



**HAL**  
open science

# Semi-Parametric Wavefront Modelling for the Point Spread Function

Tobias Liaudat, Jean-Luc Starck, Martin Kilbinger

► **To cite this version:**

Tobias Liaudat, Jean-Luc Starck, Martin Kilbinger. Semi-Parametric Wavefront Modelling for the Point Spread Function. 52ème Journées de Statistique de la Société Française de Statistique (SFdS), Jun 2021, Nice, France. hal-03444576

**HAL Id: hal-03444576**

**<https://hal.science/hal-03444576>**

Submitted on 23 Nov 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# SEMI-PARAMETRIC WAVEFRONT MODELLING FOR THE POINT SPREAD FUNCTION

Tobias Liaudat <sup>1</sup>, Jean-Luc Starck <sup>1</sup> & Martin Kilbinger <sup>1</sup>

<sup>1</sup> *AIM, CEA, CNRS, Université Paris-Saclay, Université de Paris,  
F-91191 Gif-sur-Yvette, France*  
{*tobias.liaudat, jean-luc.starck, martin.kilbinger*}@cea.fr

**Résumé.** Nous présentons une nouvelle approche pour estimer le champ de la fonction d'étalement du point d'un télescope optique en construisant un modèle semi-paramétrique de son erreur de front d'onde. Cette méthode est particulièrement avantageuse car elle ne nécessite pas d'observations de calibration pour récupérer l'erreur de front d'onde et elle prend naturellement en compte la chromaticité du système optique. Le modèle est différentiable de bout en bout et s'appuie sur un opérateur de diffraction qui nous permet de calculer les fonction d'étalement du point monochromatiques à partir des informations du front d'onde.

**Mots-clés.** Modélisation de la Fonction d'Étalement du Point, Traitement d'Images, Optique, Lentille gravitationnelle Faible.

**Abstract.** We introduce a new approach to estimate the point spread function (PSF) field of an optical telescope by building a semi-parametric model of its wavefront error. This method is particularly advantageous because it does not require calibration observations to recover the wavefront error and it naturally takes into account the chromaticity of the optical system. The model is end-to-end differentiable and relies on a diffraction operator that allows us to compute monochromatic PSFs from the wavefront information.

**Keywords.** Point Spread Function Modelling, Image Processing, Optics, Weak Lensing.

## 1 Introduction

Future cosmological surveys will require to measure the shape of galaxies with high accuracy. However, the optical instruments used inevitably affect the observations with the point spread function (PSF). If we do not correct the images for the PSF, we will have significantly biased shape measurements resulting in biased cosmological analyses. Thus, the crucial importance of building a reliable and precise PSF model that allows us to take into account the PSF effects in the shape measurement. Next-generation imaging surveys, like the Euclid space mission [14], are pushing the limits of weak gravitational lensing experiments [10] providing the motivation of this study.

There exist two main approaches to PSF modelling, parametric and non-parametric. The first one consists in building an optical model of the telescope. The model is described

by a reduced number of parameters that express a high variability. This approach requires a good knowledge of the optical system physics and normally is based on the reconstruction of the wavefront error (WFE). The recovery of the WFE from in-focus images is a very degenerate problem with respect to the reduced number of parameters used. This explains the need of calibration information like observations of out-of-focus stars that help to break this degeneracy. The model’s parameters can be modified slightly to fit the observation of in-focus stars, but the core of the model is defined previous to the observations. These models are usually reserved for space missions as the randomness added by the atmosphere would make an approach of this type unpractical. The most known example is the Tiny Tim algorithm for the Hubble space telescope [13]. The second approach relies in imaging-data to build the model with a minimal use of *a priori* information. The observed in-focus stars are considered to be samples of PSF field in the field-of-view (FoV) and are used to constrain the model. These techniques are generally based on learning features of the PSF field with a dimensionality reduction method. Then followed by an interpolation method to recover the PSF at galaxy positions [9, 5, 3, 15, 8]. While the parametric models are capable of generating chromatic PSF models with complex shapes they are prone to have considerable errors if there is a mismatch between the model and the observations [7]. The non-parametric models do not suffer from the same issue as they are build on the observations. However, they experience difficulties to model the PSF chromaticity and complex PSF variations in the FoV.

