XP LSF modelling

prepared by: Montegriffo, P.
approved by: F. van Leeuwen
authorized by: M.G. Lattanzi
reference: GAIA-C5-TN-OABO-PMN-012-1
issue: 1
revision: 5
date: 2017-09-05
status: Issued

Abstract

This note describes the parameterisation of the monochromatic LSF as function of AL position and wavelength used in the external calibration model for XP instruments. The model is based on a set of 2D basis functions derived with generalised principal component analysis of a large set of synthetic PSF/LSF.

Document History

Issue	Revision	Date	Author	Comment
I	5	2017-09-05	FVL	Final corrections.
D	4	2017-05-17	PMN	Comments by CC.
D	4	2017-04-15	PMN	Added introduction and Sect.2.
D	3	2017-04-12	PMN	Added data model section.
D	2	2017-03-31	PMN	Completed up to Mean LSF representation section.
D	1	2017-03-24	PMN	Completed up to Basis Functions Interpolation section.
D	0	2017-03-17	PMN	First draft

Contents

1 Introduction **4**

2 The LSF-084 scheme **4**

3 PSF modelling and wavelength sampling **5**

3.1 Simulations for current analysis 7

3.2 LSF normalisation 9

4 Generalised Principal Component Analysis (GPCA) **12**

4.1 LSF modelling and GPCA 13

5 Basis Functions interpolation **20**

5.1 Left Basis Functions 20

5.2 Right Basis Functions 23

6 Mean LSF representation **23**

7 Data Model **26**

8 Acknowledgements **27**

9 References **27**

.1 Acronyms 28

1 Introduction

The external calibration model for XP spectra has been described in Montegriffo (PMN-007). However the details related to the LSF model therein contained were never properly described in a dedicated technical note. That model was based on a parameterisation done through a set of 2D basis functions that were assembled starting from 1D basis functions derived with Principal Component Analysis (PCA) on a large set of theoretically computed monochromatic PSFs: PCA was run several times on PSF sets at different wavelengths and then, using a complex interpolation scheme, a grid of basis function sets for monochromatic LSFs corresponding to wavelengths ranging from 300 to 1100 nm with 1 nm spacing was finally produced (some more details can be found on the presentation given at the eleventh CU5 plenary meeting held in Bologna on April 2012, see *link*). However, to properly model the variation of the LSF with wavelength it was necessary to introduce an interpolation scheme based on CF-015 suggestion, where the coefficients multiplying the basis functions were modelled by low order polynomials, as can be seen in Eq. 6 in PMN-007. When we started the review of the whole procedure for the present technical note we discovered few subtle flaws in the PSF simulation algorithm that left their sign into the basis functions; moreover the implementation of the model into PhotPipe resulted into excessively bulky code that lead us to the decision of rethinking the whole implant of the model. The two main consequences were the recalculation of the whole set of random PSF/LSF on which the basis functions are built, and the implementation in MATLAB of the Generalised Principal Component Analysis (GPCA) algorithm instead of PCA to derive our model.

2 The LSF-084 scheme

In a series of studies (LL-084, LL-088) Lindegren proposes a minimum-dimension LSF model suitable to describe AF polychromatic LSFs: the present technical note follows closely the protocol defined by these works changing only few aspects to better fit our goals, e.g. building up a proper model for XP monochromatic LSF.

Lindegren (LL-084) model the LSF as a linear combination of basis functions that have been derived using PCA of a large number of theoretically computed and physically plausible LSFs. The physical model for the PSF is built including/modelling the following elements:

- the optical wavefront error (WFE) map which is modelled by a linear combination of normalised Legendre polynomials up to the 6^{th} order: these maps are randomly generated and scaled to have a RMS uniformly distributed between 40 and 60 nm;
- the pixel binning and four-phase TDI charge transfer;
- the charge diffusion;

Polychromatic LSF/PSFs are computed as photon-weighted averages of the monochromatic LSF/PSFs using random spectral energy distributions (SED) as weighting function. LSFs are derived by integrating PSFs in the AC direction and then normalised to unit area. AL and AC smearing due to image motion relative to the mean charge motion during exposure is not taken into account because this effect can be simulated applying a smearing *a posteriori*. A large number of polychromatic LSFs is obtained by combining 200 random WFE maps with 50 random SEDs (giving thus 10000 different realisations of LSFs). Each LSF is then considered twice by mirroring it around $u = 0$ for symmetry reasons. The mean LSF is then calculated and subtracted from each of the 20000 LSFs. PCA is then run to compute the optimal basis functions for the LSF representation.

This scheme has been followed also to build up the present model but with a few remarkable differences:

- while LL-084 model depends on a single variable (the AL position), the present one aims to model monochromatic LSFs as function of AL position and wavelength, thus we are looking for a 2D representation;
- we have found a fault in the PSF simulation code that introduces wavelength dependent errors in the LSF shape that is largely discussed in Sect. 3;
- PCA is a well known dimension reduction scheme but it cannot take into account the spatial locality of features in 2D images: hence we have implemented GPCA to compute the optimal set of basis functions, as discussed in Sect. 4.

The original MATLAB code written by Lindegren and available in SVN was used as starting framework and then integrated with few original routines specifically developed for the present purposes. This code is available under SVN at this *link*.

3 PSF modelling and wavelength sampling

The numerical calculation of the PSF and LSF has been discussed by Lindegren in various technical notes (see Lindegren (LL-046)). Here we report for convenience a brief summary of all relevant equations.

Let (x, y) be linear coordinates in the pupil plane in units of [m], and (u, v) the corresponding angular coordinates in the image plane in units of [rad]. For a given WFE map $w(x, y)$ the normalised monochromatic optical PSF is given by:

$$P_\lambda^O(u,v) = \frac{1}{\lambda^2 D_x D_y}\left[\int\int_{-\infty}^{\infty} A(x,y)\, exp\left[i\frac{2\pi}{\lambda}(xu+yv)\right]\, dx\, dy\right]^2 \quad (1)$$

where λ is the wavelength, D_x and D_y the pupil dimensions along x and y and A the complex amplitude of the incident wavefront in the pupil plane:

$$A(x,y) = \begin{cases} exp\left[i\frac{2\pi}{\lambda}\, w\,(x,y)\right] & \text{for } (x,y) \in \text{pupil} \\ \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

Eq. 1 is computed through a discrete Fourier transform (DFT) that requires to evaluate the pupil function A on a discrete grid of $N_x \times N_y$ pixels with some steps Δx and Δy. This corresponds to mapping the image plane on a corresponding grid with steps Δu and Δv respectively. The DFT algorithm requires that $N_x \Delta x \Delta u = \lambda$ and $N_y \Delta y \Delta v = \lambda$. The PSF image is computed on a sub-pixel grid with a sampling distance of s_u and s_v in each coordinate. If p_u and p_v are the pixel size (in radians), then $\Delta u = s_u\, p_u$ and $\Delta v = s_v\, p_v$ and consequently

$$\Delta x = \frac{\lambda}{N_x s_u p_u}, \quad \Delta y = \frac{\lambda}{N_y s_v p_v}, \tag{3}$$

The sampling steps s_u and s_v are free parameters and must be small enough to sample the image at least to the Nyquist frequency, which means that:

$$s_u \leq \frac{\lambda_{min}}{2 D_x p_u}, \quad s_v \leq \frac{\lambda_{min}}{2 D_y p_v}, \tag{4}$$

For the Gaia case, $\lambda_{min} = 300$ nm, $D_x = 1.4510$ m, $D_y = 0.5016$ m, $p_u = (10\mu m)/(35m)$ and $p_v = (30\mu m)/(35m)$ resulting in $s_u \leq 0.36182$ and $s_v \leq 0.34888$. The PSF image computed by DFT is a periodic function of u and v with periods $N_x \Delta u$ and $N_y \Delta v$ respectively. Since at some distance from the centre the intensity decreases with u^{-2} and v^{-2}, in order to reach a relative accuracy r at some distance u, v from the centre it is necessary that

$$N_x > \frac{u}{\Delta u} r^{-1/2}, \quad N_y > \frac{v}{\Delta v} r^{-1/2} \tag{5}$$

With $s_u = s_v = 1/8$, to have an accuracy of 10% at $u = \pm 10$ pixel it is necessary to set $N_x > 800$. Lindegren (LL-084) assumed for his computations $N_x = 1024$, $N_y = 512$: monochromatic LSFs were computed for 39 wavelengths, geometrically spaced from 330 to 1015 nm (i.e., with a constant factor 1.03 between successive wavelengths). However, despite the fact that the provided equations are valid for any wavelength value, the implementation based on DFT introduces some subtle and insidious wavelength dependent systematic error that seems to have gone unnoticed so far. This systematic can be easily seen in Fig. 1 where monochromatic LSFs for a given WFE map have been computed on a fine wavelength grid with a sampling of 1 nm. The diffraction patterns should vary linearly with wavelength and the LSF should have some smooth appearance while the computed one seems to be assembled from ribbons resulting in a discontinuous surface. Some ill behaviour is seen also in the right panel representing on a linear scale the crest of the LSF as function of wavelength.

