

1 Introduction

We aim to build a model for the HSC field-dependent optics wavefront that can be trained using out-of-focus “donut” images of stars and then used to fit the optics contribution of in-focus PSFs. Developing this model is part of our ongoing effort to factor the PSF into independent contributions from the atmosphere, optics and sensors. Such a factorization has many potential benefits, the greatest of which is likely the ability to capture all of the effects of CCD-to-CCD discontinuities into the optical component, allowing the sensor and particularly the atmospheric component to be smoothly interpolated across the entire field-of-view. Additional benefits may also include enabling a significant part of the PSF variation across the field of view to be captured using just a few parameters, better characterization of high-frequency PSF components that are not well constrained empirically using finitely sampled data, and increased ability to characterize the wavelength dependence of the PSF by characterizing the disparate wavelength dependencies of the individual components.

2 Model

Our model for the optical part of a single star’s PSF is Fourier optics. Specifically, the optical PSF is described by a pupil-obscuration function and wavefront function. (See DMTN-064 for more details). Our primary concern for this note is the spatial variability of the optical wavefront and therefore the optical PSF over the field of view. We break down the wavefront model of the i th exposure into three constituent components as follows:

$$W^i(\vec{u}; \vec{\theta}) = W_{\text{tel}}(\vec{u}; \vec{\theta}) + W_{\text{visit}}^i(\vec{u}; \vec{\theta}) + W_{\text{CCD}}^{\phi_i}(\vec{u}; \vec{\theta}) \quad (1)$$

In the above, i is an index over exposures, ϕ_i is the rotator angle for the i th exposure, \vec{u} is a coordinate in the entrance pupil, $\vec{\theta}$ is the field angle (or equivalently, a location on the focal plane). The coordinate system for \vec{u} and $\vec{\theta}$ are understood to align with the altitude and azimuth axes of the telescope. Note that this implies that in the presence of a rotator, they are not fixed with respect to focal plane or CCDs themselves.

The individual terms of the model are:

- $W_{\text{tel}}(\vec{u}; \vec{\theta})$ This term is intended to model the wavefront perturbations that are present in the otherwise unperturbed optical design of the telescope. It is independent of the exposure index i , and continuous over the pupil and field of view.
- $W_{\text{CCD}}^{\phi_i}(\vec{u}; \vec{\theta})$ This term is intended to model wavefront perturbations that are fixed to the CCD array, such as may originate from displacements in the heights of each CCD. It is continuous in the pupil, but we allow discontinuities between different CCDs across the focal plane. Because this term is constant in focal-plane coordinates, but not in our alt-az aligned \vec{u} and $\vec{\theta}$ coordinates, it also depends on the rotator angle ϕ_i . Aside from the dependence on ϕ , this term is constant in time.
- $W_{\text{visit}}^i(\vec{u}; \vec{\theta})$ This dynamic term is intended to capture exposure-to-exposure differences (beyond the rotator angle), that may be due to flexure, temperature variations, or any other time-dependent continuous perturbation to the wavefront. It explicitly depends on the exposure index i .

In DMTN-064, we described how we measured individual wavefronts for a limited number of pairs of intra- and extra-focal images and specific locations within the field of view. For these “donut” images, we used a forward model that decomposed the delivered wavefront into a Zernike polynomial series. I.e.:

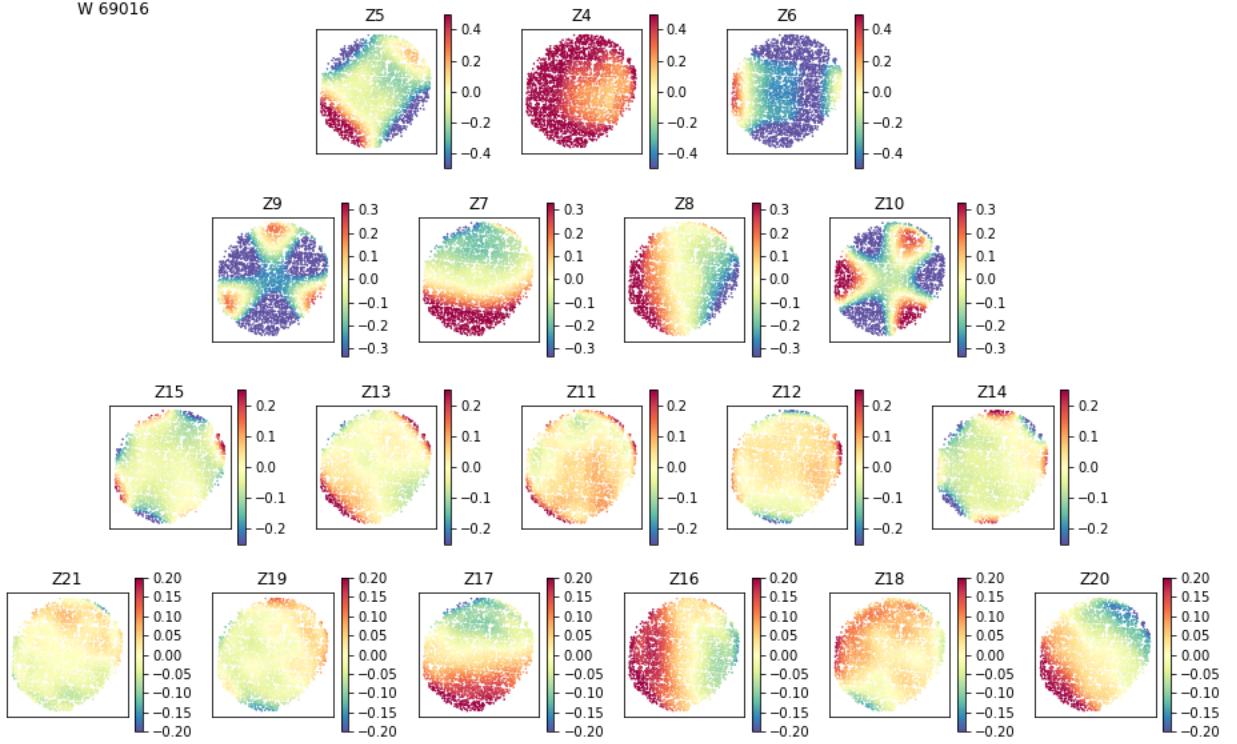


Figure 1: Zernike coefficient fits for exposure pair 69016/69018.

$$W^i \left(\vec{u}; \vec{\theta}_* \right) = \sum_{j=4}^{j_{\max}} a_j^i(\vec{\theta}_*) Z_j(\vec{u}) \quad (2)$$

In Figure 1 we show the field-of-view dependence of such fits to one exposure.