In this work we present a new family of PSF modelling methods that intend to bridge the gap between the two classical approaches. We propose to build a semi-parametric model of the WFE by using an end-to-end differentiable diffraction operator. Our model is able to account for the PSF chromaticity as well as complex variations in the FoV without relying on the *a priori* information of the optical system. The non-parametric part of the WFE is able to correct for the mismatches of the parametric part and help to regularise the inverse problem of recovering the WFE from in-focus observations. We take advantage of the automatic differentiation provided by frameworks like Tensorflow [1] to build our model. As our model is completely differentiable, it is well adapted for gradient-based optimisation techniques. We expect that our model can serve as building block for novel PSF modelling methods that can profit from the power of deep neural networks (DNN) in a more physics-motivated environment.

## 2 Wavefront PSF modelling

Let us define the PSF field for a specific image exposure as a function that inputs a position in the instrument’s focal plane (FP) coordinates and a specific wavelength, and outputs a monochromatic PSF. We denote this function  $\mathcal{H} : F_P \times \mathbb{R}_+ \rightarrow I$ , where  $F_P \in \mathbb{R}^2$  is a FP position, and  $I \in \mathbb{R}^{N \times N}$  is a square PSF postage stamp. PSF modelling consists in building an estimator,  $\hat{\mathcal{H}}$ , of the PSF field from observations at a set of FP positions

$\{u_i\}_{i=1,\dots,m_{obs}}$ . Then, use the model to output monochromatic PSFs at another set of target FP positions  $\{u_j\}_{j=1,\dots,m_{target}}$ . The image  $\mathcal{H}_{u_j}^\lambda$  corresponds to the PSF at location  $u_j$  and wavelength  $\lambda$ . We can model a polychromatic object observation as

$$\bar{G}_{u_j} = \int B(\lambda) (G_{u_j}^\lambda * \mathcal{H}_{u_j}^\lambda) d\lambda, \quad (1)$$

where  $\bar{G}_{u_j}$  is the polychromatic observation of the object  $G_{u_j}^\lambda$  convolved with the PSF  $\mathcal{H}_{u_j}^\lambda$ , and  $B(\lambda) \in \mathbb{R}$  is an indicator function of the instrument's passband. To model the PSF we consider certain star observations as point sources, hence samples of the PSF field. Therefore, we use them to constrain the PSF model  $\hat{\mathcal{H}}$ . The star observations can be approximated as

$$\bar{\mathcal{H}}_{u_i} \approx \sum_{b=1}^{N_\lambda} S_{u_i}(\lambda_b) \mathcal{H}_{u_i}^{\lambda_b}, \quad (2)$$

where we discretize the integral in (1) with  $N_\lambda$  wavelength bins,  $S_{u_i}(\lambda_b) \in \mathbb{R}_+$  is the normalised Spectral Energy Distribution (SED) of the star at position  $u_i$  and wavelength  $\lambda_b$ , and we approximate  $B(\lambda)$  as being an ideal band-pass filter in the instrument's wavelength range. In practice, the problem's inputs are a set of positions  $\{u_i\}_{i=1,\dots,m_{obs}}$  with their SED and the corresponding observed PSF postage stamps. We have to use these inputs to constrain our PSF model that we define as follows

$$\hat{\mathcal{H}}_{u_i}^\lambda \propto \mathcal{D}_\theta \left\{ P \odot \exp \left[ 2\pi i \left( \frac{\Phi_{u_i}}{\lambda} + C_{u_i}^\lambda \right) \right] \right\}, \quad (3)$$

where  $\odot$  is the Hadamard or element-wise product,  $\mathcal{D}_\theta : \mathbb{R}^{p \times p} \rightarrow \mathbb{R}^{N \times N}$  denotes the diffraction operator with telescope-specific parameters  $\theta$ , and  $P \in [0, 1]^{p \times p}$  and  $\hat{\mathcal{H}}_{u_i}^\lambda \in \mathbb{R}^{N \times N}$  are the pupil function and the PSF postage stamp at wavelength  $\lambda$  and FP position  $u_i$ , respectively. The pupil function represents the obscurations encountered in the pupil plane, where our wavefront model is defined. The model definition is proportional as we need to rescale it so that its flux, the sum of the pixels in the postage stamp, is one. We base the diffraction operator on Fraunhofer's approximation [6] that allow us to have a rapid calculation based on the Fast Fourier Transform (FFT) algorithm. The operator is proportional to the squared absolute value of the FFT of the WFE. We take care of the variable sampling in the WFE as a function of wavelength by using a variable zero-padding of the estimated WFE before applying the FFT.