This behaviour can be understood by noticing that the pupil area is sampled on a grid that changes with wavelength as seen by Eq. 3; the maximum sampled coordinate on the pupil for a

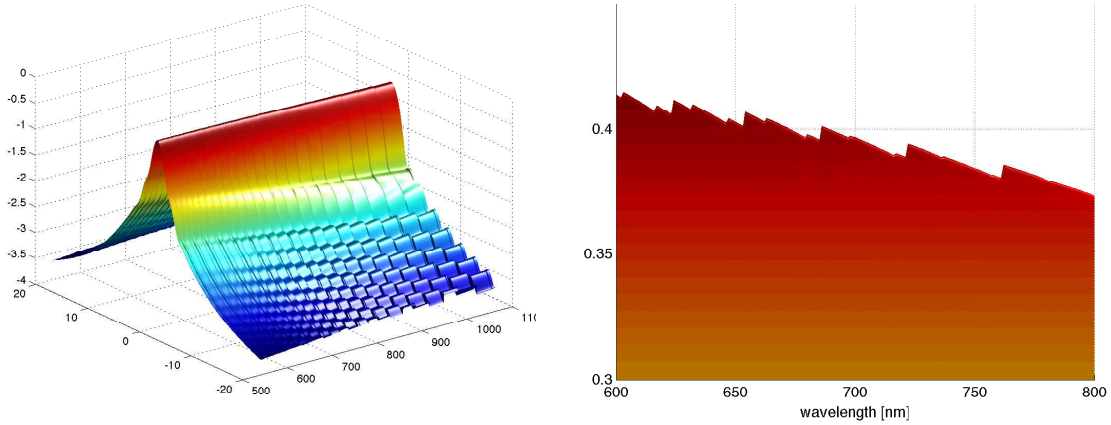

FIGURE 1: Complex effect of the incomplete pupil coverage on a randomly generated monochromatic LSF as function of wavelength. *Left*: logarithmic representation of the LSF in the wavelength range $[550-1015]$ nm ; *Right*: linear representation of the LSF value computed at $u = 0$ in the wavelength range $[600 - 800]$ nm.

given wavelength λ, normalised in the range $[-1, +1]$ is given by:

$$x_{max} = floor\left(\frac{N_x s_u p_u D_x}{2\lambda}\right) \Big/ \left(\frac{N_x s_u p_u D_x}{2\lambda}\right) \tag{6}$$

This means that the effective pupil of the instrument is a function of the wavelength as can be seen in Fig. 2 where the fraction of the pupil coverage is plotted against the wavelength for the AL and AC sampling scheme: having the pupil area not fully covered by the sampling is like having an instrument with a smaller diameter. The scale of the LSF (the FWHM as well as the angular dimension of the diffraction patters) should increase linearly with λ/D but also D locally increases with λ and this causes the overall pattern seen in Fig. 1. This effect may be minimal at shorter wavelength but becomes progressively larger at the red end. An easy way to ensure a proper pupil coverage is to select only those wavelengths where this naturally happens:

$$\lambda_i = \frac{N_x s_u p_u D_x}{2i} \quad \text{for } 1 \leq x \leq \frac{N_x}{2}; \tag{7}$$

However this wavelength sampling scheme ensures a correct coverage for the AL scan direction only as can be seen from Fig. 2. The only way to remove effects due to AC irregular sampling is to change the AC sampling step s_v accordingly to the following relation:

$$s_v = s_u \frac{N_x}{N_y} \frac{D_x}{D_y} \frac{1}{3}; \tag{8}$$

where the last factor comes from the ratio between AL and AC pixel size. In this way the sampling scheme given by Eq.7 is optimal also for the AC pupil coverage.

3.1 Simulations for current analysis

The modulation transfer function (MTF) to compute the effective PSF models the along-scan pixel integration, the along-scan four-phase TDI charge transfer, the across scan pixel integra-

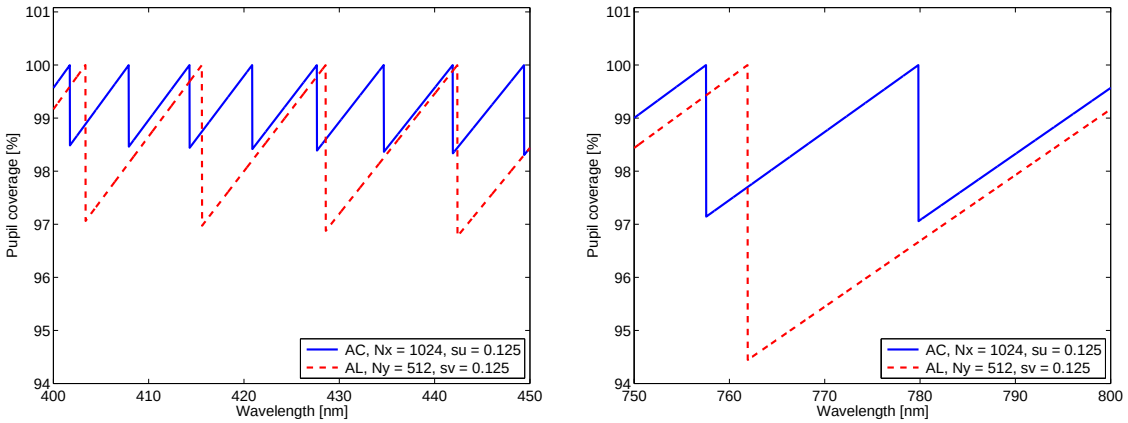

FIGURE 2: Maximum pupil coverage in AL and AC as a function of wavelength obtained for sampling steps $s_u = s_v = 1/8$ pixel and for a number of discretisation points $N_x = 1024, N_y = 512$. The pupil coverage is shown for two different wavelength ranges.

tion and the charge diffusion (modelled as a bivariate normal distribution with diffusion width $\sigma_u = \sigma_v = 4\ \mu m$).

The requirements for the current analysis are the following:

- sampling step s_u = 1/16;
- accuracy of the simulation $r = 1\%$ up to sample $u = \pm 25$ pixels from the centre;
- wavelength sampling scheme to ensure homogeneous pupil coverage both in AL and AC directions.

The 1^{st} requirement comes from the fact that the AL sampling grid of the simulated LSFs will constitute the sampling grid for the basis functions which, as will be shown in following sections, exhibit features with spatial frequencies higher than those present in a typical LSF, so a higher sampling density is advisable for a more accurate analytical representation of the derived basis itself. The 2^{nd} requirement was set to minimise any possible effect of the small discontinuity present at the boundaries of the tails approximation, especially at longer wavelengths. Eq. 5 resulted in $N_x > 4000$ pixel. Since DFT algorithms are highly efficient for array lengths that are powers of 2, we set $N_x = 4096$. The 3^{rd} requirement has been achieved by setting the AC sampling step s_v accordingly to Eq. 8.

The final settings adopted for the simulations are: $N_x = 4096$, $N_y = 1024$, $s_u = 0.0625$, $s_v = 0.2411$, $u_{max} = 26$ pixel; each LSF is then sampled in the AL direction on a grid of $n_u = 833$ samples.

The achieved s_v value is well below the limit set by the Nyquist frequency; the chosen AC image dimension N_y ensures to compute the PSFs with an accuracy in the AC direction of 1% up to ± 24.6 pixel from the centre. The wavelength grid resulting from Eq.7 was used to

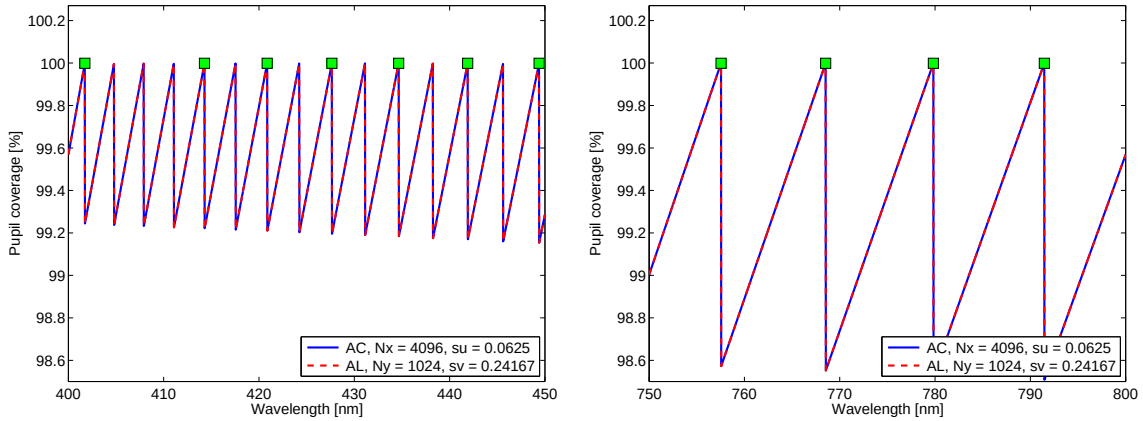

FIGURE 3: Maximum pupil coverage in AL and AC as a function of wavelength obtained for the final sampling steps and number of discretisation points. The pupil coverage is shown for two different wavelength ranges. Green squared points represent the chosen wavelength sampling values.

build up the final wavelength sampling grid taking 1 value every 4 for wavelengths in the range $288-410$ nm, 1 value every 2 for wavelengths between 410 and 540 nm and all remaining values up to 1154 nm. The final grid contains 100 values distributed between 288.4 and 1153.6 nm. Fig. 3 shows the maximum pupil coverage in AL and AC direction obtained with the current settings and the assumed wavelength grid for the simulations.

3.2 LSF normalisation

A total number of 5000 WFE have been randomly generated providing an equal number of LSF sets. The AL LSF is obtained by summing the PSF in the across scan direction and multiplying by s_v. At this stage the LSF needs to be properly normalised to ensure that $\int_{-\infty}^{\infty} L(u)\,\mathrm{d}u = 1$. However the numerical LSF spans a limited interval up to $\pm u_{max} = 26$ pixel so a proper model for the wings must be introduced.

To model the wings for polychromatic and quasi-monochromatic LSF Lindegren (LL-084) proposed the tail function:

$$t(u) = \begin{cases} 0 & \text{if } u \leq \alpha \\ \gamma_1 (u-a)^3 + \gamma_2 (u-\alpha)^4 & \text{if } \alpha < u \leq \beta \\ \gamma_3 u^{-2} & \text{if } \beta < u \end{cases} \tag{9}$$

This function is strictly zero for $u < \alpha$, decreases as u^{-2} for $u > \beta$ thus reproducing the expected behaviour of a true LSF, and is a fourth-degree polynomial in the interval $[\alpha, \beta]$. Coefficients γ can be computed to make $t(u)$ ans $t'(u)$ continuous at $u = \beta$ and $\int_{-\infty}^{\infty} t(u)\,\mathrm{d}u = 1$. Setting $\alpha = 22$, $\beta = 26$ it turns out that $\gamma_1 = 35/14384$, $\gamma_2 = -8/17127$, $\gamma_3 = 3959/162$.
To make the analytical LSF representation continuous at $u = \pm\beta$, the positive and negative tail functions should be scaled respectively for the quantities

$$s^+ = L(\beta)/t(\beta), \quad s^- = L(-\beta)/t(\beta) \tag{10}$$

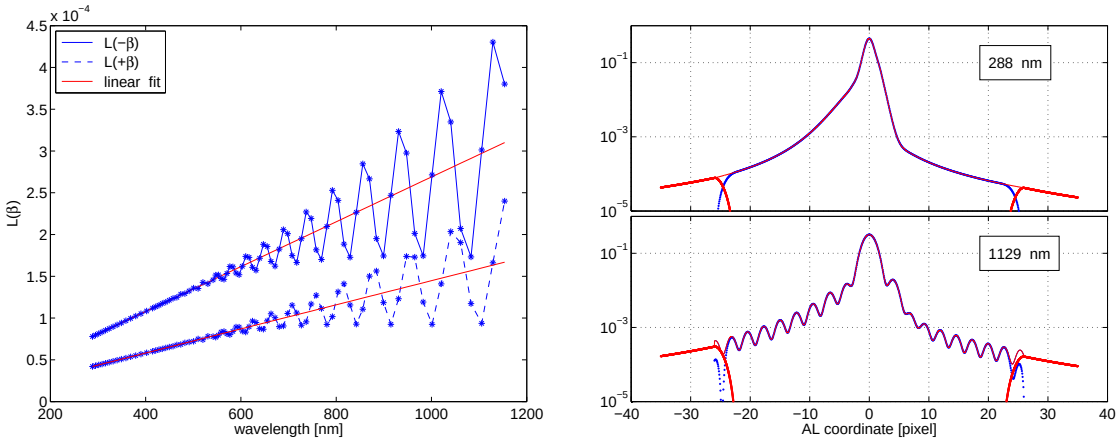

FIGURE 4: *Left*: example of computation for the LSF tail scaling factors: blue dots are the LSF data values at $u = \pm\beta$; red line is the linear fit to data used to evaluate the correct scaling factors s_λ^+ and s_λ^-. *Right*: blue dots represents the L_λ^{mod} term of Eq. 14, red dots are the two scaled tail functions, the red line is the semi-analytical LSF model given by the sum of the other components. Top panel refers to $\lambda = 288$ nm, bottom panel to $\lambda = 1129$ nm.

