Machine Learning Determination of Wavefront Perturbations

Jun Yin, Chris Stubbs
Harvard University

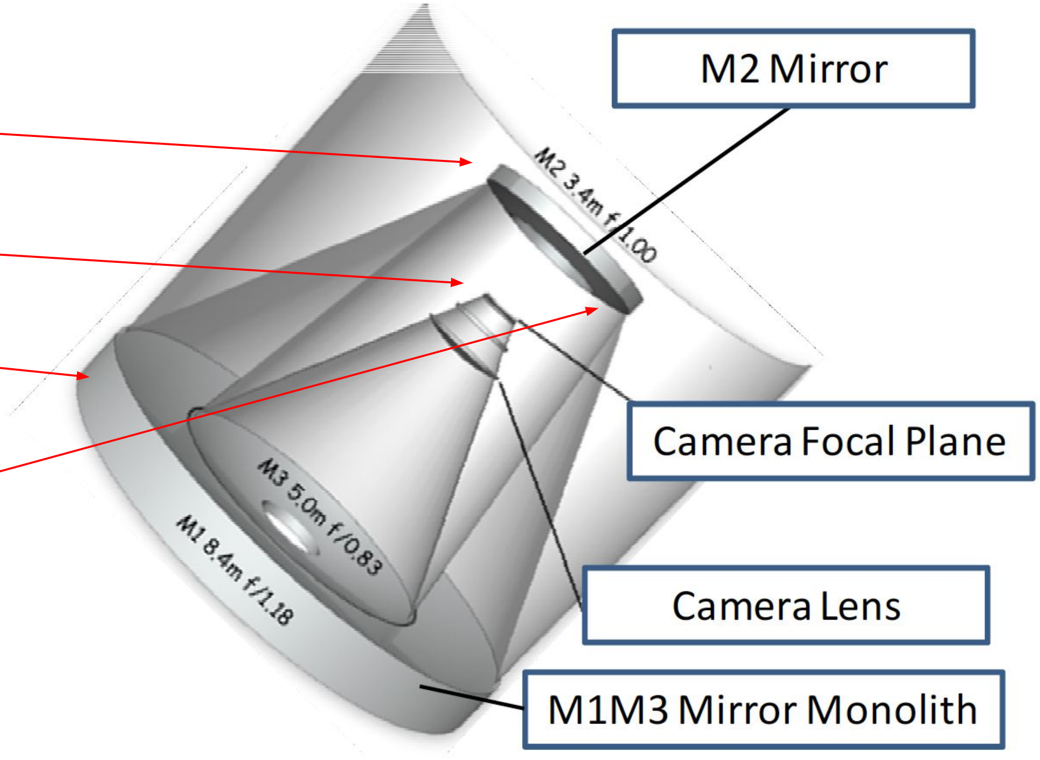
Control parameters

M2: dz, dx, dy, Rx, Ry

Cam: dz, dx, dy, Rx, Ry

M1M3 bending: 20 DOFs

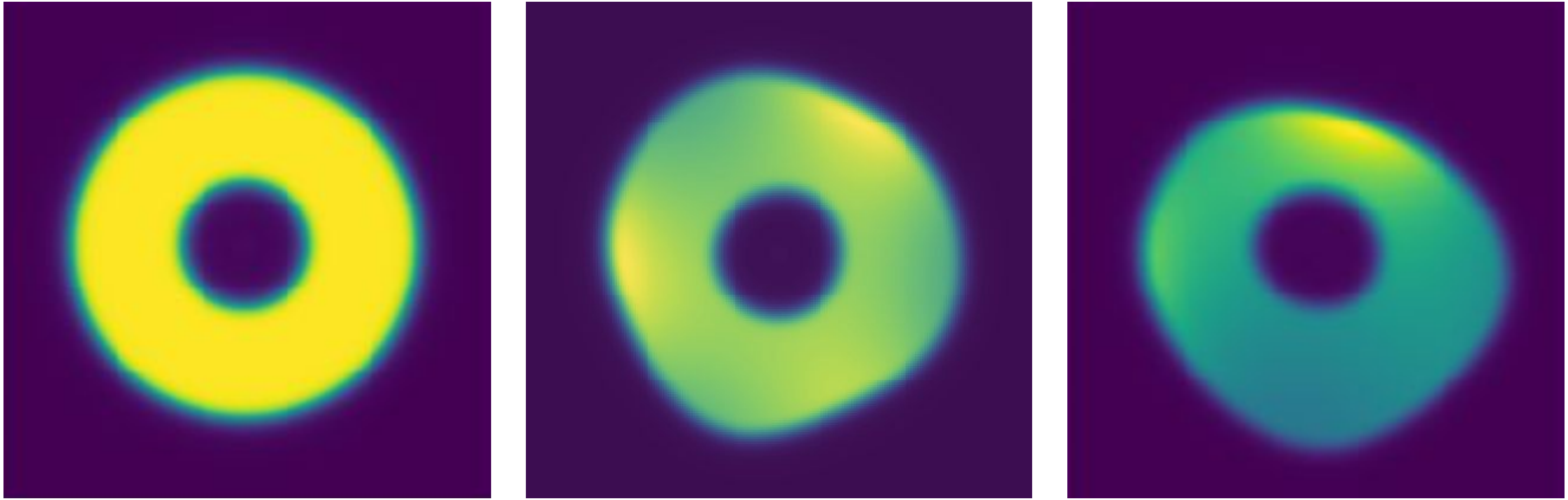
M2 bending: 20 DOFs



Proc. of SPIE Vol. 9150 91500G-1

Goal: predicting the control parameters by analyzing the donut images

No Zernikes, No matrix inversions



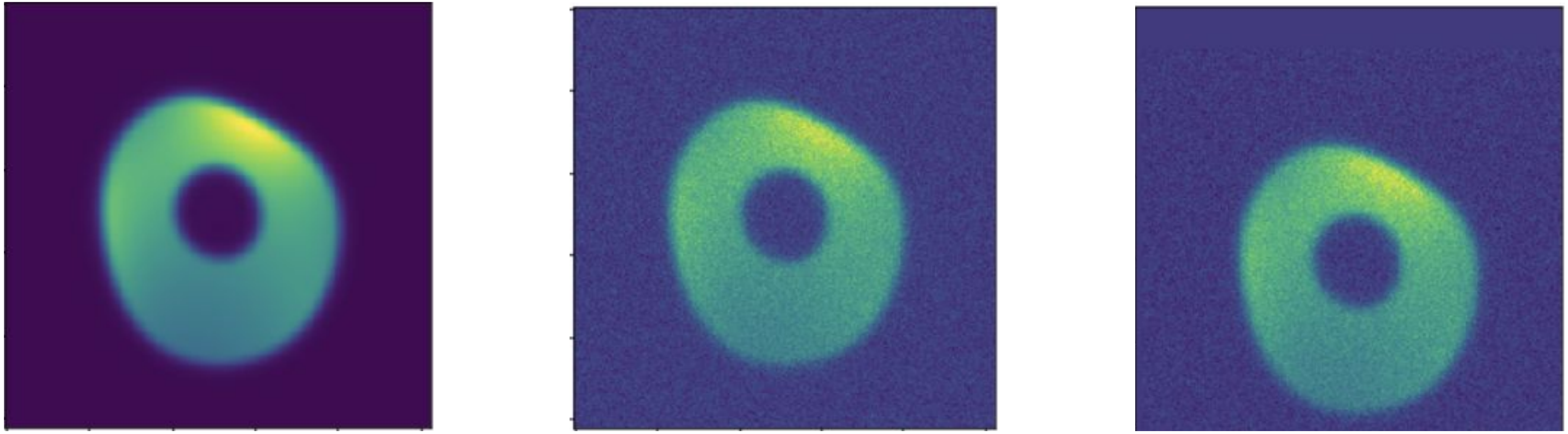
All CPs = 0

50 CPs with non zero values

Current methods and performance.

- The current methods analyze the donut images and extract the Zernike amplitudes. Information Loss
- To adjust the telescope, we need to convert Zernike amplitude to Control Parameters. The conversion from Zernike amplitudes to Control Parameters is not always uniquely defined or invertible. Ill conditioned
- A new approach was developed to obtain the Control Parameters directly from the donut images. **This has been the main focus of my work.**
- The current methods can take from a few seconds up to 1 minute to extract the Zernike amplitudes. Time consuming

Donut Generation



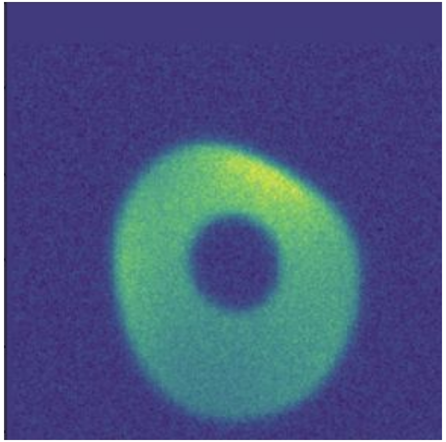
No noise and no shift   With noise but no shift   With noise and shift

Donut images were produced using the *makedonut* code provided by A. Roodman.

