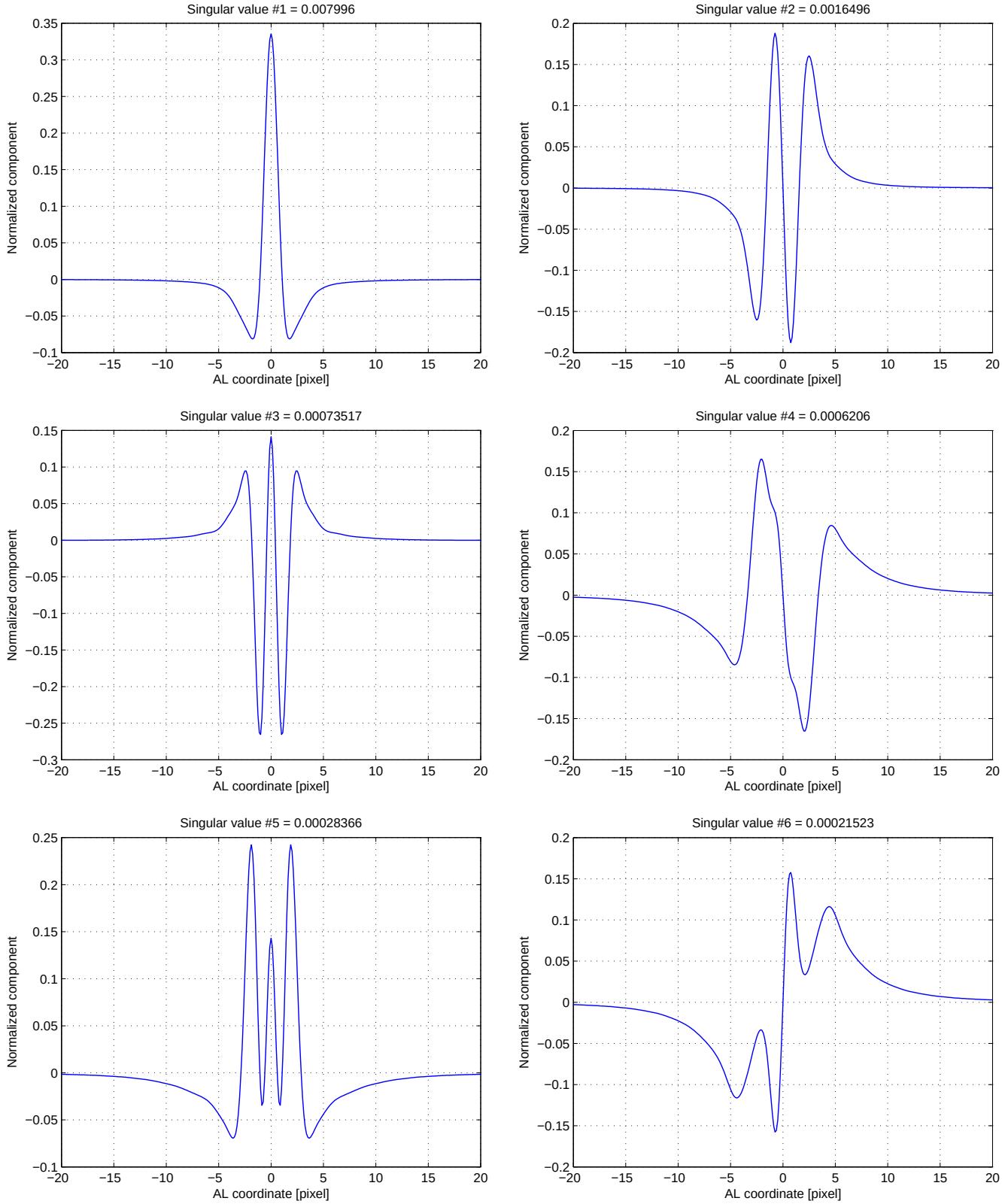

FIGURE 5: The first six basis functions $B_1(u)$ through $B_6(u)$. The corresponding singular values are shown above each graph.

