Fast Active Optics Control of Wide-Field Telescopes based on Science Image Analysis

R. Holzlöhner[a], A. Rakich[a], L. Noethe[a], K. Kuijken[b] and P. Schipani[c]

[a]European Southern Observatory (ESO), Karl-Schwarzschild-Str. 2, D-85748 Garching, Germany;
[b]Leiden Observatory, Niels Bohrweg 2, NL-2333 CA Leiden, the Netherlands;
[c]Osservatorio Astronomico di Capodimonte, Salita Moiariello 16, I-80131 Naples, Italy

ABSTRACT

We study a novel active optics control scheme at the VST on Cerro Paranal, an $f/5.5$ survey telescope with a 1×1 degree field of view and a 2.6 m primary mirror. This scheme analyzes the elongation pattern of the star PSFs across the full science image (267 Mpixels) and compares their second moments with an analytical model based on 5th-order geometrical optics, comprising 9 degrees of freedom in mirror misalignments and deformations. Using a numerical optimization method, we can complete the star extraction and fitting process in under one minute, fast enough for effective closed-loop active optics control in survey observing cadences.

Keywords: active optics, wide-field telescopes, point spread function, elongation, aberration, optical plate diagram

1. INTRODUCTION

Closed-loop active optics (ActO) control of wide-field telescopes is demanding both because of the tight alignment tolerances in fast optics, and also because the field dependence of aberrations in misaligned wide-field systems places tighter constraints on various system degrees of freedom than is the case with more conventional narrow field-of-view telescopes.

We report on studies carried out at the VST on Cerro Paranal, a modified Ritchey-Chrétien $f/5.5$ survey telescope with a 1×1 degree field of view and a 2.6 m primary mirror (M1), currently the largest telescope designed to exclusively survey the skies in visible light.[1] M1 consists of Astrositall, is 140 mm thick and has a conic constant of -1.14. It has 84 axial supports which are push-only and non-astatic and three of which are hard points with actuators.[2] The secondary mirror has a diameter of 0.9 m and is mounted on a hexapod. The mirrors are followed by two spherical meniscus lenses made of silica, a filter and a curved dewar window. The telescope totals ten optical surfaces until the surface of the detector. The VST also possesses an atmospheric dispersion compensator (ADC) which, however, has not been used in our studies.

The VST has only one instrument which is installed on the Cassegrain focus, the high-resolution visible camera OmegaCAM.[3] The science detector is an 8×4 mosaic of $2\text{k}\times4\text{k}$ visible blue-sensitive thinned CCDs by e2v with a pixel scale of 0.21".[4] On either side of the quadratic science field, there are two technical chips, namely a guider chip and an out-of-focus chip (2 mm above the focal plane on one side and 2 mm below it on the other). These chips are employed for curvature sensing (donut method) to control the active optics in closed loop, independently of the science exposure.[5,6] This scheme has proven to be robust, but it suffers from some fundamental limitations due to sparse coverage of the focal plane and due to detector and atmospheric noise.

One of the science drivers is dark energy search through gravitational weak lensing analysis (KiDS survey[7,8]) which would benefit from uniform high-quality point-spread functions (PSFs) across the field in multiple ways:

Sharpness:

- separating stars from galaxies becomes easier

Further author information: (Send correspondence to R.H. E-mail: rholzloe[at]eso.org, Tel. +49 89 3200 6671)

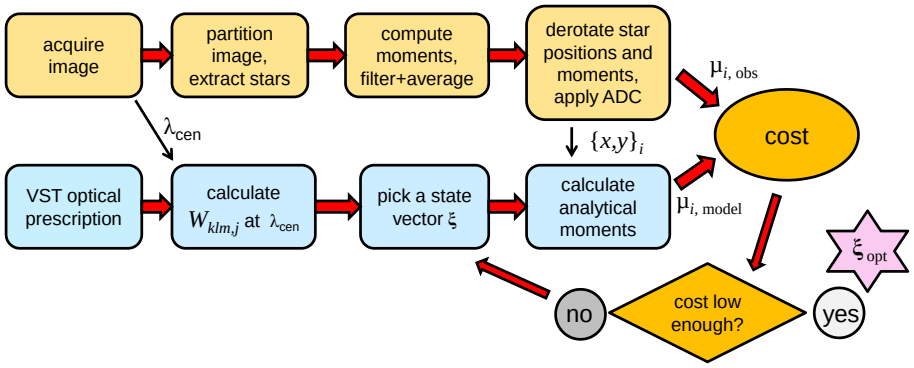

Figure 1. Flow diagram of the presented method to deduce telescope misalignments and mirror deformations from the science image, resulting in the telescope aberration state vector $\boldsymbol{\xi}_{\rm opt}$. The steps are described in Section 2.

- faint galaxies are very small, so distinguishing them from stars requires good image quality (IQ) — at a given integration time one can see fainter sources when the IQ is better (less sky noise in the image area)

Low ellipticity:

- KiDS measures the weak gravitational lensing effect, which coherently distorts galaxies with a shear. This manifests itself as a slight average ellipticity on the same scale as the field of view of the telescope, making it hard to distinguish from telescope aberrations. If the PSF is elliptical, a coherent ellipticity effect is caused which is not straightforward to correct on noisy galaxy images. The smaller the PSF ellipticity, the less correction needs to be applied and the more robustly the weak lensing signal can be detected.
- in fact a smaller PSF helps here as well, as it means the observed shape of each galaxy is closer to its true one (again a correction that needs to be applied, and the smaller this effect the better)

Our goal is to supplement the current ActO control scheme by a real-time analysis of the PSF elongation pattern in the full science image (550 MB), characterizing all relevant optical aberrations with their field dependence, while smoothing the atmospheric turbulence with typical exposure times of one to a few minutes.

We model the aberrations using an analytical 5th-order geometrical optics approximation, with coefficients derived using an extended version of the Optical Plate Diagram. This approach yields the relationship between mirror displacements/deformations and the wavefront error in the exit pupil and its associated field dependence. We then derive the second moments of the seeing-free PSFs across the field, which we express as three ellipse parameters (the major and minor axes plus the orientation angle). We consider nine scalar degrees of freedom in the telescope perturbation (M1 figure astigmatism and trefoil with two lateral components each, M2 $x/y/z$ vertex offset and M2 tip/tilt).

2. METHOD

Figure 1 shows a flow diagram of our proposed method. The upper row of boxes with the beige background pertains to the observed image analysis and the lower row to the analytical model. The rest of this section describes the individual steps denoted by the diagram items.

2.1 Processing of the Observed Image

2.1.1 Star extraction and star moment computation

The image analysis commences by partitioning the total frame of 16k×16k pixels into up to 20×20 equal tiles and identifying the 5–8 brightest objects in each tile. Objects that lie too close to another object or are too close to the image boundary, or those with saturated pixels, are rejected. Selecting only the brightest objects maximizes the sign-to-noise ratio with respect to shot noise and sky background, but also tends to favor stars

versus extended objects such as galaxies. Galaxies shall not be used for our method because besides having an unknown intrinsic elongated shape, they also tend to be subject to weak lensing due to their large distance from Earth. It is thus instructive to think of bright stars as foreground beacons and galaxies as background sources.

We thus extract up to 3200 stars, uniformly covering the field. The next step consists of estimating the three normalized second central moments of the star images $I(x, y)$ defined as

$$\mu_{ij}^{(2)} = I_{\text{tot}}^{-1} \int (x - x_0)^{\,i} (y - y_0)^{\,j} I(x, y)\, dxdy, \quad i, j \in \{0,\, 1,\, 2\},\ \ i + j = 2, \tag{1}$$
$$x_0 = I_{\text{tot}}^{-1} \int x\, I(x, y)\, dxdy, \quad y_0 = I_{\text{tot}}^{-1} \int y\, I(x, y)\, dxdy, \quad I_{\text{tot}} = \int I(x, y)\, dxdy,$$

where x_0, y_0 denote the image centroid and I_{tot} is the total flux. If I/I_{tot} is considered a probability distribution then the $\mu_{ij}^{(2)}$ equal the covariances of x and y. We will henceforth assume that all extracted objects pertain to (point-like) stars and that

$$I(x, y)/I_{\text{tot}} = g(x, y, \lambda) * s(\lambda), \tag{2}$$

where the right hand side is the convolution of $g(x, y)$, the geometrical optics telescope point spread function (PSF) at the star location (x, y) and wavelength λ, and s, the seeing PSF, which is centrosymmetric (isotropic in the field) and independent of (x, y). More precisely, s is the long-exposure star image in an ideal 2.6 m aberration-free telescope at λ and thus is also influenced by telescope diffraction. The telescope PSF in the VST has a non-negligible dependence on wavelength due to the chromatic glass elements. However, setting aside chromaticity for the moment, we thus assume that the flux normalized image of a star, and in particular the moments $\mu_{ij}^{(2)}$, only depend on the telescope aberrations at the field point(x, y) and on a centrosymmetric and field constant seeing blur. We will drop the superscript in $\mu_{ij}^{(2)}$ from now on.