The parametric model,  $\Phi_{u_i}$ , is based on a weighted sum of Zernike polynomials [17]. An interesting property of these polynomials is that they are orthogonal in the unit disk making them well adapted to describe mirror aberrations. Our model reads

$$\Phi_{u_i}[x, y] = \sum_{j=1}^{N_Z} a_{j,u_i} Z_j[x, y], \quad (4)$$

where  $a_{j,u_i} \in \mathbb{R}$  corresponds to the coefficient for the Zernike polynomial of Noll’s single-index  $j$  [17],  $Z_j \in \mathbb{R}^{p \times p}$  to the disk of the Zernike polynomial of index  $j$ , and  $N_Z$  is the maximum index considered. The  $[x, y]$  are coordinates of the pupil plane and are not to be confused with the FP coordinates  $u_i$ . To model the variations on the FP we define  $a_{j,u_i}$  as a polynomial of the FP positions. Each coefficient with Zernike index  $j$  has an independent position polynomial of maximum order  $d_P$ . For example, a maximum degree of 1 gives us  $a_{j,u_i} = c_0^j + c_1^j u_i[0] + c_2^j u_i[1]$ , where  $u_i[0]$  and  $u_i[1]$  are the first and the second coordinates of the FP position, respectively. The number of parameters of  $\Phi$  to estimate is  $N_Z(d_P + 1)(d_P + 2)/2$ . The chromatic variation of the parametric model is encoded in the  $\lambda$  divisor of  $\Phi_{u_i}$  in (3) which corresponds to the natural chromatic variations due to diffraction.

The non-parametric model,  $C_{u_i}^\lambda$ , is defined with a simple, but successful, model in this work. We use a matrix factorisation scheme where each FP position is a weighted sum of learned WFE features. The weights are constructed as polynomials of the FP position with a maximum degree  $d_{NP}$  providing  $r$  coefficients. The model reads

$$C_{u_i}^\lambda = \frac{1}{\lambda} (\Pi_{u_i} \times_1 S)_{i_2, i_3} = \frac{1}{\lambda} \sum_{i_1=1}^r (\Pi_{u_i})_{i_1} S_{i_1, i_2, i_3}, \quad (5)$$

where  $C_{u_i}^\lambda \in \mathbb{R}^{p \times p}$ ,  $\times_1$  is the 1-mode product of a tensor with a vector [11],  $S \in \mathbb{R}^{r \times p \times p}$  is the tensor containing  $r$  wavefront feature images, and  $\Pi_{u_i} \in \mathbb{R}^{1 \times r}$  contains the position polynomials. For example, if we set  $d_{NP}$  to 1 we obtain  $\Pi_{u_i} = [1, u_i[0], u_i[1]]$ . For simplicity, we only consider chromatic variations due to diffraction but the model allows to include more complex models.

## 3 Numerical experiment

### 3.1 Data

We simulate a polychromatic PSF field using the proposed model seen in (3) with its non-parametric part set to zero. The optical parameters are taken from Euclid’s visible instrument characteristics [14]. We draw 200 uniformly distributed positions in the square FoV, where 70% are used for training the model and the rest for testing it. Each star is randomly assigned one of 13 stellar SED from the Pickles library [18] as done in [19, §5.3]. The maximum order of Zernike polynomial,  $N_Z$ , is set to 45, the number of wavelength bins,  $N_\lambda$ , to 20, the polynomial maximum degree,  $d_P$ , to 2, and the dimensions  $p$  and  $N$  to 256 and 32, respectively. The instrument’s passband is considered as ideal in the range of 550 nm to 900 nm. We draw random values for the coefficients of the Zernike polynomials with the constraint that of having a RMS value of 0.1  $\mu\text{m}$  to limit the optical system’s aberrations. This allows to have a randomly varying PSF field. Finally, we add a random white Gaussian noise to each training star so that each one has a random SNR value

in the range  $[10, 70]$ . We use an Euclid-like obscuration that is composed of a circular centred obscuration with three supporting arms that can be seen in the study [20].