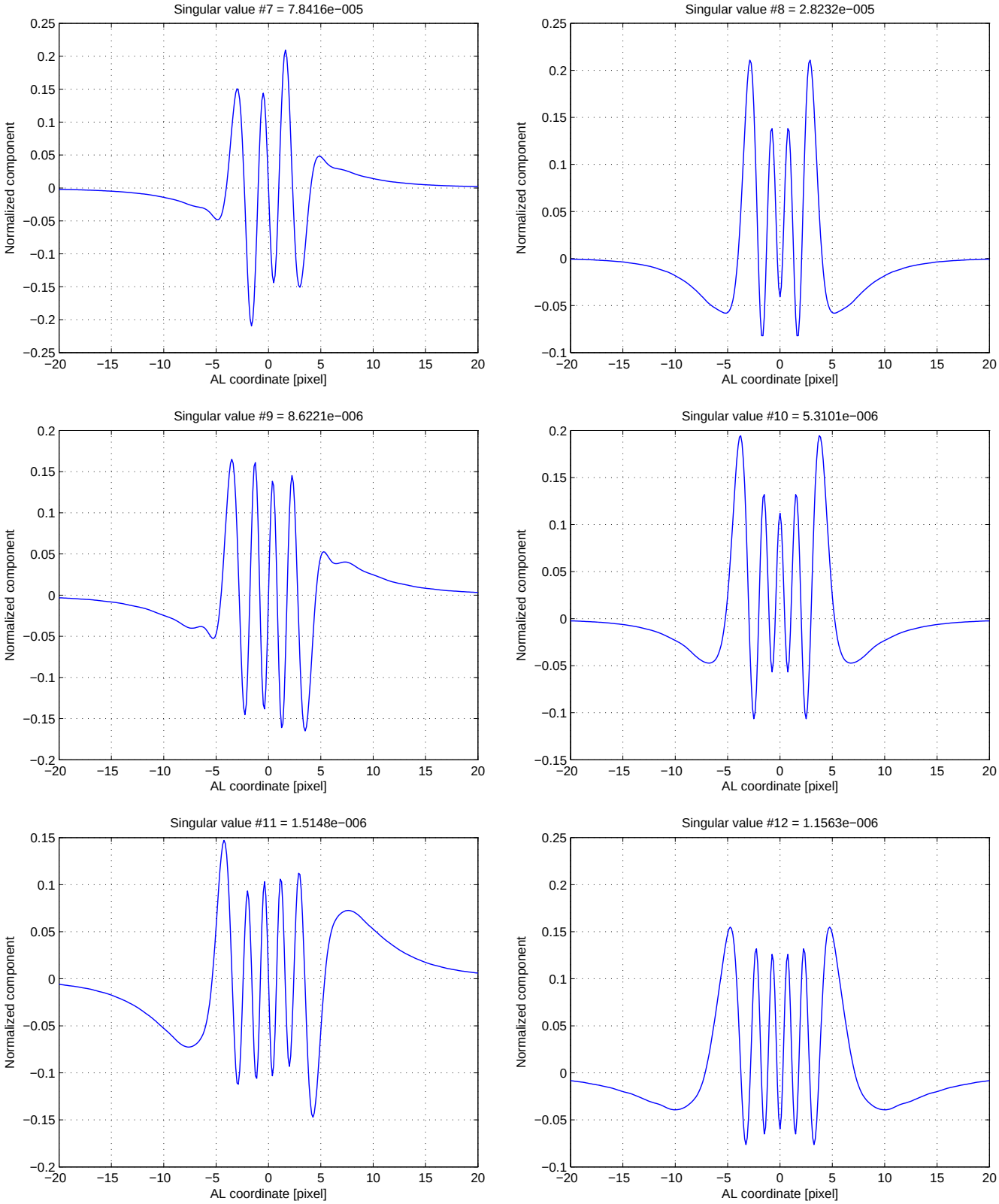

FIGURE 6: The next six basis functions $B_7(u)$ through $B_{12}(u)$. The corresponding singular values are shown above each graph.