Equation (2) possesses the convenient property that the second central moments of $g * s$ equals the sum of the second central moments of g and s,[9] a relationship that we will apply extensively in this work. *

While the accurate computation of (1) is trivial in the absence of noise, it becomes quite involved when adding sky background and detector noise and is a subject of current research.[10] The problem is that the magnitude of $(x - x_0)^{\,i} (y - y_0)^{\,j}$ grows quadratically with distance from the star centroid so that the integrand becomes increasingly susceptible to noise. A widely adopted approach consists of multiplying $I(x, y)$ by a bell-shaped weighting function to reduce the uncertainty of the integration result. In this work, we have employed an elliptical rotated Gaussian weighting function whose second moments equal the current best estimate of μ_{ij}. The estimate of μ_{ij} with updated weighting is then iterated until convergence occurs. Details will be published elsewhere.

We have implemented the star extraction algorithm, as the rest of the numerical code described below, in Mathematica v.9. We note that we have first considered employing the widely used freeware utility SExtractor,[11] but were dissatisfied with its implementation of the moments calculation. In particular, the μ_{ij} computed in SExtractor (X2WIN, Y2WIN and XWWIN) must be multiplied by the factor 2 and are strongly correlated with I_{tot}. We have posted this issue on the Astromatic online forum.

2.1.2 Filtering, derotation, ADC

We reject stars in the moment computation whose total flux is below a certain threshold and those whose moments deviate too strongly from the statistical distribution of the μ_{ij}. Tiles containing less than three valid stars are considered vacant and are be disregarded in the further process. In sky regions with sparse star coverage, it may be advantageous to reduce the number of tiles below 20×20.

In order to mitigate the uncertainty due to the chromatic variation of the star images and due to noise, we take the median of the μ_{ij} within each tile. The median is a statistically more robust function than the mean value with respect to outliers such as galaxies or nearly overlapping stars that were mistaken for a single object.

*The theorem does not require that the function s be Gaussian or near-Gaussian. It is not even necessary that s be centrosymmetric, however, if s becomes asymmetric, its second central moments in x and y will differ in general and hence $\mu_{20}^{(2)} - \mu_{02}^{(2)}$ will depend on s, which compromises our method.

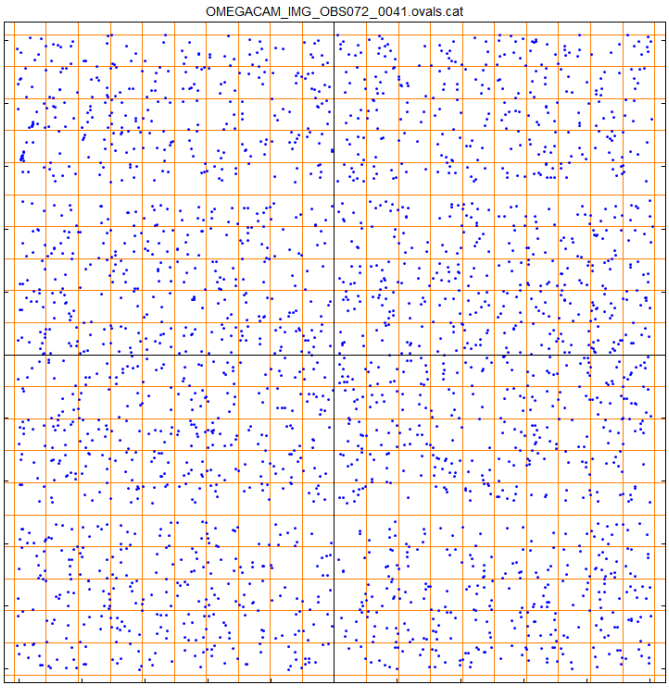

Figure 2. Star locations in an extracted image. The tile boundaries are denoted by the orange lines. At about 1/4 and 3/4 of the image diameter from the bottom, there are rectangular horizontal regions without stars due to the two readout lines in the CCD mosaic.

We also take the median of the valid star centroids within each tile as its average field point instead of the geometric tile center. An example of the extracted star locations in a typical OmegaCAM image is shown in Fig. (2); the tile boundaries are indicated by the orange grid lines.

OmegaCAM tracks the sky such that the top of the detector always points towards North. In order to transform from image coordinates to telescope coordinates, where the vertical axis points toward zenith when the telescope is pointing at the horizon, we begin by derotating all average star centroids about the image center (variable ABSROT in OmegaCAM FITS files). Since a Cassegrain telescope additionally inverts the image, we rotate by the total angle of $\psi = \mathrm{ABSROT}{+}180^{\mathrm{o}}$. Finally, we invert the sign of the x coordinate to convert to a right-hand Cartesian coordinate system in which the z-axis points from M2 towards the detector, y grows from the ground to zenith (telescope pointing to horizon) and x increases towards the left when looking along the z-axis. This convention has the advantage that it agrees with the coordinate system in Zemax, a very convenient circumstance when validating our analytical optical model with ray tracing results. Besides the star locations, we also rotate the covariance matrices as follows

$$\begin{bmatrix} x \\ y \end{bmatrix}_{\mathrm{tel}} = \mathbf{MR} \begin{bmatrix} x \\ y \end{bmatrix}_{\mathrm{det}}, \quad \begin{bmatrix} \mu_{20} & \mu_{11} \\ \mu_{11} & \mu_{02} \end{bmatrix}_{\mathrm{tel}} = \mathbf{MR} \begin{bmatrix} \mu_{20} & \mu_{11} \\ \mu_{11} & \mu_{02} \end{bmatrix}_{\mathrm{det}} \mathbf{R}^t \mathbf{M}^t, \tag{3}$$

$$\mathbf{R} = \begin{bmatrix} \cos\psi & -\sin\psi \\ \sin\psi & \cos\psi \end{bmatrix}, \qquad \mathbf{M} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}, \tag{4}$$

where $\mathbf{R}$ is a rotation matrix by the angle ψ and $\mathbf{M}$ applies a mirror operation about the y-axis (note that the two operations do not commute). The indices tel and det denote telescope and detector coordinates, respectively. In the remainder of this paper, we will only work in telescope coordinates and drop the tel subscript.

Once the second moments are rotated into the telescope frame in which the y-axis is directed orthogonal to the horizon, we remove the influence of atmospheric dispersion (analytical atmospheric dispersion compensation,

ADC) from the covariance matrix by correcting μ_{02}, while μ_{20} and μ_{11} remain unchanged

$$\begin{aligned} \mu'_{02} &= \mu_{02} - \gamma^2\delta^2, & (5)\\ \delta &= (n(\lambda_{\min},p,T) - n(\lambda_{\max},p,T))\tan\zeta, & (6)\\ \gamma &= \frac{\sqrt{\mu_{\rm filter}}}{\lambda_{\max}-\lambda_{\min}} \approx 0.294, & (7) \end{aligned}$$

where δ is the atmospheric angular dispersion, $n(\lambda,p,T)$ is the path-averaged refractive index at wavelength λ, air pressure p and temperature T (the latter two measured on the ground), ζ is the zenith angle of the telescope pointing, $\lambda_{\min}$ and $\lambda_{\max}$ are the corner wavelengths where the filter reaches 50% if its peak transmission, and $\mu_{\rm filter}$is the second moment of the filter transmission profile, which we model as a fourth-order super-Gaussian (for a standard Gaussian, $\gamma \approx 0.425$ and for a top hat profile, $\gamma \approx 0.289$). Strictly speaking, by filter profile we mean the chromatic transmission curve of the filter plus the rest of the telescope times the quantum efficiency curve of the detector. We use an approximation of $n(\lambda,p,T)$ from[12] that is also implemented in the VST and VLT control software.[13] We note that δ can become significant at large zenith angles and/or short wavelength, e.g. $\delta = 0.7\,$arcsec at $\zeta = 45^\circ$ when using the green SDSS filter ($\lambda = 408\ldots 542\,$nm).