However modelling a monochromatic LSF is less trivial because of the presence of diffraction patterns in the PSF profile, especially at longer wavelengths.

In absence of aberrations the monochromatic optical PSF can be written as:

$$P_\lambda^O(u,v) = \frac{D_x D_y}{\lambda^2}\ sinc^2\left(\frac{\pi u D_x}{\lambda}\right)\ sinc^2\left(\frac{\pi v D_y}{\lambda}\right) \tag{11}$$

Therefore the optical LSF, obtained by integrating the previous relation along the v direction, yields:

$$L_\lambda^O(u) = \frac{1}{\pi^2 D_x}\ \frac{\lambda}{u^2}\ sin^2\left(\frac{\pi u D_x}{\lambda}\right) = \frac{1}{2\pi^2 D_x}\ \frac{\lambda}{u^2}\ \left(1 - cos\left(\frac{2\pi u D_x}{\lambda}\right)\right) \tag{12}$$

This equation shows two interesting properties:

- the wings of the LSF can be described as a function decreasing with u^{-2} added to a cosine function;
- at a given u, the LSF central value (i.e. excluding the cosine term) depends linearly from λ.

The first point suggests that the area below the wings can be computed modelling the wing itself with the tail function of Eq. 9 if properly scaled, because the cosine contribution to the area is null. These properties are still valid in presence of aberrations at a proper distance from the LSF centre, as can be seen in Fig.4. The scaling factor of the tail function can be computed by a linear fit of the values of the numerical LSF at $u = \pm\beta$ as shown in the left panel. The resulting analytical LSF representation is almost continuous for shorter wavelengths because the

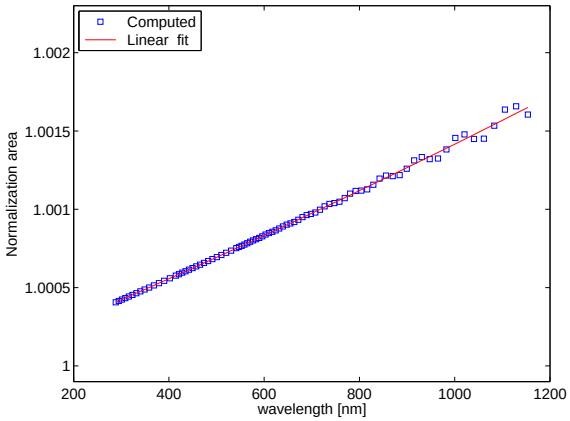

FIGURE 5: Normalisation factor as a function of wavelength: blue squares are the numerical values computed via Eq. 16; red line is a linear fit to the data used to compute the LSF normalisation function.

amplitude of the diffraction patterns is practically zero below 500 nm. At longer wavelength a discontinuity is present and its magnitude changes with the wavelength because the amplitude of the pattern increases and the phase of the cosine function at $u = \pm\beta$ changes with λ as well. The resulting semi-analytical model for the the LSF is then:

$$L_\lambda(u) = L_\lambda^{mod}(u) + s_\lambda^+ t(u) + s_\lambda^- t(-u) \tag{13}$$

where

$$L_\lambda^{mod}(u) = L_\lambda(u) - s_\lambda^+ t(u) - s_\lambda^- t(-u) \tag{14}$$

is numerically computed and defined only over a discrete grid on the interval $[-\beta, \beta]$;

$$\begin{aligned} s_\lambda^+ &= \left(c_0^+ + c_1^+ \lambda\right) / t(\beta) \\ s_\lambda^- &= \left(c_0^- + c_1^- \lambda\right) / t(\beta) \end{aligned} \tag{15}$$

where c_0^+, c_1^+ and c_0^-, c_1^- are the coefficients of the fit of $L_\lambda(\beta)$, and $L_\lambda(-\beta)$ respectively. The integral of the LSF is then approximated by:

$$\int_{-\infty}^{\infty} L_\lambda(u)\, du \simeq \sum_{-\beta}^{+\beta} L_\lambda^{mod}(u)\, du + s^+ + s^- \tag{16}$$

However, as can be seen in Fig. 5, while the approximation is good at shorter wavelengths, at longer ones the effect introduced by the truncation of the diffraction patterns in the transition between the numerical LSF and the analytical tail function becomes rather evident. To avoid the inclusion of this numerical error into the LSF we have then decided to compute the final normalisation factor as a linear fit of the values numerically computed through Eq. 16; an example of such normalisation function is shown in Fig. 5.

4 Generalised Principal Component Analysis (GPCA)

The Generalised Principal Component Analysis (hereafter GPCA) (Ye et al., 2004) is a fast and efficient algorithm for 2D image compression and retrieval: this algorithm is used to concentrate relevant information of a given data set in a small number of dimensions, but unlike the PCA, it is able to preserve the spatial locality of pixels into an image by projecting the images to a vector space that is the tensor product of two lower-dimensional vector spaces.

Consider a matrix $X \in \mathbb{R}^{r\times c}$ and a (ℓ_1, ℓ_2)-dimensional axis system $u_i \otimes w_j$ for $i = 1, \ldots, \ell_1$ and $j = 1, \ldots, \ell_2$ (the symbol $\otimes$ denotes the tensor product), where $\ell_1 < r$, $\ell_2 < c$, $u_i \in \mathbb{R}^{r\times 1}$ and $w_j \in \mathbb{R}^{c\times 1}$. The projection of X on to the (i, j)-th coordinate $u_i \otimes w_j$ is given by $u_i^T \cdot X \cdot w_j$.

Let $A_i \in \mathbb{R}^{r\times c}$, with $i = 1, \ldots, n$, denotes a set of matrices, $\overline{A} = \frac{1}{n}\sum_{i=1}^{n} A_i$ their mean and $\widetilde{A_i} = A_i - \overline{A}$: GPCA task is to derive an optimal (ℓ_1, ℓ_2)-dimensional axis system, designed by two matrices $U \in \mathbb{R}^{r\times\ell_1}$ and $W \in \mathbb{R}^{c\times\ell_2}$ with orthonormal columns, such that the projections of the data points $\widetilde{A_i}$ onto this axis system have the maximum variance over all the possible (ℓ_1, ℓ_2)-dimensional axis systems. The variance is defined as

$$var(U, W) = \frac{1}{n-1}\sum_{i=1}^{n} \|U^T \cdot \widetilde{A_i} \cdot W\|_F \tag{17}$$

the symbol $\|.\|_F$ denotes the Frobenius norm of a matrix.
Two interesting properties of matrices U and W are that:

- for a given W, left matrix U consists of the ℓ_1 eigenvectors of the matrix $M_U = \sum_{i=1}^{n} \widetilde{A_i} \cdot W \cdot W^T \cdot \widetilde{A_i}^T$ corresponding to the largest ℓ_1 eigenvalues;
- for a given U, right matrix W consists of the ℓ_2 eigenvectors of the matrix $M_W = \sum_{i=1}^{n} \widetilde{A_i}^T \cdot U \cdot U^T \cdot \widetilde{A_i}$ corresponding to the largest ℓ_2 eigenvalues.

These properties provide an iterative method to effectively compute U and W: for a fixed U we can compute W through Singular Value Decomposition (SVD) of M_W:

$$M_W = \phi_W \cdot D_W \cdot \phi_W^T \tag{18}$$

$D_W \in \mathbb{R}^{c\times c}$ is a diagonal matrix containing the eigenvalues of M_W while the columns of $\phi_W \in \mathbb{R}^{c\times c}$ are the corresponding eigenvectors; we can then update the W matrix by selecting the first ℓ_2 columns:

$$W = \left[\phi_{W_1}, \phi_{W_2}, \ldots, \phi_{W_{\ell_2}}\right] \tag{19}$$

After this we can then update the U matrix by computing the ℓ_1 eigenvectors of M_U corresponding to the largest ℓ_1 eigenvalues

$$M_U = \phi_U \cdot D_U \cdot \phi_U^T \tag{20}$$

with $D_U \in \mathbb{R}^{n\times n}$ and $\phi_U \in \mathbb{R}^{n\times n}$

$$U = \left[\phi_{U_1}, \phi_{U_2}, \ldots, \phi_{U_{\ell_1}}\right] \tag{21}$$

and so on until the result converges. Experience shows that if the process is initialised by setting $U_0 = \{I_d, 0\}^T$, where I_d is the $\ell_1 \times \ell_1$ identity matrix, the derived solution is satisfactory and the process usually converges in very few iterations.
For given U and W matrices, the projection of $\widetilde{A_i}$ can be computed as

$$D_i = U^T \cdot \widetilde{A_i} \cdot W \qquad \text{with } D_i \in \mathbb{R}^{\ell_1\times\ell_2} \tag{22}$$

From D_i we can reconstruct $\widetilde{A_i}$ by setting

$$\widetilde{A_i} \approx U \cdot D_i \cdot W^T \tag{23}$$

then

$$A_i \approx U \cdot D_i \cdot W^T + \overline{A} \tag{24}$$

The reconstruction error for A_i is then

$$E_i = \|\widetilde{A_i} - U \cdot D_i \cdot W^T\|_F = \|\widetilde{A_i} - U \cdot U^T \cdot \widetilde{A_i} \cdot W \cdot W^T\|_F \tag{25}$$

The convergence can be measured by monitoring the root mean square error (RMSE) defined as

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n} {E_i}^2} \tag{26}$$

RMSE measures the average reconstruction error.