## 3.2 Experience

The PSF field modelling problem consists in estimating the simulated test stars having as input the noisy training stars. The SED information of both is available. We train and compare three models: i) Parametric model with  $N_Z = 15$  and  $d_P = 2$ ; ii) Parametric model with  $N_Z = 45$  and  $d_P = 2$ ; iii) Semi-parametric model with  $N_Z = 15$ ,  $d_P = 2$ , and  $d_{NP} = 3$ . The objective function used to train the model is the Mean Square Error (MSE) between the observed stars and the PSF model reconstruction. All the models use a gradient-based optimization method based on the Rectified Adam optimiser [16] with a batch size of 16. Both parametric models use a learning rate of  $10^{-2}$  and a number of epochs of 30. For the semi-parametric model, we first train the parametric part using a learning rate of  $10^{-2}$  for 15 epochs with the non-parametric part set to zero. Then we fix the parametric part and start to train the non-parametric part using a learning rate of  $1^{-1}$  during 100 epochs. The evaluation metric we use is the root mean square error (RMSE) between the test stars and the model estimations. Let us point out that our implementation follows the Tensorflow 2.1 framework [1] but it is used as an automatic differentiation library to perform optimisation.

## 3.3 Results

Table 1 presents the quantitative results. We see that the parametric model with  $N_Z = 15$  is the worst performing, as expected, due to a misspecification of the parametric model with respect to the underlying data model. The parametric model with  $N_Z = 45$  has a perfect match with the data model, and we observe a 35% performance gain with respect to the previous one. However, the interesting part is that the semi-parametric model with a wrongly-specified parametric part is able to obtain the best performance. Thanks to its non-parametric part the proposed model can recover the performance gap between the two parametric models, and even go beyond by obtaining a gain of almost 65% with respect to the parametric  $N_Z = 45$  model.

The estimation of the Zernike coefficients using in-focus stars is a difficult problem, known to be ill-posed, and possibly presents multi-modalities. Even if we use all the training stars in our data-set we are prone to get stuck at a local minimum due to the non-convexity of the problem. The non-parametric part of the model acts as a regularizer and helps to escape from local minima by providing an easier optimization landscape. This behaviour has been observed in the optimization of overparametrized DNN and is currently a subject of study [2]. The metrics are calculated on the test stars that are not available for the training of the model, showing that the non-parametric part is not over-fitting the training stars.

Model	RMSE	Relative RMSE
Parametric $N_Z = 15$	$8.69 \times 10^{-4}$ ( $7.90 \times 10^{-4}$ )	11.4% (9.9%)
Parametric $N_Z = 45$	$5.59 \times 10^{-4}$ ( $4.95 \times 10^{-4}$ )	7.3% (6.2%)
Semi-parametric $N_Z = 15, d_{NP} = 3$	$2.03 \times 10^{-4}$ ( $1.71 \times 10^{-4}$ )	2.6% (2.1%)

Table 1: Pixel RMSE results on the test dataset, in parenthesis for the train dataset, for the different PSF models.

## 4 Discussion and conclusions

We have presented the building blocks that open the way to a brand-new family of PSF models that aims to improve the two classical families, which are the parametric and the non-parametric methods. The proposed semi-parametric approach allows us to reconstruct the wavefront error and derive the chromatic point spread function at any position in the field-of-view. One of the novelties of our approach is that it uses only in-focus stars to characterise the wavefront error. This is in contrast with other wavefront-based models [4, 12] that need out-of-focus stars, sometimes referred as donuts, to estimate its parameters. Nonetheless, our model could also profit from these rich observations if they are available.

The proposed PSF model provides a way to estimate the PSF field taking into account the PSF chromaticity and field-of-view variations. The semi-parametric approach is able to learn features that a parametric model cannot capture. The non-parametric part of the model corrects for a mismatch between the parametric model and the underlying observed data model. It also helps to regularize the challenging ill-posed inverse problem of estimating the wavefront error from in-focus stars. We show on a realistic data-set that it can even improve the results over a perfectly specified parametric model and that the model is not over-fitting the testing data. Our model is based on an end-to-end differentiable approach which allows for gradient-based optimization techniques which, in practice, give good results with the semi-parametric approach. Its framework allows us to easily incorporate known physical models, while the non-parametric part learns to correct for its imperfections. In this way, we can continuously incorporate physical knowledge to the model.