2.2 Analytical Model

2.2.1 Some aberration theory

Our goal is to deduce the telescope aberration state from the star images. Starting from the optical prescription of the unperturbed telescope, the first step is to find the relationship between misalignments and deformations of the optical elements (henceforth called the telescope perturbation state $\boldsymbol{\xi}$) and the wavefront error in the exit pupil. In a second step, we will describe the effect of the wavefront error on the μ_{ij}.

We follow Hopkins' [14] expansion of the wavefront error $W = W(\rho,H)$, where ρ is the normalized pupil coordinate ($0 \le |\rho| \le 1$) and H is the normalized field coordinate. For convenience, we choose to express both $\rho = r\,\exp(i\varphi)$ and $H = (x + i\,y)/\alpha$ as complex numbers (i is the imaginary unit here), where $\alpha = 0.509^\circ$ is half the field of view of the VST. The total wavefront error of a rotationally symmetric optical system with a (near) circular aperture can then be written as a sum over all optical surfaces j [15,16]

$$\begin{aligned} W &= W\left(H\cdot H,\,\rho\cdot H,\,\rho\cdot\rho\right)\\ &= \sum_j\sum_p^\infty\sum_n^\infty\sum_m^\infty (W_{klm})_j\,(H\cdot H)^p\,(\rho\cdot H)^n\,(\rho\cdot\rho)^m, \qquad k=2p+m,\ l=2n+m, \end{aligned} \tag{8}$$

where the expression $a\cdot b = b\cdot a = \mathrm{Re}(a\,b^*) \neq a\,b$ denotes the scalar product between two complex numbers a and b and the asterisk means complex conjugation. In (8), the wavefront error contribution originating from surface j is expanded in polynomials which, due to rotational symmetry, only depend on the three scalars $H\cdot H = |H|^2$, $\rho\cdot\rho = |\rho|^2$ and $\rho\cdot H = |\rho|\,|H|\cos\theta$, with θ being the angle between H and ρ in the complex plane. The $(W_{klm})_j$ are real coefficients whose derivation we will describe in the next subsection. We ignore piston, tip and tilt of the wavefront and thus the lowest order term that we consider is defocus $W_{020}(\rho\cdot\rho)$. The five so-called 3rd-order aberrations (meaning $k+l-1=3$) are

$$W^{(3)} = \underbrace{W_{040}(\rho\cdot\rho)^2}_{\text{spherical abberation}} + \underbrace{W_{131}(\rho\cdot H)(\rho\cdot\rho)}_{\text{coma}} + \underbrace{W_{222}(H\cdot\rho)^2}_{\text{astigmatism}} + \underbrace{W_{220}(H\cdot H)(\rho\cdot\rho)}_{\text{field curvature}} + \underbrace{W_{311}(H\cdot H)(H\cdot\rho)}_{\text{displacement}}. \tag{9}$$

In the 1970s, Buchroeder, Shack and Thompson[17] realized that the expression in (8) can be generalized for nonzero decenters and tilts of the optical surfaces by introducing the aberration field centers σ_j, projected in the image plane, that shift the field coordinate at each surface according to $H_{Aj} = H - \sigma_j$. Those shifts break the degeneracy of the multiple zeros at the field center by introducing aberration field nodes; for example one can then generalize the coma term in (9) to $W_{131}\,[\rho\cdot(H-a_{131})]\,(\rho\cdot\rho)$ with the 3rd-order coma node a_{131}. The details can be found in.[15] The aberration nodes are themselves shifted in the focal plane by the vector $H_{\rm ref}$ which is the field point where the principal ray of the axial pencil traced through the displaced optical system

intersects with the focal plane. We compute $H_{\rm ref}$ using paraxial ABCD matrices, where all traced angles and ray heights are expressed as complex numbers in the $x + i\,y$ plane.

The Hopkins wavefront error expansion can be complemented by the addition of field-independent aberrations that arise directly in the pupil (which is the primary mirror in our case); specifically we define

$$W_{\rm pupil} = \underbrace{W_{040,\rm M1}(\rho\cdot\rho)^2}_{\text{M1 spherical abb.}} + \underbrace{(W_{022}\cdot\rho^2)/2}_{\text{M1 astigmatism}} + \underbrace{(W_{033}\cdot\rho^3)/4}_{\text{M1 trefoil}}, \tag{10}$$

where the effect of M1 spherical aberration is identical to the first term in (9) and the difference lies only in the way we derive the coefficients. However, the M1 astigmatism term in (10) is different from the astigmatism term in (9) in that it does not contain a field dependence. Moreover, the coefficients W_{022} and W_{033} are complex and their real and imaginary parts are four elements of the telescope perturbation vector $\boldsymbol{\xi}$, in contrast to the other W_{klm} that are real and not part of $\boldsymbol{\xi}$ (note also that $\rho^2 \neq \rho\cdot\rho$ and ρ^3 are complex numbers).

2.2.2 Optical Plate Diagram

We employ the Optical Plate Diagram to analytically calculate the coefficients W_{klm}. Formulated by C. R. Burch in the 1940s,[18] it is a particularly elegant approach to the third-order analysis of optical systems. In setting up the Plate Diagram a system of Schmidt-like plates (plates that contribute spherical aberration without contributing optical power) is used to replace all spherical or aspheric surfaces in the optical system; the resultant system has identical third-order characteristics to the original system but it does not focus the light; so collimated light at the input will generate collimated light out (to the first-order). A summary of equations for the third-order surface coefficients and aberration field centers, yielding the third-order field nodes defined by Thompson[15] is given in.[19] The W_{klm} derived in the latter reference are valid for aspheric surfaces with arbitrary conic constants.

When comparing spot diagrams generated at different field points from Zemax with our analytical third-order model, we found unsatisfactory agreement. We thus extended the model to the next order and added the seven fifth-order terms with coefficients $W_{242}, W_{422}, W_{240}, W_{420}, W_{331}, W_{151}, W_{060}$. In this work, we have obtained the numerical values of the $(W_{klm})_j$ by comparison with Zernike wavefront error terms and also with analytical coefficients from a macro, both evaluated by Zemax. We further set all fifth-order field nodes to zero. With this approximation, we find that our model agrees with the wavefront error computed in Zemax to within 20 nm at the field corners ($\lambda/30$). In future development, we will employ analytical formulas to calculate the fifth-order surface coefficients[20] and derive from those the field nodes.[21] Figures 3–5 show comparisons between spot diagrams generated by nonseqential Zemax with about 160,000 rays (left) and by our analytical method (right).