It is useful to apply the same decomposition for the \vec{u} dependence of the model terms. I.e., for the telescope and visit terms we will write:

$$W_{\text{tel}} \left(\vec{u}; \vec{\theta} \right) = \sum_j b_j^{\text{tel}}(\vec{\theta}) Z_j(\vec{u}) \quad (3)$$

$$W_{\text{visit}}^i \left(\vec{u}; \vec{\theta} \right) = \sum_j c_j^i(\vec{\theta}) Z_j(\vec{u}) \quad (4)$$

We will write the CCD term using focal plane coordinates \vec{u}' and $\vec{\theta}'$ (instead of alt-az aligned coordinates), since these are the coordinates in which there is no dependence on the rotator angle ϕ , and hence no dependence on the exposure index i :

$$W'_{\text{CCD}} \left(\vec{u}'; \vec{\theta}' \right) = \sum_{j'} d_{j'}(\vec{\theta}') Z_{j'}(\vec{u}') \quad (5)$$

The trick now is to concretely model the $\vec{\theta}$ dependence which we left purely abstract above. For this, we will again use Zernike polynomials, forming a double zernike basis. For the telescope and visit terms this becomes:

$$W_{\text{tel}}(\vec{u}; \vec{\theta}) = \sum_{jk} b_{jk}^{\text{tel}} Z_k(\vec{\theta}) Z_j(\vec{u}) \quad (6)$$

$$W_{\text{visit}}^i(\vec{u}; \vec{\theta}) = \sum_{jk} c_{jk}^i Z_k(\vec{\theta}) Z_j(\vec{u}) \quad (7)$$

For the CCD term, we additionally indicate which CCD n the field angle $\vec{\theta}'$ projects to (using the indicator function $\mathbb{1}_n(\vec{\theta}')$):

$$W'_{\text{CCD}}(\vec{u}'; \vec{\theta}') = \sum_{nj'k} d_{nj'k} \mathbb{1}_n(\vec{\theta}') Z_k(\vec{\theta}') Z_{j'}(\vec{u}') \quad (8)$$

To incorporate the rotator angle into the last equation, we assert

$$W_{\text{CCD}}^\phi(\vec{u}; \vec{\theta}) = W'_{\text{CCD}}(R^\phi \vec{u}; R^\phi \vec{\theta}) \quad (9)$$

where R^ϕ is the normal 2D rotation matrix for angle ϕ . For a given donut or star observation at $\vec{\theta}_*$ (in alt-az aligned field angle), this leads to

$$W_{\text{CCD}}^\phi(\vec{u}; \vec{\theta}_*) = \sum_{nj'k} d_{nj'k} \mathbb{1}_n(R^\phi \vec{\theta}_*) Z_{j'}(R^\phi \vec{u}) Z_k(R^\phi \vec{\theta}_*) \quad (10)$$

The $Z_k(R^\phi \vec{\theta}_*)$ is simply a real number we can compute. The $Z_{j'}(R^\phi \vec{u})$ factor is as yet a function here. We will find it convenient to further decompose it into a series in $Z_j(\vec{u})$ (eliminating the R^ϕ) using the mathematics presented in (Tatulli (2013) arXiv:1302.7106v1). Namely

$$Z_{j'}(R^\phi \vec{u}) = \sum_j M_{jj'}^\phi Z_j(\vec{u}) \quad (11)$$

where

$$M_{jj'}^\phi = \delta_{nn'} \times \begin{cases} \cos(m\phi), & m = m' \\ \sin(m\phi), & m = -m' \\ 0 & |m| \neq |m'| \end{cases} \quad (12)$$

where n (n') and m (m') are the radial and azimuthal indices of the Zernike polynomial with Noll index j (j'). Note that rows or columns of $M_{jj'}^\phi$ only ever have at most 2 non-zero entries (when $n = n'$ and $m = \pm m'$). With this decomposition, we obtain for each of W_{tel} , W_{visit} , and W_{CCD} a series in alt-az aligned pupil Zernikes $Z_j(\vec{u})$.

3 Donut fits

3.1 Formalism

To determine the unknown coefficients b_{jk}^{tel} , c_{jk}^i , and $d_{nj'k}$ of our model from a set of pupil Zernike measurements $\{a_j^*\}$, we minimize the sum of the least squares residuals R over all of the wavefront measurements:

$$\begin{aligned} R \propto & \sum_* \int d\vec{u} \{ \text{data} - \text{model} \}^2 \\ & \times \sum_* \int d\vec{u} \left\{ \sum_j a_j^* Z_j(\vec{u}) - \sum_{jk} b_{jk}^{\text{tel}} Z_k(\vec{\theta}_*) Z_j(\vec{u}) - \sum_{ijk} \mathbb{1}_i(*) c_{jk}^i Z_k(\vec{\theta}_*) Z_j(\vec{u}) - \sum_{ink} \mathbb{1}_n(R^{\phi_i} \vec{\theta}_*) \mathbb{1}_i(*) \sum_{j'} d_{nj'k} M_{jj'}^{\phi_i} Z_k(R^{\phi_i} \vec{\theta}_*) Z_j(\vec{u}) \right\}^2 \\ & \times \sum_* \int d\vec{u} \left\{ \sum_j \left(a_j^* - \sum_k b_{jk}^{\text{tel}} Z_k(\vec{\theta}_*) - \sum_{ik} \mathbb{1}_i(*) c_{jk}^i Z_k(\vec{\theta}_*) - \sum_{ink} \mathbb{1}_n(R^{\phi_i} \vec{\theta}_*) \mathbb{1}_i(*) \sum_{j'} d_{nj'k} M_{jj'}^{\phi_i} Z_k(R^{\phi_i} \vec{\theta}_*) \right) Z_j(\vec{u}) \right\}^2 \end{aligned} \quad (13)$$

We use the notation $\mathbb{1}_i(*)$ to indicate if a given star was observed on the i th exposure.

Expanding the $\{\cdot\}^2$ above, we can perform the integral over $d\vec{u}$ by taking advantage of the orthogonality of Zernike polynomials:

$$\int Z_j(\vec{u}) Z_{j'}(\vec{u}) d\vec{u} = \pi \delta_{jj'} \quad (14)$$

Essentially, only the squares of the terms being indexed by j survive, and all the cross terms vanish. This leaves us with

$$\sum_j \sum_* \left\{ a_j^* - \sum_k b_{jk}^{\text{tel}} Z_k(\vec{\theta}_*) - \sum_{ik} \mathbb{1}_i(*) c_{jk}^i Z_k(\vec{\theta}_*) - \sum_{ink} \mathbb{1}_n(R^{\phi_i} \vec{\theta}_*) \mathbb{1}_i(*) \sum_{j'} d_{nj'k} M_{jj'}^{\phi_i} Z_k(R^{\phi_i} \vec{\theta}_*) \right\}^2 \quad (15)$$

(with the square now occurring *inside* the summation over j).

Although Eq. 15 may look complicated, since the terms corresponding to the model are linear in each of the unknown variables of the model (b_{jk}^{tel} , c_{jk}^i , and $d_{nj'k}$), we can solve it via standard linear algebra routines. Figuratively, we are looking for the least squares solution to:

$$\begin{bmatrix} \dots & \dots & \dots \\ \dots & \text{geometry} & \dots \\ \dots & \dots & \dots \end{bmatrix} \cdot \begin{bmatrix} \text{b}'s \\ \text{c}'s \\ \text{d}'s \\ \dots \end{bmatrix} = \begin{bmatrix} \dots \\ \text{a}'s \\ \dots \end{bmatrix} \quad (16)$$

Also note that the $\{b_{jk}^{\text{tel}}\}$, and $\{c_{jk}^i\}$ unknown coefficients only occur in terms that include a_j^* , i.e., they never occur in a term that includes $a_{j'}^*$ with $j \neq j'$. Similarly, the unknown $\{d_{nj'k}\}$ coefficients only occur in terms with at most one of two different values of a_j^* , which follows from our earlier comment that $M_{jj'}^\phi$ only ever has at most 2 non-zero entries per row or column. These facts enable us to find the minimum of Eq. 15 in blocks of either 1 or 2 j values at a time.

