Centroid definition for the Astro Line Spread Function

L. Lindegren

GAIA–C3–TN–LU–LL–068-1 (11 July 2006)

Abstract. Issues related to the geometric and photometric calibration of the Astro Line Spread Function (LSF) are briefly discussed, in particular the separation of shape, position and flux, and the precise definition of the centroid. A convention for the centroid definition is proposed, based on a simple analytical weighting function, which appears close to optimal (for Gaia-3 parameters and typical WFE) from the viewpoint of statistical precision and interpolation accuracy. It is furthermore proposed that this weighting function is used for the very initial centroiding on AF samples, before the LSF has been calibrated.

1 Definition of the Line Spread Function

The Line Spread Function (LSF) is a very central concept for both the astrometric and photometric processing of Astro data. It provides the basis for the probabilistic modelling the sample data, and hence for any reasonably efficient estimators of the image location (needed for astrometry) and total flux in the image (needed for the photometry). Accurate calibration of the LSF is also needed in order disentangle partially overlapping images (as in double stars and crowded regions), and as a means to monitor the status of the instrument (in terms of wavefront aberrations, CTI distortion, etc).

In this document we are only concerned with the along-scan (AL) LSF as seen in the non-dispersive part of the instrument (SM and AF). Strictly speaking, this LSF is the marginal density of the Point Spread Function (PSF), integrated over an infinite range in the AC direction, but for the present discussion it will be sufficient to consider the LSF as the part of the PSF falling inside the observed window. To a certain approximation, the fraction of the PSF contained within a given finite range in AC (as defined by the window) is constant along the AL coordinate [3], and the variation of this factor with AC window size and decentering, wavelength, etc., is presently ignored.

The Line Spread Function $L(x)$ is a continuous function such that, for a point-like source of constant intensity on a uniform background, the expected number of photoelectron counts in sample k is given by

$$\mathrm{E}[n_k] = \alpha L(k - \xi) + \beta \tag{1}$$

Here, α is the total flux of the image (in electrons), β is the background level (in electrons per sample), and ξ is the location of the image centroid in the sample stream. The location parameter ξ is a real variable expressed in samples, i.e., using the same origin and unit

as k, but in general non-integer. For example, $\xi = 1862557.333$ would mean that the image centroid is located between samples $k = 1862557$ and $k = 1862558$, one third of a sample from the former. Knowing the datation of the samples and the gate used for the integration of the image, this value ξ can immediately be transformed into an observed CCD transit time.

The definition of the image parameters α, β and ξ are intimately linked to the definition of $L(x)$, in the sense that the latter must satisfy certain constraints to render the parameter values unique for a given observation. These constraints are:

- that the background β is the asymptotic intensity level far away from the image centre, i.e.
$$L(x) \to 0 \quad \text{for} \quad x \to \pm\infty \tag{2}$$
- that the flux parameter α equals the expected total number of counts in the image, i.e.
$$\int_{-\infty}^{\infty} L(x)\,\mathrm{d}x = 1 \tag{3}$$
- that the origin of x in $L(x)$ corresponds to a suitably defined centre of the LSF; we adopt the constraint
$$\int_{-\infty}^{\infty} L(x)w(x)\,\mathrm{d}x = 0 \tag{4}$$
where $w(x)$ is a *weighting function* to be defined below: this defines ξ as the location of the LSF centre among the samples.

For physical reasons, $L(x)$ must also be non-negative.

Of these constraints, only (4) is less than obvious and requires further specification. The problem arises because of the non-symmetric shape of the general LSF. While a symmetric LSF has an evident centre in its point of symmetry, such that $L(-x) = L(x)$ for all x, there is no such unambiguous choice for an asymmetric LSF. In fact, different proposals have been made in the past (e.g., [1, 2, 3, 4]), but their relative merits have not been assessed and it is now time to settle this question more definitely.

It should be noted that an 'obvious' centroid definition such as the centre of gravity, or first moment (mean) of the distribution, corresponding to the weighting function $w(x) = x$, is highly unsuitable for several reasons. First of all, there are mathematical difficulties in that the integral in (4) does not converge for the Fraunhofer diffraction of a rectangular pupil [5]. Even if this difficulty can be overcome, the resulting centroid is mainly sensitive to the wings of the image, which contribute little or nothing to the astrometric information.

A proposal for the practical choice of weighting function is given in Sect. 3. The numerical results shown below use LSF calculations based on Gaia-3 parameters (see also [6]) and two representative WFE maps provided by Astrium, denoted F32 (for AF5, CCD row 4) and F55 (AF1, CCD row 1). The RMS WFE, excluding piston and tilt, is 37.0 nm for F32 and 51.5 nm for F55. The resulting LSFs for an unreddened G2V star (from the Pickles library [7]) is shown in Fig. 1.

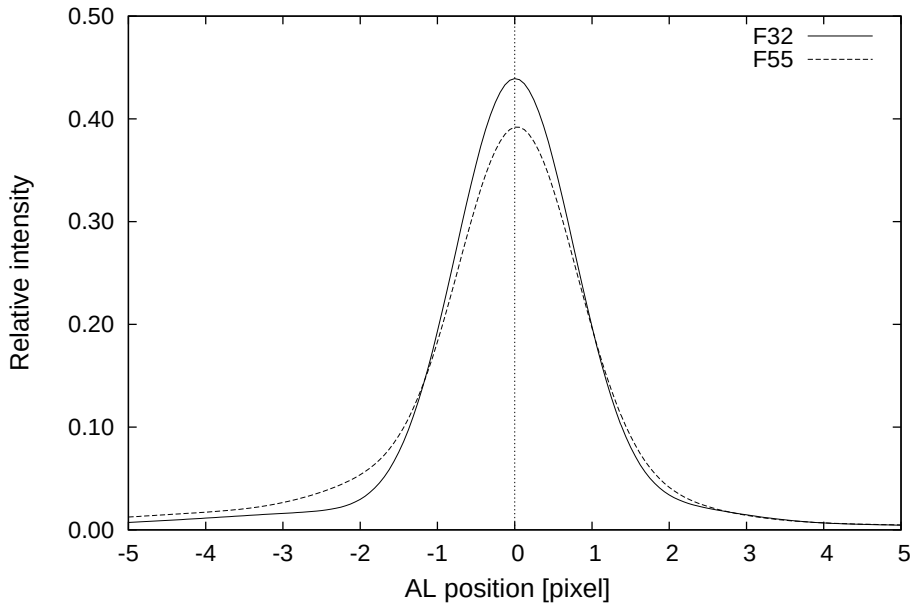

FIGURE 1: Polychromatic LSF calculated for an unreddened G2V star using Gaia-3 parameters and two representative Astrium WFE maps, F32 and F55.