2.2.3 Spot Ellipses

Once we have derived an analytical expression for $W(\rho, H)$, we can calculate the analytical approximation $\mu_{ij,\rm model}$ to the extracted μ_{ij} by averaging the complex wavefront gradient $\nabla W = dW/d\rho$ over the normalized pupil annulus as follows

$$p_{\rm model} = y_{\rm Ep}^{-2}\left\langle \left|\nabla W - \langle\nabla W\rangle\right|^2 \right\rangle = \mu_{20,\rm model} + \mu_{02,\rm model}, \tag{11a}$$

$$q_{\rm model} = y_{\rm Ep}^{-2}\left\langle \left(\nabla W - \langle\nabla W\rangle\right)^2 \right\rangle = \mu_{20,\rm model} - \mu_{02,\rm model} + 2i\,\mu_{11,\rm model}, \tag{11b}$$

$$\langle u\rangle = \left(\pi(1-\eta^2)\right)^{-1} \int_\eta^1 \int_{-\pi}^{\pi} u(r,\varphi)\, d\varphi\, r dr, \tag{11c}$$

where $y_{\rm Ep}$ is the radius of the entrance pupil ($y_{\rm Ep} = 1305$ mm in the VST), η is the ratio of the radius of the central pupil aberration due to the shadow of the M2 mirror cell onto M1 to $y_{\rm Ep}$ (in the VST $\eta = 0.461$) and u is an arbitrary function defined in the polar pupil coordinates (r, φ). About sign conventions and units: $W(\rho, H)$ denotes the distance in meters by which the exit wavefront of a parallel incoming wavefront arriving from the (normalized) direction H is advanced at the normalized exit pupil point ρ with respect to the unaberrated

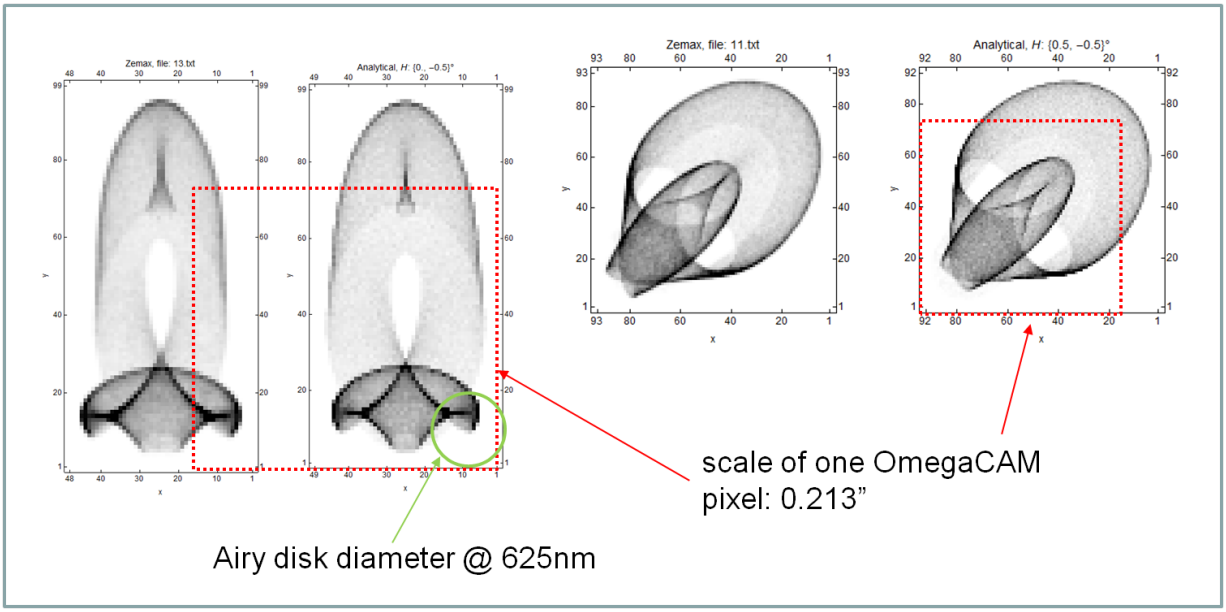

Figure 3. Comparison of spot diagrams from nonseqential Zemax (left) and our analytical method (right) at the center of the field edge and field corner for the well-collimated telescope. Red dotted square: VST pixel size.

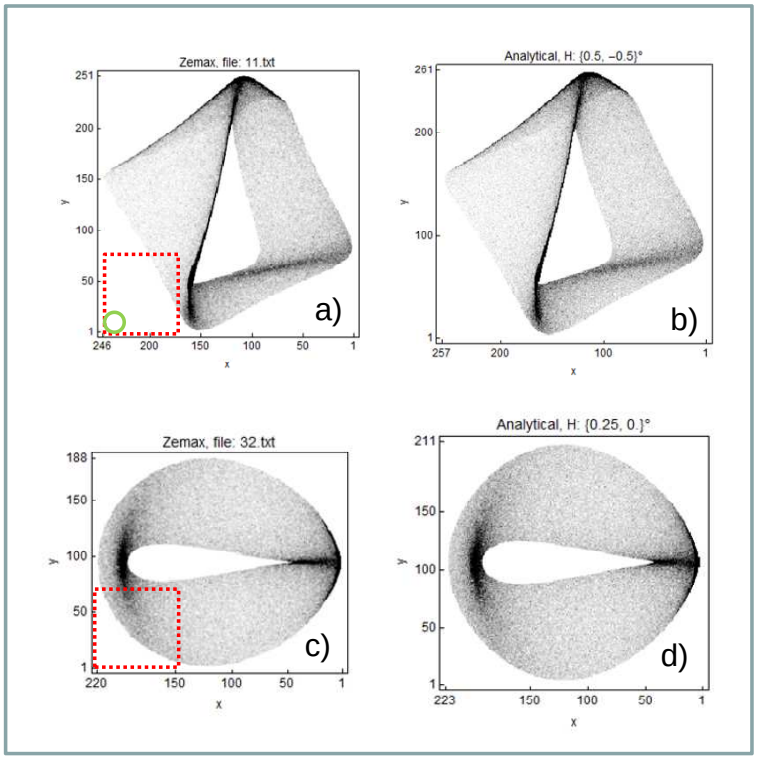

Figure 4. Comparison of spot diagrams from nonseqential Zemax (left) and our analytical method (right) at the field center (top) and field corner (bottom) for 400 nm rms of Zernike Z6 (astigmatism). Red dotted square: VST pixel size.

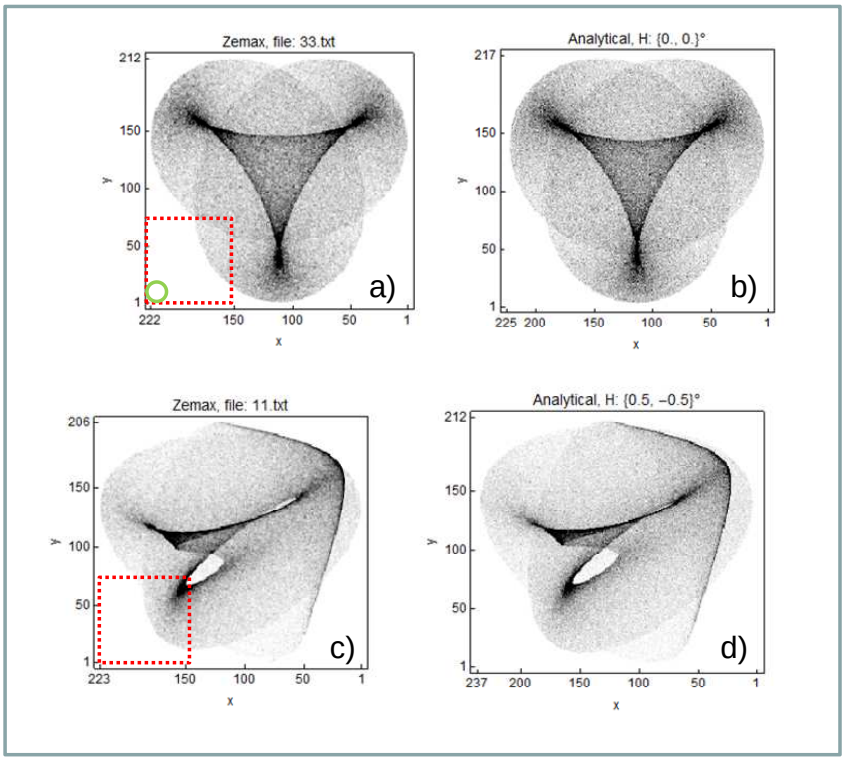

Figure 5. Comparison of spot diagrams from nonseqential Zemax (left) and our analytical method (right) at the field center (top) and field corner (bottom) for 200 nm rms of Zernike Z9 (trefoil). Red dotted square: VST pixel size.

wavefront and thus p and q have dimensions of angle squared. The term $-(1/y_{\mathrm{Ep}})\nabla W$ yields the angle between the wavefront normal and the optical axis and $-(1/y_{\mathrm{Ep}})\,\langle\nabla W\rangle$ represents the total centroid displacement of a star. The analytical expression for the gradient in the complex plane are given in,[15,16] Eq. (A14). The integrals in (11a) and (11b) can be solved analytically and yield a weighted sum of products of pairs of the W_{klm}. We will present the detailed expressions elsewhere.