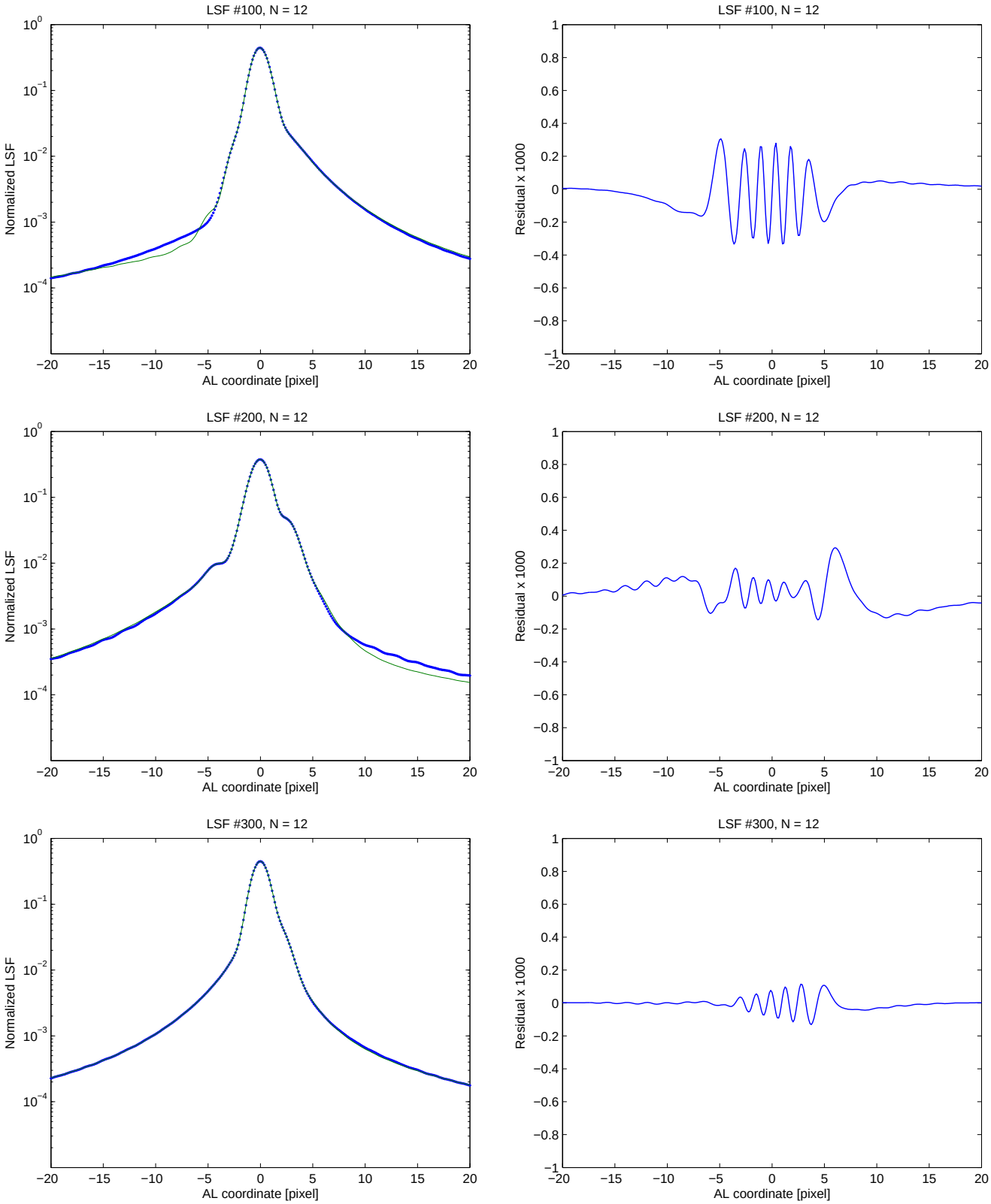

FIGURE 7: Random examples of the LSF approximation with $N = 12$ components: to the left, the approximation $\tilde{L}(u)$ (line) superposed on the actual LSF $L(u)$ (points); to the right, the residuals of the fit magnified a factor 1000.

6 Practical considerations and an application to test data

In this section we consider some aspects of the practical implementation of the basis function approach to the LSF fitting, and demonstrate its feasibility on Astrium test data.

6.1 Spline representation of basis functions

The numerically computed basis functions $B_m(u)$ (i.e., the vectors $\boldsymbol{B}_m$) span the interval $-20 \le u \le +20$ pixel in steps of 0.125 pixel. This discretization is not fine enough to be used with linear interpolation. A more sophisticated interpolation method must therefore be used. For some purposes it is also desirable to be able to extrapolate the basis functions beyond ± 20 pixels, e.g., in order to compute the normalization factors b_m in Eq. (4).

Here, the bi-quartic B-splines introduced in [4] come in handy. They are sufficiently flexible to allow reasonable approximation of all spatial frequencies expected in the diffraction images of Gaia, an satisfy the 'shift-invariant sum' condition important for the LSF model. However, while the spline provides a good approximation within the span of the data, it is not useful for extrapolation. For this purpose, a Lorentzian function (Cauchy distribution) superposed on the spline was introduced in [4], with the spline gently attaching to the Lorentzian at suitable transition points, making the function and first three derivatives continuous everywhere. The Lorentzian was chosen because (i) for large u it falls off as u^{-2}, as theoretically expected for the diffraction wings of a rectangular pupil; and (ii) is has no singularity at $u = 0$. However, for the present basis functions the Lorentzian is less suitable. It would probably work for the even basis functions (including B_0), but what about the odd functions? One could think of using the derivative of the Lorentzian for them. However, since this function has a certain slope at $u = 0$, it would also affect the centroid of the resulting synthesized LSF in an undesirable way.

The proposed solution is to use a special 'tail function' $t(u)$ to represent each wing, with the following properties:

- For large u, $t(u) \propto u^{-2}$. Specifically, this should hold for $u > \beta$, where β is the maximum $|u|$ for which the basis functions have been computed.
- For small u, $t(u) = 0$. We take this to hold for $u < \alpha$, where α is at least a few pixels (e.g., $\alpha > 3$ is required to avoid affecting the Tukey biweight centroid).
- $t(u)$ and its derivative are everywhere continuous.

The tail function used in the present example (Fig. 8) is

$$t(u) = \begin{cases} 0 & \text{if } u \le \alpha \\ \gamma_1 (u-\alpha)^3 + \gamma_2 (u-\alpha)^4 & \text{if } \alpha \le u \le \beta \\ \gamma_3 u^{-2} & \text{if } \beta \le u \end{cases} \tag{14}$$

with $\alpha = 5$, $\beta = 20$, $\gamma_1 = 44/732375$, $\gamma_2 = -4/1220625$, and $\gamma_3 = 3200/217$. It is strictly zero for $u < 5$, decreases as u^{-2} for $u > 20$, and in the interval $[5, 20]$ it is a fourth-degree polynomial. The γ coefficients have been computed to make $t(u)$ and $t'(u)$ continuous at $u = 20$, and $\int_{-\infty}^{\infty} t(u)\,\mathrm{d}u = 1$. At $u = 5$ also the second derivative is continuous.

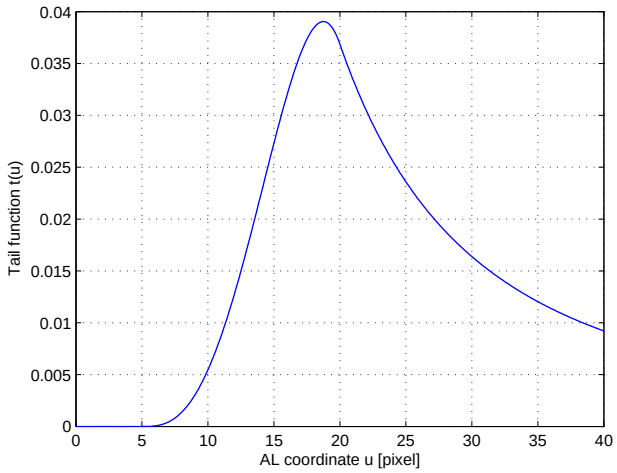

FIGURE 8: The tail function, Eq. (14).