3.2 Configuration

In the above formalism, we never explicitly mentioned the upper indices of the various summations. In some cases, these may be fairly obvious, like the upper index over i is the number of exposure pairs being analyzed, and the upper index over n is the number of CCDs being analyzed. In other cases, we consciously omitted the indices because they are part of the model design – they’re flexible. For example, the sum over j and k in W_{tel} (Eq. 6) indicates how much complexity in the pupil (the j index) and across the field of view (the k index) we want to allow in the constant-continuous part of the model. Our intuition for this term from observing donut fits and from modeling the HSC optics is to allow a large amount of freedom in the field-of-view dependence of W_{tel} , and set $k_{\text{tel}}^{\max} = 55$, which is an arbitrary 10th order 2D polynomial. The visit term should probably have less freedom than this, as we expect flexure and temperature fluctuations to induce mostly slowly varying changes to the wavefront. We choose $k_{\text{visit}}^{\max} = 10$, which is a 3rd degree polynomial. For W_{CCD} , we only use $k_{\text{CCD}}^{\max} = 1$, indicating that each CCD only has freedom to change the wavefront by a constant amount over the CCD. For each of these terms, we also choose $j^{\max} = 21$, simply because that’s the largest value of j for which we made donut fits a_j .

With the above index upper limits, and using 10 visits and 104 CCDs, our complete model has 4230 parameters to fit. We have 20,984 donuts images to constrain the model, or 377,712 individual a_j^* Zernike measurements.

Before we can build the “geometry” matrix of Eqn 16 and fit the model, we must address one final point, which is that even with many times more data points than parameters, our model currently has too much freedom to have a unique best fit solution – it’s degenerate. For example, increasing the value of $b_{4,1}^{\text{tel}}$ is degenerate with either decreasing the values of each of $c_{4,1}^i$, or decreasing the values of each of d_{n41} . To fix this, we add the constraint functions

$$\sum_i c_{jk}^i = 0 \quad (17)$$

and

$$\sum_n d_{nj'k} = 0 \quad (18)$$

which select a unique solution from the otherwise infinite set of degenerate solutions. The above equations are simply additional rows in the “geometry” matrix and corresponding 0’s added to the “a’s” matrix of Eqn. 16.

In figure 2, we show the design matrix for the $j = 4$ block of Eqn. 15. Each row corresponds to an individual donut measurement a_4^* . The columns of this design matrix show 3 broad distinct regions, which correspond from left-to-right with W_{tel} , W_{CCD} , and W_{visit} . Within the W_{CCD} region, we can see further subdivisions that correspond to individual CCDs and visits, and within the W_{visit} section, we can see individual visits. In short, the structure of the design matrix is that every donut constraints W_{tel} , constrains the W_{CCD} parameters corresponding to the 1 CCD that it lands on in the 1 visit, and constrains W_{visit} parameters that correspond to its particular visit.

In Figure 3, we show the design matrix for the combined $j = 5$ and $j = 6$ block of Eqn. 15. The same subblocks as in Figure 2 are apparent, but doubled. The doubling is a consequence of the mixing of Z_5 and Z_6 under a coordinate system rotation, as discussed earlier.

3.3 Fit

Figures 4, 5, and 6 show the results of fitting the model for W_{tel} , W_{CCD} , and for one particular visit (69016) of W_{visit} .

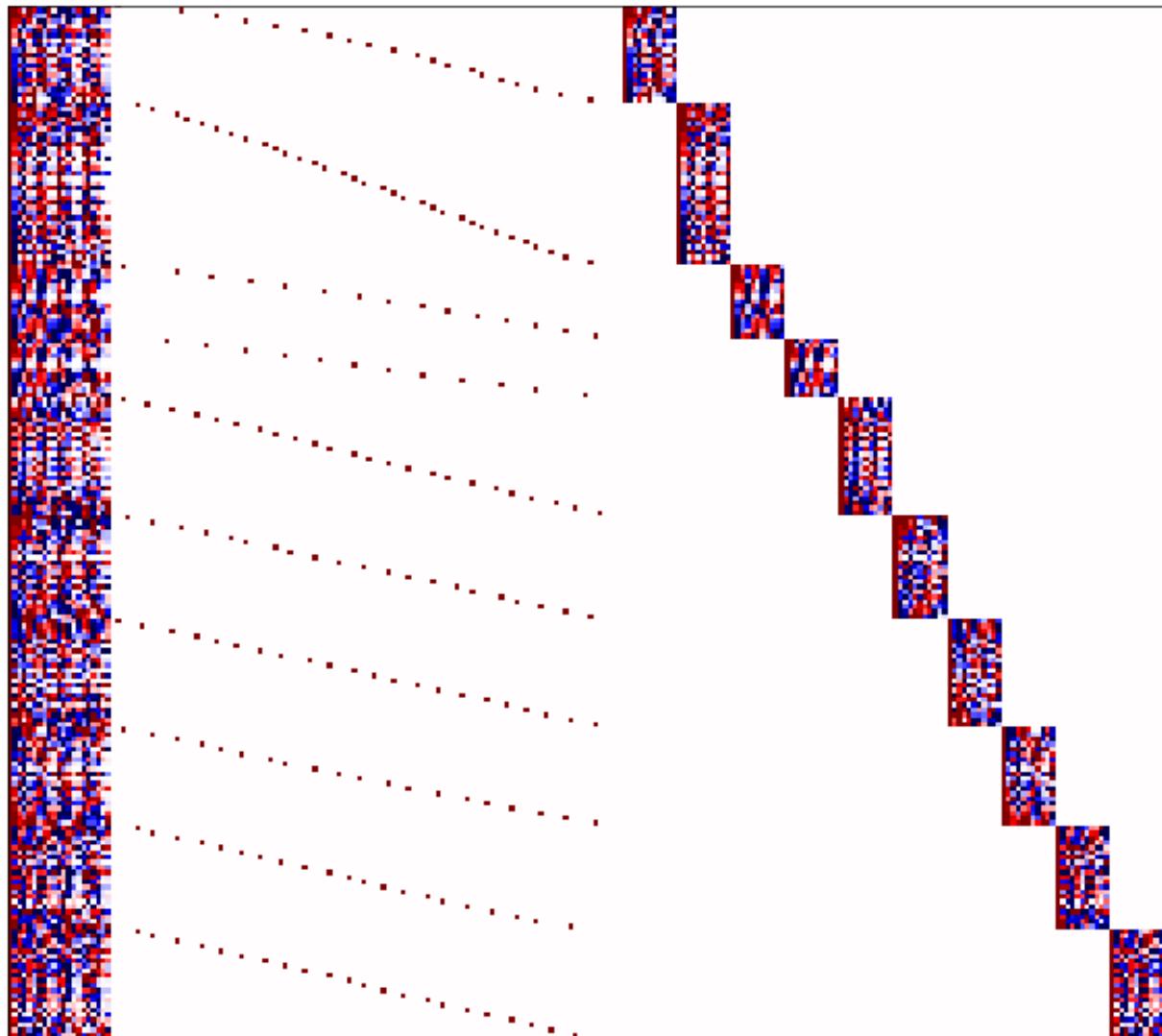


Figure 2: Design matrix for Z_4 showing sparse structure.