4.1 LSF modelling and GPCA

GPCA fits quite well the needs for a proper LSF modelling: each simulated LSF consists of a matrix with $n_u = 833$ rows and $n_w = 100$ columns; as we will see, each LSF can be accurately described with a typical $(\ell_1, \ell_2) = (9, 3)$-dimensional axis system (or possibly with even lower dimensionality). The columns of the left matrix U are the basis functions that model the LSF dependencies in the AL direction while those of the right matrix W model the wavelength dependencies. If $\overline{L}$ is the mean LSF, then Eq. 24 states that a LSF can be approximated as:

$$L \approx \overline{L} + U \cdot D \cdot W^T \tag{27}$$

or, in extended notation:

$$L(u_i, \lambda_j) = \overline{L}_{u_i,\lambda_j} + \sum_{m=1}^{\ell_1}\sum_{n=1}^{\ell_2} d_{m,n} \cdot U_{i,m} \cdot W_{j,n} \tag{28}$$

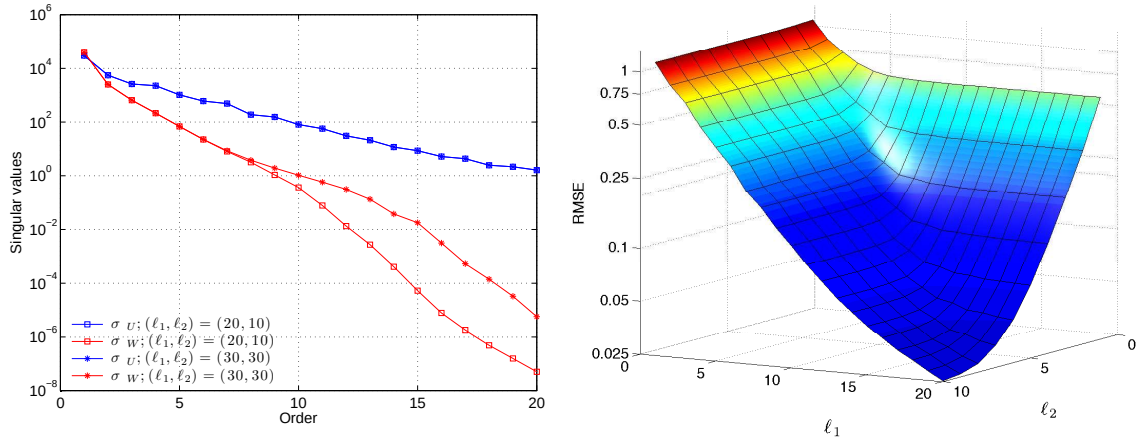

FIGURE 6: *Left*: singular values for the matrix U (blue lines) and W (red lines) computed for two different subspaces (ℓ_1, ℓ_2) = $(20, 10)$ and $(30, 30)$ respectively.*Right*: empirical RMSE for the set of 10000 LSF measured as function of order (ℓ_1, ℓ_2).

i.e., the LSF can effectively be rendered as a linear model depending on less than 27 parameters. Moreover the LSF can be easily 2D-interpolated to continuous variables (u, λ) by 1D-interpolation of U and W bases separately (see Sect. 5). The GPCA algorithm is so efficient that the whole set of $(5000 \times 2) \times (833 \times 100)$ matrices can be processed in few minutes on a common laptop. Each LSF is considered twice by reversing the array columns (equivalent to swapping the AL axis), to preserve the intrinsic AL symmetry of the problem. We have run the algorithm by setting the number of dimensions (ℓ_1, ℓ_2) to many different pairs $(90/30, 40/30, 40/20, 30/30, 30/20, 30/10, 20/20, 20/10)$ obtaining every time virtually identical results. Fig. 6 (left panel) shows the resulting first 20 singular values σ_U and σ_W for the two cases $(\ell_1, \ell_2) = (30, 30)$ and $(\ell_1, \ell_2) = (20, 10)$; while the left eigenvalues σ_U don't depend on the number of corresponding dimensions ℓ_1, the eigenvalues σ_W exhibit some drop off around the order $\ell_2 \simeq \ell_1/2$; moreover the left eigenvalues σ_W keep approximately the same value of σ_U for half of the corresponding order: this behaviour suggests that the LSF wavelength modelling should require barely half dimensions with respect to the AL dependency modelling. This is confirmed also by the RMSE that has been measured on a grid of (ℓ_1, ℓ_2) pairs as shown in Fig. 6, right panel: note how the function is steeper along the ℓ_2 axis than the ℓ_1 one. For the present analysis we have decided to fix the subspace dimensions for GPCA to $(\ell_1, \ell_2) = (20, 10)$. Table 4.1 reports the (ℓ_1, ℓ_2) pairs corresponding to the minimal number

Rank	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
$\ell_1 \cdot \ell_2$	1	2	3	4	8	10	12	18	21	24	27	36	40	44	55	60	65	78
ℓ_1	1	2	3	4	4	5	6	6	7	8	9	9	10	11	11	12	13	13
ℓ_2	1	1	1	1	2	2	2	3	3	3	3	4	4	4	5	5	5	6
RMSE	1.1702	0.9253	0.7869	0.6950	0.5479	0.4553	0.3975	0.3360	0.2636	0.2314	0.2006	0.1625	0.1409	0.1213	0.1020	0.0887	0.0767	0.0678

TABLE 2: Steepest descent path in the RMSE surface with minimal number of parameters.

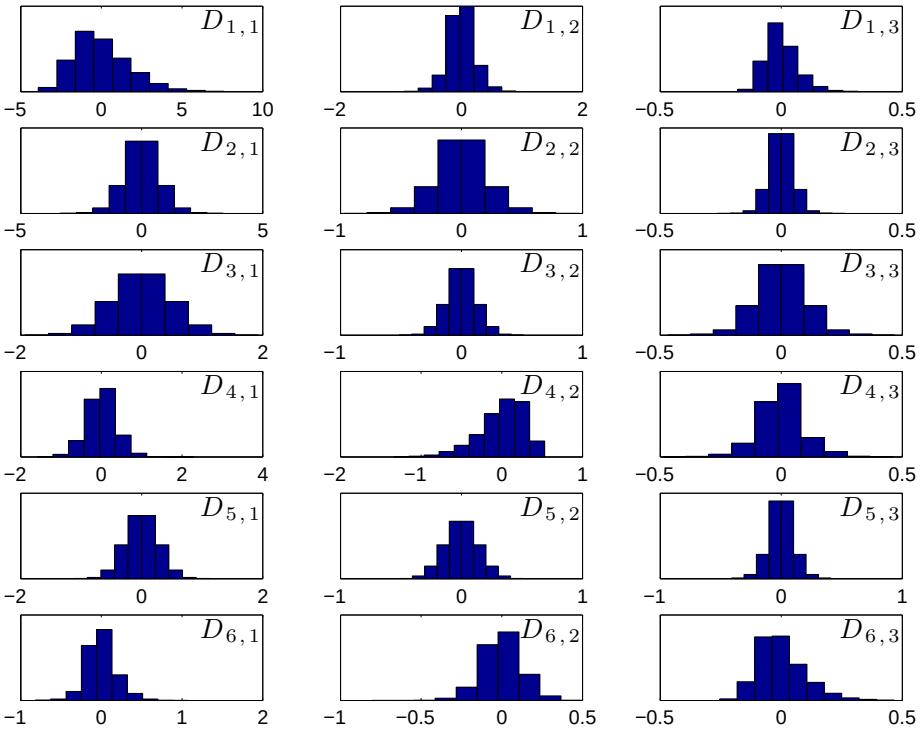

FIGURE 7: Distribution of the coefficients of the projection of the 10000 LSF onto the (U, W) axis system. Coefficients for the ith LSF are computed as $D_i = U^T \cdot (L_i - \overline{L}) \cdot W$.

of parameters required to get the steepest descent path through RMSE values. This sequence provides a useful scheme to be followed in the source update process where a compromise between number of model parameters and quality of the data will need to be found.

The distributions of the coefficients of the projection for the 10000 LSF onto the (U, W) axis system are plotted in Fig. 7 . The coefficients for the ith LSF are computed following Eq. 22: $D_i = U^T \cdot (L_i - \overline{L}) \cdot W$. These distributions may provide useful information for delimiting the starting search region for the true LSF parameters during the source update process.

Fig. 8 shows the first four basis functions for the left matrix U and for the right matrix W: left basis functions are plotted as a function of the AL coordinate, while the right ones as a function of wavelength. Unsurprisingly the U_i bases resemble basis functions from Lindegren (LL-084), although the present ones are a bit more complex, especially at higher orders, for the presence of small fluctuations due to diffraction patterns in the LSF wings. Figs. 9-10 show the tensor product between the first left bases $U_1, \ldots, U_6$ and the first two right bases W_1, W_2. Only the central AL coordinate range $(-10, 10)$ is plotted for clearness.

Finally two examples of the approximations obtained with three different configurations of model parameter numbers is shown in Fig. 11. The two cases have been chosen to represent the