The basis function $B_m(u)$ is then approximated as the sum of a bi-quartic spline $S_m(u)$, which is non-zero only for $|u| < \beta$, and two tail functions,

$$B_m(u) \simeq S_m(u) + s_m^- t(-u) + s_m^+ t(u) \tag{15}$$

The coefficients of the tail function are chosen to make the approximation continuous at $u = \pm\beta$:

$$s_m^- = B_m(-\beta)/t(\beta)\,, \quad s_m^+ = B_m(\beta)/t(\beta) \tag{16}$$

From the properties of $t(u)$ and the spline, it is seen that the approximating function has a continuous derivative at $u = \pm\beta$, and continuous first and second derivatives at all other points, including $u = \pm\alpha$.

$S_m(u)$ is a bi-quartic spline defined on the knot sequence $\tau_k = -\beta + 0.5k$, $k = 0, 1, \ldots, 4\beta$. A bi-quartic spline can formally be handled as a 6th order spline; thus, for any u in the interval $[\tau_\ell, \tau_{\ell+1})$, so that ℓ is the 'left index' of u, there are at most 6 non-zero B-splines, which are denoted $\mathcal{B}_{\ell-5}(u)$, $\mathcal{B}_{\ell-4}(u)$, $\ldots$, $\mathcal{B}_\ell(u)$.[3] The function $S_m(u)$ therefore has $4\beta - 5$ degrees of freedom (75 in the present example), namely the coefficients s_{mk} of $\mathcal{B}_k(u)$ for $k = 0 \ldots 4\beta - 6$. However, for any given u at most 6 B-splines contribute to $S_m(u)$, which can therefore be written as

$$S_m(u) = \sum_{k=\max(0,\ell-5)}^{\min(4\beta-6,\ell)} s_{mk} \mathcal{B}_k(u) \tag{17}$$

The spline coefficients s_{mk} are determined by least squares fitting of Eq. (17) to $B_m(u) - s_m^- t(-u) - s_m^+ t(u)$. The integral of the approximating function is

$$b_m = \sum_{k=0}^{4\beta-6} s_{mk} + s_m^- + s_m^+ \tag{18}$$

For any linear combination of the basis functions, the LSF is obtained in a similar form, i.e., as a sum of a bi-quartic spline and two tail functions. Their coefficients are obtained by linear combination of the corresponding coefficients for the basis functions. The numerical evaluation of the LSF and its derivative will therefore be fast and simple.

[3] The notation $\mathcal{B}$ for the B-splines is temporarily used here to distinguish from the LSF basis functions.

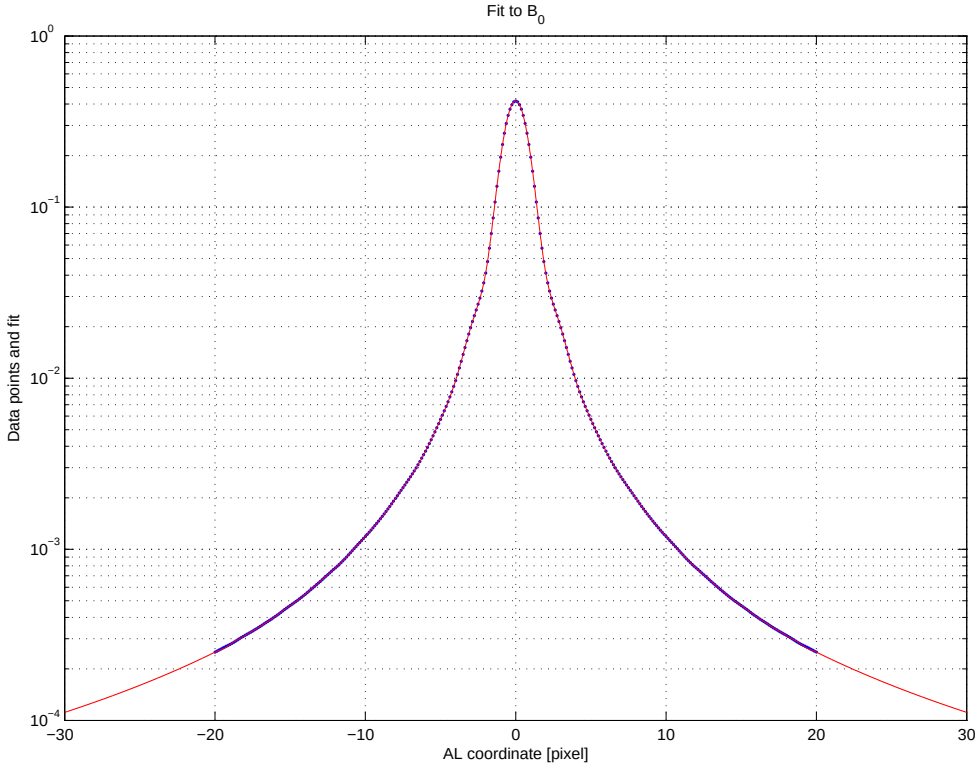

FIGURE 9: The dots show the numerically calculated basis function $B_0(u)$ for $|u| \leq 20$ pixel (i.e., $\boldsymbol{B}_0$), the line is the fitted approximation, Eq. (15). The residuals of the fit are shown in Fig. 10.