Starting from the observed science image and the extracted μ_{ij} instead of the analytical model, we can define $p_{\mathrm{obs}} = \mu_{20} + \mu_{02}$ and $q_{\mathrm{obs}} = \mu_{20} - \mu_{02} + 2i\,\mu_{11}$ analogously to (11a) and (11b).

The quantities p and q, which may also be interpreted as the invariants trace and (complex valued) determinant of the covariance matrix, respectively, are related to the spot ellipse parameters as

$$\sigma_l^2 = \frac{p + |q|}{2}, \quad \sigma_s^2 = \frac{p - |q|}{2}, \quad \phi = \frac{\arg(q)}{2} = \frac{1}{2}\arctan\frac{\Im q}{\Re q}, \tag{12}$$

where σ_l and σ_s are the rms major and minor ellipse half axes and $\phi \in \{-\pi/2, \pi/2\}$ is the ellipse orientation angle in the field.

2.3 Cost Function

As shown in Fig. (1), we quantify the difference between the second moments extracted from the observed image $\mu_{ij,\mathrm{obs}}$ and those from the model $\mu_{ij,\mathrm{model}}$ by a scalar cost function. The goal is to characterize the pattern of the μ_{ij} in a meaningful and seeing-independent way. There are numerous examples in the literature of whisker diagrams in which the major ellipse orientations ϕ are indicated by little sticks centered on the star centroids (due to the 180^{o} symmetry of ellipses, it is not useful to use vector diagrams). With a bit of training, it is often possible to distinguish basic aberration types from each other by eye from whisker diagrams, in which all sticks have equal length. The next refinement step is to scale the whisker lengths by a measure of the ellipticity, e.g. the metric

$$\epsilon = 1 - \frac{\sigma_s}{\sigma_l}, \tag{13}$$

which can vary between 0 and 1. Unfortunately, ϵ is not seeing independent. The simplest seeing-independent ellipticity definition is $\epsilon_m = |q| = \sigma_l^2 - \sigma_s^2$.[22] We have chosen to emphasize differences between smaller ellipticities by employing the function

$$\epsilon'_m = \sigma_0^2 \log\left(1 + \frac{|q|}{\sigma_0^2}\right), \tag{14}$$

with the empirical scaling constant $\sigma_0 = 0.183$ arcsec (note that $\lim_{|q|\to 0} d\epsilon'_m/d|q| = 1$). Both ϵ and ϵ'_m are invariant under rotations of the ellipse and we will introduce the orientation angle of the ellipse by multiplying with the unit vector $\{\cos\phi, \sin\phi\}$.

We have so far only considered one of the covariance matrix invariants, namely $|q|$, and the angle ϕ. In order to further distinguish different science images, we need to include the quantity $p = \sigma_l^2 + \sigma_s^2$ in the cost function, which yields information about the spot size variation. However, p is obviously not seeing independent since (isotropic) seeing raises both σ_l^2 and σ_s^2 by the same amount. Under the (strong) assumption that the seeing PSF s is uniform across the field in long-exposure images, we may instead consider the variation of p about its field average. We now define

$$\boldsymbol{p} = \{p^{(1)}, p^{(2)}, \ldots, p^{(n)}\}, \tag{15}$$

$$h = \epsilon'_m\{\cos\phi, \sin\phi\}, \tag{16}$$

$$\boldsymbol{v} = \{h^{(1)}, h^{(2)}, \ldots, h^{(n)}\}, \tag{17}$$

$$c = \frac{10^6}{n}\left(\frac{1}{4}\left|(\boldsymbol{p}_{\rm model} - \hat{\boldsymbol{p}}_{\rm model}) - (\boldsymbol{p}_{\rm obs} - \hat{\boldsymbol{p}}_{\rm obs})\right|^2 + \left|\boldsymbol{v}_{\rm model} - \boldsymbol{v}_{\rm obs}\right|^2\right), \tag{18}$$

where $\boldsymbol{p}$ is the vector of all $p^{(k)}$ and $\hat{\boldsymbol{p}}$ denotes the median of the $p^{(k)}$ times the n-vector of 1's, where $n \le 400$ is the number of valid tiles in the image (again, the median is a statistically more robust quantity than the mean). Further, $h^{(k)}$ is a pair of real numbers, characterizing the star elongations in tile k, which we concatenate in the vector $\boldsymbol{v}$ of length $2n$. Finally, we have defined the cost c as the sum of the squared differences between model and observation. We scale the cost so that it becomes independent of the number of tiles and yields typical values of $10\ldots1000$ if p and q are entered in units of arcsec^2. In the beginning of our work, we have additionally applied a radial weighting to limit the cost contributions stemming from the field corners, but later we found this weighting to be unnecessary and practically disabled it in the results presented in Section 3.

Our choice of the cost function, based directly on ellipticity information from up to 400 image tiles, differs from the approach adopted in,[22] where we first expanded the ellipticity pattern in slope Zernike polynomials in the field and then ran the minimization algorithm in the space of the Zernike coefficients.

2.4 Numerical Optimization

The cost $c = c(\boldsymbol{\xi})$ is a function of the nine-dimensional telescope perturbation vector $\boldsymbol{\xi}$ and our goal is to find its global minimum. We apply the built-in Mathematica function `NMinimize[]` based on the Nelder-Mead algorithm,[23] which is a highly efficient and robust multidimensional simplex minimization method (it does not rely on the gradient of the cost function). In order to raise the chances of finding the global minimum, we start 12–16 parallel optimization threads with different random initial values and finally pick the lowest cost solution.

In order to further accelerate convergence and lower the probability of getting stuck in local minima, we transform to a new basis $\boldsymbol{\Xi}$, in which the Hessian (the matrix of all second derivatives) $\partial^2 c'/\partial\Xi_k\partial\Xi_l$ of the transformed cost function $c'(\boldsymbol{\Xi})$ becomes orthonormal at a given (local) minimum. This orthonormalization procedure also allowed us to identify the useful set of nine degrees of freedom, excluding the M1 vertex decenter and tip/tilt (which are fixed at zero). If we instead included the M1 vertex position and tip/tilt in $\boldsymbol{\xi}$, we would introduce 'soft cost modes' that correspond to rotations of M1 and M2 about a common point, hardly varying the cost and thus leading to somewhat arbitrary optimization results. We have validated in Zemax that common rotations can be ignored, even when rotating by 60 arcsec. Finally, we found that the M1 figure spherical aberration coefficient $W_{040,\rm M1}$ defined in 10 is difficult to distinguish from defocus arising from M2 piston and therefore we also exclude it from $\boldsymbol{\xi}$ and set it to zero. We will elaborate on the orthonormalization procedure in a forthcoming publication.

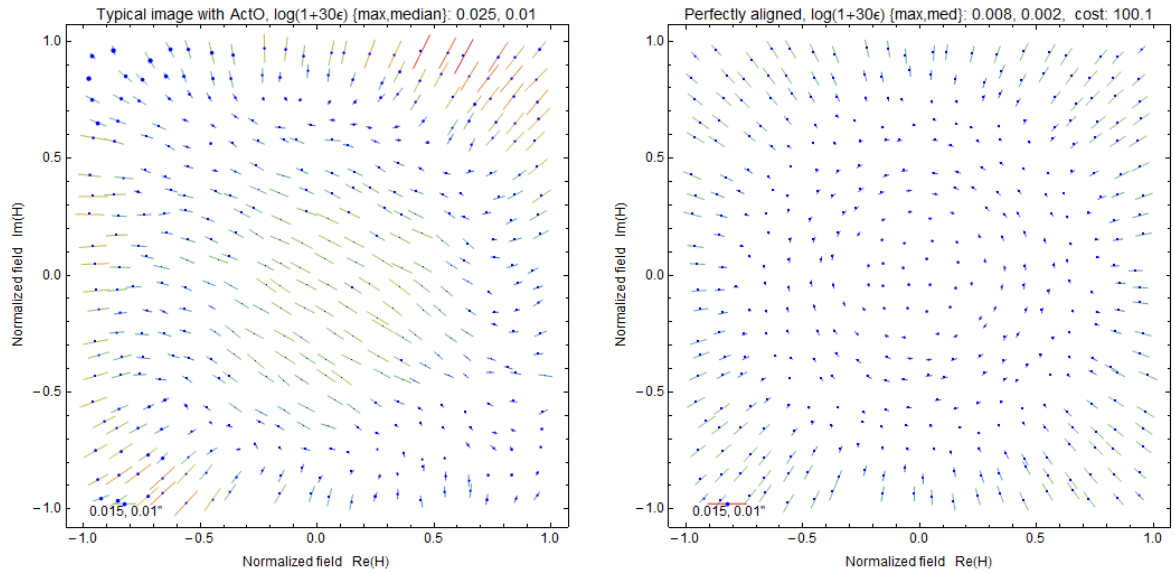

Figure 6. Left: Typical VST PSF whisker plot on survey image with ActO correction. Right: Expected ellipticity pattern under perfect alignment; the residual ellipticities are inherent to the VST optical design (note the scale difference of the whiskers). The diameters of the dots indicate variations of the spot size.