2 Choice of weighting function

2.1 Candidate functions

Several candidate weighting functions are described and compared below. They are given as functions of the normalised argument $z = x/s$, where s is a scale parameter determining the width (in x) over which the function operates. Thus the centroid definition is actually

$$\int_{-\infty}^{\infty} L(x)w(x/s)\,\mathrm{d}x = 0 \tag{5}$$

The use of weighting functions for robust estimation of location is briefly discussed in *Numerical Recipes* [8], Ch. 15.7, from which two of the candidate functions are taken.

Gaussian derivative: this weighting function is proportional to the derivative of the centred normal probability density for unit standard deviation, $f(z) = (2\pi)^{-1/2}\exp(-z^2/2)$, namely

$$w(z) = z\exp(-z^2/2) \tag{6}$$

(Fig. 2). Using this weighting function corresponds to least-squares fitting the scaled normal probability function $f(z/s)s^{-1}$ to the LSF. $w(z)$ is approximately linear for small $|z|$ and has its extreme values at $z = \pm 1$, after which it gently decreases to zero. Although the support is in principle infinite, the function can in practice be regarded as zero for

$|z| \gtrsim 4$. This weighting function has been used (with $s \simeq 1$ pixel) in several investigations of chromaticity.

In contrast to the gaussian derivative, the following weighting functions all have limited support, and the scale parameter is defined such that the functions are strictly zero for $|z| > 1$. These weighting functions are depicted in Fig. 3. It can be seen that they represent, to various degrees, an approximation of the gaussian derivative for a finite support.

Truncated mean:

$$w(z) = \begin{cases} z & \text{if } |z| < 1 \\ 0 & \text{otherwise} \end{cases} \tag{7}$$

Truncated median:

$$w(z) = \begin{cases} -1 & \text{if } -1 < z < 0 \\ +1 & \text{if } 0 < z < +1 \\ 0 & \text{otherwise} \end{cases} \tag{8}$$

Andrew's sine:

$$w(z) = \begin{cases} \sin(\pi z) & \text{if } |z| < 1 \\ 0 & \text{otherwise} \end{cases} \tag{9}$$

This is adapted from [8]. The extreme values are at $z = \pm 1/2$.

Tukey's biweight:

$$w(z) = \begin{cases} z(1-z^2)^2 & \text{if } |z| < 1 \\ 0 & \text{otherwise} \end{cases} \tag{10}$$

This is adapted from [8]. The extreme values are at $z = 1/\sqrt{5} \simeq 0.447$.

Cubic spline:

$$w(z) = \begin{cases} 2z(1-3z^2) & \text{if } |z| \le 1/3 \\ 3(1-z)^3 - 12(2/3-z)^3 & \text{if } 1/3 < |z| \le 2/3 \\ 3(1-z)^3 & \text{if } 2/3 < |z| \le 1 \\ 0 & \text{otherwise} \end{cases} \tag{11}$$

This is the cubic spline uniquely defined (up to a constant of proportionality) by the following conditions: the knots are at $z = 0,\ \pm 1/3,\ \pm 2/3,\ \pm 1$; it is an odd function; the extreme values are at $z = \pm 1/3$; and the function and its first and second derivatives are everywhere continuous.[1]

[1] As it happens, this weighting function (scaled to $s = 3/2$) is proportional to the derivative of the bi-quartic spline defined in [3]. No special significance should be attached to this coincidence.

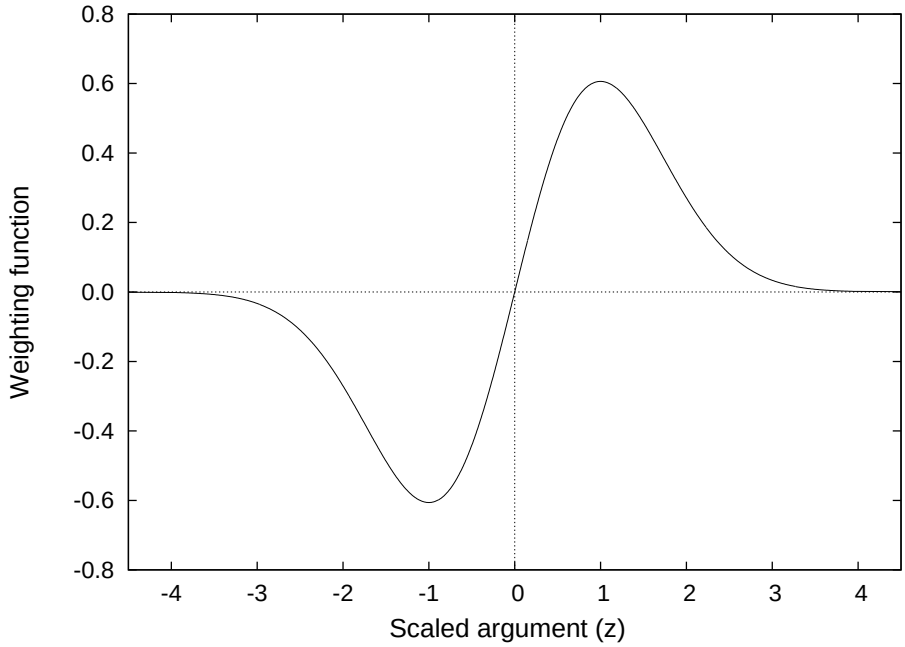

FIGURE 2: Weighting function $w(z)$ of the 'gaussian derivative' type, Eq. (6).