## References

- [1] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mané, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Tal-

- war, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viégas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu, and X. Zheng. TensorFlow: Large-scale machine learning on heterogeneous systems, 2015. Software available from tensorflow.org.
- [2] Z. Allen-Zhu, Y. Li, and Y. Liang. Learning and generalization in overparameterized neural networks, going beyond two layers. In H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 32, pages 6158–6169. Curran Associates, Inc., 2019.
- [3] E. Bertin. Automated Morphometry with SExtractor and PSFEx. In I. N. Evans, A. Accomazzi, D. J. Mink, and A. H. Rots, editors, *Astronomical Data Analysis Software and Systems XX*, volume 442 of *Astronomical Society of the Pacific Conference Series*, page 435, July 2011.
- [4] C. P. Davis, J. Rodriguez, and A. Roodman. Wavefront-based PSF estimation. In H. J. Hall, R. Gilmozzi, and H. K. Marshall, editors, *Ground-based and Airborne Telescopes VI*, volume 9906, pages 2156 – 2168. International Society for Optics and Photonics, SPIE, 2016.
- [5] M. Gentile, F. Courbin, and G. Meylan. Interpolating point spread function anisotropy. *A&A*, 549:A1, 2013.
- [6] J. W. Goodman. Introduction to fourier optics. *Introduction to Fourier optics, 3rd ed.*, by JW Goodman. Englewood, CO: Roberts & Co. Publishers, 2005, 1, 2005.
- [7] S. L. Hoffmann and J. Anderson. A Study of PSF Models for ACS/WFC. Instrument Science Report ACS 2017-8, Oct. 2017.
- [8] M. Jarvis et al. Dark Energy Survey Year 3 Results: Point-Spread Function Modeling. 11 2020.
- [9] M. J. Jee, J. P. Blakeslee, M. Sirianni, A. R. Martel, R. L. White, and H. C. Ford. Principal component analysis of the time- and position-dependent point-spread function of the advanced camera for surveys. *Publications of the Astronomical Society of the Pacific*, 119(862):1403–1419, dec 2007.
- [10] M. Kilbinger. Cosmology with cosmic shear observations: a review. *Reports on Progress in Physics*, 78(8):086901, 2015.
- [11] T. G. Kolda and B. W. Bader. Tensor decompositions and applications. *SIAM Review*, 51(3):455–500, 2009.
- [12] J. E. Krist and C. J. Burrows. Phase-retrieval analysis of pre- and post-repair hubble space telescope images. *Appl. Opt.*, 34(22):4951–4964, Aug 1995.



- [13] J. E. Krist, R. N. Hook, and F. Stoehr. 20 years of Hubble Space Telescope optical modeling using Tiny Tim. In M. A. Kahan, editor, *Optical Modeling and Performance Predictions V*, volume 8127, pages 166 – 181. International Society for Optics and Photonics, SPIE, 2011.
- [14] R. Laureijs, J. Amiaux, S. Arduini, J.-L. Augueres, J. Brinchmann, R. Cole, M. Cropper, C. Dabin, L. Duvet, A. Ealet, et al. Euclid definition study report. *ArXiv e-prints*, 2011.
- [15] Liaudat, T., Bonnin, J., Starck, J.-L., Schmitz, M. A., Guinot, A., Kilbinger, M., and Gwyn, S. D. J. Multi-ccd modelling of the point spread function. *A&A*, 646:A27, 2021.
- [16] L. Liu, H. Jiang, P. He, W. Chen, X. Liu, J. Gao, and J. Han. On the variance of the adaptive learning rate and beyond, 2020.
- [17] R. J. Noll. Zernike polynomials and atmospheric turbulence\*. *J. Opt. Soc. Am.*, 66(3):207–211, 1976.
- [18] A. J. Pickles. A stellar spectral flux library: 1150 - 25000 a. *Publications of the Astronomical Society of the Pacific*, 110(749):863–878, jul 1998.
- [19] M. A. Schmitz. *Euclid weak lensing : PSF field estimation*. Theses, Université Paris Saclay (COMUE), Oct. 2019.
- [20] L. M. G. Venancio, L. Carminati, J. Amiaux, L. Bonino, G. Racca, R. Vavrek, R. Laureijs, A. Short, T. Boenke, and P. Strada. Status of the performance of the Euclid spacecraft. In M. Lystrup, M. D. Perrin, N. Batalha, N. Siegler, and E. C. Tong, editors, *Space Telescopes and Instrumentation 2020: Optical, Infrared, and Millimeter Wave*, volume 11443, pages 45 – 60. International Society for Optics and Photonics, SPIE, 2020.