Observations

Donut is not always in the center of the postage stamp

-> We want the neural networks to be translation-invariant



There is noise

-> We want the neural networks robust to noise

L2 loss is inferior

-> We want to add or learn priors to compute loss

Review on the problem set up

- Goal: recover 50 control parameters from donut images
- Data: 30,000 x 8 donuts images (64 x 64) with S/N = 10 to 100
- Machine learning problem setup:

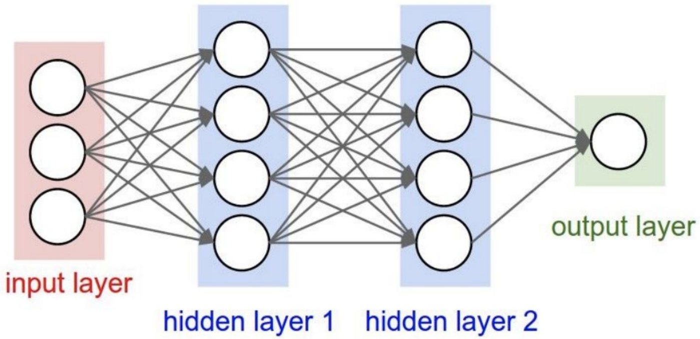
  Train on the pair of donut images with the control parameters as an array.

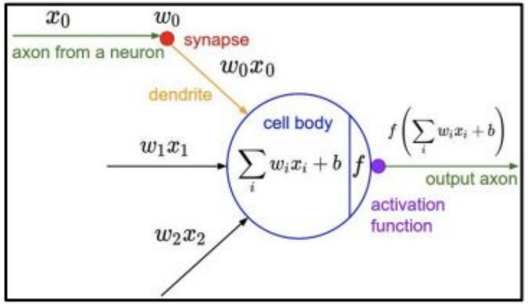
  Test on how accurately the algorithm is able to recover the control parameters.

  Validate the result.

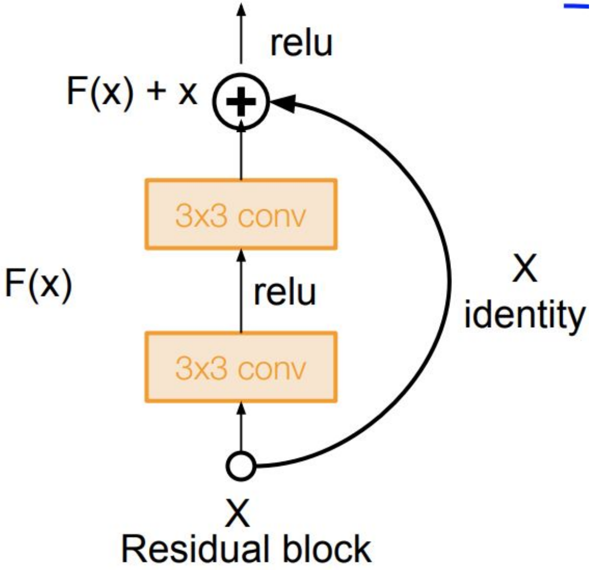
Regression!

Review on Machine Learning (ML)



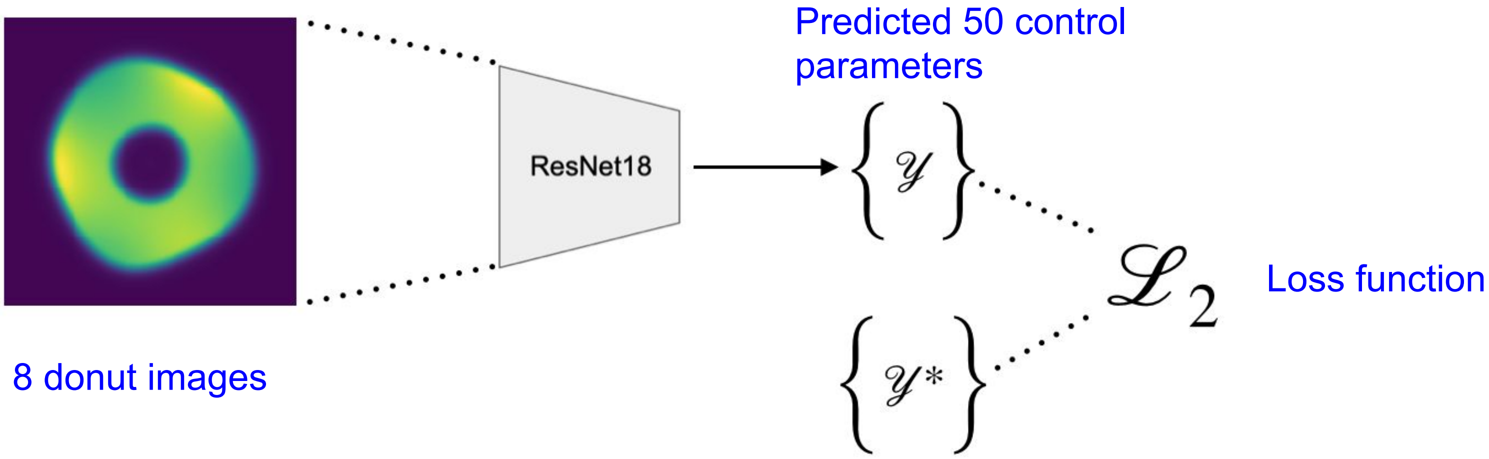


- Input
- Weights
- Bias
- Activation function
- Hidden layers
- Output
- Loss function
- Backpropagation
- Parameters update



Training details

- Deep convolutional neural nets: ResNet18
- Self attention module: donoise
- Anti aliasing max pooling: shift invariant
- Loss function: scaled L2 + PSF



Results:

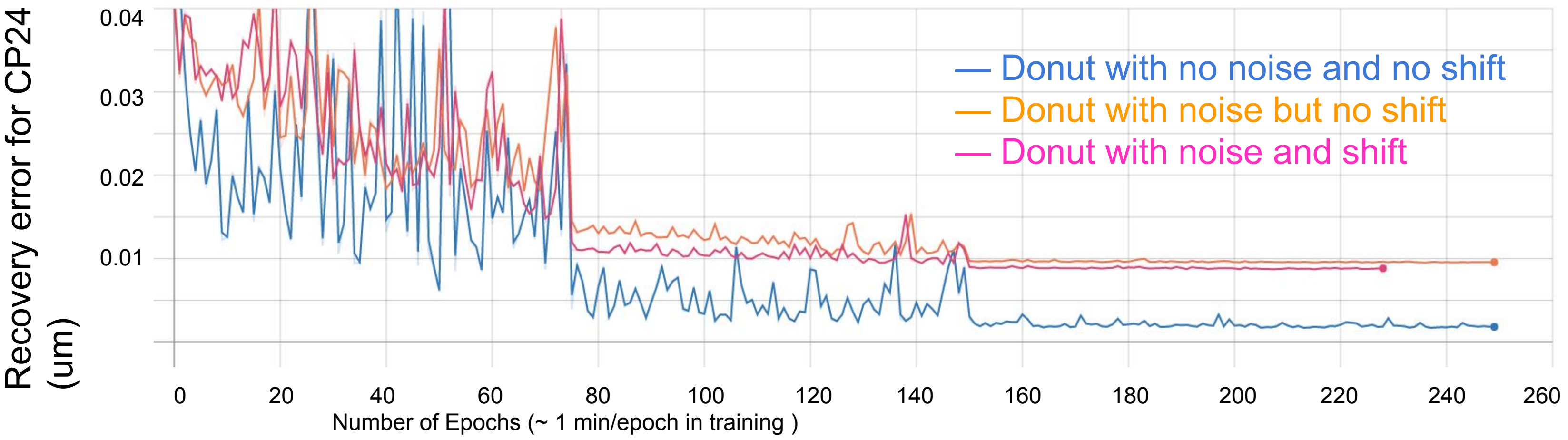
CP	M2dZ (um)	M2dX (um)	M2dY (um)	M2tiltX (arcsec)	M2tiltY (arcsec)	CamdZ (um)	CamdX (um)	CamdY (um)	CamtiltX (arcsec)	CamtiltY (arcsec)	R2
Error budget	**16**	**210**	**210**	**6**	**6**	**15**	**770**	**770**	**13**	**13**	
No noise No shift											
With noise no shift											
With noise with shift											

Results:

CP	M2dZ (um)	M2dX (um)	M2dY (um)	M2tiltX (arcsec)	M2tiltY (arcsec)	Camd Z (um)	CamdX (um)	CamdY (um)	CamtiltX (arcsec)	CamtiltY (arcsec)	R2 (best=1)
Error budget	**16**	**210**	**210**	**6**	**6**	**15**	**770**	**770**	**13**	**13**	
No noise No shift	3.72	21.01	22.03	1.60	2.15	2.65	18.03	7.82	0.95	0.86	0.89
With noise no shift	11.19	23.04	23.07	1.87	2.25	9.06	20.05	18.72	1.23	1.26	0.61
With noise with shift	11.19	25.01	24.09	2.11	2.18	9.07	23.45	19.56	1.27	1.27	0.59

Results for bending parameters

- Error budget: 0.07 um to 0.24 um
- Recovery error < 0.01 um for donuts without noise and shift
- Recovery error < 0.02 um for donuts with noise and shift



Conclusion

- It’s accurate.
- It’s fast. (~a millisecond)
- It has short prediction time. It prepays the computational time in donut generating and training.
- It directly recovers the perturbations. No zernikes. No inversion.
- The machine learning method could determine the wavefront perturbations well.