Z5/6 design matrix

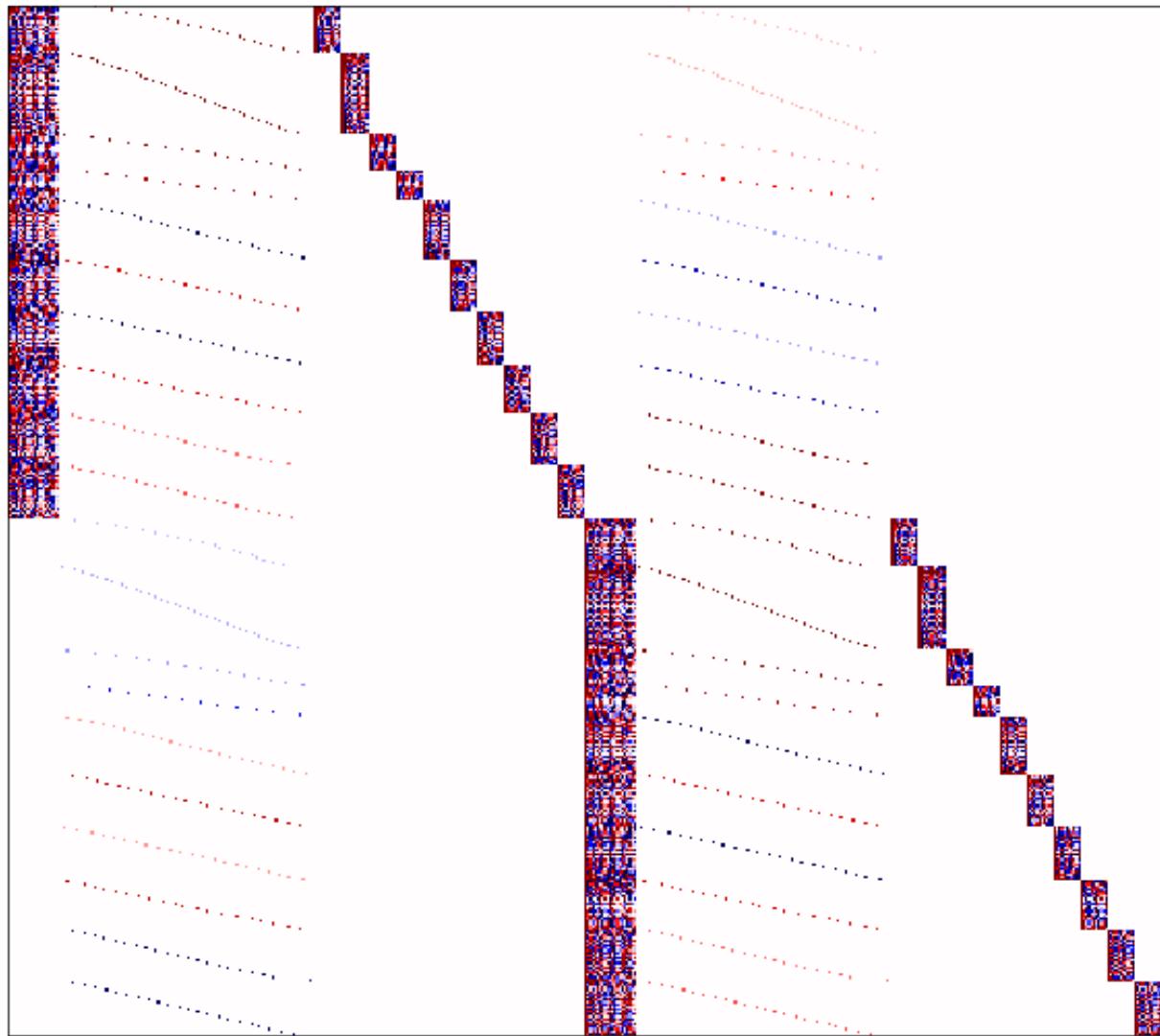


Figure 3: Design matrix for Z_5 and Z_6 showing how these terms mix together.

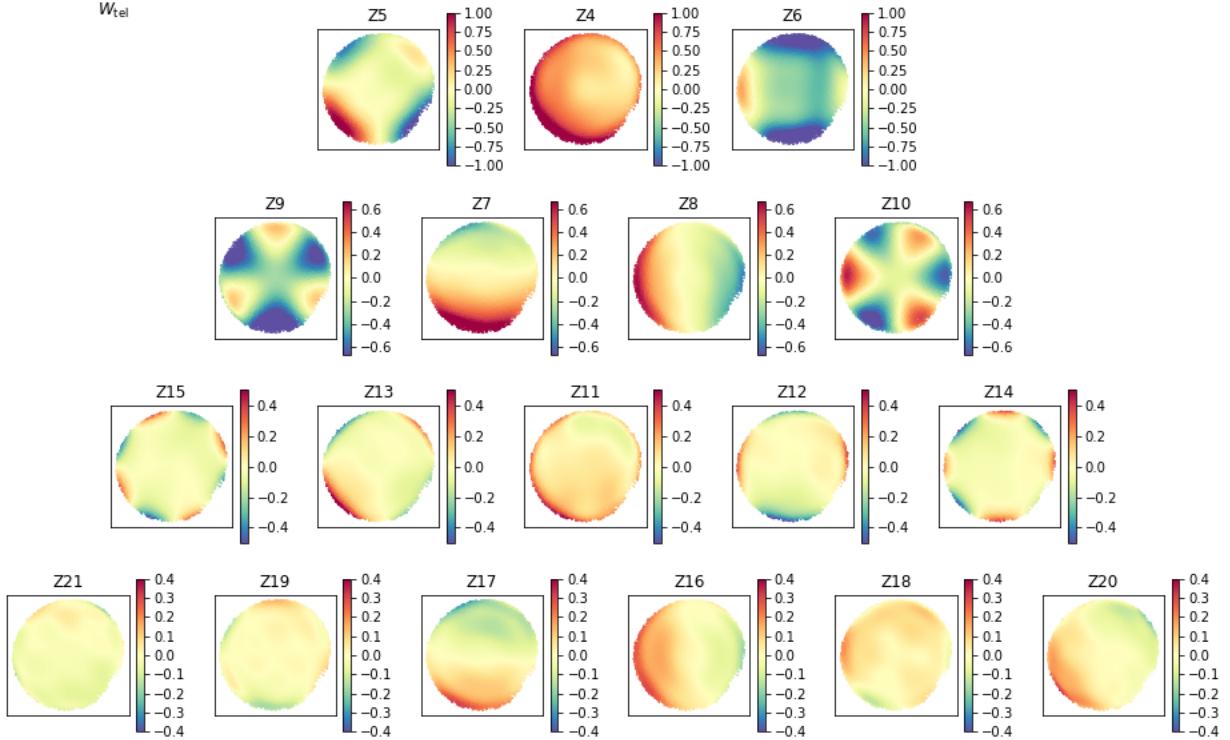


Figure 4: W_{tel}

W_{tel} broadly shows the kinds of patterns that we expect for the telescope design. For instance, Zernike terms that vary like $\cos(m\theta)$ in the pupil azimuthal angle θ vary like $\cos(m\phi)$ in the field azimuthal angle ϕ , as must be the case for a circularly symmetric optical system. Higher order pupil Zernikes show less amplitude than lower order.

In W_{CCD} , discontinuities between CCDs are clearly visible (although note that the amplitudes are pretty small compared to the design amplitudes). It's worth pointing out that for the circularly symmetric terms (Z4 and Z11), we expect some degeneracy between W_{tel} and W_{CCD} . In the other terms, this degeneracy can largely be broken by simultaneously fitting exposures taken at different rotator angles.

In Figures 7 and 8 we see the complete model and the residual. From the residual plot, it's clear that we've captured most of the variation in the data.

4 In-focus fits

To apply our model to in-focus exposures we A) Fix the values of W_{tel} and W_{CCD} found in the preceding analysis (i.e., the values b_{jk}^{tel} and $d_{nj'k}$), and B) apply a dimensionality reduction to W_{visit} (i.e., the values of c_{jk}^i). For B), we specifically use principle component analysis to derive eigen-visits from the 10 available W_{visit} data that we have. We find that keeping 3, 4, 5 eigen-visits captures 75%, 84%, and 89% of the variance amongst all W_{visit} terms, respectively. For the subsequent analysis, we use 4 eigen-visits.