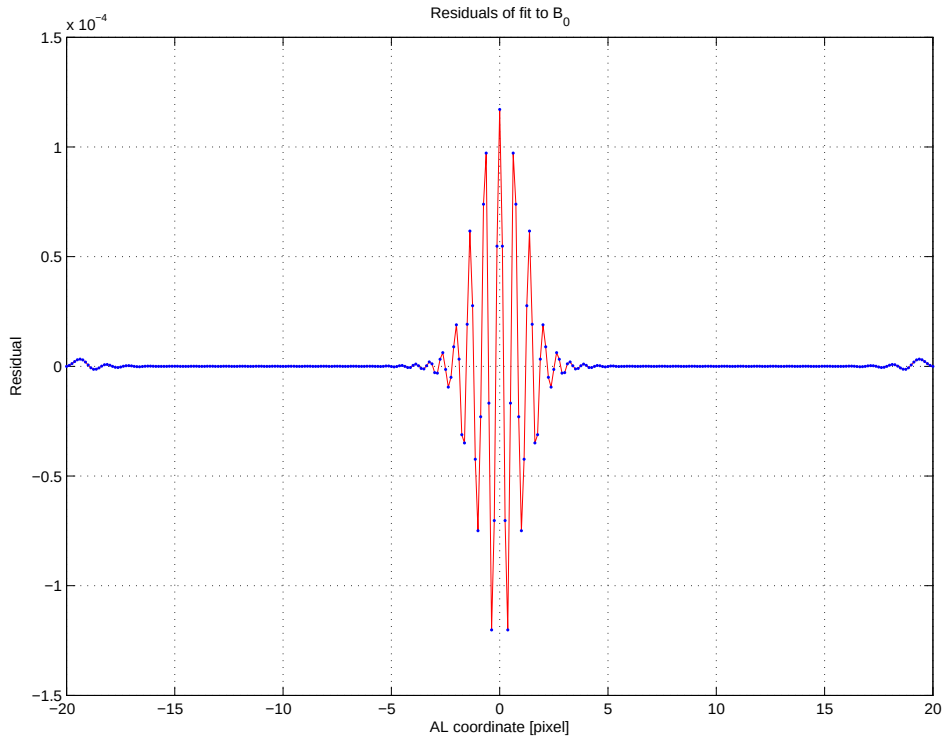

FIGURE 10: Residuals of the fit in Fig. 9. The residuals are in units of 10^{-4}.

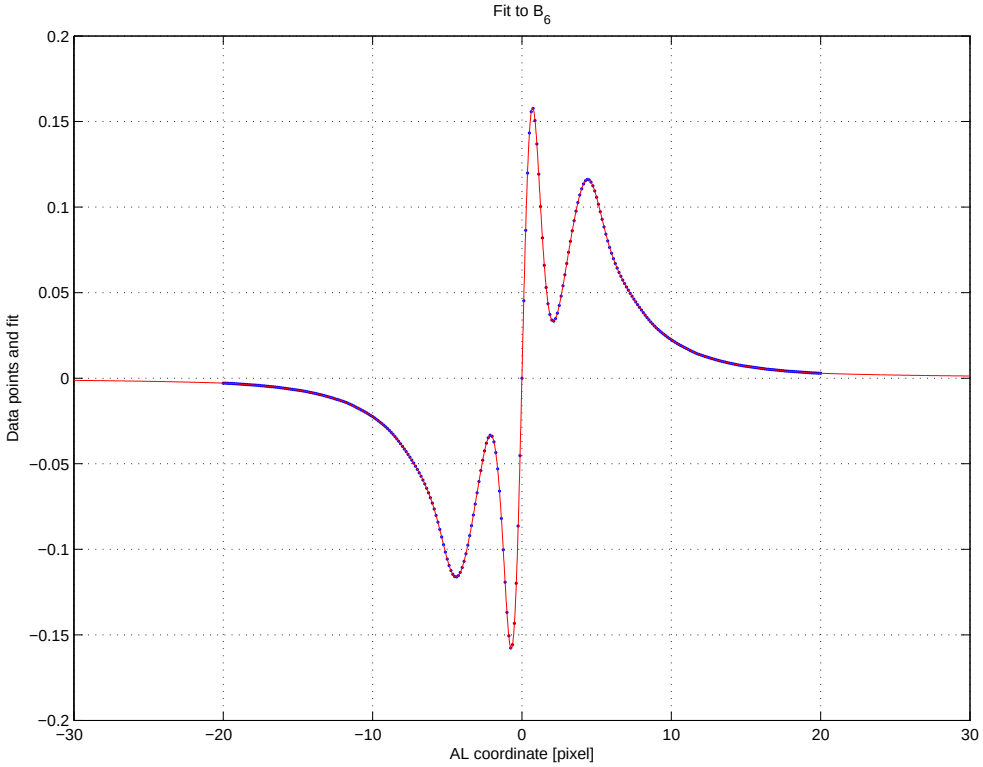

FIGURE 11: The dots show the numerically calculated basis function $B_6(u)$ for $|u| \leq 20$ pixel (i.e., $\boldsymbol{B}_6$), the line is the fitted approximation, Eq. (15). The residuals of the fit are shown in Fig. 12.