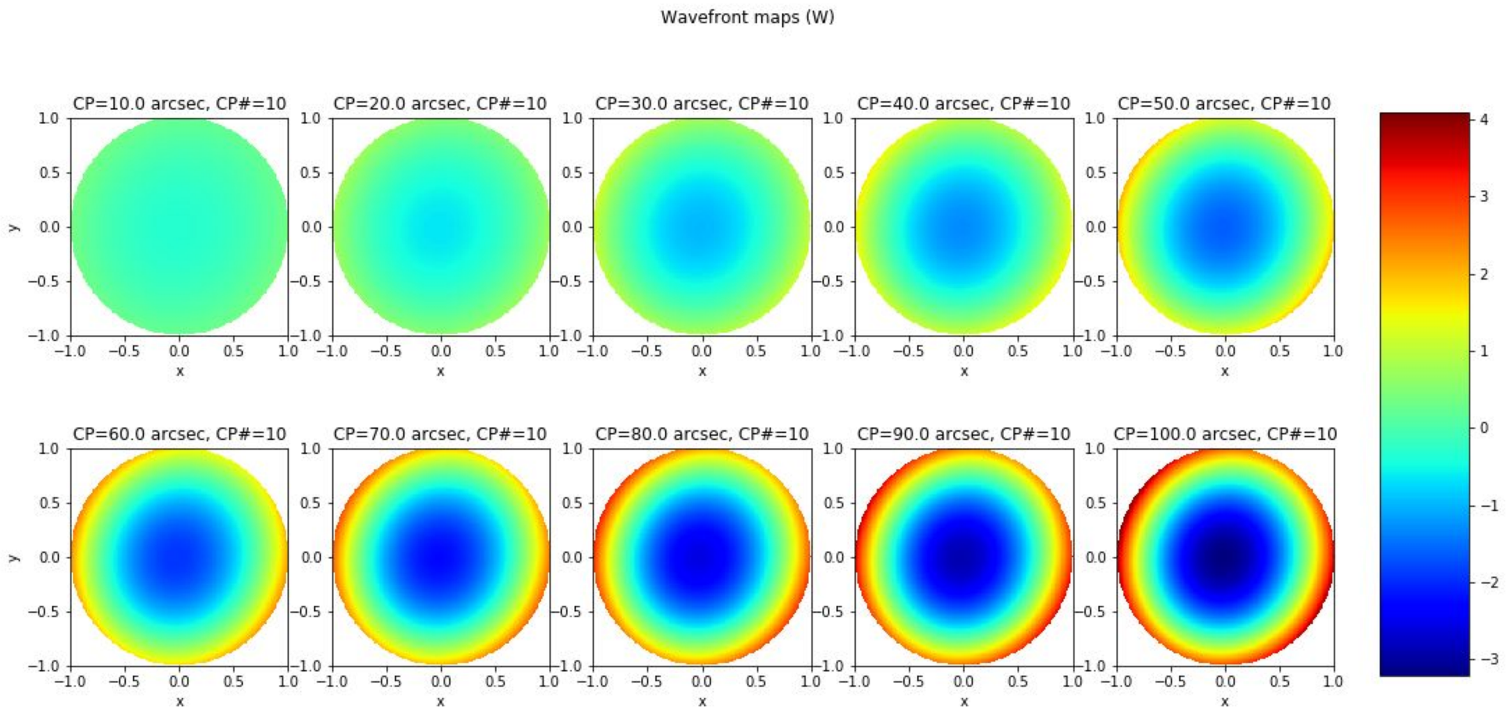
Thank You!

Backup slides

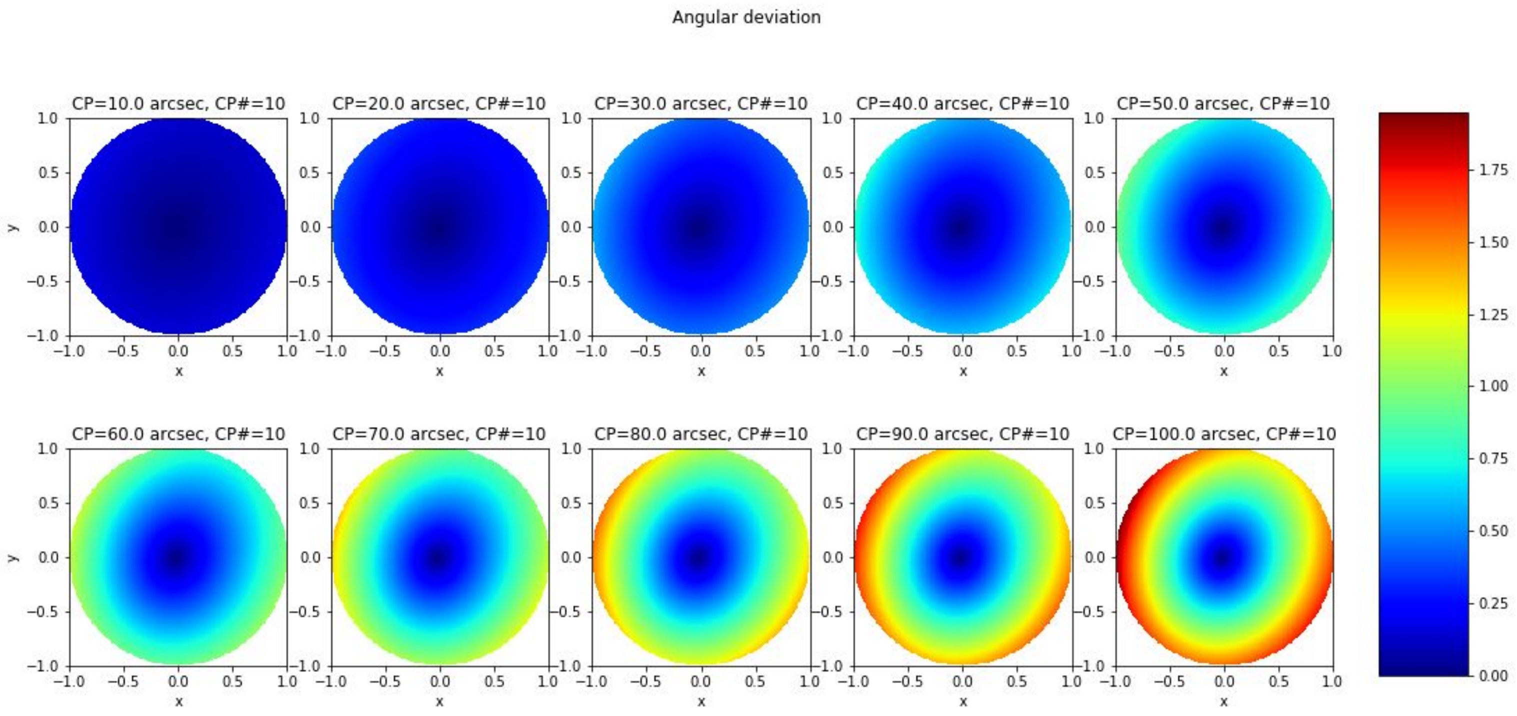
Constraining the Control Parameters (CPs).

- Define the *xy* plane to be plane of the wavefront when all CPs are 0 (all Zernike amplitudes are 0).
- For each CP, calculate the corresponding Zernike amplitudes and construct the “distorted” wavefront.
- Use the gradient of the “distorted” wavefront to determine normal directions to the wavefront across the pupil.
- Calculate the opening angle between the normal and the *z* axis (the normal direction to the undistorted wavefront).
- Define the opening angle q_{80} that contains 80% of all normal vectors. This angle is related to the 80% encircled energy diameter (EE80).
- Find the CP that corresponds to q_{80} = 0.4 urad = 36 mas.

Example: wavefront maps for CP 10 (CAM Ry).