With the 4 eigen-visits and fixed values of W_{tel} and W_{CCD} , we can construct a model for the optical part of the PSF across an entire field of view that only depends on 4 numbers, which are coefficients of the eigen-visits.

To compare this model to in-focus data, however, we also need to account for the sensor and atmospheric part of the in-focus PSF (otherwise we'd be comparing relatively small optics-only PSFs to relatively large

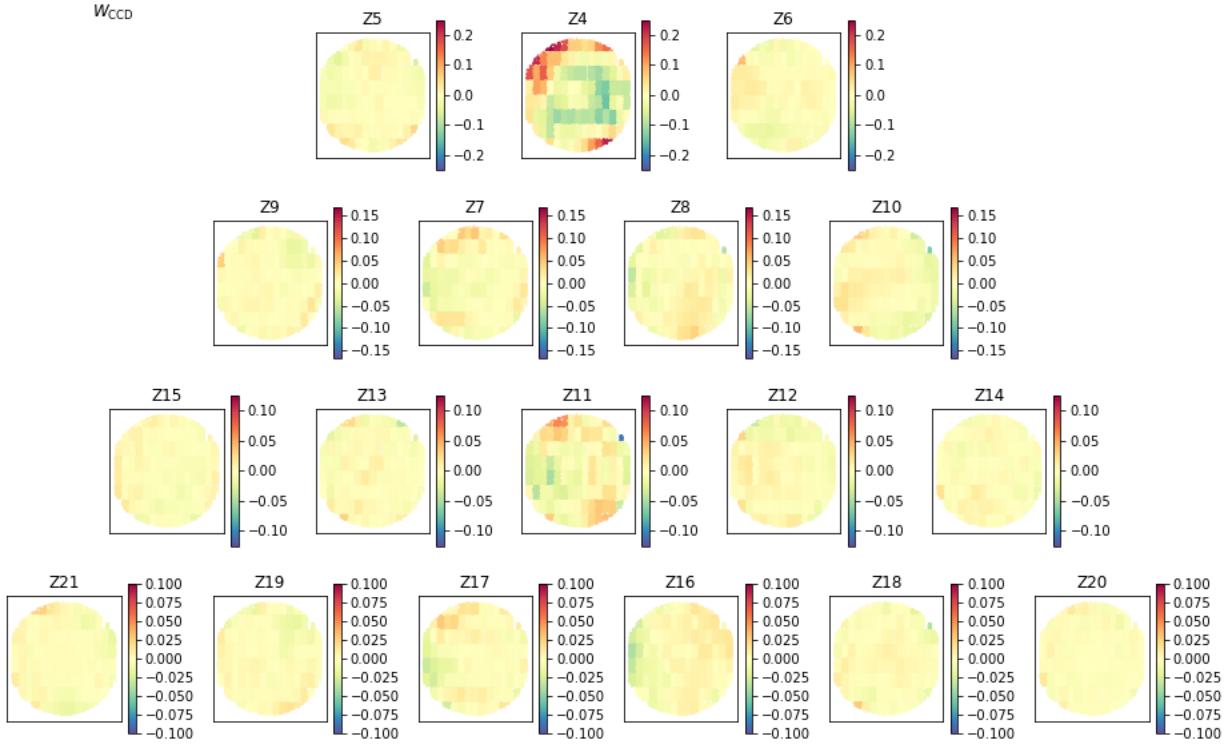


Figure 5: W_{CCD} Note that the coordinate axes are along the rows and columns of the focal plane here. (And therefore this wavefront figure is not aligned with the others in this note.)

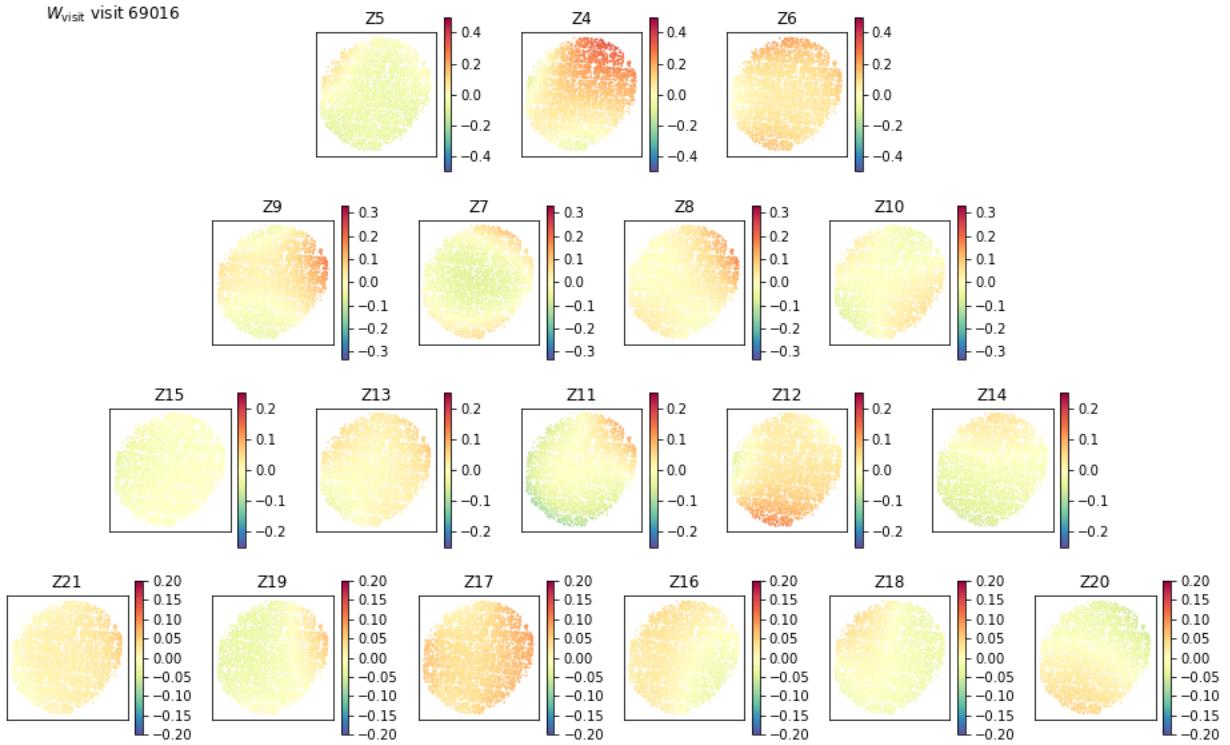


Figure 6: W_{visit} for exposure pair 69016/69018.