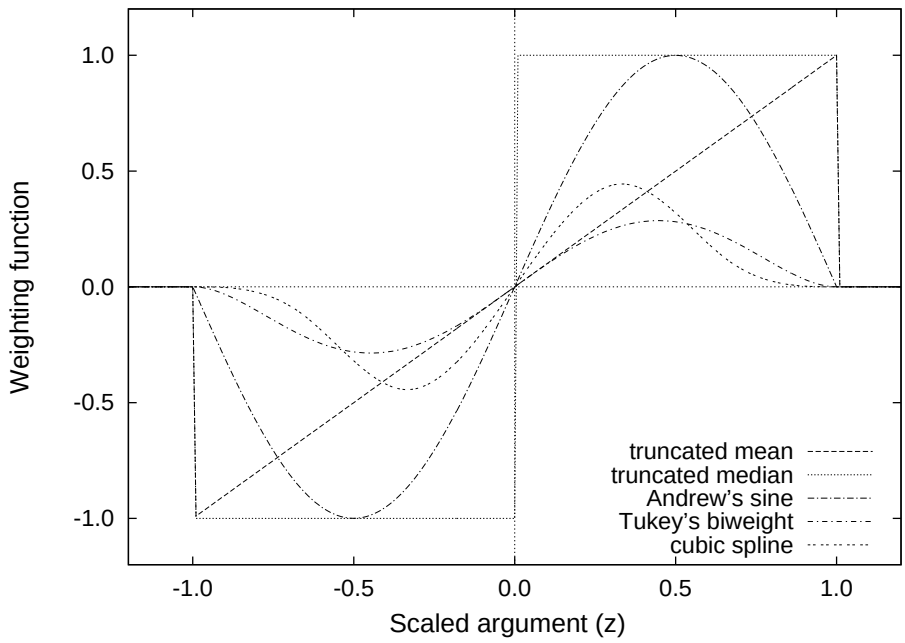

FIGURE 3: Five considered weighting functions $w(z)$ with finite support; see Eqs. (7)–(11).

2.2 Criteria for selecting the 'best' weighting function

Inasmuch as the definition of the LSF centre is simply a matter of convention, any of the above weighting functions might be acceptable, with a wide range of choices for the scale parameter s. However, in practice the definition must be applied to a set of data values representing a finite portion of the LSF, to finite precision, and it is highly desirable that the definition is not overly sensitive to such limitations.

Moreover, with a sensible choice of weighting function and scale parameter s, the initial centroiding on the samples $\{n_k\}$ could be usefully done by direct application of the centroid definition, i.e., by solving the equation

$$\sum_k n_k \, w\left(\frac{k-\xi}{s}\right) = 0 \tag{12}$$

The resulting centroid estimate $\hat{\xi}$ is in general less precise than what a full LSF fitting will give, but has the significant advantage that the LSF need not be calibrated before doing the centroiding. However, the limited size of the AF windows (normally six samples) puts a constraint on the support of $w(z)$. Assuming that the centroid is located between the 3rd and 4th sample of the window, it is seen that (12) can only be correctly solved from the six available samples if $w(x/s) = 0$ for $|x| > 3$ pixels. For the gaussian derivative this implies $s \lesssim 0.75$ pixel, while for the other weighting functions it implies $s \leq 3$ pixel.

When (12) is used to centre of real samples, we should be concerned both with the statistical precision of the resulting estimate, caused by the Poissonian nature of the counts and the additional readout noise, as well as systematic errors depending on the sub-pixel position of the centroid. For given LSFs, such as the two representative cases shown in Fig. 1, these errors can easily be evaluated for the different weighting functions.

The choice of weighting function has a strong impact on the size and behaviour of the chromaticity effect [5]. This can be studied, at least in a first approximation, by considering the variation, as function of wavelength, of the centroid location for the monochromatic LSF. A small variation is preferable to a large variation, and a linear variation is better – or at least more easily calibrated – than a strongly non-linear one.

In summary, the criteria that will be used to select the 'best' weighting function and scale parameter are:

1. Length of support: the weighting function should be effectively zero for $|x| > 3$ pixel in order to be useful for centering on windowed samples
2. Statistical precision of centroiding on real (noisy) samples
3. Systematic errors (depending on sub-pixel position) when centering on real samples
4. Total size of chromaticity
5. Non-linearity of chromaticity

2.3 Comparison of candidate weighting functions

2.3.1 Length of support

As we have seen, all the candidate weighting functions satisfy this criterion provided that the scale parameter is small enough. For the gaussian derivative (6) we require $s \lesssim 0.75$, while for the other weighting functions (7)–(11), $s \leq 3$ is required.

2.3.2 Statistical precision

The expected statistical precision of $\hat{\xi}$ obtained by solving (12) can be evaluated as

$$\sigma_\xi \propto \left(\int w(x/s)^2 V(x)\,\mathrm{d}x \right)^{1/2} \left| \frac{1}{s} \int w'(x/s) L(x)\,\mathrm{d}x \right|^{-1} \tag{13}$$

where $V(x)$ gives the total noise variance in the image. Two extreme cases are considered: In the signal-dominated case, $V(x) \propto L(x)$ due to the photon noise. In the background-dominated case, $V(x) \simeq$ constant. The relative performances of the different weighting functions versus s are shown in Figs. (4)–(5) for the signal-dominated case, and in Figs. (6)–(7) for the background-dominated case. Note that it will be necessary to select a weighting function and an s value that performs well both in the signal-dominated and background-dominated case.

The curves in Figs. (4)–(7) are very similar for F32 and F55, so this criterion is relatively insensitive to the detailed WFE.

For the gaussian derivative, the maximum scale parameter compatible with criterion 1 ($s = 0.75$ pix) should be selected for precision.

The truncated mean and median perform significantly worse in the background-dominated case than the other weighting functions. The cubic spline tends to require a scale parameter greater than 3 pixels in the signal-dominated case – in other words, it does not decline fast enough to zero.

The best weighting functions, according to this criterion, are Andrew's sine and Tukey's biweight. They have similar performances both in the signal-dominated and background-dominated cases, and ideally require $s \simeq 2$–2.5 as a compromise between the two cases.

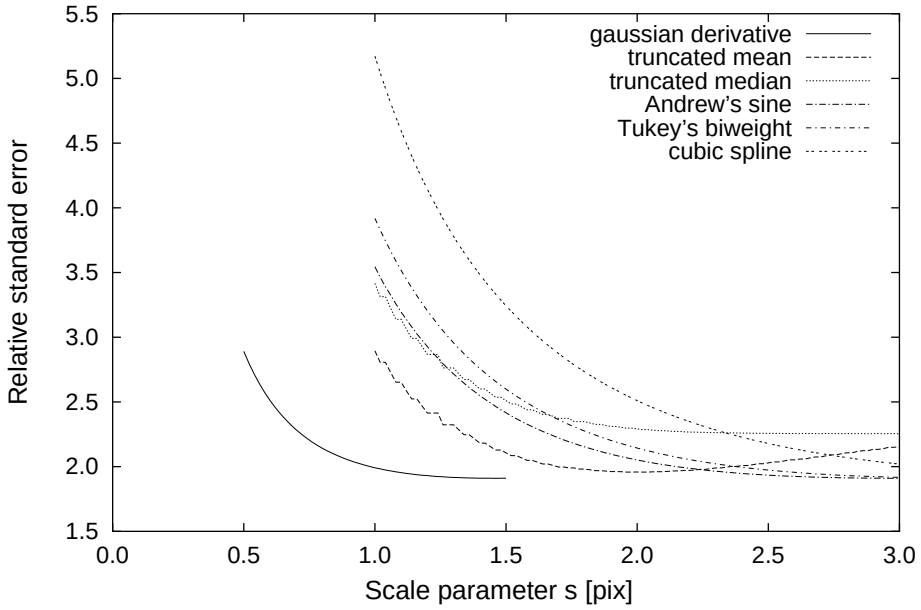

FIGURE 4: Relative precision of the location estimate $\hat{\xi}$ from Eq. (12) in the case of strong signals (dominated by the Poisson noise of the star image), evaluated for WFE map F32.