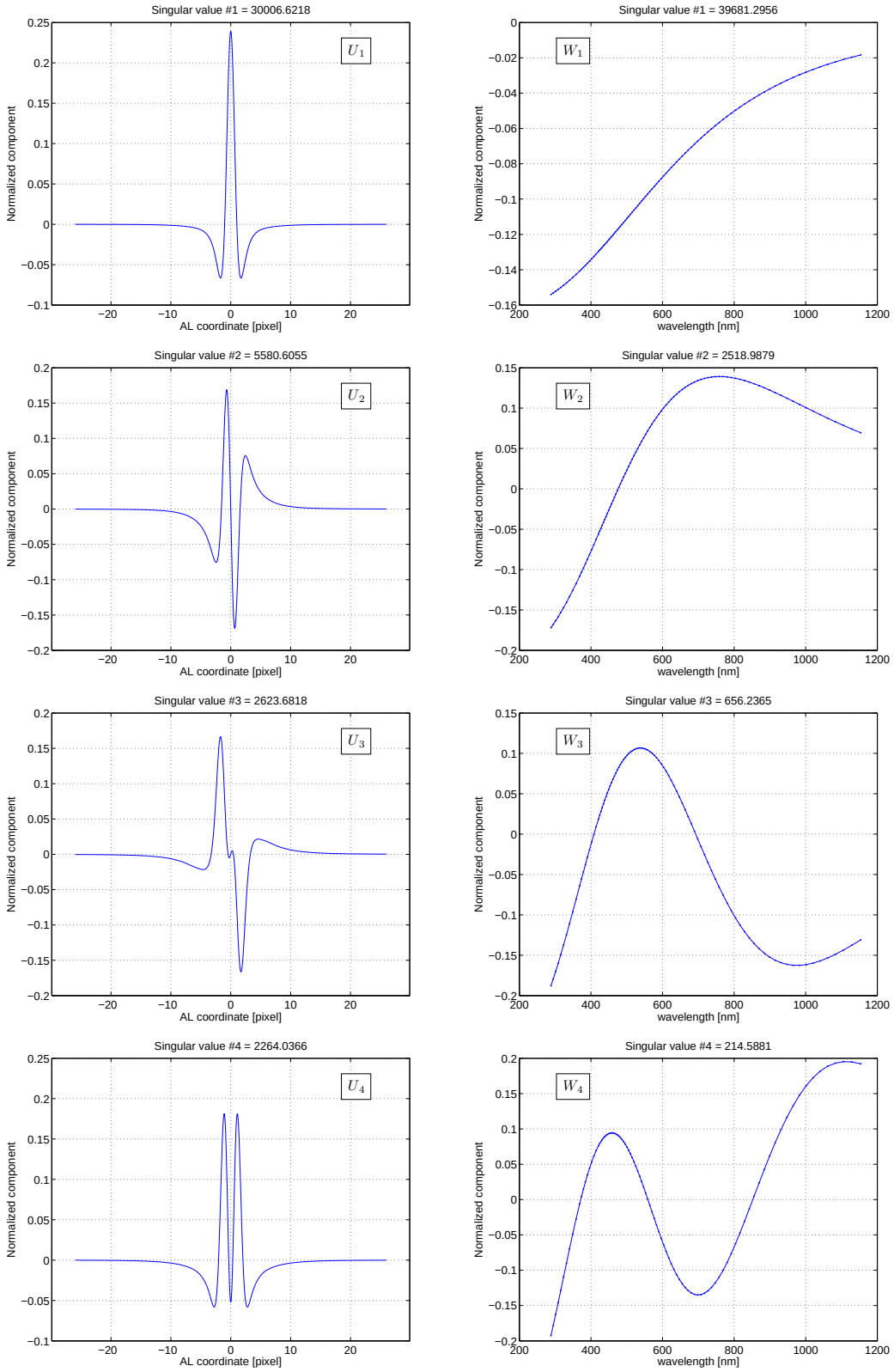

FIGURE 8: *Left*: first 4 left basis functions (matrix U columns) as function of the AL coordinate. *Right*: first 4 right basis functions (matrix W columns) as function of the wavelength.

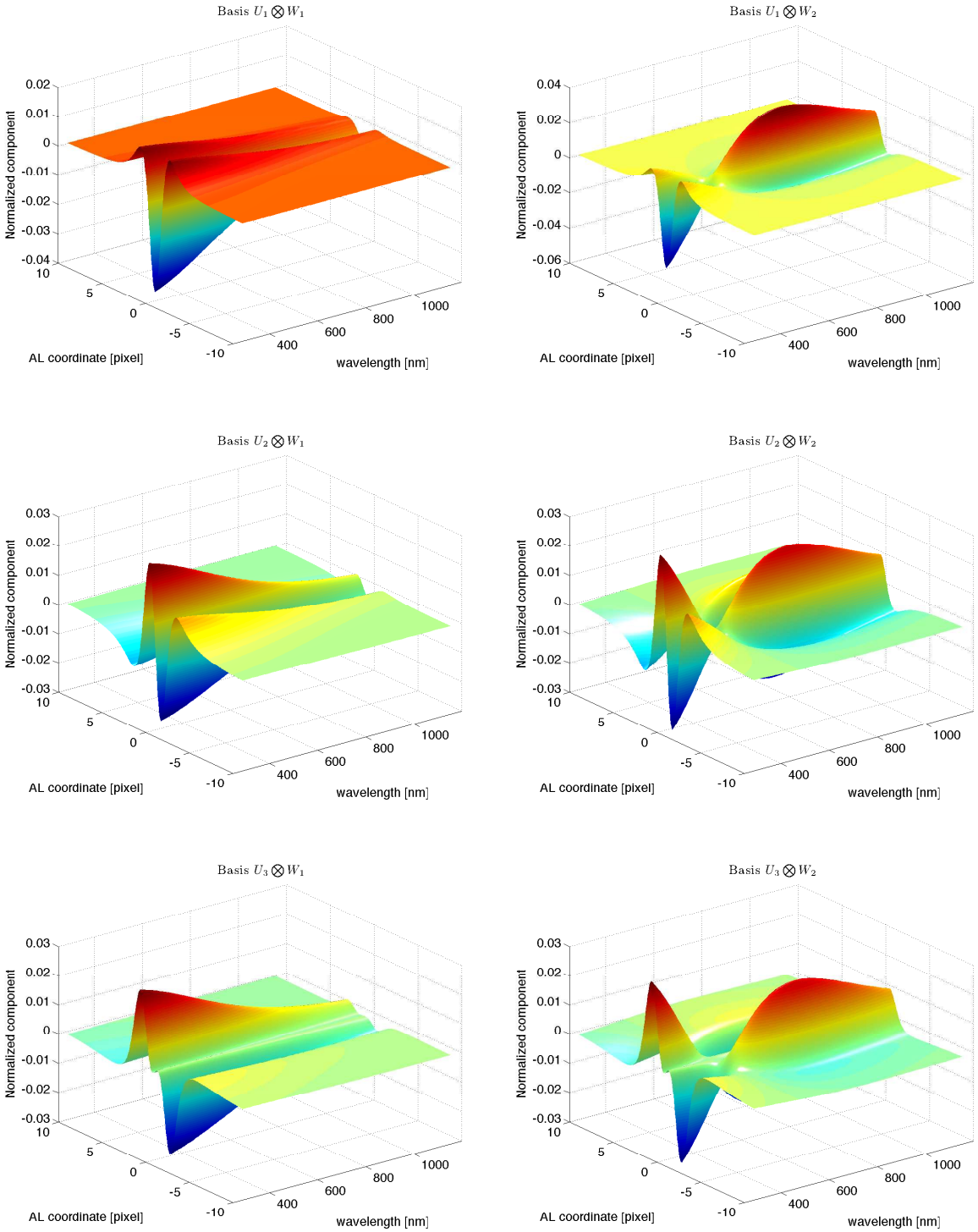

FIGURE 9: Tensor product between the first 3 left basis functions (U_1, U_2, U_3) and the first 2 right basis functions (W_1, W_2). Only the central range $(-10, 10)$ pixel in the AL axis is shown for clearness.

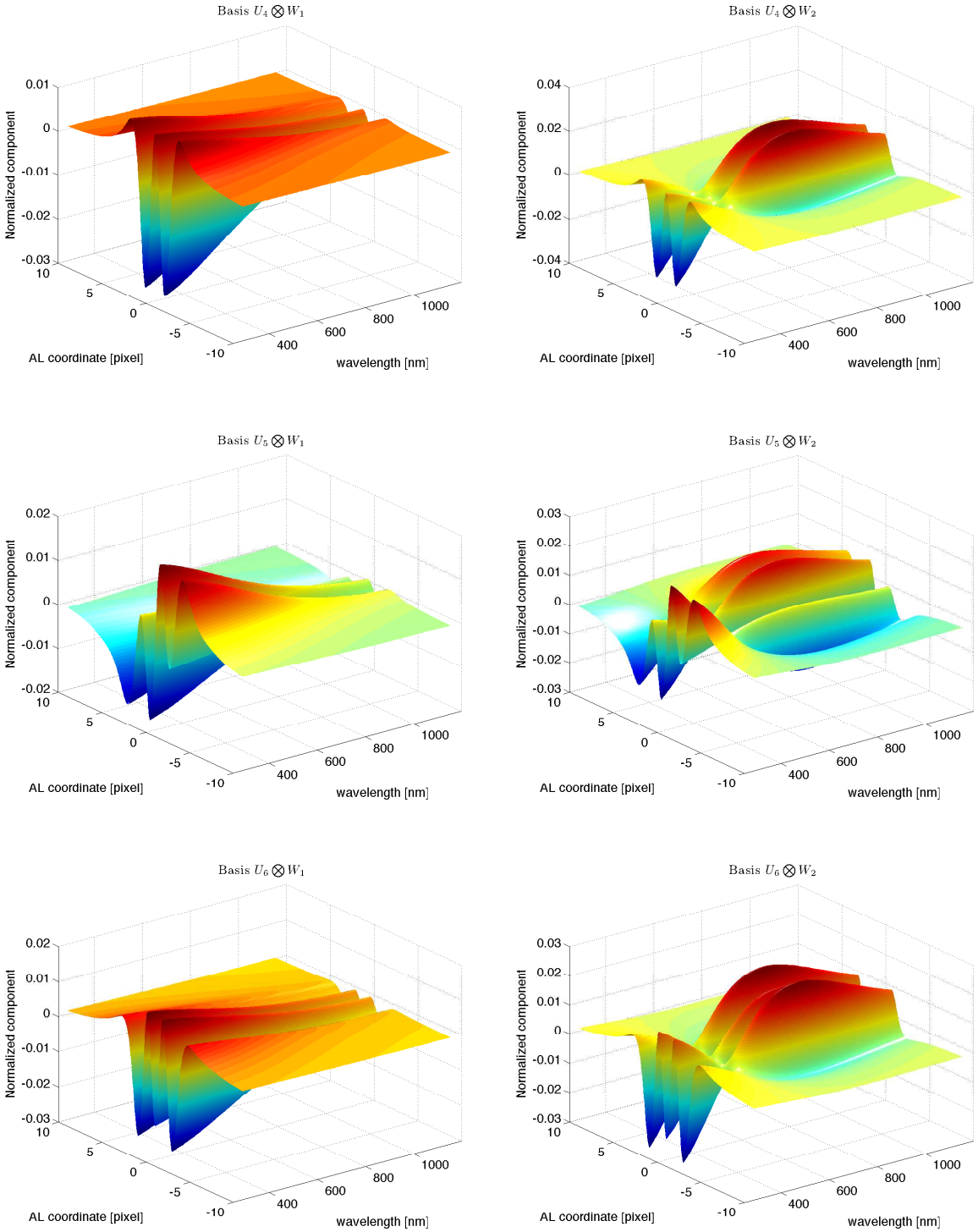

FIGURE 10: Tensor product between the (U_4, U_5, U_6) left basis functions and the first 2 right basis functions (W_1, W_2).

interquartile range in the RMSE distributions: these distributions are shown in Fig. 12 for each