3. RESULTS

Figure 6 shows a comparison of the whisker plot of a typical VST image under closed-loop ActO correction (left) with a plot of whiskers generated from our analytical method with perfect alignment. The residual ellipticities in the right plot are inherent to the VST optical design, which balances different aberration terms against each other.[24] The ellipticity pattern of the left plot, with the whiskers predominantly pointing in one direction near the middle of the image and orthogonal to that direction in the upper right and lower left corners, shows a clear signature of uncorrected M1 astigmatism (the amount of uncorrected aberration may be reduced after future control software updates).

In both plots, the ellipticity magnitude vanishes near a circle whose diameter equals about half of the field diameter, often called 'circle of least confusion'. Inside this circle, defocus, astigmatism and coma combine in a way that the ellipses are normally directed tangentially and outside the circle, they turn in a radial direction. Note the different scale of the whiskers, indicated by the reference stick and the dot in the lower left corners of the plots. Figure 7 shows an array of 9×9 spot diagrams (blue) superimposed by the spot ellipses given by σ_l, σ_s and ϕ (red) from the analytical method for the perfectly aligned VST (left) and with M2 moved towards M1 by 5 μm, underlining how narrow the alignment tolerances can be in wide-field telescopes. The circle of least confusion clearly expands due to the added defocus.

We now apply the method depicted in Fig. 1 to calibration images that have been recorded during the VST commissioning in 2011[25]. In this run, a typical set of images would start from a well collimated telescope with the ActO algorithm running in closed loop, then the ActO (but not the tracking/guiding) is turned off, a mirror displacement or surface deformation is applied (hence, one of the ξ_k is changed from near zero to a defined finite value), and a series 3–5 images is recorded with an exposure time of 60 seconds each (all images were taken with the red SDSS filter centered at 625 nm and 135 nm wide). Finally, the displacement/deformation is reversed and another image is taken (the VST also includes a deployable Shack-Hartmann wavefront sensor, but it has not been used in this run). The applied perturbations are rather large (e.g., 60–180 arcsec rotation about the coma-neutral point or 500 nm rms M1 astigmatism). We process all images in the set with our method, obtaining the most likely perturbation perturbation $\boldsymbol{\xi}_{\rm opt}$ for each, and finally compute the differences of subsequent $\boldsymbol{\xi}_{\rm opt}$.

Figure 8 shows the whisker diagram from a test image with added 500 nm of astigmatism on M1. This example is a relatively basic test case since the aberration is applied in the pupil (coefficient W_{022}) and thus

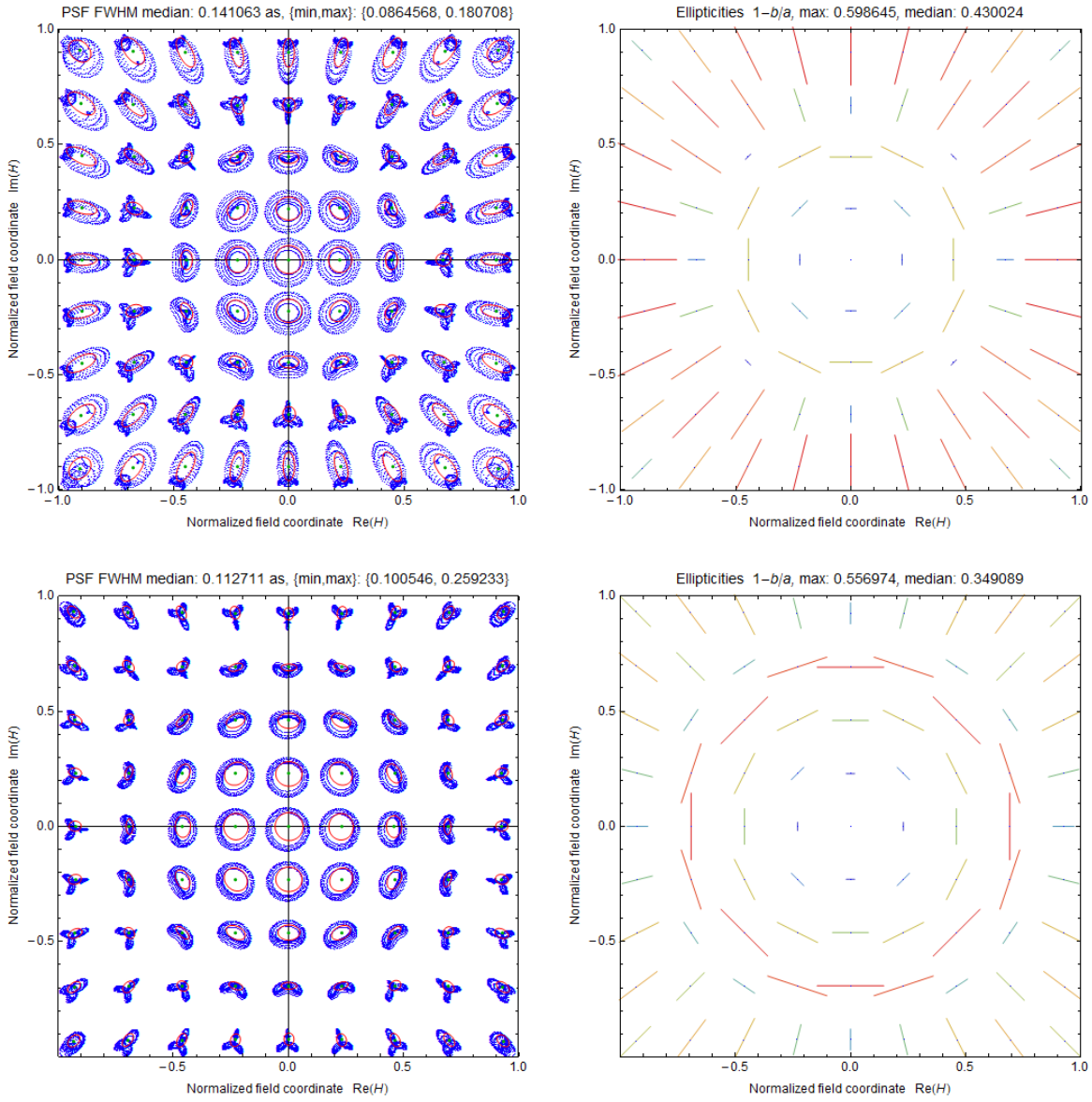

Figure 7. Left: Spot diagrams (blue) with spot ellipses (red) and corresponding whisker plots for the perfectly aligned VST, Right: Same, but M2 has been moved towards M1 by 5 μm.