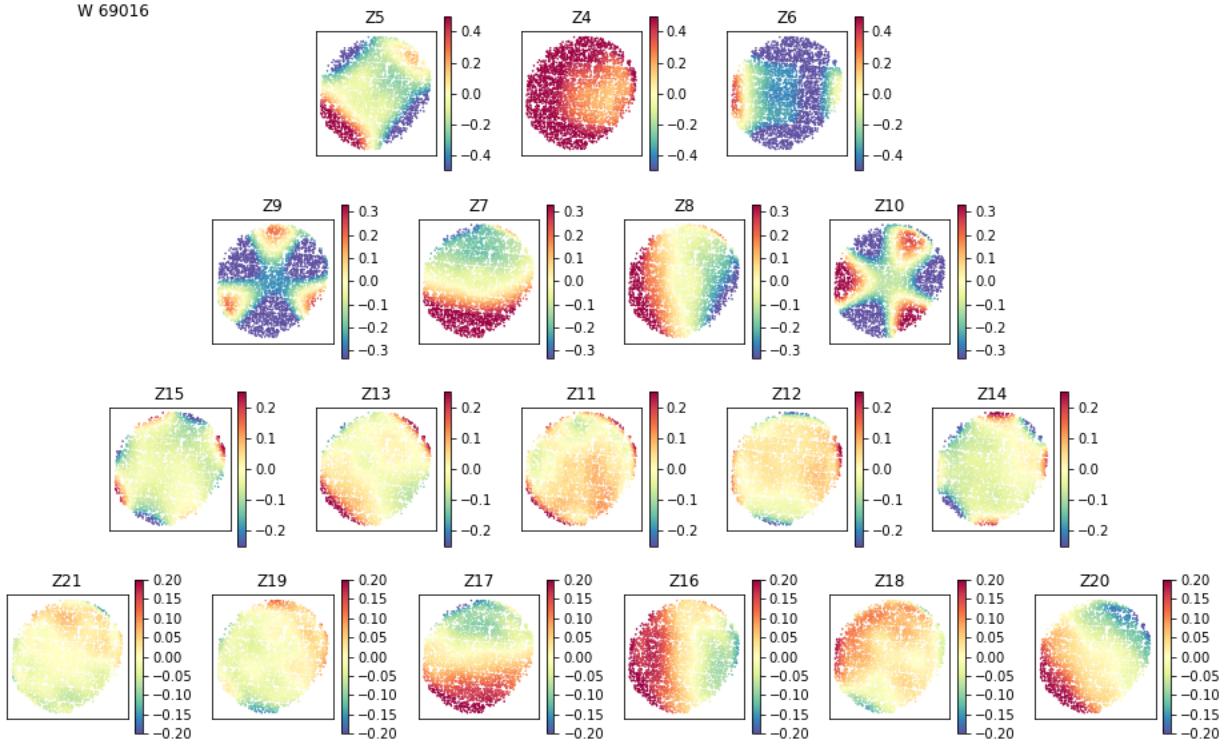


Figure 7: The complete model for exposure pair 69016/69018. Note that this is *slightly* different than the straight sum of the data in Figures ??, ??, and ??, since the W_{CCD} term has been rotated to the particular rotator angle appropriate for this exposure pair.

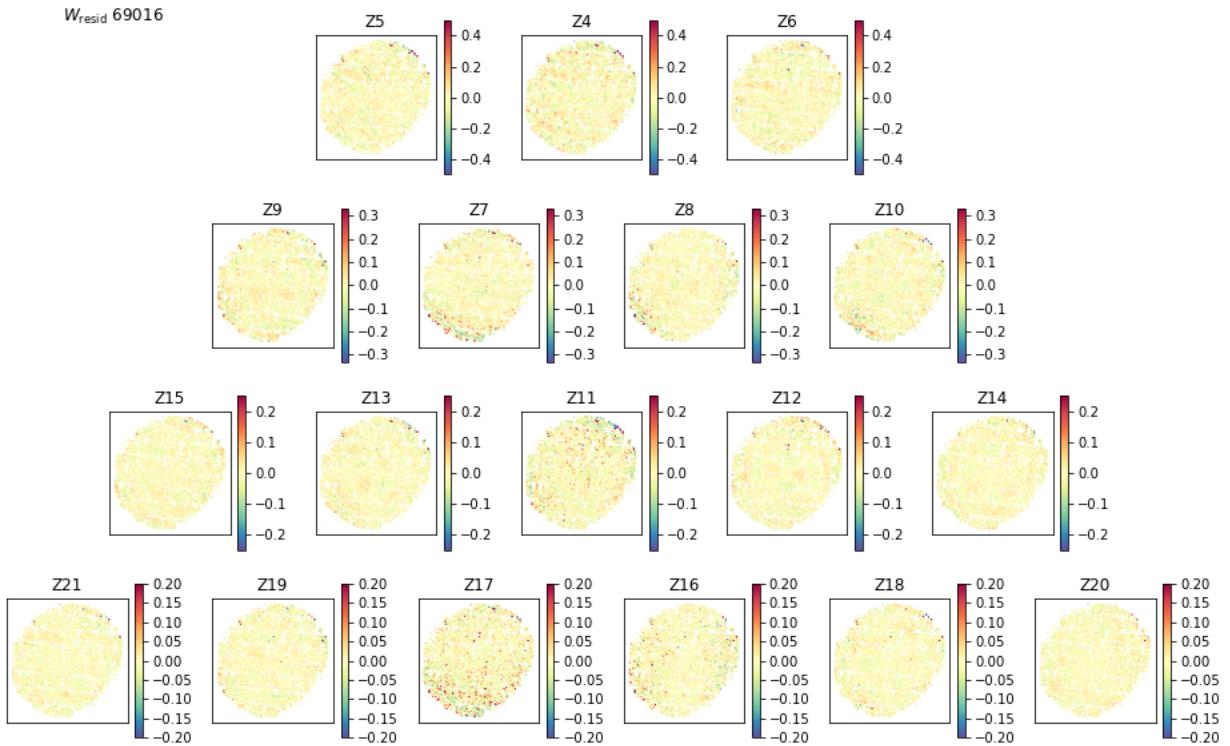


Figure 8: Residual for exposure pair 69016/69018. I.e., data (Figure 1) minus model (Figure 7). Aside from individual point outliers, there is some barely visible structure in the centers of the trefoil (Z9 and Z10) and second astigmatism (Z12 and Z13) panels.

optics+atmosphere+sensor PSFs). Since we are principally interested in the optics part of the PSF here, we take a fairly simple model for the atmospheric+sensor PSF, which is a single elliptical Kolmogorov profile. This model has three parameters: 1 for the overall size and 2 for the ellipticity. Our intention is that this model will grossly capture the effects of not only the atmosphere, but also the sensors (charge diffusion) and other PSF contributions not part of the optics, for example wind shake or tracking errors during the exposure.

The optics + atmosphere + sensor model then contains 7 parameters to optimize for each exposure: 4 controlling relative amplitudes of the dominant eigen-visits, and 3 controlling the atmospheric+sensors+etc PSF.

We test this model on a few in-focus exposures. To fit the model to the data, we randomly select 100 stars over the field of view and reduce them to their second moments as measured by HSM. We do the same for the PSFs predicted by the model. We then minimize the sum of the squares of the second moments of all 100 stars.

The results of this in-focus PSF fitting procedure for three different exposures are presented in Figures 9, 10, and 11. The PSF sizes and e1 components are relatively well fit by the model. The e2 components are not particularly well fit; in fact, it appears that there may be a minus sign error somewhere that is affecting the e2 results. Overall though, this approach seems promising in that it can explain a significant part of the large-field variation of PSF parameters using a very small number of parameters.