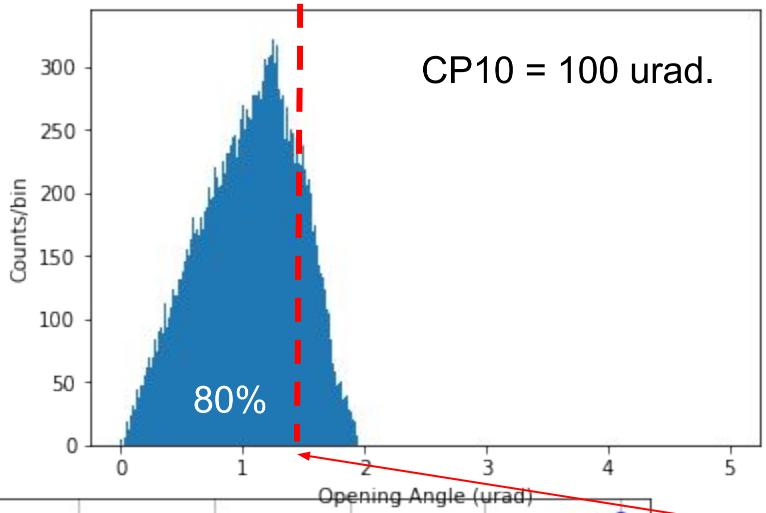


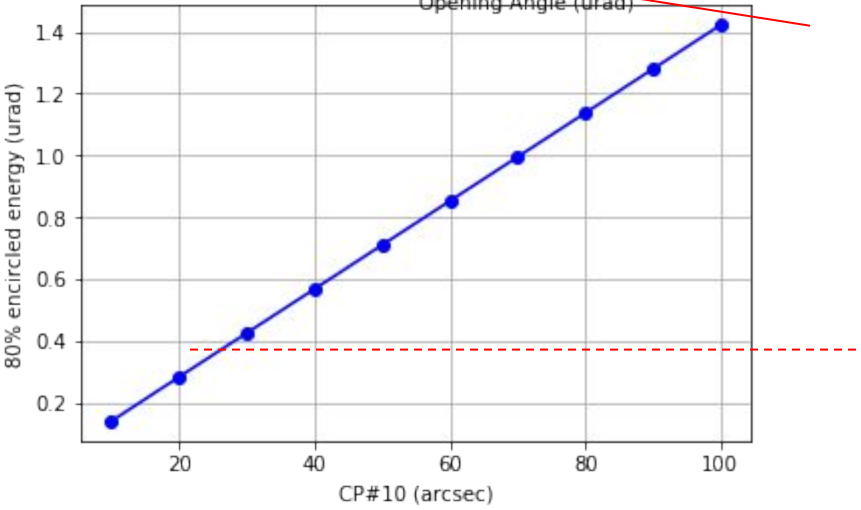
Example: opening angles for CP 10 (CAM Ry).



Example: CP 10 (Cam RY).

- CP 10: CAM y−tilt (in arcsecond).
- Calculate the opening angle at all grid points on the wavefront.
- Find the 80% angle.
- Plot the 80% angle vs the tilt angle.
- Determine the slope and use it to determine when the 80% angle is 0.4 urad: for CP10, the angle is 28 arcsec.





Summary of 80% limits for CP1 - CP10

CP	Type	80% limit	CP	Type	80% limit
1	M2 piston	16 um	6	CAM piston	15 um
2	M2 x−decenter	210 um	7	CAM x−decenter	770 um
3	M2 y−decenter	210 um	8	CAM y−decenter	770 um
4	M2 x−tilt	6 arcsec	9	CAM x−tilt	13 arcsec
5	M2 y−tilt	6 arcsec	10	CAM y−tilt	13 arcsec

The limit is for if there is only one control parameter varies..

Translation-invariant neural networks (BlurPool)

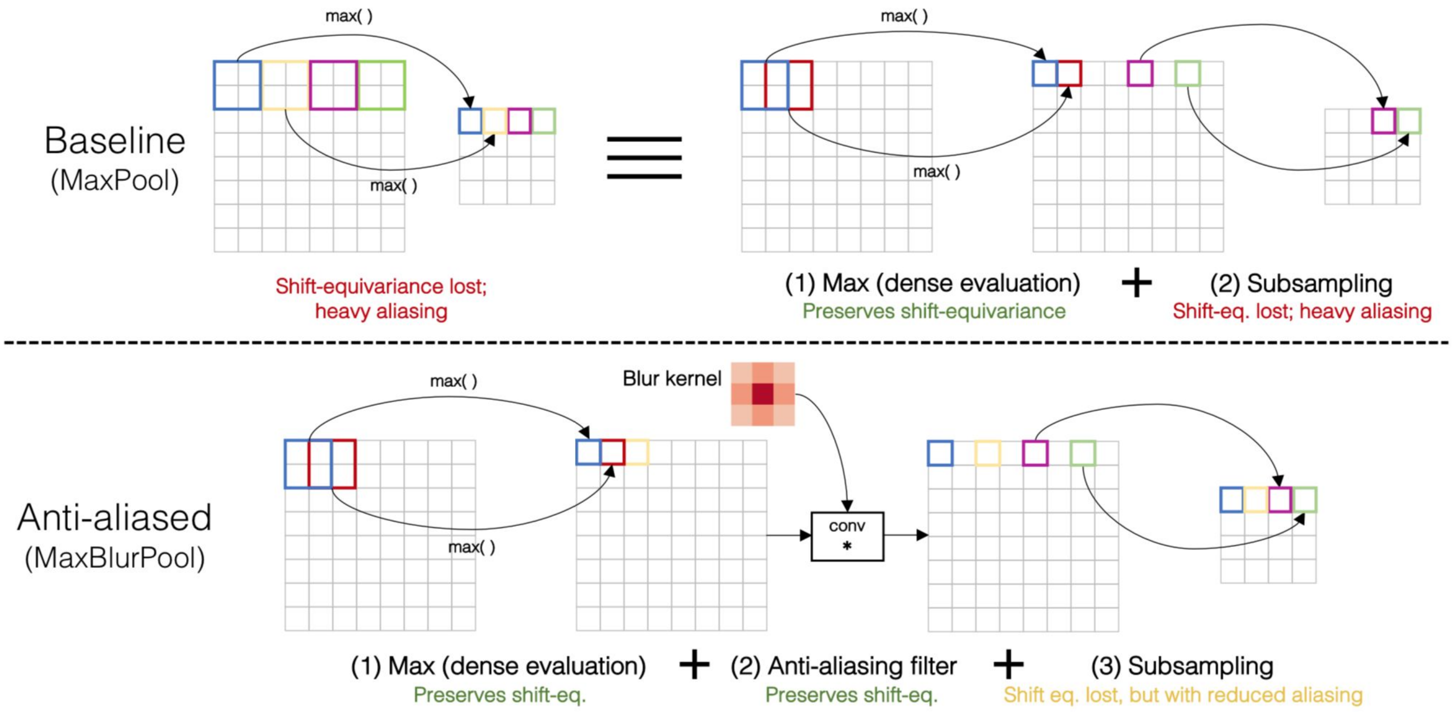


Figure credit: Making Convolutional Networks Shift-invariant Again, https://arxiv.org/abs/1904.11486

Feature denoising using non-local neural networks

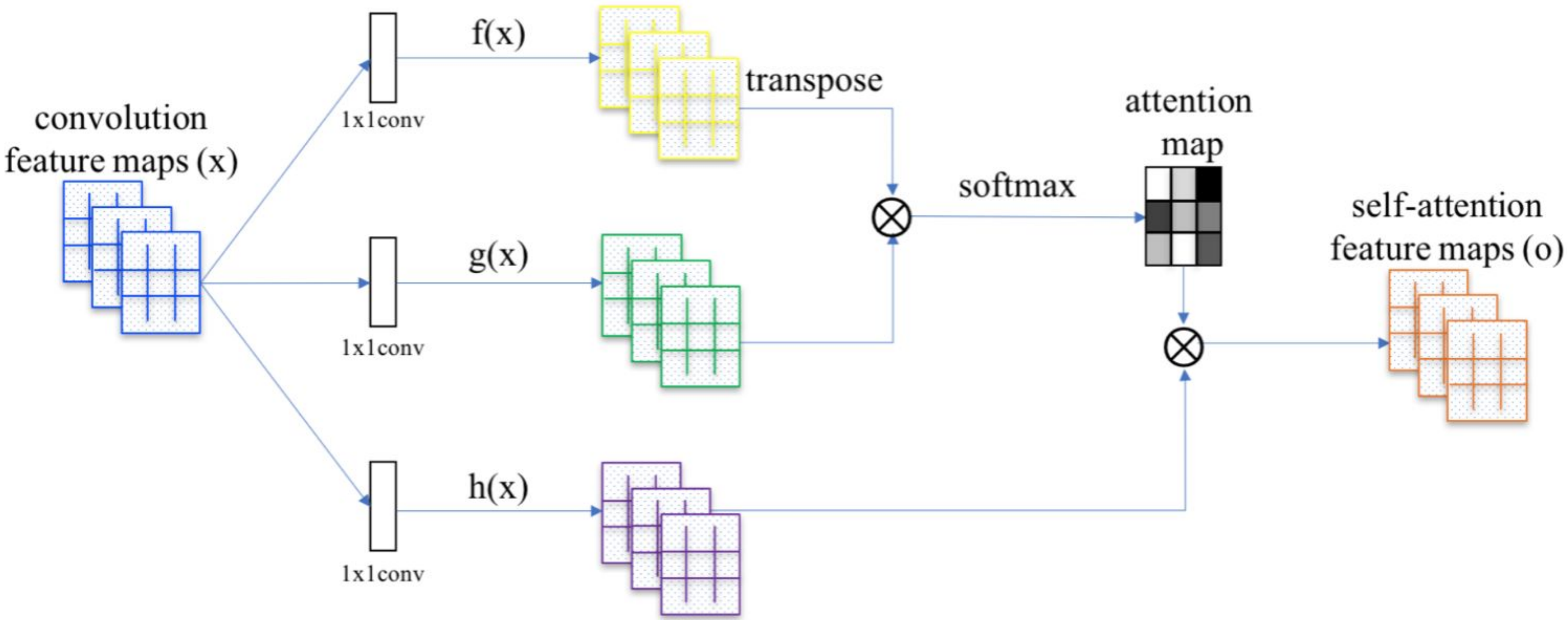
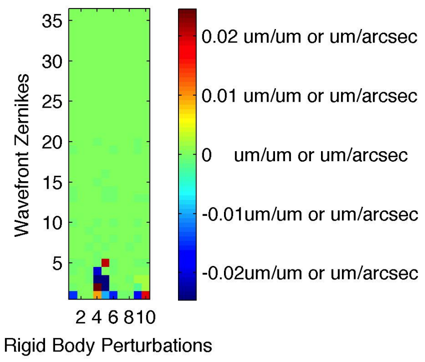