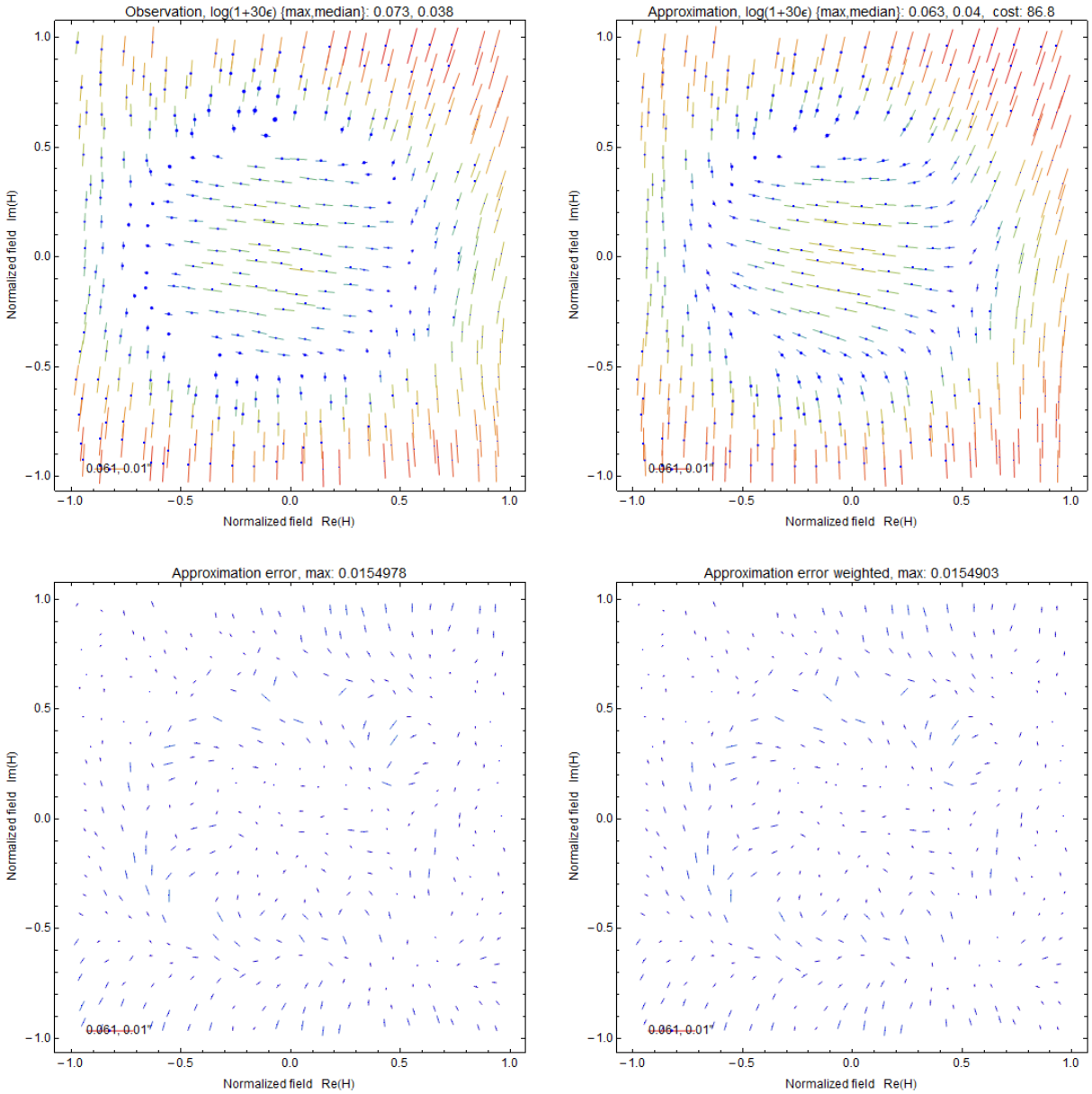

Figure 8. Upper left: whisker diagram from test image with added 500 nm of astigmatism on M1. Upper right: analytical approximation, lower right: differences between the two upper figures.

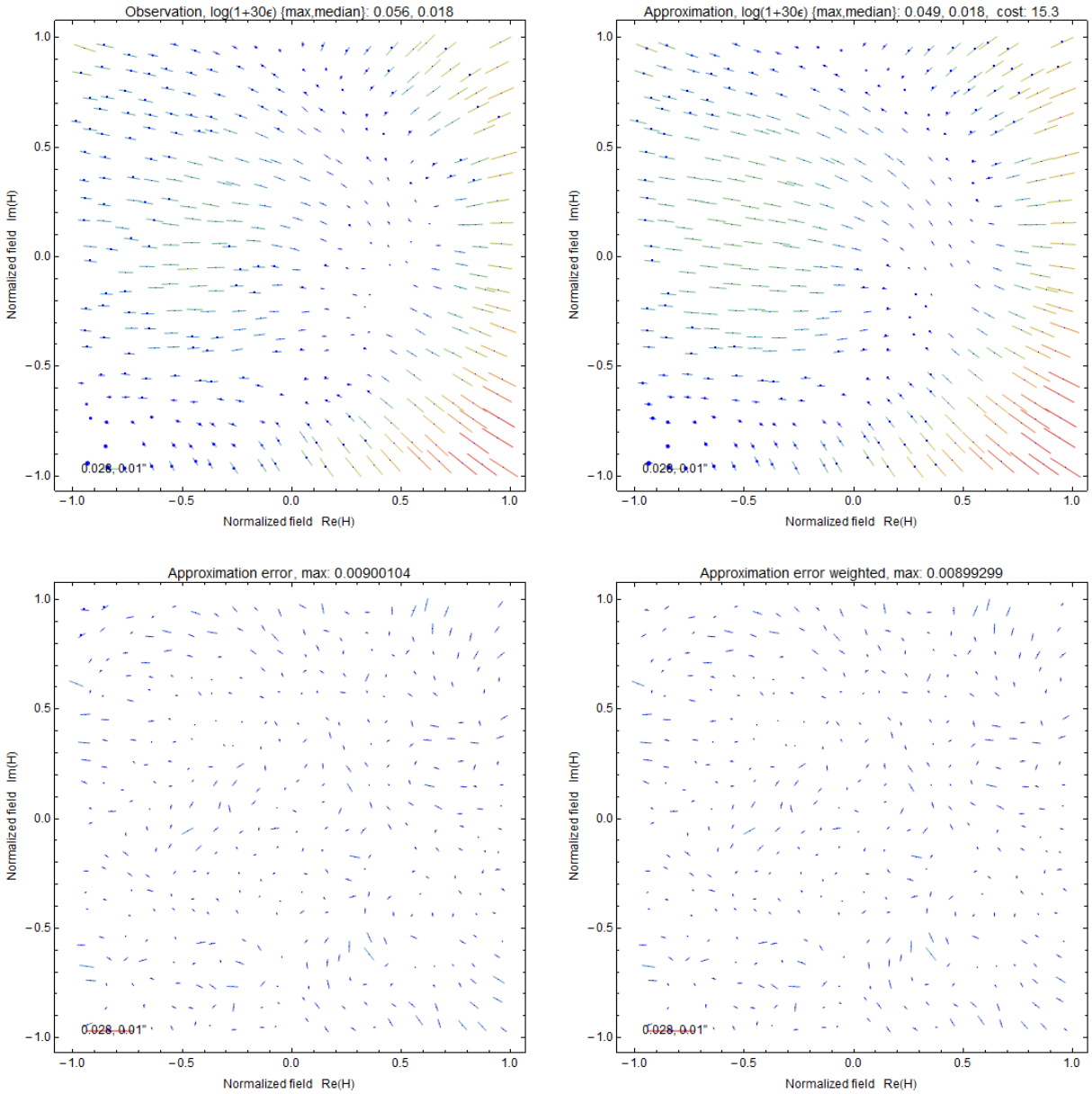

Figure 9. Upper left: Whisker diagram from test image with added 60 arcsec rotation of M2 about its coma-neutral point, upper right: analytical approximation, lower right: differences between the two upper figures.

does not alter the other W_{klm} or the field nodes whose calculation require a correct implementation of the Plate Diagram formulas. The lower right plot shows the residual errors (differences between the two upper plots). An important test is to check whether the errors show a systematic pattern, which would be an indication that the correct perturbation state has been missed. The errors tend to be a bit larger near the circle of least confusion, however, this region is also where the ellipticity changes most rapidly, hence one would expect the errors to peak there. The extracted M1 astigmatism difference from the previous image in the observed set (not shown) was 570 nm, which we consider a good agreement, given the error sources discussed in Section 2.