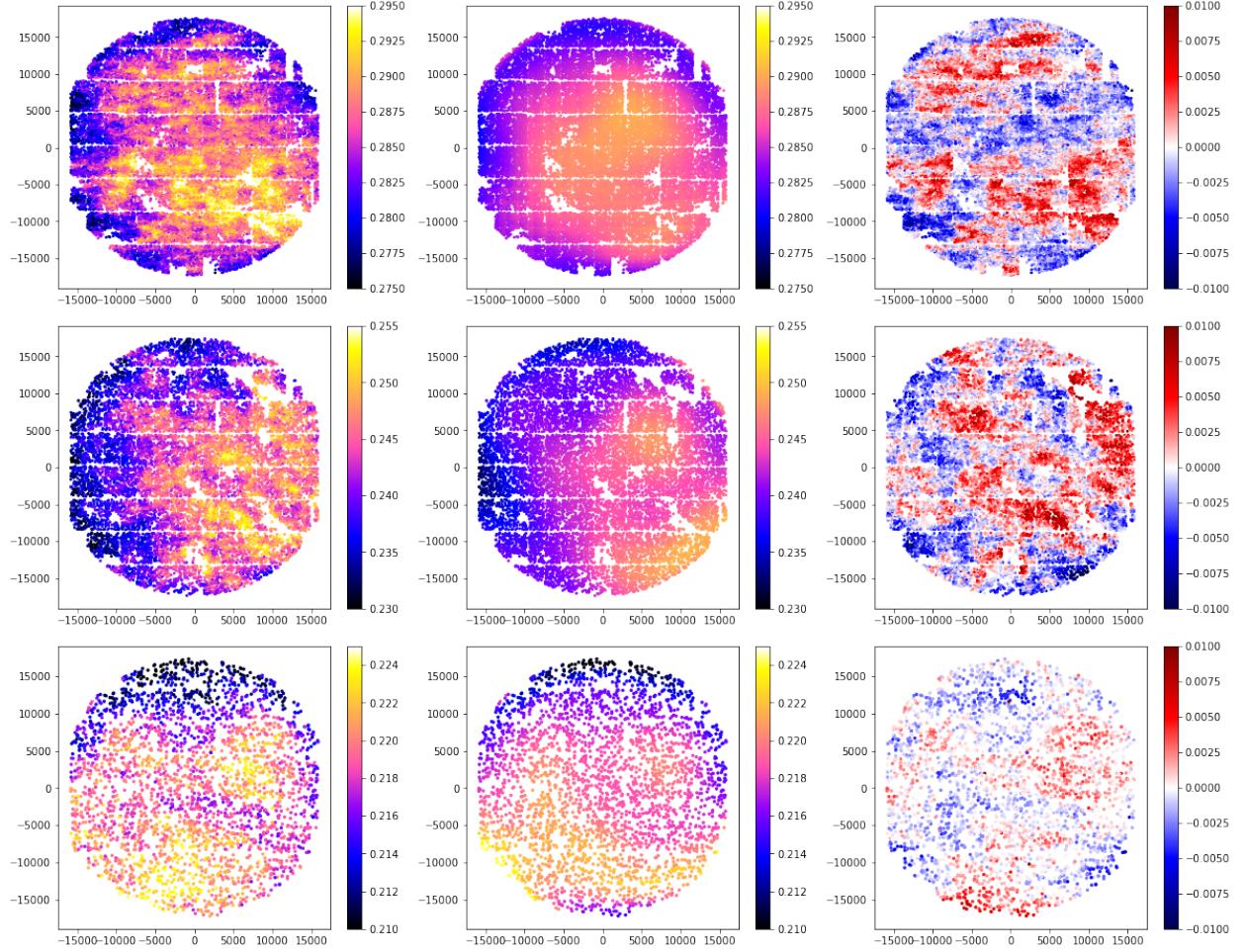


Figure 9: PSF sizes. Different rows correspond to different exposures. The first column are the HSM-measured PSF sizes in the data. The middle column are the best-fit model predictions. The right column shows the residuals.