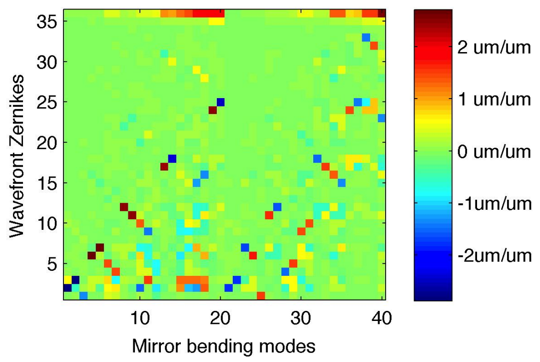
Figure credit: Self-attention GAN: https://arxiv.org/abs/1805.08318

Evaluation

- Predict 10 control parameters fast and accurately
- Compare to traditional method, the ML methods taking
- CP2,3 5,6:

  Optical system is very insensitive to translation motions





The broad landscape of PSF model requirements for LSST dark energy

- Weak lensers want to use the PSF estimated from bright stars, interpolated to the positions of faint galaxies, to estimate WL shear.
- WL will likely generate the most stringent requirements (of all DESC probes) on PSF model quality:

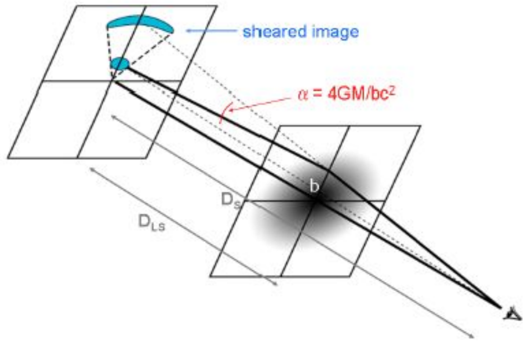


Figure 14.1: The primary observable effect of weak lensing is to impose an apparent tangential ellipticity on background galaxies which are truly randomly aligned.

$$\langle \tilde{g}\tilde{g} \rangle(\theta) = (1+m)\langle gg \rangle(\theta) + f(\theta)$$

PSF model size errors

PSF model shape errors

The amount shearing depends on: Projected mass, separation on sky Separation between us, lens, source

Producing Donut Images

- Donut images were produced using the *makedonut* code provided by A. Roodman.
- For each set of control parameters, the Zernike amplitudes were calculated and used as input for *makedonut*.
- The following parameters were used:
    - bkgr = 0.0. The background is assumed to be zero.
    - xCam = 0.00 and yCam = 0.00. The center of the camera is assumed to coincide with the center of the focal plane.
    - nPixels = 224. Each donut image is 224 by 224 pixels.
    - nbin = 1792. This is the number of bins of the arrays that are used for the Fourier analysis.
    - eEle = 1E6 or 1E8 This is the total number of electrons in the image, distributed across all pixels. This corresponds to 100 or 10,000 electrons/pixel (S/N = 10 or S/N=100).
    - Wavelength = 700 nm.
    - Poisson statistics is used to determine the actual observed number of electrons in each pixel

PROCEEDINGS OF SPIE

SPIEDigitalLibrary.org/conference-proceedings-of-spie

Wavefront sensing and the active optics system of the dark energy camera

Aaron Roodman, Kevin Reil, Chris Davis

In Ground-based and Airborne Telescopes V, volume 9145 of Proc. SPIE, page 914516, July 2014.

Current methods and performance.

- The paper by Angeli *et al*. provide examples of the performance of current methods.
- For example for a ±1 mm sensor separation:
  - Maximum non-linear behavior in reconstructed Z7 is 300 nm.
  - Maximum non-linear behavior in reconstructed Z4 is -50 nm.

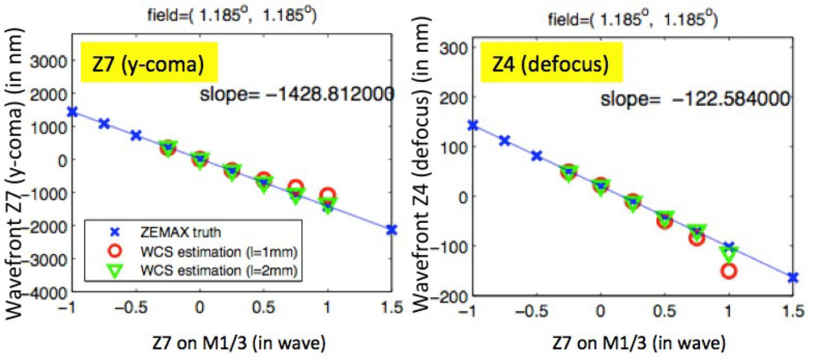


Figure 7 Representative examples of the WCS linearity tests validating functionality and performance.

Proc. of SPIE Vol. 9150 91500H-1

A few remarks from "*LSST Active Optics System Software Architecture*", Thomas *et al*.

- The overall system image quality budget for the LSST is 0.4 arcsec FWHM with 0.25 arcsec allocated to the telescope and 0.3 arcsec associated to the camera.
- The corrections given by the look-up-table (LUT), however, do not allow the system to meet the LSST image quality requirements due to non-repeatable and/or unpredictable effects. These effects include temperature errors, wind loads and hysteresis.
- LSST plans to use real-time wavefront sensor measurements to compensate for these errors to achieve the LSST required image quality.
- In order to maintain the required image quality, the AOS output consists of the bending modes sent to the mirror support system of M1M3 and M2 to control the mirror shapes, and positions sent to the M2 and camera hexapods.
- The observatory cadence of the LSST telescope is relatively fast since the telescope pointing changes every 39 seconds. These 39 seconds are decomposed into two 16s visits (1s of which is to allow the shutter to open and close), a 2s readout and a 5s slew time. The current baseline plan uses the data from the first visit only and performs the calculation during the second visit.

Alternative methods and performance.

PUBLICATIONS OF THE ASTRONOMICAL SOCIETY OF THE PACIFIC, **118**: 1165–1175, 2006 August

Donut: Measuring Optical Aberrations from a Single Extrafocal Image

A. TOKOVININ AND S. HEATHCOTE

Cerro Tololo Inter-American Observatory, Casilla 603, La Serena, Chile; atokovinin@ctio.noao.edu, sheathcote@ctio.noao.edu

Received 2005 December 20; accepted 2006 June 16; published 2006 July 20

ABSTRACT. We propose a practical method to calculate Zernike aberrations from analysis of a single long-exposure defocused stellar image. It consists of fitting the aberration coefficients and the seeing blur directly to a realistic image binned into detector pixels. This "Donut" method is different from curvature sensing in that it does not make the usual approximation of linearity. We calculate the sensitivity of this technique to detector and photon noise and determine optimal parameters for some representative cases. Aliasing of high-order unmodeled aberrations is evaluated and shown to be similar to a low-order Shack-Hartmann sensor. The method has been tested with real data from the SOAR and Blanco 4 m telescopes.

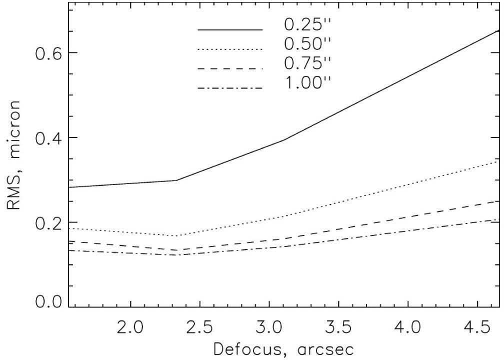


FIG. 6.—The rms noise of the astigmatism coefficient a_5 for various diameters of the donut and different CCD pixel scales (indicated on the plot) under 1″ seeing. Readout noise is 10 e^-, and $N_{\rm ph} = 1000$.