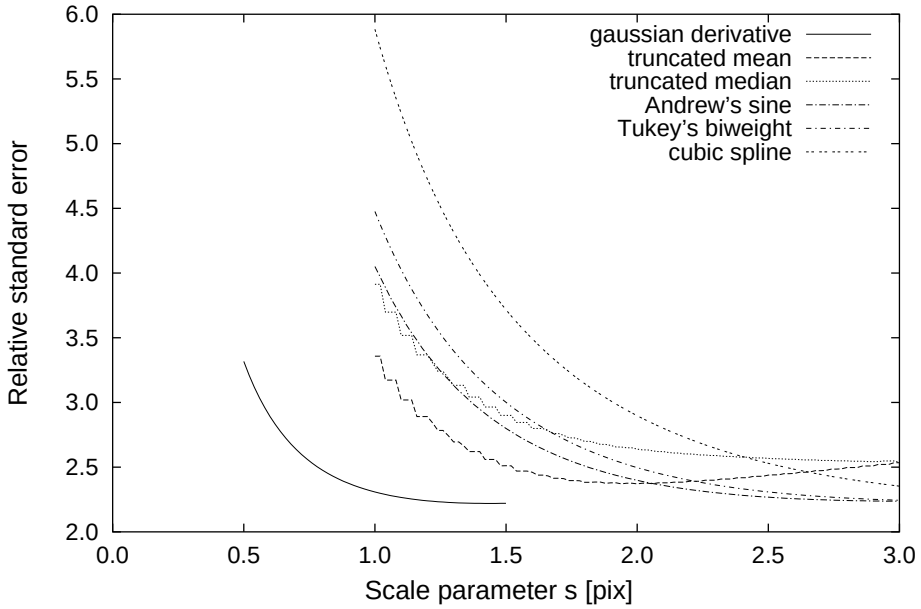

FIGURE 5: Relative precision of the location estimate $\hat{\xi}$ from Eq. (12) in the case of strong signals (dominated by the Poisson noise of the star image), evaluated for WFE map F55.

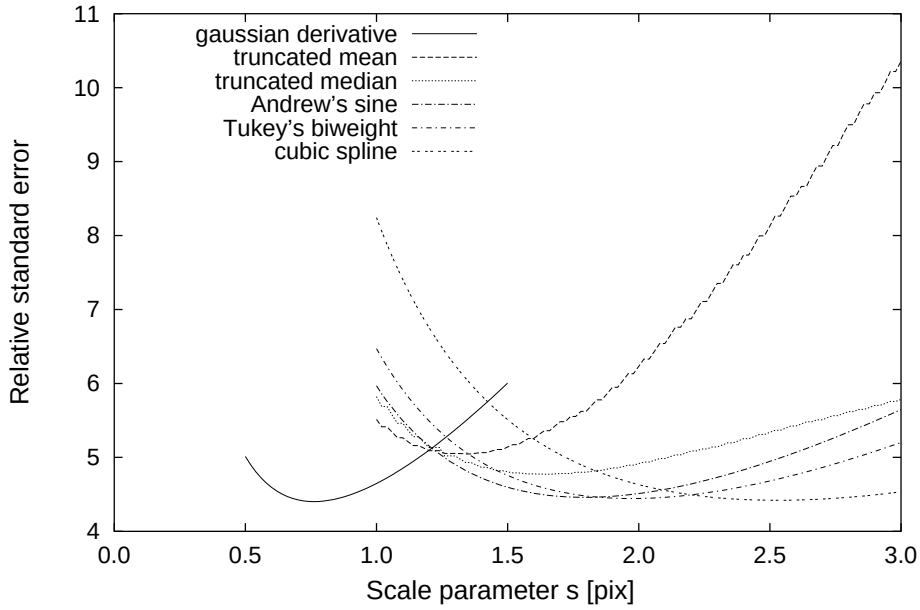

FIGURE 6: Relative precision of the location estimate $\hat{\xi}$ from Eq. (12) in the case of weak signals (dominated by the readout noise and/or Poisson noise of background), evaluated for WFE map F32.

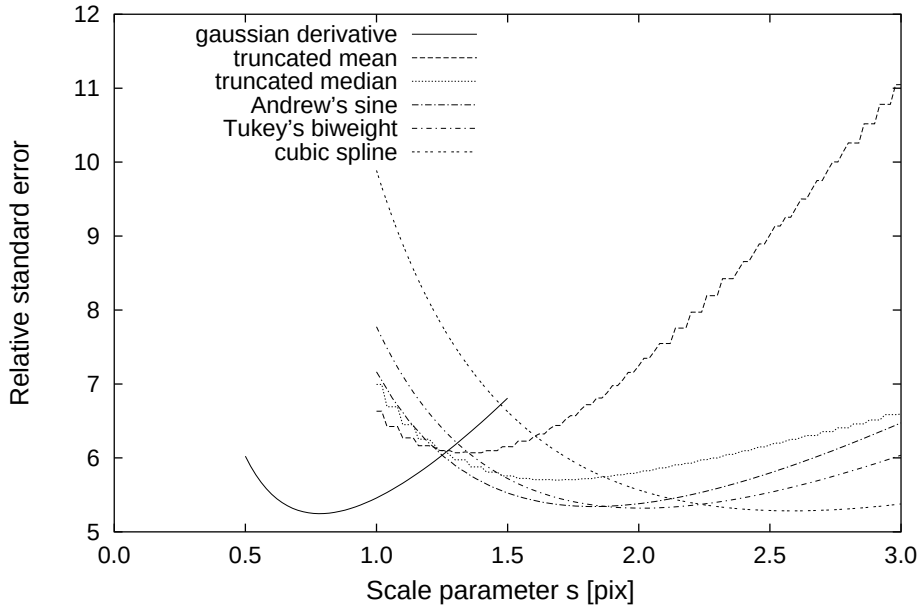

FIGURE 7: Relative precision of the location estimate $\hat{\xi}$ from Eq. (12) in the case of weak signals (dominated by the readout noise and/or Poisson noise of background), evaluated for WFE map F55.