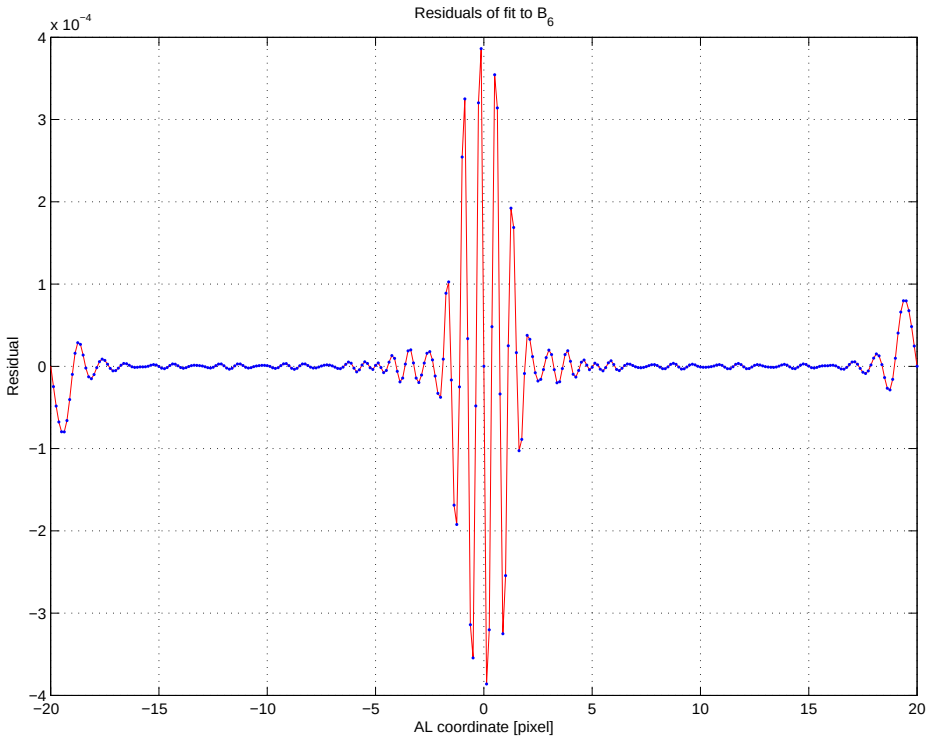

FIGURE 12: Residuals of the fit in Fig. 11. The residuals are in units of 10^{-4}.

6.2 Illustration of the spline+tail approximation

Fitting the approximation in Eq. (15) to the previously calculated basis functions with $m = 0, 1, \ldots, 12$ does not give a perfect fit, but certainly one that is good enough for the purpose of LSF modelling. Figure 9 shows the fit to $B_0(u)$, including the extrapolation to $u = \pm 25$ pixels; the residuals of the fit are shown in Fig. 10. The maximum residual is 1.2×10^{-4} and the RMS residual 2.2×10^{-5}.

Figure 11 shows the fit to the odd basis function $B_6(u)$, illustrating a case where the two tails have opposite signs. The residuals of the fit are shown in Fig. 12. The maximum residual in this case is 3.9×10^{-4} and the RMS residual 7.9×10^{-5}. However, these residuals should be multiplied by c_6 to give the corresponding errors in $\tilde{L}$. The RMS expectation of any of the coefficients c_m (for $m > 0$) is given by $\sigma_m^{1/2}$, where σ_m is the singular value (see Fig. 2 and Figs. 5–6). With $\sigma_6 = 0.000215$ we find that the expected RMS error in the LSF due to the spline approximation of $B_6(u)$ is about 10^{-6}. Considering all the basis functions with $m > 0$, the expected approximation error in $\tilde{L}$ is always at least a factor 4 smaller than the approximation error due to $B_0(u)$. Thus, the spline+tail approximation is accurate enough for all basis functions.

The residual plots in Figs. 10 and 12 show some wiggles at the end, which are due to a slight mismatch between the slope of $B_m(u)$ at $u = \pm 20$ pixel and the slope of the tail function. Since the wiggles are much smaller than the residuals in the LSF core, no attempt was made to model more accurately the transition from the spline to the tail function.

6.3 Application to Astrium test data

The Astrium Campaign #3 astrometric tests provide lots of realistic data on which the LSF model can be tested. Only one very limited test is reported here. It is based on the test data on the regular astrometric mask [1], using 20 scans (runs) of the first hole (lowest AC and AL index) across the non-irradiated zone of the CCD.[4] For each scan, a linear background was fitted to samples before and after the mask and subtracted from the samples. The provisional location of each hole was determined using Tukey's biweight as described in Sect. 4 of [5]. Then, 51 AL samples were extracted for each hole, centred on the sample nearest to the estimated location. The flux (estimated as the sum of the 51 samples) was normalized, and the standard error on each normalized sample calculated from the gain and readout noise values provided by Astrium. Using only data for the first of the 1100 holes thus resulted in $20 \times 51 = 1020$ samples covering about ± 25 pixels around the estimated location, with a fair distribution of sub-sample positions.

A weighted least-squares fitting of Eq. (1) to the data points was then made, using the spline+tail representation of the basis functions as described above. Only data within ± 10 pixels from the origin were used in the fitting. Separate fits were made with $N = 0 \ldots 12$ basis functions and the unit weight error $U = (\chi^2/\nu)^{1/2}$ was calculated in each case (χ^2 is the usual chi-square, $\nu = 1020 - N$ the degrees of freedom), see Table 1. The table suggests that some 10 basis functions are needed for a good representation of the data. An inspection of the residuals of the fit confirms this conclusion. Figure 13 shows the test data together with the fit on a logarithmic scale. The residuals are shown in Fig. 14, and the normalized residuals (residuals divided by their standard errors) in Fig. 15.

[4] The dataset used was 20080626/16h30_TDIMult_Mag15_ZNI_MaskCalibration.