Figure 8 shows the whisker diagram from a test image with added 60 arcsec of M2 rotation about the coma-neutral point, which lies 1075 mm behind M2, rotating about the x-axis, which distorts the ellipticity pattern in a characteristic U-shape. The ellipticity method finds a rotation of 69.1 arcsec about x and 1.4 arcsec about y, and further the rotation center is 790 mm away from M2, which again appears as an acceptable to good agreement. We note that this test image does probe more deeply the accuracy of the optical model. The reference ray lies at the field point $\alpha H_{\mathrm{ref}} = -0.00209^\circ + i\,0.0009^\circ$.

In total, we have processed 46 calibration files. The method reproduced the applied perturbations in most cases to within ±20%, although the errors were larger in the case of massive aberrations (180 arcsec rotation about the coma neutral point or 58 arcsec rotation about the M2 center of curvature). However, the sign of the reproduced aberrations were always correct. We note that our moment extraction algorithm shows significant systematic errors at strong ellipticities such as $\epsilon \approx 0.5$ and moreover the optimization algorithm may require using an initial aberration state vector that lies closer to the result and thus far away from zero, both issues that would seldom arise in closed-loop operation. Also, the turbulent atmosphere causes seeing variations across the pupil and across the field with nonzero residuals even after 60 s exposure time.[26] We are studying this error source at present.

4. CONCLUSIONS

We have described an method to deduce misalignments and optical surface deformations of a telescope from its science image and applied it to the VST in a process depicted in Fig. 1. The method relies on an analytical fifth-order aberration theory rather than on optical ray tracing and is thus very fast. We have validated that the spot diagrams generated by our optical model agree with those from commercial optical design software (Zemax) to far below the seeing limit for both the unaberrated and the aberrated telescope. From the optical aberrations, we are computing the second moments of the star PSFs using elliptical Gaussian window functions, followed by median-averaging in a 20×20 grid of tiles. We quantify the difference between observation and model by a cost function, which is corrected for atmospheric dispersion and independent of the contribution due to the turbulent atmosphere. Finally, we employ a multidimensional simplex optimization algorithm to find the minimum cost solution in a 9-dimensional space of VST mirror misalignments (M2) and deformations (M1).

Currently, we are able to reproduce the misalignment state of VST calibration images with deliberately introduced moderate perturbations with an accuracy of about ±20%, and we are investigating limitations of the method imposed by variable conditions such as high sky background, bad seeing and sparse star fields. The CPU processing time is about 50 seconds on a workstation, on which we run on eight cores in parallel, enabling closed-loop ActO control much faster than the typical telescope drift time of several minutes.

Future development will focus on the improvement of the extraction algorithm of star image ellipticities and implementing the analytical calculation of the fifth-order surface aberration coefficients and the fifth-order field nodes. We are further planning to apply the method to the VST in closed loop, which is the only method to ultimately validate its accuracy in a real-world implementation.

We believe that our method could directly benefit a number of wide-field telescopes such as SkyMapper or accelerate the alignment of 8 m-class telescopes such as the Very Large Telescope (VLT). It may ultimately become a viable active optics control option for the PanSTARRS telescope, DECam and the even the LSST.

REFERENCES

1. M. Capaccioli, D. Mancini, and G. Sedmak, "The VLT Survey Telescope: A Status Report," *ESO Messenger* **120**, pp. 10–13, Jun 2003.
2. P. Schipani, S. D'Orsi, L. Ferragina, D. Fierro, L. Marty, C. Molfese, and F. Perrotta, "Active optics primary mirror support system for the 2.6 m VST telescope," *Appl. Opt.* **49**, pp. 1234–1241, Mar 2010.
3. K. Kuijken, R. Bender, and E. Cappellaro, *et al.*, "OmegaCAM: wide-field imaging with fine spatial resolution," in *Ground-based Instrumentation for Astronomy*, A. F. M. Moorwood and M. Iye, eds., *Proc. SPIE* **5492**, pp. 484–493, Sept. 2004.
4. O. Iwert, D. Baade, and A. Balestra, *et al.*, "The OmegaCAM 16k×16k CCD detector system for the ESO VLT Survey Telescope (VST)," *Proc. SPIE* **6276**, June 2006.
5. A. Tokovinin and S. Heathcote, "Donut: measuring optical aberrations from a single extra-focal image," *Pub. ASP* **118**, 2006.
6. P. Schipani, S. D'Orsi, D. Fierro, and L. Marty, "Active optics control of VST telescope secondary mirror," *Appl. Opt.* **49**, pp. 3199–3207, Jun 2010.
7. K. Kuijken, "Lensing with KiDS," in *KITP Program: Applications of Gravitational Lensing: Unique Insights into Galaxy Formation and Evolution*, Oct. 2006.
8. J. T. A. de Jong, G. A. Verdoes Kleijn, K. H. Kuijken, and E. A. Valentijn, "The Kilo-Degree Survey," *Exp. Astron.* **35**, pp. 25–44, Jan. 2013.
9. C. A. Laury-Micoulaut, "The n-th centered moment of a multiple convolution and its applications to an intercloud gas model," *Astron. Astrophys.* **51**, pp. 343–346, Sept. 1976.
10. N. Kaiser, G. Squires, and T. Broadhurst, "A Method for Weak Lensing Observations," *Astrophys. J.* **449**, p. 460, Aug. 1995.
11. E. Bertin and S. Arnouts, "SExtractor: Software for source extraction," *Astron. Astrophys. Supp.* **117**, pp. 393–404, June 1996.
12. A. V. Filippenko, "The importance of atmospheric differential refraction in spectrophotometry," *Pub. of the ASP* **94**, pp. 715–721, Aug. 1982.
13. P. Schipani, J. Farinato, and C. Arcidiacono, *et al.*, "The ADC for the VST Telescope: theory and preliminary test of the electromechanical system," *Proc. SPIE* **7739**, July 2010.
14. H. H. Hopkins, *The Wave Theory of Aberrations*, Clarendon Press, Oxford, UK, 1950.
15. K. P. Thompson, "Description of the third-order optical aberrations of near-circular pupil optical systems without symmetry," *J. Opt. Soc. Am. A* **22**, pp. 1389–1401, Jul 2005.
16. K. P. Thompson, "Description of the third-order optical aberrations of near-circular pupil optical systems without symmetry: errata," *J. Opt. Soc. Am. A* **26**, pp. 699–699, Mar 2009.
17. R. A. Buchroeder, *Tilted component optical systems.* PhD thesis, Univ. of Arizona, Tucson, Arizona, 1976.
18. C. R. Burch, "On the optical see-saw diagram," *MNRAS* **102**(6), 1942.
19. A. Rakich, "Calculation of third-order misalignment aberrations with the optical plate diagram," in *Proc. SPIE 7652, International Optical Design Conference 2010, 765230*, 2010.
20. J. Sasián, "Theory of sixth-order wave aberrations," *Appl. Opt.* **49**, pp. D69–D95, Jun 2010.
21. K. P. Thompson, "Multinodal fifth-order optical aberrations of optical systems without rotational symmetry: spherical aberration," *J. Opt. Soc. Am. A* **26**, pp. 1090–1100, May 2009.
22. L. Noethe, P. Schipani, R. Holzlöhner, and A. Rakich, "A method for the use of ellipticities and spot diameters for the measurement of aberrations in wide-field telescopes," *Adv. Opt. Tech.* **3**, pp. 315–333, Jun 2014.
23. J. A. Nelder and R. Mead, "A simplex method for function minimization," *Comp. J.* **7**(4), pp. 308–313, 1965.
24. P. Schipani, L. Noethe, and C. Arcidiacono, *et al.*, "Removing static aberrations from the active optics system of a wide-field telescope," *J. Opt. Soc. Am. A* **29**, pp. 1359–1366, Jul 2012.
25. P. Schipani, "Image elongations caused by the introduction of error sources," ESO internal report VST-COM-OAC-20000-2022-1.0.DOC, Issue1, European Southern Observatory, Mar 2011.
26. F. Roddier, D. Roddier, M. J. Northcott, J. E. Graves, and D. L. McKenna, "One-dimensional spectra of turbulence-induced Zernike aberrations: time-delay and isoplanicity error in partial adaptive compensation," *J. Opt. Soc. Am. A* **10**, pp. 957–965, May 1993.