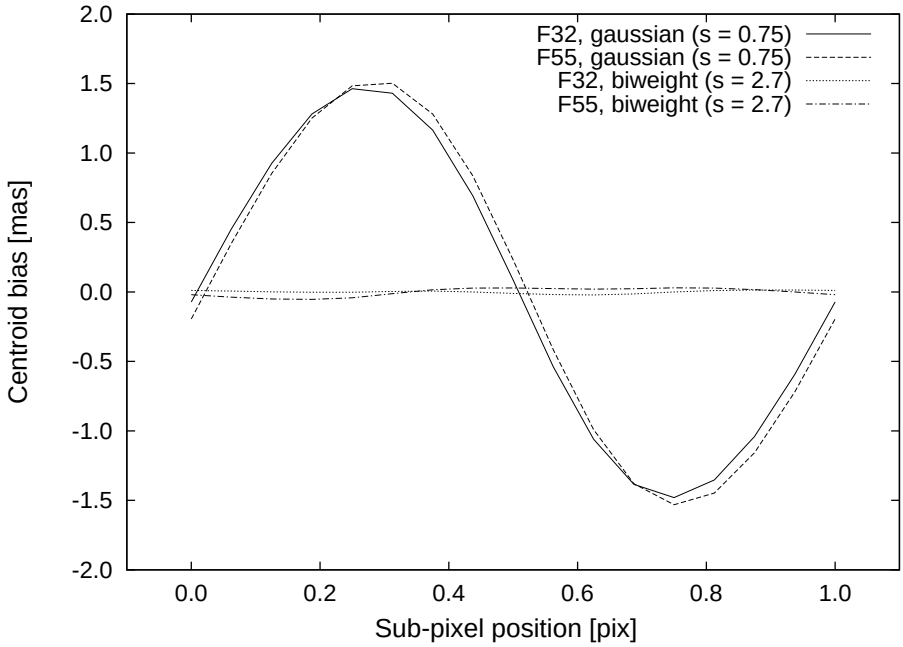

FIGURE 8: Bias in the location estimate, as function of sub-pixel position, when Eq. (12) is used to calculate the centroid of sampled data. Curves are shown for the two different WFE maps (F32 and F55) and for two selected weighting functions: the gaussian derivative with $s = 0.75$, and Tukey's biweight with $s = 2.7$.

2.3.3 Systematic errors

The bias resulting from using (12) to centre on real samples can be evaluated by solving ξ in

$$\sum_k L(k - \xi_0)\, w\left(\frac{k - \xi}{s}\right) = 0 \tag{14}$$

where ξ_0 is the true centroid, and plotting the bias $\xi - \xi_0$ as function of the fractional part of ξ_0. Fig. 8 shows a few examples of the resulting bias curve.

The RMS value of the bias was evaluated for the different candidate weighting functions and a range of s values. Results are shown in Figs. 9 and 10. Note that for the gaussian derivative (solid curve) the largest value of s is 0.75, at which point this weighting function performs much worse than several of the others at their appropriate s values.

It appears that Tukey's biweight, with $s \simeq 2.7$, is particularly good for both WFE maps, with RMS bias values in the 10–25 μas range. A test with a third WFE map verified that this was no coincidence. The low-amplitude curves in Fig. 8 show the merit of this weighting function in striking contrast to the gaussian derivative.

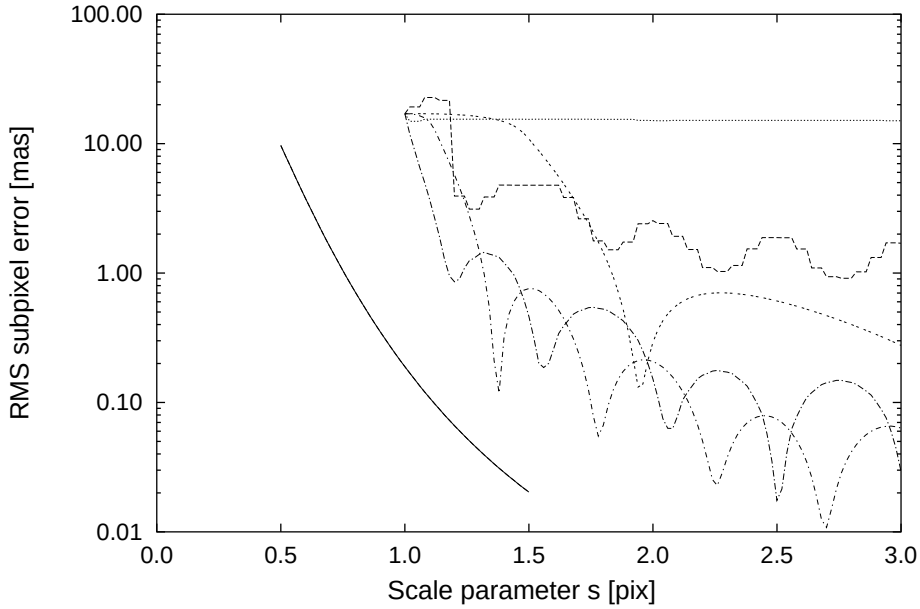

FIGURE 9: RMS bias in the centroid position when the different weighting functions are used to centre on data samples separated by one pixel (WFE map F32). The line types represent the different weighting functions, using the same keys as in Figs. 4–7.