$$I_0 = \sum I_{ij}, \quad x_c = I_0^{-1} \sum x_{ij} I_{ij}, \quad y_c = I_0^{-1} \sum y_{ij} I_{ij},$$

$$M_x = I_0^{-1} \sum (x_{ij} - x_c)^2 I_{ij}, \quad M_y = I_0^{-1} \sum (y_{ij} - y_c)^2 I_{ij},$$

$$M_{xy} = I_0^{-1} \sum (x_{ij} - x_c)(y_{ij} - y_c) I_{ij}.$$

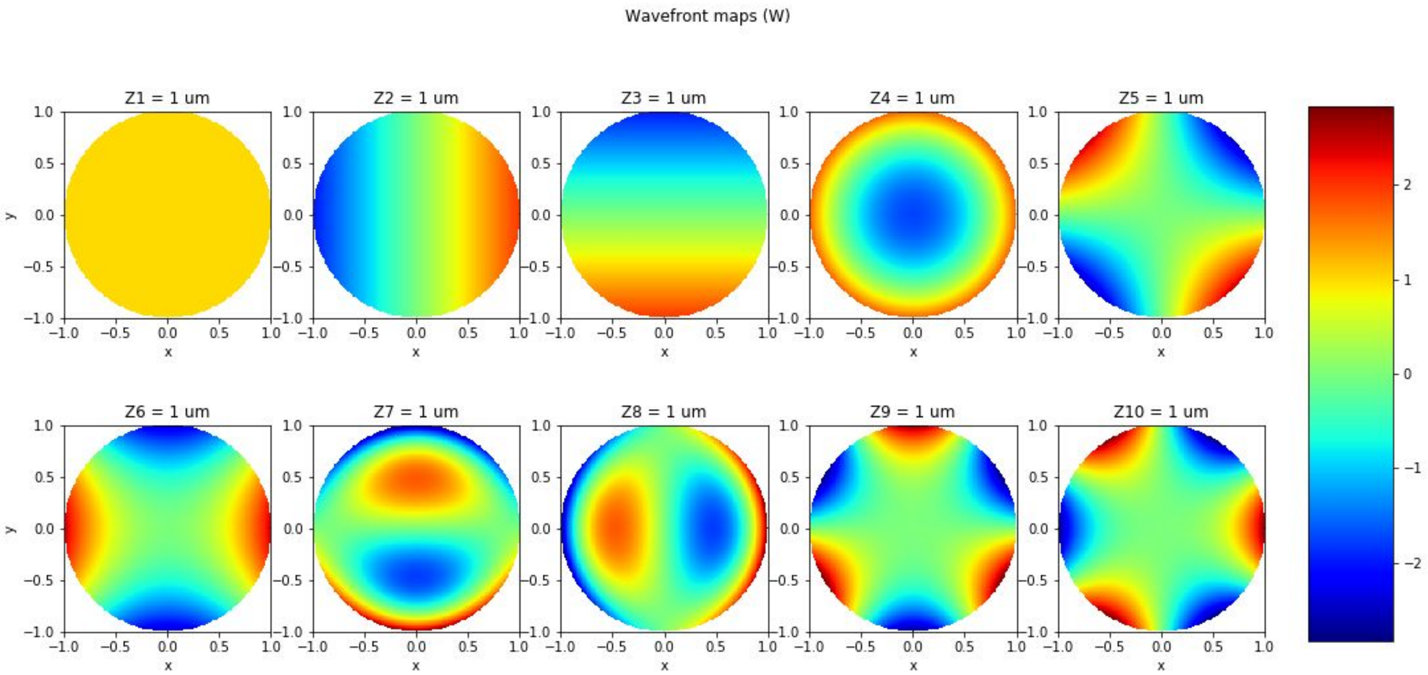
$$\boxed{A_4} = p\sqrt{(M_x + M_y)/2},$$

$$\boxed{A_5} = pM_{xy}(M_x M_y)^{-1/4},$$

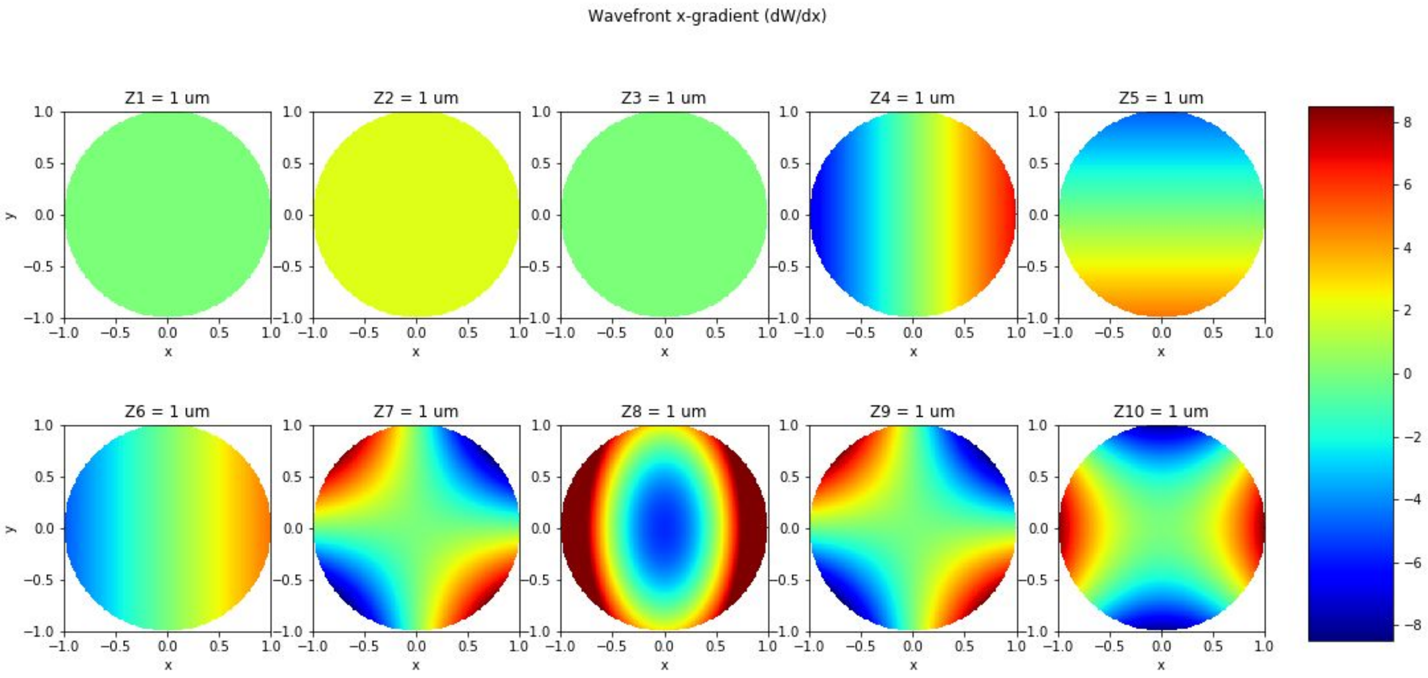
$$\boxed{A_6} = 0.5p(M_x - M_y)(M_x M_y)^{-1/4}.$$

Zernike amplitudes

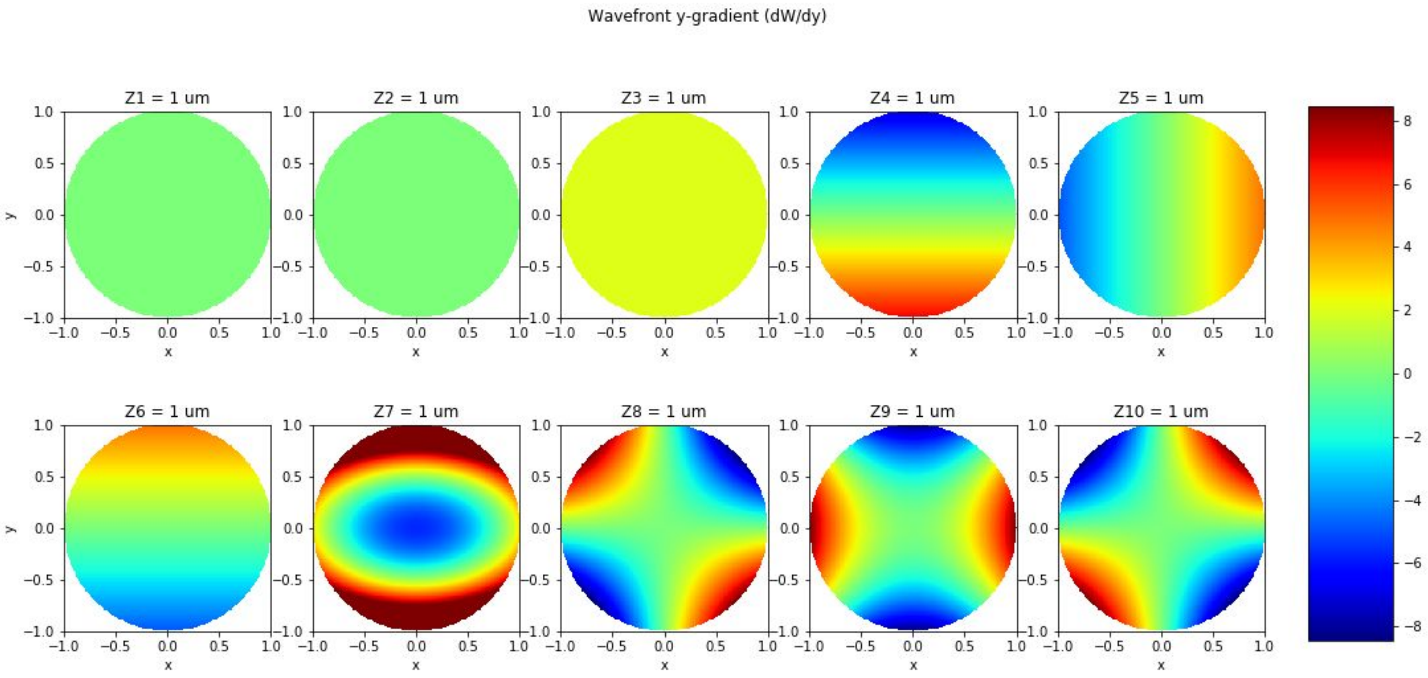
Wavefront for Zernikes Z1 - Z10.



X-component of wavefront gradient for Z1 - Z10.



Y-component of wavefront gradient for Z1 - Z10



Angular deviation across the wavefront for Z1 - Z10.