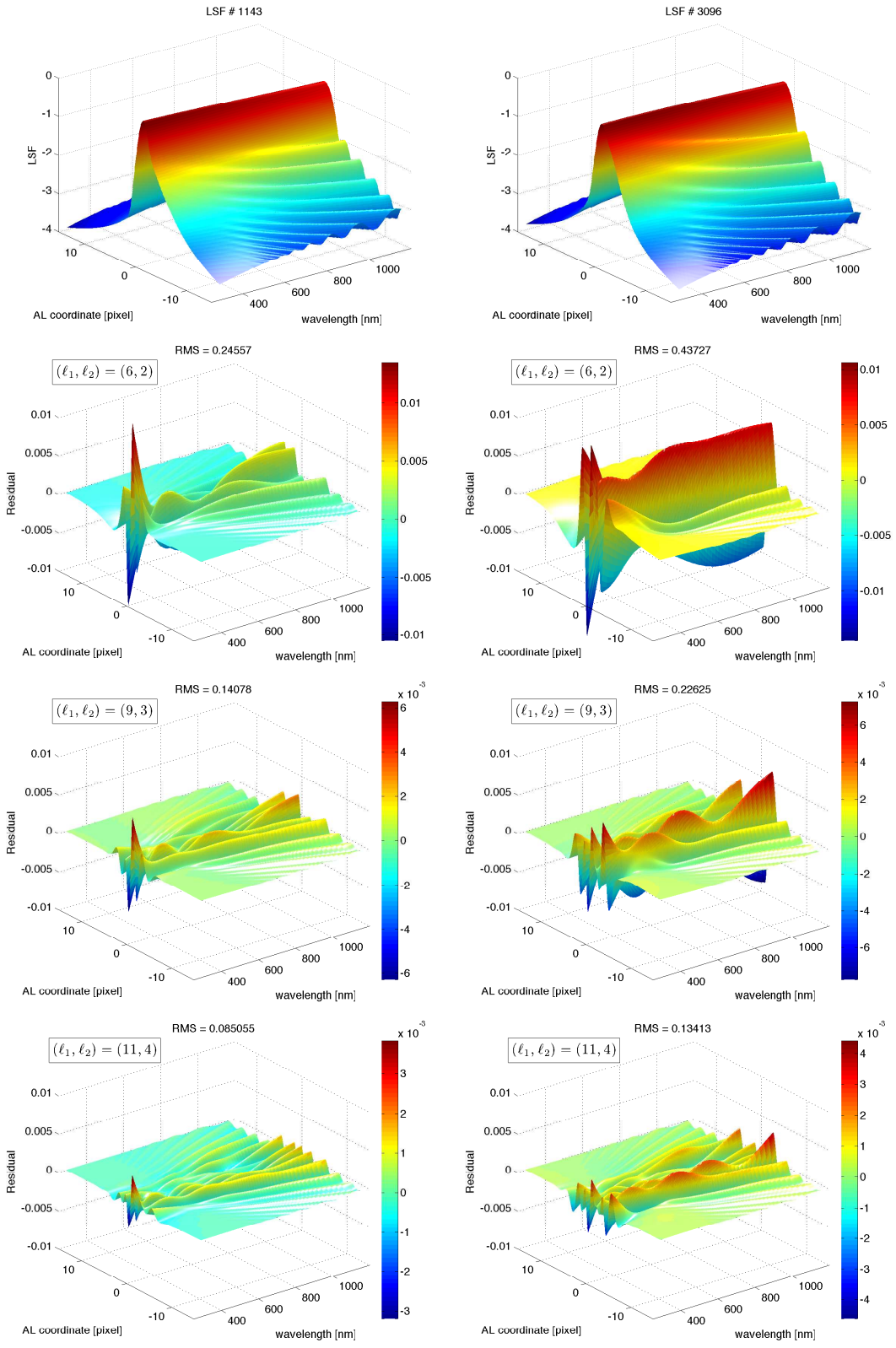

FIGURE 11: Typical reconstruction error as function of the number of basis functions used. The two examples represent the interquartile range in the RMSE distribution.

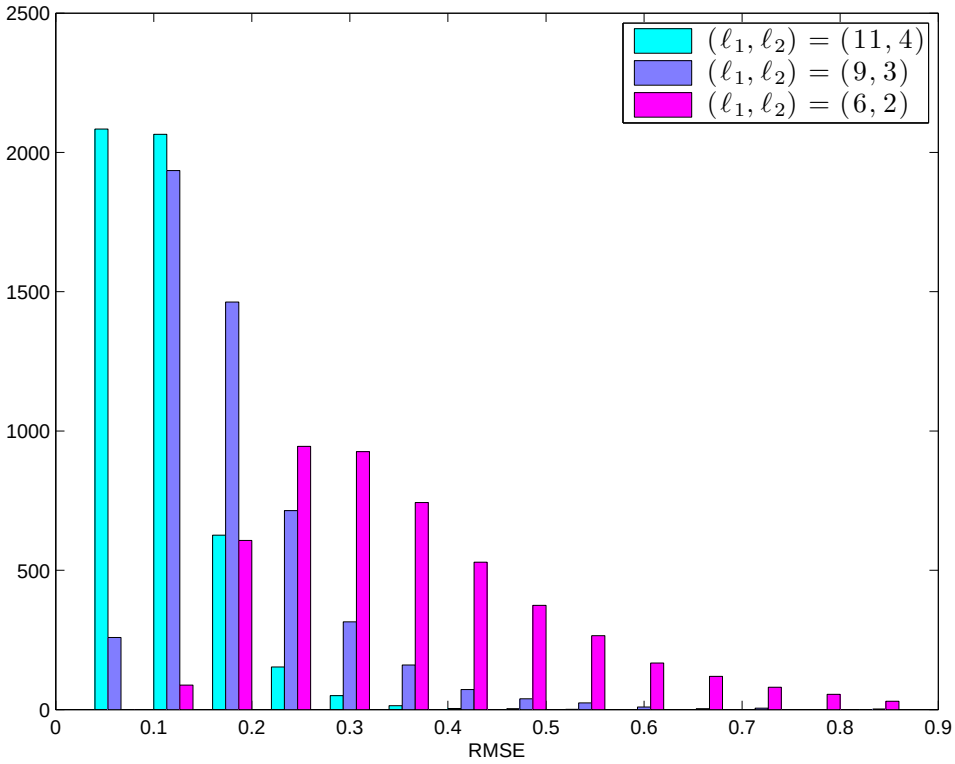

FIGURE 12: Distribution of the average reconstruction error obtained with three different model (ℓ_1, ℓ_2) order pairs.

model configuration.

5 Basis Functions interpolation

5.1 Left Basis Functions

The numerical basis functions contained in matrix U span the interval $-26 \leq u \leq +26$ pixel in steps of 0.0625 pixel while those contained in the right matrix W span the interval $288.4 \leq \lambda \leq 1153.6$ nm with a sampling scheme set by Eq. 7. The most suitable interpolation method for the set of U basis functions has been described by Lindegren (LL-046): the bi-quartic B-spline function consists of a normalised cubic B-spline on a knot sequence with interval 0.5 pixels and convolved with a centred rectangular function of unit area and width. Alternatively it can be seen as the sum of two adjacent quartic B-splines on a regular knot sequence. This function has the remarkable property to satisfy the 'shift-invariant sum' condition (i.e. preserve the underlying function normalisation independently from the sub-pixel position of the sampling

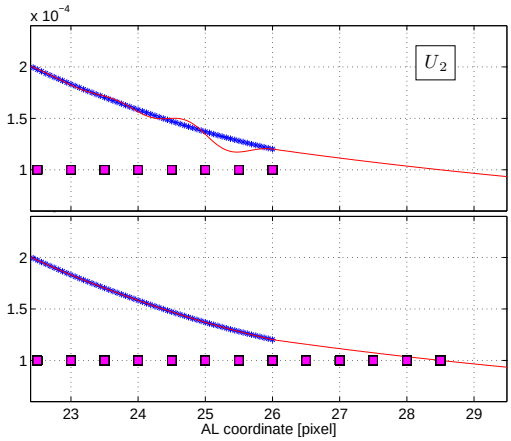

FIGURE 13: Blue points represent a numerical basis function, red line is the spline+tail fit, magenta squares are the spline knots. *Top*: the knots cover the whole function domain; *bottom*: 5 extra-knots have been added to provide full support in the basis function domain.

grid). Formally a bi-quartic B-spline can be considered as a 6^{th} order spline, thus if $x_k = 0.5\,k + \delta$ is the knot sequence (being δ a fixed number), the kth bi-quartic B-spline term $B_k(x)$ is non-zero only on the interval $[x_{k-3}, x_{k+3}]$. Moreover for any $u \in [x_\ell, x_{\ell+1})$ there are at most 6 non-zero B-splines, which are denoted by $B_{\ell-5}(u), B_{\ell-4}(u), \ldots, B_\ell(u)$. Lindegren (LL-084) assumed the knots sequence to be $x_k = -\beta + 0.5\,k$, with $k = 0, 1, \ldots, 4\,\beta$, so that the first and last knots are set equal to the pixel coordinate of the first and last basis function element. However B-splines have a full support interval that is narrower than the knots interval (B-splines naturally tend to zero at the edges of the support interval). As a consequence the spline approximation done by Lindegren on his basis functions showed some residual 'wiggles' at the edges of the fitting interval that was erroneously interpreted in LL-084 as a mismatch between the slopes of the basis function and that of the tail function. To overcome this problem one can add a number of extra-knots equal to the spline degree at both ends of the required support interval so that $x_{k'} = -\beta + 0.5\,k'$, with $k' = -5, -4, \ldots, 4\,\beta + 5$; the apex in k' is set to remark the fact that this index can assume negative values, while we would have used $k = k' + 5$ to designate the traditional array index. Fig. 13 shows the effect of adding these extra-knots on the residuals of the fit to the basis functions. The bi-quartic B-spline for the ith basis function can then be written as:

$$S_i(u) = \sum_{k'=\ell-5}^{\ell} s_{ik'}\, B_{k'}(u) \tag{29}$$

which is strictly valid in the full interval support $-\beta \le u \le \beta$: the 'left' index ℓ spans the range $(0, 4\beta - 1)$. To summarise, since in the present study we have set $\beta = 26$, the total number of spline knots is $n_{knots} = 4\beta + 1 + 10 = 115$, the number of spline coefficients $s_{ik'}$ is equal to the number of degrees of freedom of $S_i(u)$, i.e. $n_{coeffs} = n_{knots} - 6 = 109$.