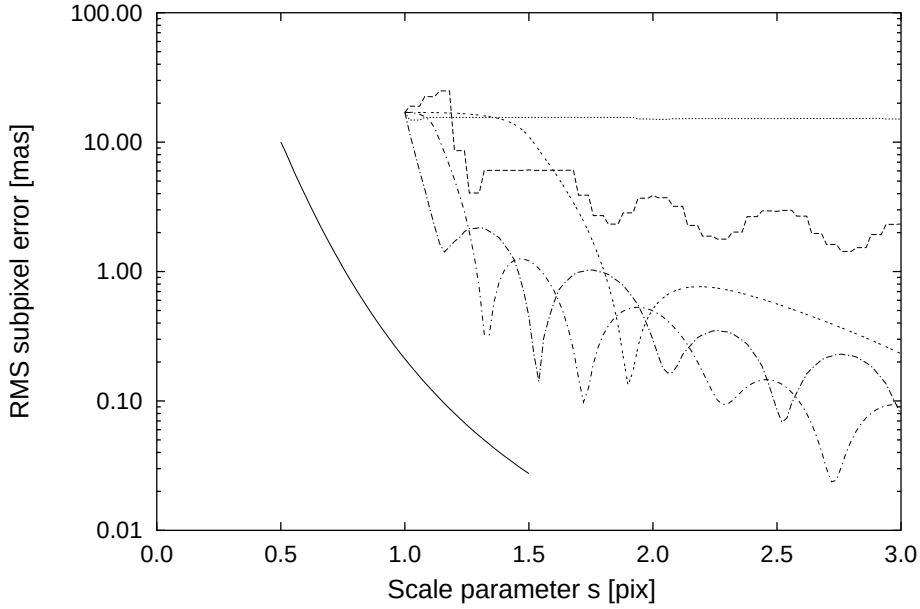

FIGURE 10: RMS bias in the centroid position when the different weighting functions are used to centre on data samples separated by one pixel (WFE map F55). The line types represent the different weighting functions, using the same keys as in Figs. 4–7.

2.3.4 Total size of chromaticity

Figures 11 (for F32) and 12 (for F55) are shown as examples of how the monochromatic centroid position may vary with wavelength. The centroid positions are plotted against the inverse wavelength, since this relation tends to be more linear than if wavelength is used. The displayed relations are for Tukey's biweight function, but the plots for other weighting functions are qualitatively similar for the WFE maps considered here. For other WFE maps, the curves may however look completely different, as suggested already by the considerable difference between the present two maps.

As a measure of the total size of the chromaticity, the RMS variation of the monochromatic centroid position for the wavelength range 300 to 1050 nm is shown in Figs. 13–14 for the two WFE maps. From these plots it is hardly possible to draw any conclusions about the relative merits of the weighting functions, except that the truncated mean is perhaps less suitable. For the other weighting functions, the dependence on s is almost opposite between the two WFE maps, so that a good choice of s in the case of F55, for example, may be practically the worst choice for F32.

2.3.5 Non-linearity of chromaticity

As a measure of the non-linearity, we use the RMS residual of a linear regression of ξ versus λ^{-1}, i.e., the RMS deviation from a straight line in plot like Figs. 11–12, taken over the whole wavelength range 300–1050 nm. Results are shown in Figs. 15–16 for the two WFE maps. The conclusions are similar as for the total chromaticity: the truncated mean is less suitable, but otherwise no clear preference can be seen due to the very different behaviour for the two WFE maps.

2.4 Conclusion on the choice of weighting function

It is clear that some properties of the weighting functions depend so critically on the actual WFE map that they cannot be used as a criterion for choosing between the functions. This applies in particular to the behaviour with respect to chromaticity, both its total size and non-linearity.

Other properties, such as the precision and sub-pixel systematics when the weighting function is used to centre on real samples, are on the other hand quite consistent and do not depend critically on the WFE. Consequently, the choice of weighting function must primarily be guided by these properties.

With respect to sub-pixel systematics, there is a very clear and strong preference for Tukey's biweight with $s = 2.7$ pix (Figs. 9–10). This choice is very good also with respect to precision, especially in the signal-dominated case (Figs. 4–5); in the background-dominated case (Figs. 6–7), the 7–10% increase in standard error seems an acceptable sacrifice.[2]

[2]Even at $G = 20$, the three central samples of the image (at $\sim 150,\ 350,\ 150\ \mathrm{e}^-$) are usually dominated by the Poisson noise of the image, rather than the background and readout noise; thus the background-dominated case actually applies only in exceptional circumstances.

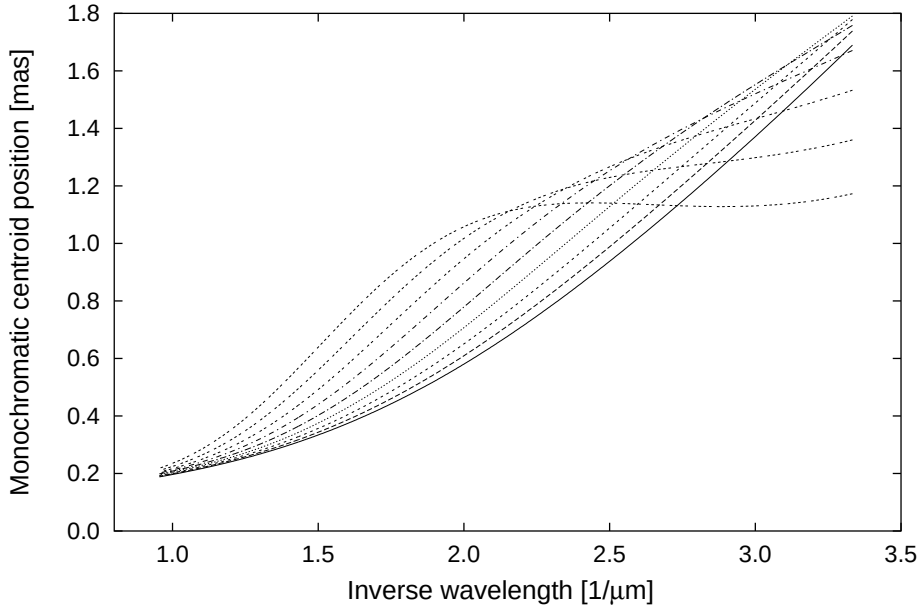

FIGURE 11: The monochromatic centroid position for WFE map F32, using Tukey's biweight, as function of the inverse wavelength. The curves are for different values of the scale parameter s, starting with $s = 1.0$ for the lowest, solid curve; $s = 1.25$ for the next, dashed curve, etc; and ending with $s = 3.0$ for the topmost curve.