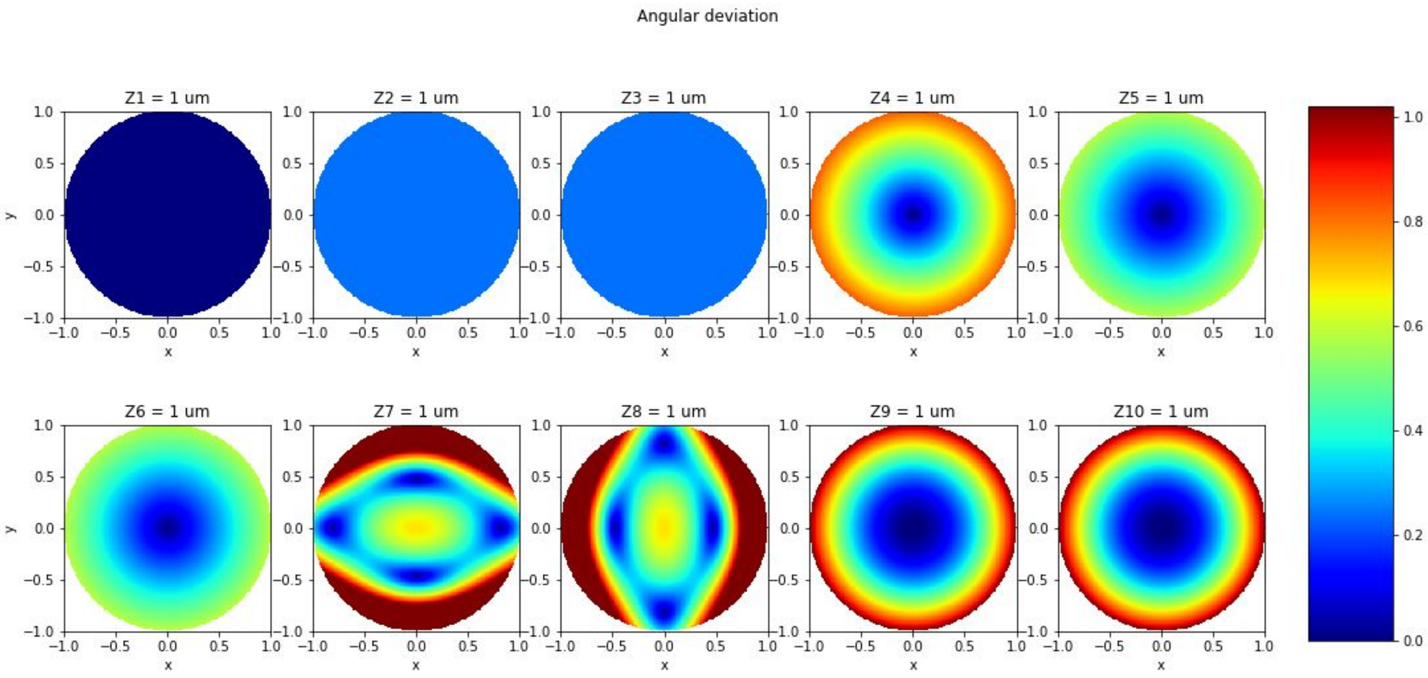
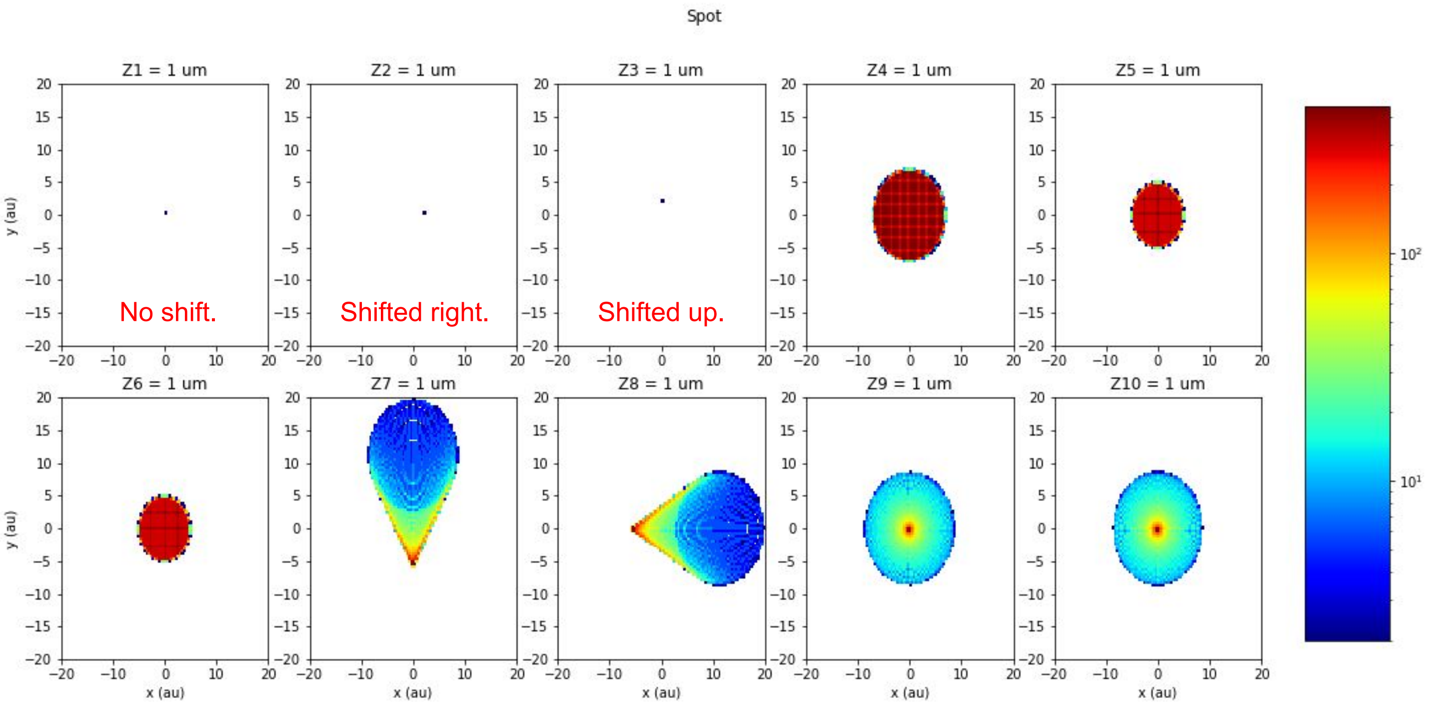


Image shape for Z1 - Z10.



Zernike functions (Noll indices):

TABLE I. Zernike polynomials. The modes, Z_j, are ordered such that even j corresponds to the symmetric modes defined by $\cos m\theta$, while odd j corresponds to the antisymmetric modes given by $\sin m\theta$. For a given n, modes with a lower value of m are ordered first.

Radial degree (n)	Azimuthal frequency (m) 0	1	2	3	4	5
0	$Z_1=1$ Constant					
1		$Z_2=2r\cos\theta$ $Z_3=2r\sin\theta$ Tilts (Lateral position)				
2	$Z_4=\sqrt{3}(2r^2-1)$ Defocus (Longitudinal position)		$Z_5=\sqrt{6}r^2\sin 2\theta$ $Z_6=\sqrt{6}r^2\cos 2\theta$ Astigmatism (3rd Order)			
3		$Z_7=\sqrt{8}(3r^3-2r)\sin\theta$ $Z_8=\sqrt{8}(3r^3-2r)\cos\theta$ Coma (3rd order)		$Z_9=\sqrt{8}r^3\sin 3\theta$ $Z_{10}=\sqrt{8}r^3\cos 3\theta$		
4	$Z_{11}=\sqrt{5}(6r^4-6r^2+1)$ 3rd order spherical		$Z_{12}=\sqrt{10}(4r^4-3r^2)\cos 2\theta$ $Z_{13}=\sqrt{10}(4r^4-3r^2)\sin 2\theta$		$Z_{14}=\sqrt{10}r^4\cos 4\theta$ $Z_{15}=\sqrt{10}r^4\sin 4\theta$	
5		$Z_{16}=\sqrt{12}(10r^5-12r^3+3r)\cos\theta$ $Z_{17}=\sqrt{12}(10r^5-12r^3+3r)\sin\theta$		$Z_{18}=\sqrt{12}(5r^5-4r^3)\cos 3\theta$ $Z_{19}=\sqrt{12}(5r^5-4r^3)\sin 3\theta$		$Z_{20}=\sqrt{12}r^5\cos 5\theta$ $Z_{21}=\sqrt{12}r^5\sin 5\theta$
6	$Z_{22}=\sqrt{7}(20r^6-30r^4+12r^2-1)$ 5th order spherical		Z_{23} Z_{24}		Z_{25} Z_{26}	

Robert J. Noll, Journal of the Optical Society of America, Vol. 66, Issue 3, pp. 207-211,(1976).

Wavefronts for CP1.