To extend the U basis functions beyond the pixel $\beta = \pm 26$, we use the special tail functions already introduced in Sect. 3.2, so that the ith basis function is approximated as:

$$U_i(u) = S_i(u) + s_i^-\, t(-u) + s_i^+\, t(u) \tag{30}$$

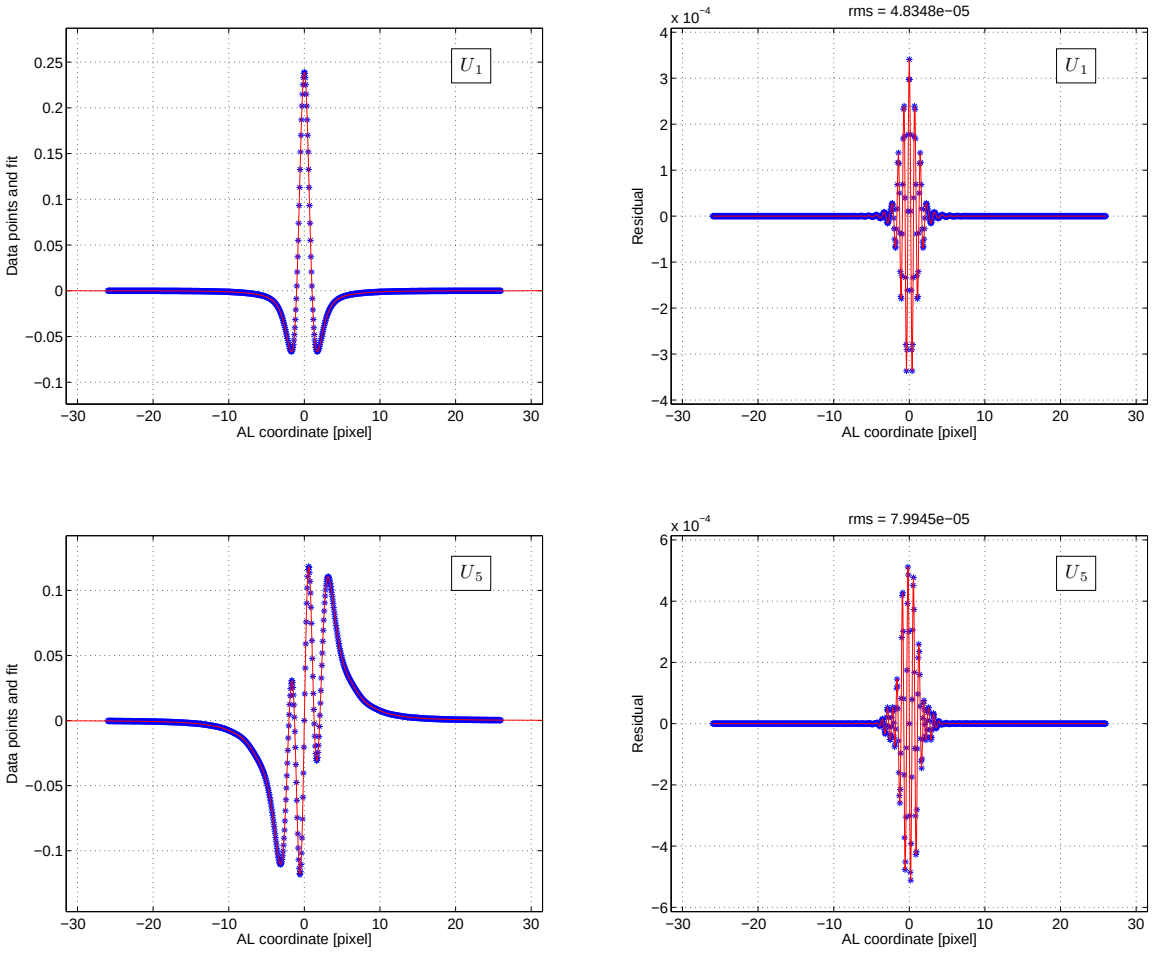

FIGURE 14: Examples for the spline+tail approximation for right U basis functions (*left*) and corresponding residuals (*right*).

where $t(u)$ is the tail function and the two coefficients are

$$s_i^- = \frac{U_i(-\beta)}{t(\beta)}, \quad s_i^+ = \frac{U_i(\beta)}{t(\beta)} \tag{31}$$

to make the approximation continuous at $u = \pm\beta$. The spline coefficients s_{ik} are computed by fitting Eq. 29 to the function $U_i(u) - s_i^-\, t(-u) - s_i^+\, t(u)$. The integral for the approximating function (Eq. 30) is given by

$$\int_{-\infty}^{\infty} U_i(u)\,\mathrm{d}u = \sum_{k'=0}^{4\beta-1} s_{ik'} + s_i^- + s_i^+ \tag{32}$$

Fig. 14 shows the spline+fit approximation for two basis functions (U_1 and U_5) and the relative residuals. RMS of the residuals ranges from 4.8×10^{-5} for U_1 to 4.8×10^{-4} for U_{10}: the total RMS due to the spline approximation is expected to be much smaller than the typical reconstruction error seen in Fig. 11. To verify this we have fitted a random numerical LSF

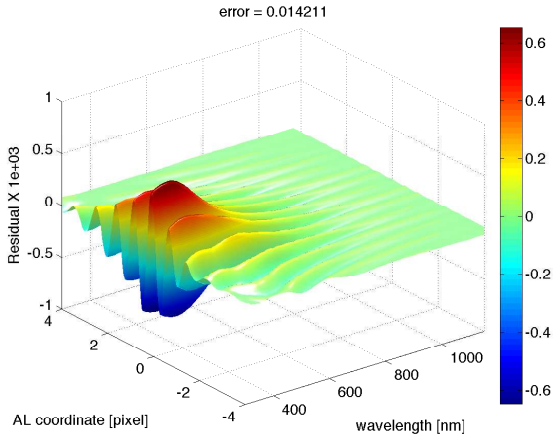

FIGURE 15: Example of the bi-quartic B-spline approximation for a random numerical LSF. Vertical scale has been magnified by a factor of 10^3; only the central 8 pixel are shown along the y-axis.

computed for the present analysis with the spline+tail approximation and the residuals of the fit are shown in Fig. 15: the maximum of the residuals and the global RMS error are both a factor 10 smaller than the typical reconstruction error shown in Fig. 11.

5.2 Right Basis Functions

The interpolation for the right basis functions W can be done using a cubic spline. Since the wavelength support interval spans the range $288.4-1153.6$ nm which is wider than that covered by XP photometers, there is no need to extrapolate the basis functions outside this interval.

6 Mean LSF representation

The mean LSF has been computed on the same AL and wavelength sampling grids of the numerical LSF, hence we have to settle a method to interpolate/extrapolate it in the AL direction and to interpolate in the wavelength direction. In principle the interpolation/extrapolation in AL coordinate could be done with the same algorithm adopted for the left U_i basis function seen in Sect. 5.1. On the other hand interpolation in the wavelength space is more problematic because of the relative coarse sampling density which makes linear or spline interpolation not enough accurate. However GPCA can come the handy to solve the problem: the idea is to use algorithm seen in Sect. 4 to express the mean LSF as $\overline{L} \approx U_L \cdot D_L \cdot {W_L}^T$ such that the error $E_{\overline{L}} = \|\overline{L} - U_L \cdot D_L \cdot {W_L}^T\|_F$ is minimal. The columns of the right matrix W_L can then be interpolated with a cubic spline to the desired wavelength grid, providing the oversampled matrix $\widetilde{W_L}$ which is used to finally reconstruct the interpolated $\overline{L}$ as

$$\overline{L}_I = U_L \cdot D_L \cdot \widetilde{W_L}^T. \tag{33}$$

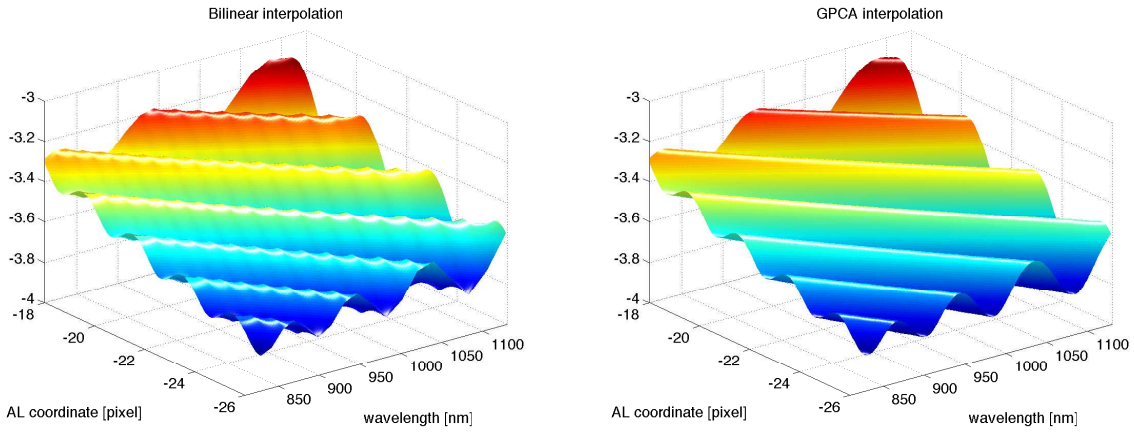

FIGURE 16: Detail of 1nm-wavelength interpolated mean LSF (logarithmic scale). *Left*: bi-linear interpolation; *right*: GPCA-based interpolation.

Fig. 16 shows a comparison between the result of bi-linear interpolation and the proposed GPCA-based interpolation scheme, where the mean LSF has been oversampled to a wavelength grid of 1 nm spacing. Only a limited range in both x and y axis is shown to emphasise the different quality of the interpolation which is more evident in the LSF diffraction patterns that appear smoother in the GPCA case. The residuals of the mean LSF reconstruction, $\overline{L} - U_L \cdot D_L \cdot {W_L}^T$, computed for subspace dimensionality (ℓ_1, ℓ_2) equal to $(20, 20)$ and $(35, 35)$ are shown in Fig. 17. The 2^{nd} configuration has been also used to produce the GPCA interpolation of Fig. 16

The bi-quartic B-spline interpolation scheme can be used to interpolate the basis functions U_L:

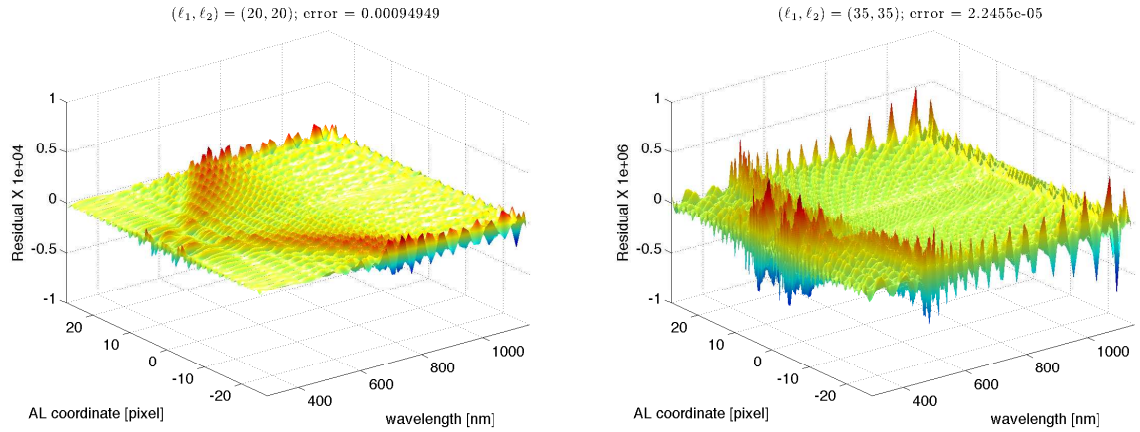

FIGURE 17: Error of the reconstructed mean LSF for a $(20, 20)$-dimensional (*left*) and for a $(35, 35)$-dimensional (*right*) representation. The two z-scales have been magnified respectively by a factor of 10^4 and 10^6.