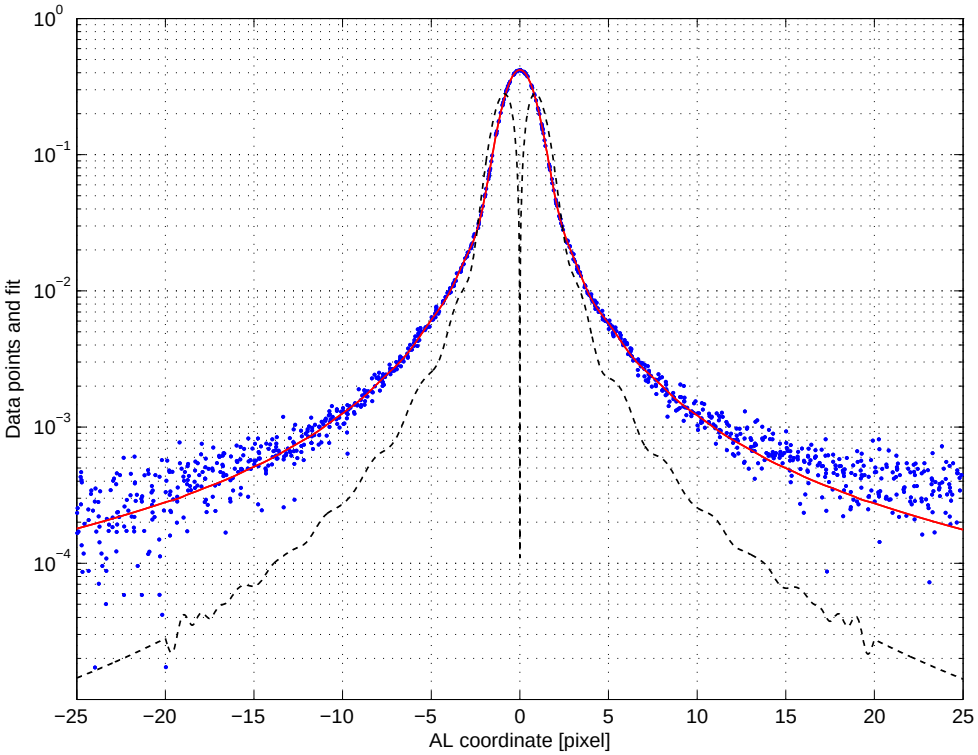

FIGURE 13: The LSF model with $N = 10$ degrees of freedom fitted to the test data. The dots are the data, the solid curve the fitted LSF model, the dashed curve the derivative of the fitted LSF (absolute value plotted).

The normalized residuals show no indication of any unmodelled component at least in the central part of the LSF (within $u = \pm 10$ pixels), where the fitting was done. In the leading wing (at $u \simeq -20$ to -15 pixels) there is a slight excess flux in the data, and in the trailing wing (for $u > 10$ pixels) the data are systematically higher than the model. The deviation in the trailing wing is probably due to the overlapping leading wing from the next hole (at $u \simeq 50$), which approximately doubles the flux at $u = 25$. Thus the model seems to be rather accurate even outside the fitted region (vertical lines in Figs. 14–15).

It is possible to improve the initial centroiding by least-squares adjustment of the initial LSF to each scan. This reduced U for $N = 10$ from 1.01566 to 1.01163. Iterating this procedure brought some further improvement (to about $U = 1.01075$), but did not seem to produce consistent results, for reasons that are not yet understood.

TABLE 1: Unit weight error (U) for fits of the LSF model to Astrium test data, as function of the number of basis functions used (N).

N	U	N	U	N	U
0	4.59788	5	1.42120	10	1.01566
1	4.56550	6	1.42043	11	1.01435
2	4.56245	7	1.41366	12	0.98923
3	1.51239	8	1.05858		
4	1.47346	9	1.04453		

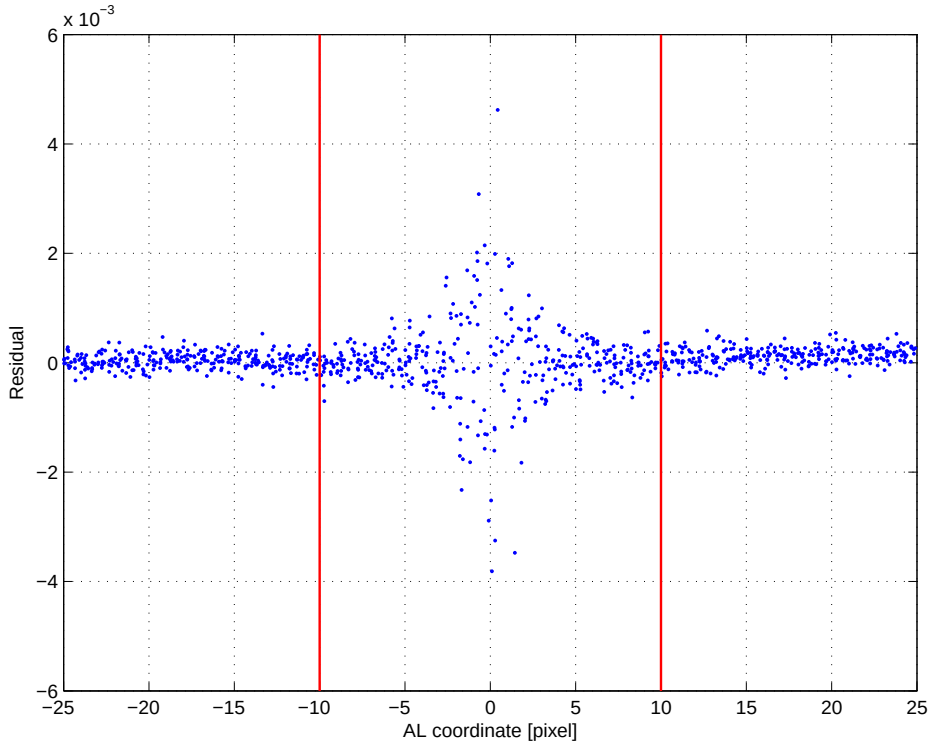

FIGURE 14: Residuals of the fit in Fig. 13, in units of 10^{-3}. Only the data between the thick vertical lines (at $u = \pm 10$ pixel) were used in the fitting. The increased spread of residuals near the origin of the AL coordinate is due to the higher level of photon noise in the central part of the image.