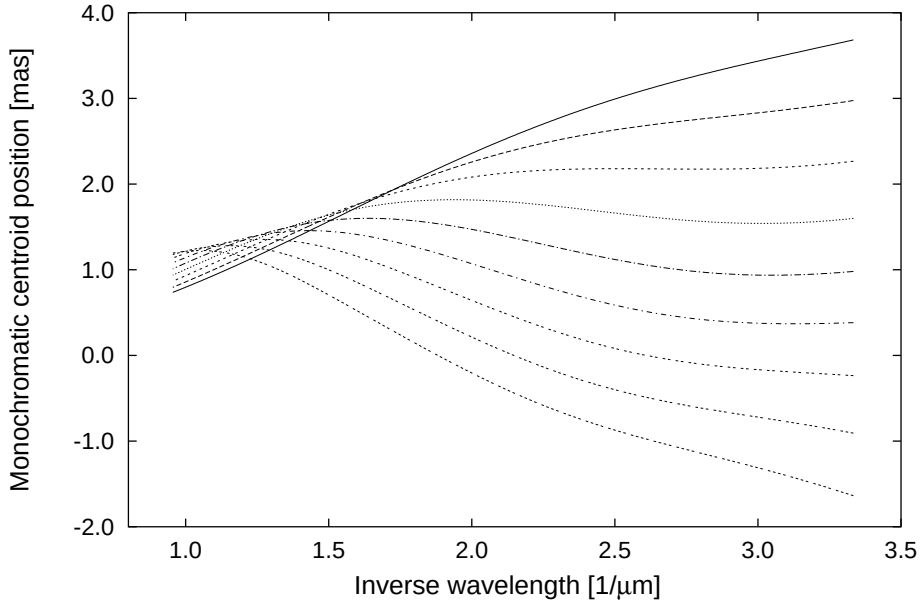

FIGURE 12: The monochromatic centroid position for WFE map F55, using Tukey's biweight, as function of the inverse wavelength. The curves are for different values of the scale parameter s, starting with $s = 1.0$ for the upper, solid curve; $s = 1.25$ for the next, dashed curve, etc; and ending with $s = 3.0$ for the lowest curve.

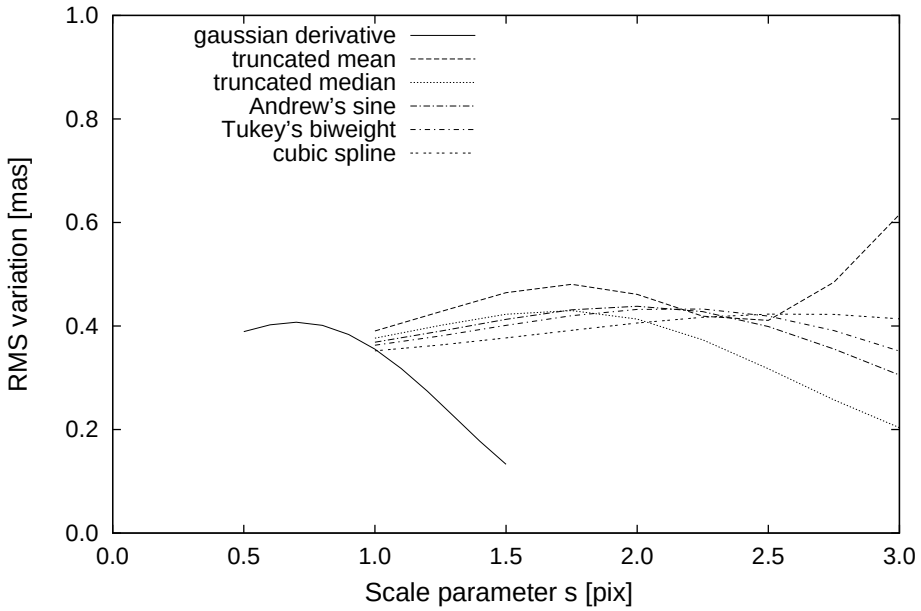

FIGURE 13: RMS variation of the monochromatic centroid position in the wavelength range 300–1050 nm for WFE map F32.

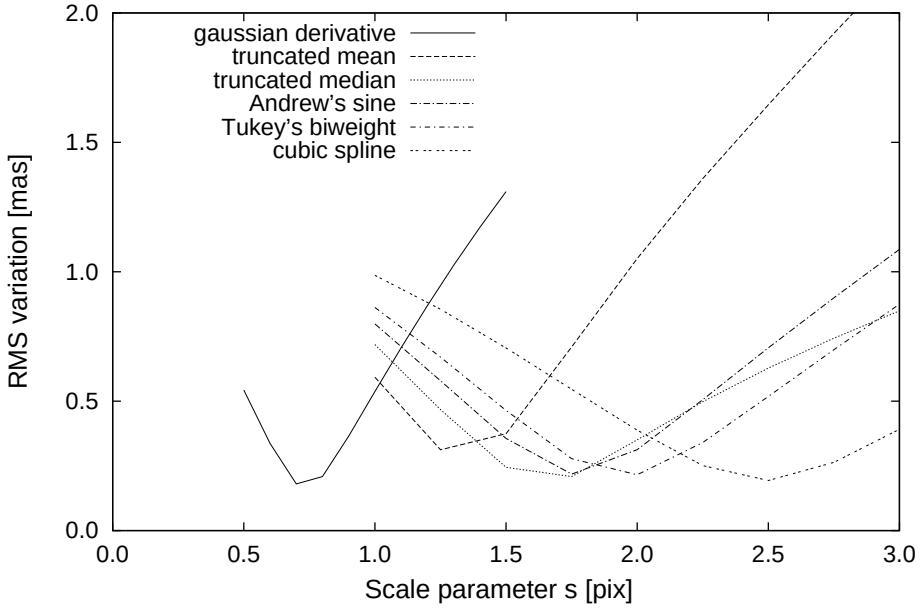

FIGURE 14: RMS variation of the monochromatic centroid position in the wavelength range 300–1050 nm for WFE map F55.