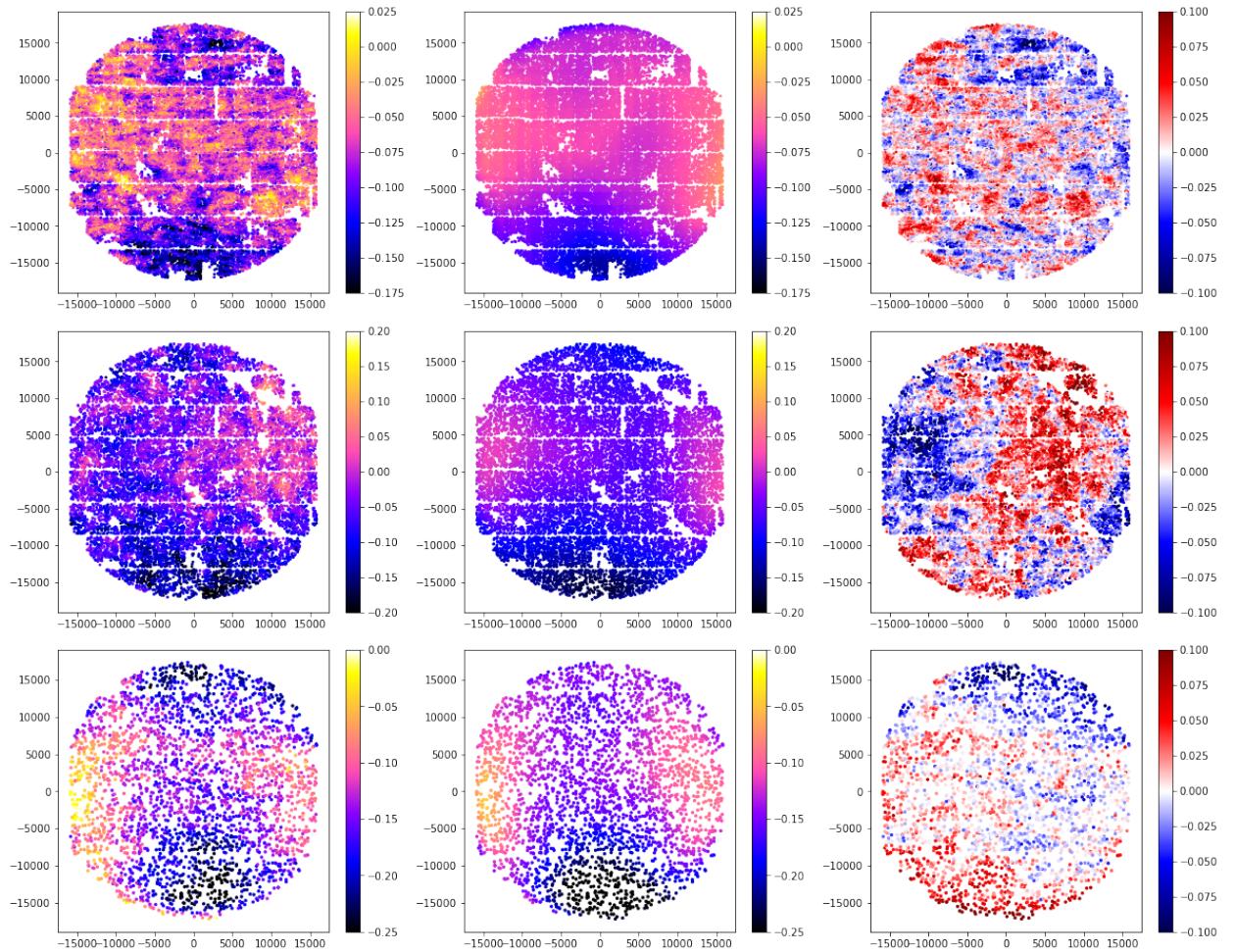


Figure 10: PSF e1 component. Rows and columns are as in 9

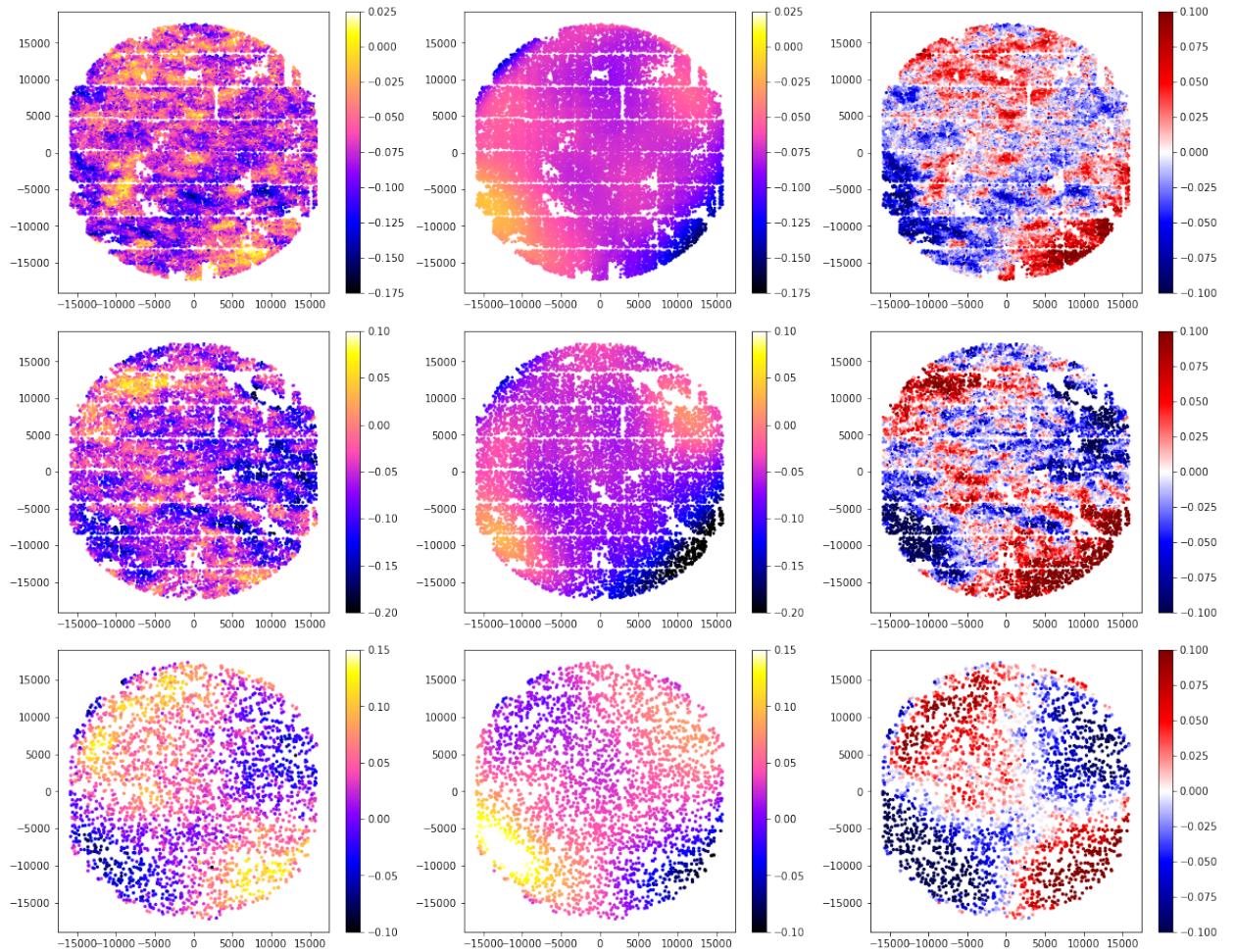


Figure 11: PSF e2 component. Rows and columns are as in 9

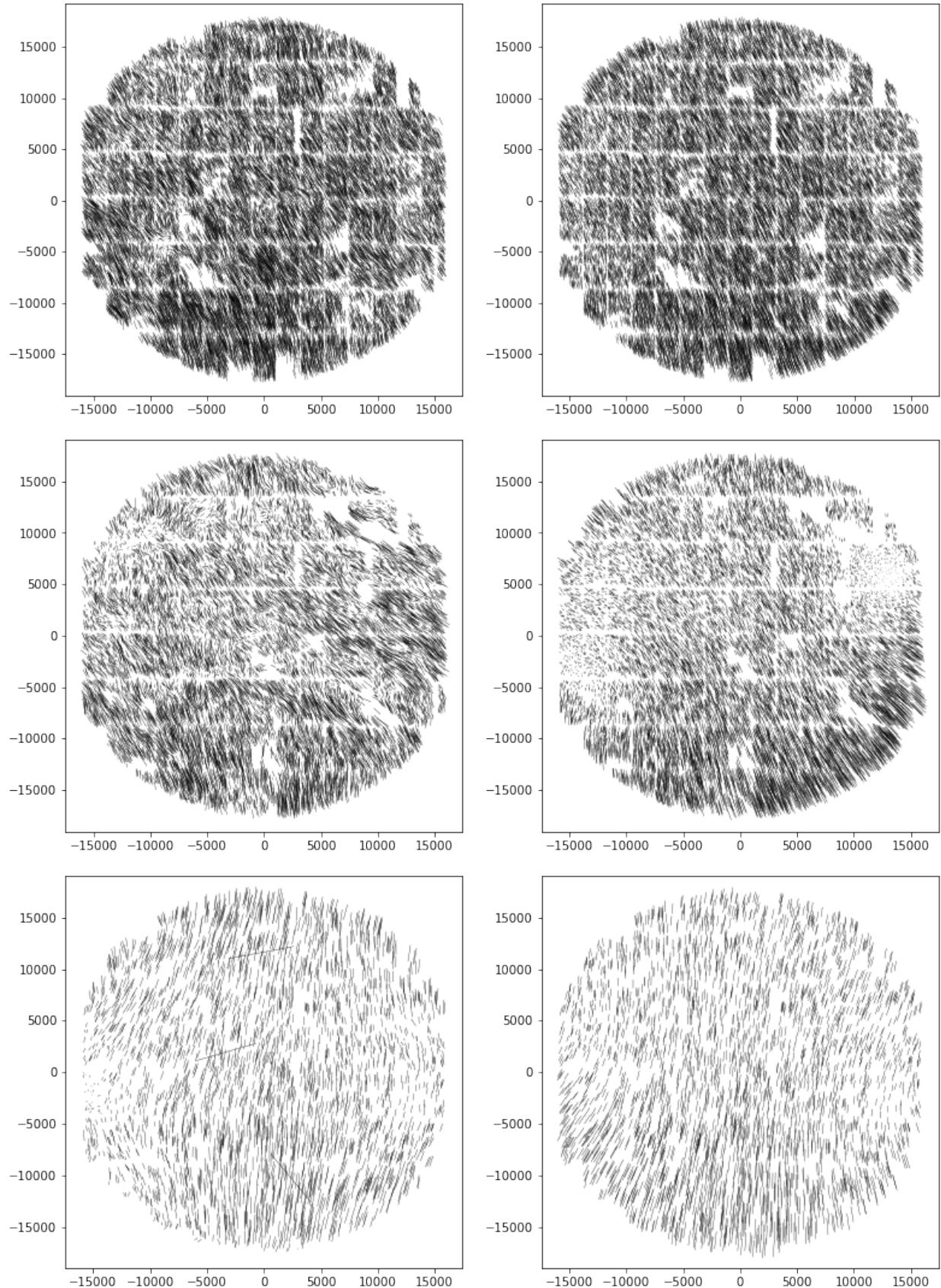


Figure 12: PSF whisker comparison. Rows indicate different exposures. The left column is data, the right column is the best fitting model.