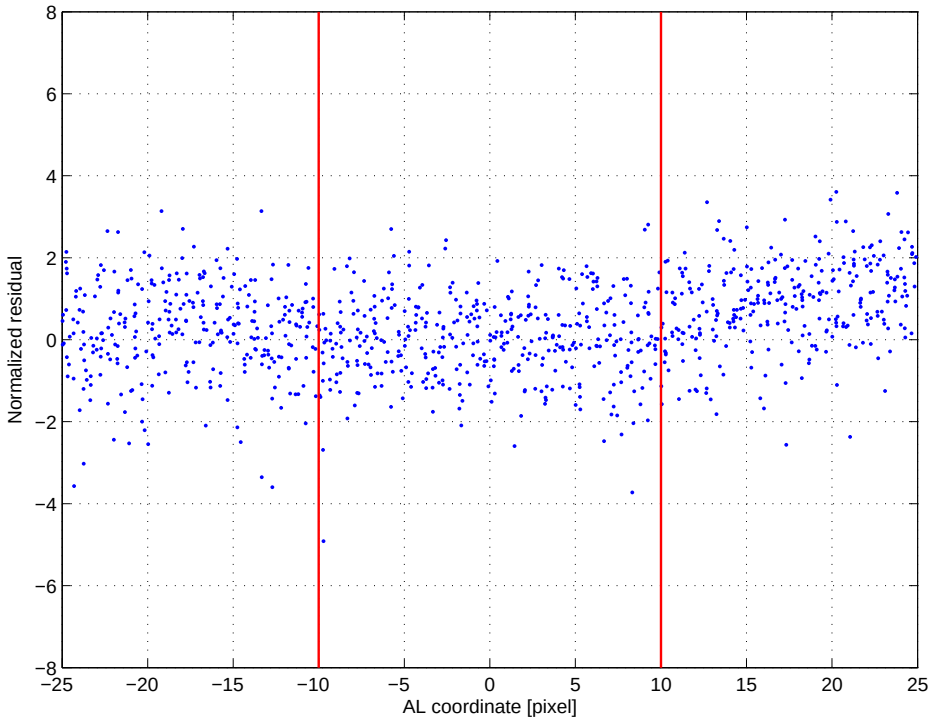

FIGURE 15: Normalized residuals of the fit in Fig. 13. Only the data between the thick vertical lines (at $u = \pm 10$ pixel) were used in the fitting.

7 Conclusions

The results of this study suggest that at least the LSF can be modelled to acceptable accuracy with a moderate (~ 10) number of free parameters. Such a model is also very flexible since it is readily adapted to the amount and quality of the data available. With more and better data available per SED bin, the number of fitted parameters (N) can be increased successively.

Using a spline+tail representation of the basis functions, the LSF is obtained in a form that allows simple and accurate interpolation and even a reasonably safe extrapolation beyond the fitted data.

Fitting the LSF model to Astrium Campaign #3 test data indicates that even the provisional set of basis functions derived in this note might work with the real data. Thus it is proposed to start using this LSF model in the IDT/IDU as soon as the planning allows.

This study should however also be extended somewhat, addressing in particular the following questions:

1. The basis functions should be evaluated for more realistic WFE maps, including in particular the high-frequency polishing errors, which may change the outer parts of the LSF in a systematic way. Also some parts of the theoretical PSF calculation could be improved (treatment of charge diffusion, pupil discretization).
2. The study should be extended to the two-dimensional PSF. In particular it should be investigated how accurate the 'outer product' model is ($P = L^{\mathrm{AL}} \times L^{\mathrm{AC}}$). E.g.: how many additional components are needed to reach different levels of accuracy?
3. The transition from spline to tail function can perhaps be improved. Although the residual wiggles are small in absolute size, they are potentially problematic if the model is used to correct data for a disturbing star about 20 pixels away.
4. Initial attempts at iterating the LSF fitting and location estimation did not seem to produce consistent results. This should be further investigated.

References

[1] *Radiation campaign #3 Astrometric regular data delivery*, GAIA.ASF.TCN.PLM.00389 (13 November 2008)

[2] Lindegren L., 2003: *Representation of LSF and PSF for GDAAS-2*, GAIA-LL-046 (2 May 2003)

[3] Lindegren L., 2005: *A theoretical investigation of chromaticity*, GAIA-C2-TN-LU-LL-064-1 (15 November 2005)

[4] Lindegren L., 2006: *PSF and LSF representation for the simulation of Gaia-3 Astro data*, GAIA-C2-TN-LU-LL-066-1 (28 April 2006)

[5] Lindegren L., 2006: *Centroid definition for the Astro Line Spread Function*, GAIA-C3-TN-LU-LL-068-1 (11 July 2006)

[6] Lindegren L., 2009: *A framework for consistent definition and use of LSFs in AF/BP/RP/RVS*, GAIA-C3-TN-LU-LL-080-02 (26 June 2009)