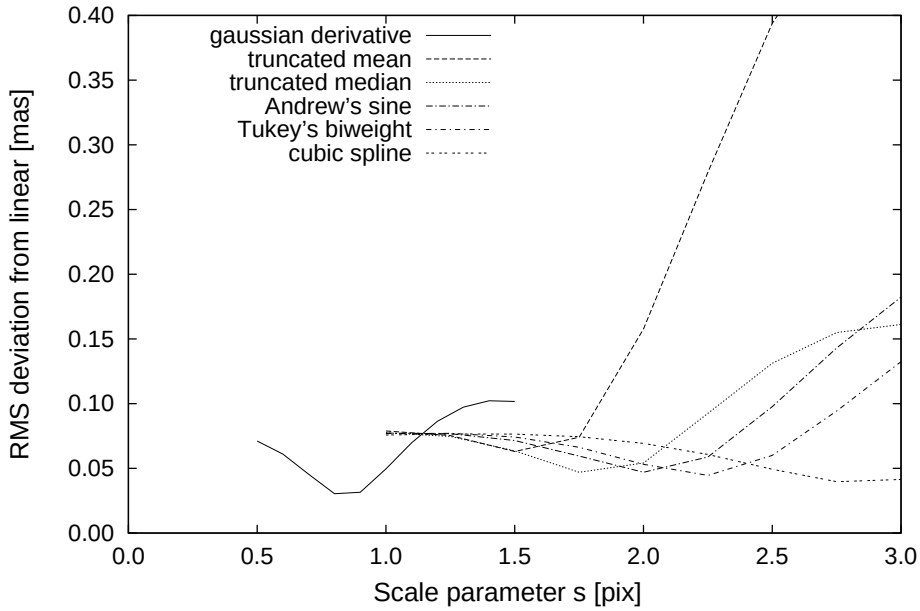

FIGURE 15: RMS deviation of the monochromatic centroid position from a linear variation with λ^{-1}, calculated in the wavelength range 300–1050 nm for WFE map F32.

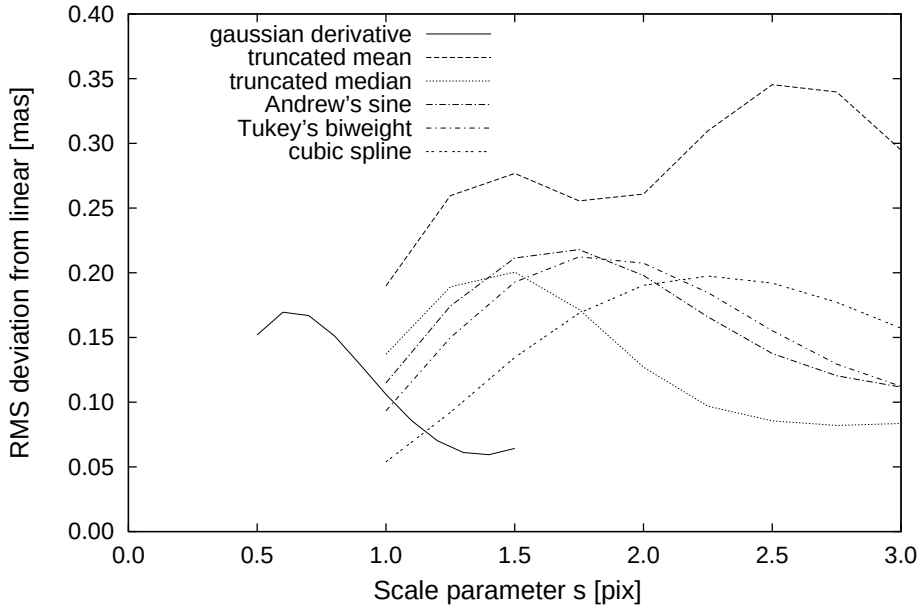

FIGURE 16: RMS deviation of the monochromatic centroid position from a linear variation with λ^{-1}, calculated in the wavelength range 300–1050 nm for WFE map F55.

3 Proposed convention

It is proposed that the centre (origin) of the Astro LSF $L(x)$ (where x is expressed in pixels relative to the origin) is defined by the condition

$$\int_{-\infty}^{\infty} L(x) w(x/s)\,\mathrm{d}x = 0 \tag{15}$$

Here, w is the following weighting function (Tukey's biweight)

$$w(z) = \begin{cases} z(1-z^2)^2 & \text{if } |z| < 1 \\ 0 & \text{otherwise} \end{cases} \tag{16}$$

and s is a scale parameter set to the fixed value $s = 2.7$ pixel.

4 Proposal for the initial centroiding of AF samples

It is furthermore proposed that the above weighting function (with $s = 2.7$) is also used to compute an initial estimate of the location of the LSF centroid among the observed AF samples $\{n_k\}$, by means of the equation

$$\sum_k n_k\, w\left(\frac{k-\xi}{s}\right) = 0 \tag{17}$$

This has the advantage that it can be applied before any LSF calibration is available, e.g., very early in the initial data treatment, and as an initial step of the LSF calibration process. It is however less accurate than a full LSF fitting, and cannot replace the latter in subsequent processing stages.

Equation (17) is non-linear but can be solved in a few Newton–Raphson iterations from a starting value of ξ equal to the index k of the largest n_k. The iteration formula is

$$\xi \leftarrow \xi + s\,\frac{\sum_k n_k\, w\left(\frac{k-\xi}{s}\right)}{\sum_k n_k\, w'\left(\frac{k-\xi}{s}\right)} \tag{18}$$

where

$$w'(z) = \begin{cases} (1-5z^2)(1-z^2) & \text{if } |z| < 1 \\ 0 & \text{otherwise} \end{cases} \tag{19}$$

An approximate error estimate is obtained as

$$\sigma_\xi = s\,\frac{\left[\sum_k V_k\, w\left(\frac{k-\xi}{s}\right)^2\right]^{1/2}}{\left|\sum_k n_k\, w'\left(\frac{k-\xi}{s}\right)\right|} \tag{20}$$

where $V_k = n_k + 1 + r^2$ is the estimated variance of n_k, with r the RMS readout noise in electrons.

References

[1] Lindegren L., 1998: *Calculation of chromatic displacement*, SAG–LL–024, (25 October 1998)

[2] Lindegren L., with input from GST and WG members, 2003: *Algorithms for GDAAS Phase II: Definition*, GAIA–LL–044, Version 4 (21 March 2003); see p. 66

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

[4] Lindegren L., 2004: *Chromaticity specification*, GAIA–LL–053, Version 1 (8 May 2004)

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

[6] 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)

[7] Pickles A.J., 1998: *A Stellar Spectral Flux Library: 1150–25000 Å*, PASP 110, 863

[8] Press W.H., Flannery B.P., Teukolsky S.A., Vetterling W.T., 1992: *Numerical recipes in Fortran*, Cambridge University Press (free on-line version available at http://www.nr.com/)

Acronyms

The following table has been generated from the on-line Gaia acronym list:

Acronym	**Description**
AC	ACross scan (direction)
AF	Astrometric Field (in Astro)
AL	ALong scan (direction)
CCD	Charge-Coupled Device
CTI	Charge Transfer Inefficiency
GDAAS	Gaia Data Access and Analysis Study
GST	Gaia Science Team
LSF	Line-Spread Function
PSF	Point-Spread Function
RMS	Root-Mean-Square
SM	Sky Mapper
WFE	WaveFront Error
WG	Working Group