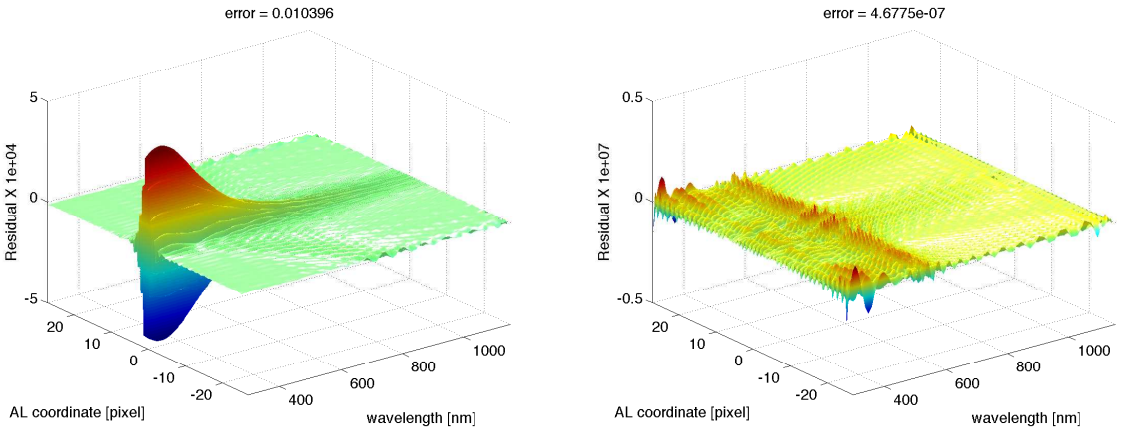

FIGURE 18: *Left*: Error in the reconstructed mean LSF produced by the bi-quartic B-spline interpolation. *Right*: Mean LSF reconstruction error introduced by the diagonalization of the coefficients matrix D_L: the vertical scale has been enlarged by a factor 10^7.

even if these are dominated by the mean LSF diffraction patterns and hence are much more complicated with respect to previous U bases, the overall reconstruction error due to the bi-quartic B-spline approximation, shown in Fig. 18, has the same magnitude as the one obtained by fitting the numerical LSF of Fig. 15. The extrapolation beyond $u = \pm\beta$ can be done with the tail function of Eq. 9 but we need to keep into account considerations done in Sect. 3.2: in particular, the tails scaling factors s^+ and s^- (which in the case of the mean LSF should be identical) must be computed with Eq. 15, where the coefficients $c_0^{\pm}, c_1^{\pm}$ are computed by linear fitting the mean LSF borders as function of λ. However this tail scaling cannot be applied to the basis functions U_L because GPCA decomposition has disentangled the wavelength dependencies in the right W_L matrix. The only possibility is to subtract the tail model scaled through Eq. 15 *before* the GPCA decomposition, i.e. rather than $\overline{L}$ we decompose the matrix $\overline{L}^{\,mod}$ whose $(i,j)^{th}$ element is given by:

$$\overline{L}_{i,j}^{\,mod} = \overline{L}_{i,j} - \frac{\left(c_0^+ + c_1^+\lambda_j\right)}{t(\beta)}\, t(u_i) - \frac{\left(c_0^- + c_1^-\lambda_j\right)}{t(\beta)}\, t(-u_i) \tag{34}$$

Once the GPCA is applied, the matrix U_L columns are fitted with a bi-quartic B-spline, while those of W_L with a cubic spline; to get the value to a given (u,λ) pairs, first the interpolated matrix rows U_u and W_λ are computed, then the corresponding mean LSF value is evaluated as:

$$\overline{L}(u,\lambda) = U_u \cdot D_L \cdot {W_\lambda}^T + \frac{\left(c_0^+ + c_1^+\lambda\right)}{t(\beta)}\, t(u) + \frac{\left(c_0^- + c_1^-\lambda\right)}{t(\beta)}\, t(-u) \tag{35}$$

Finally it is worth noticing that the matrix D_L has off-diagonal elements that are nearly zero: if we set these elements to zero and compute the error in the reconstructed $\overline{L}$ produced by this approximation we obtain the residuals shown in the right panel of Fig. 18. The resulting error is completely negligible with respect to the reconstruction error budget seen so far, thus this kind of optimisation is acceptable and useful to speed up the mean LSF evaluation process.

Moreover columns of W_L matrix can be scaled once for corresponding diagonal elements of D_L matrix thus simplifying further the evaluation process and the data model because of the needlessness to store separately the D_L array.

7 Data Model

The basis functions AL sampling grid ranges from $\beta = -26$ to $\beta = +26$, hence the knot sequence for bi-quartic spline interpolations ranges from $-\beta-2.5$ to $+\beta+2.5$ and is composed of 115 points. The only quantity needed for the reconstruction of the knots sequence to be stored in the data model (DM) is the β value while the other tail function parameter α is necessary to completely characterise the LSF tails model. The number of degrees of freedom for the bi-quartic spline is then $115-6=109$ and this corresponds to the number of spline coefficients to be saved for each left basis function (i.e. both U and U_L columns). Each U basis function needs also two more parameters which are the positive and the negative tail scaling factors, while, in the case of the mean LSF model, these tail factors are computed through the polynomial relations of Eq. 15: in this case the DM will store only two coefficients for the positive tail and two for the negative one. The total number of left basis functions to be stored are 20 for U and 35 for U_L. The wavelength sampling for the right W and W_L matrices is made of 100 points, hence the corresponding number of cubic spline coefficients is 99×4 for each basis function (10 bases for W and 35 for W_L). Finally the mean LSF model requires also the diagonal of the coefficients matrix D, hence an array of 35 elements. The following tables contains a description of the required data types to be added to the current PhotPipe DM; corresponding types will be needed for the MDB DM to store the LSF basis functions in the MDB.

PhotPipeDm:ExtBqsSet		
alpha	double	Tail function parameter.
beta	double	Tail function parameter.
dof	int	Degrees of freedom.
dim	int	Number of basis functions.
scalingFactor	double	Optional scaling parameter.
coeffs	Array of double [byte]	bi-quartic spline coefficients.
tailsPos	Array of double [byte]	Positive tail functions scaling parameter.
tailsNeg	Array of double [byte]	Negative tail functions scaling parameter.

TABLE 3: Data type to hold necessary quantities for the bi-quartic B-spline modelling of a GPCA left matrix (AL dependencies): for the mean LSF case tailsPos and tailsNeg host the two coefficients for the linear fit of the scaling factor as function of wavelength; in the other case they are two arrays of length equal to the number of bases.

PhotPipeDm:ExtCsSet		
pieces	int	Number of spline pieces.
order	int	Order of the spline.
dim	int	Number of basis functions.
knots	Array of double [100]	Knots positions.
coeffs	Array of double [short × byte]	Spline coefficients.

TABLE 4: Data type to hold quantities for the cubic spline modelling of a GPCA right matrix (wavelength dependencies).

PhotPipeDm:ExtGpcaSet		
left	ExtBqsSet	Left bases.
right	ExtCsSet	Right bases.

TABLE 5: Data type to collect left and right GPCA matrices.

PhotPipeDm:ExtLsfBases		
mean	ExtGpcaSet	Mean LSF bases.
bases	ExtGpcaSet	LSF bases.

TABLE 6: Data type to collect all required quantities for the XP LSF model.

8 Acknowledgements

This research was supported by the ASI (Agenzia Spaziale Italiana) through grant ASI I/058/10/0 and by the INAF (Istituto Nazionale di AstroFisica). I thank Michael Weiler who suggested long time ago the usage of GPCA algorithm for the XP LSF modeling.

9 References

[CF-015], Fabricius, C., Voss, H., 2009, *An interpolation scheme for XP LSFs*, GAIA-C5-TN-UB-CF-015, URL http://www.rssd.esa.int/cs/livelink/open/2963659

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

[LL-084], Lindegren, L., 2009, *Minimum-dimension LSF modelling*, GAIA-C3-TN-LU-LL-084, URL http://www.rssd.esa.int/cs/livelink/open/2915742

[LL-088], Lindegren, L., 2010, *LSF modelling with zero to many parameters*,

GAIA-C3-TN-LU-LL-088,
URL http://www.rssd.esa.int/cs/livelink/open/3042524

[PMN-007], Montegriffo, P., 2014, *A forward model for the external calibration of XP mean spectra. I. The Instrument Update Process.*,
GAIA-C5-TN-OABO-PMN-007,
URL http://www.rssd.esa.int/cs/livelink/open/3255263

Ye, J., Janardan, R., Li, Q., 2004, In: ACM (ed.) Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining, KDD04, Seattle, Washington, USA., 354–363,
URL http://dl.acm.org/citation.cfm?id=1014092

.1 Acronyms

The following is a complete list of acronyms used in this document. The following table has been generated from the on-line Gaia acronym list:

Acronym	Description
AC	Across scan
AF	Astrometric Field (in Astro)
AL	Along scan
ASI	Agenzia Spaziale Italiana
BP	Blue Photometer
DFT	Discrete Fourier Transform
DM	Data Model
FWHM	Full Width at Half-Maximum
INAF	Instituto Nazionale di Astrofisica (Italy)
LSF	Line Spread Function
MDB	Main DataBase
MTF	Modulation Transfer Function
PCA	Principle Component Analysis
PSF	Point Spread Function
PhotPipe	Photometric Pipeline (CU5 / DPCI)
RMS	Root-Mean-Square
RP	Red Photometer
SED	Spectral Energy Distribution
SVD	Singular Value Decomposition
SVN	SubVersioN
TDI	Time-Delayed Integration (CCD)

WFE	WaveFront Error
XP	Gaia photometers BP and RP