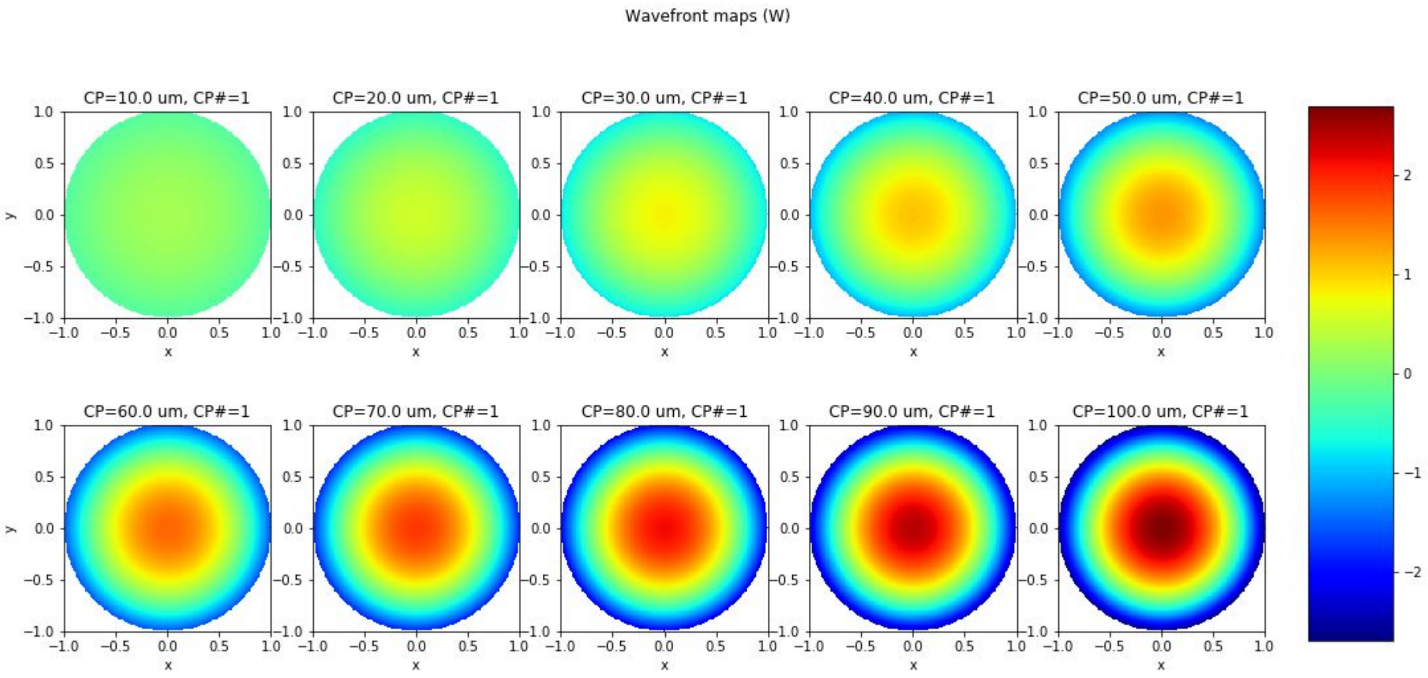
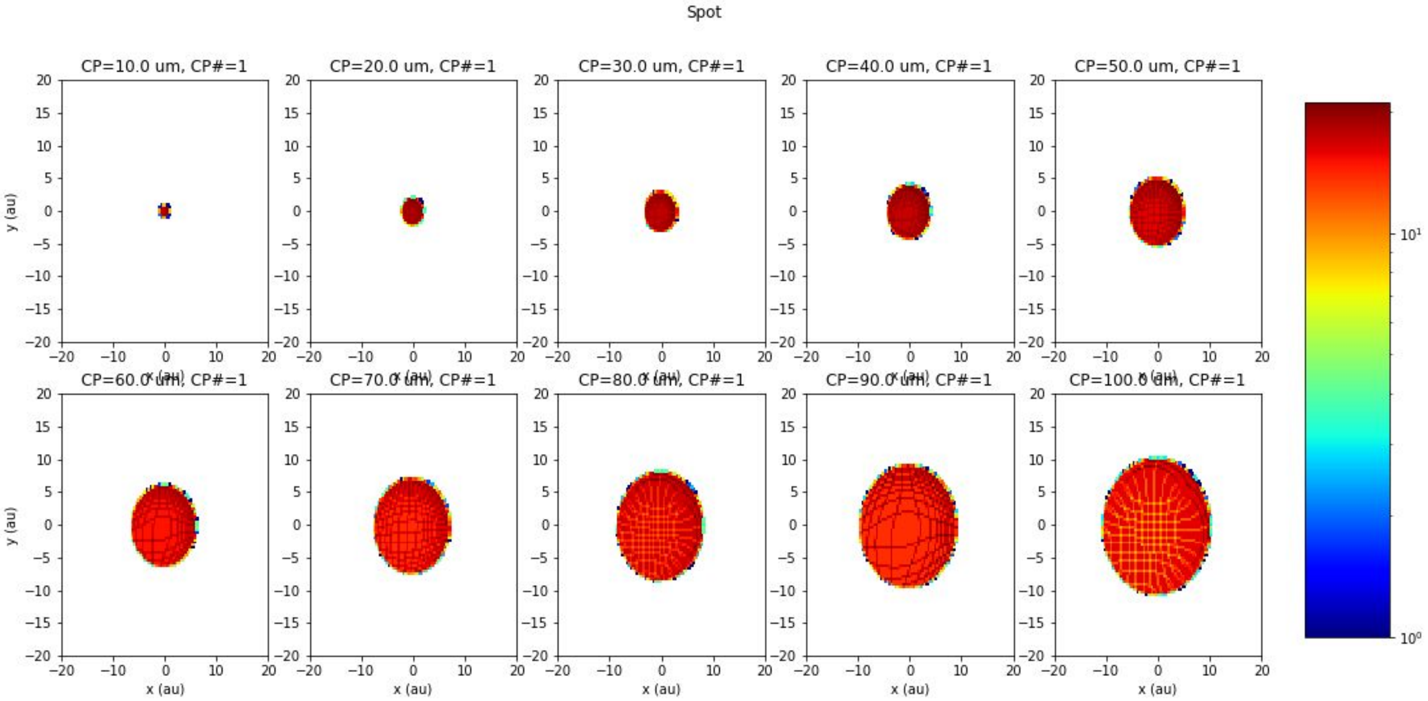
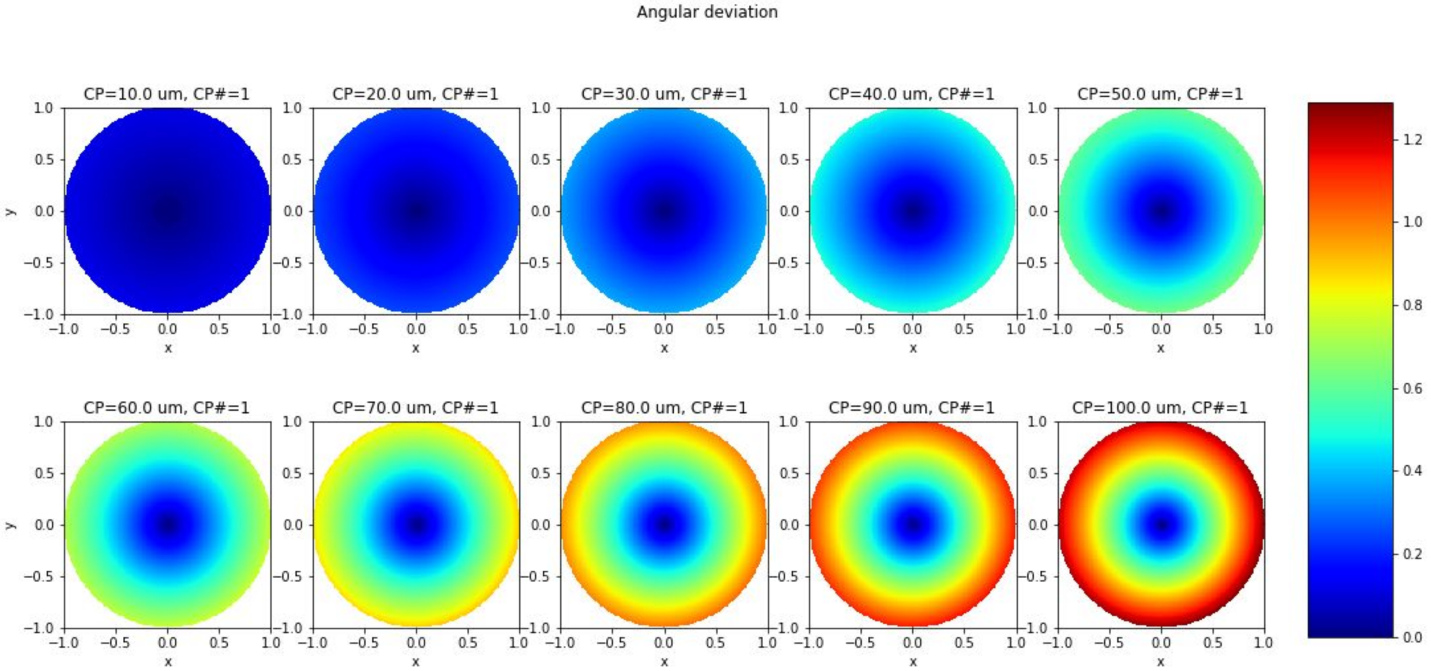


Image Spots for CP1.



Angular Deviation for CP1.



Wavefronts for CP2.

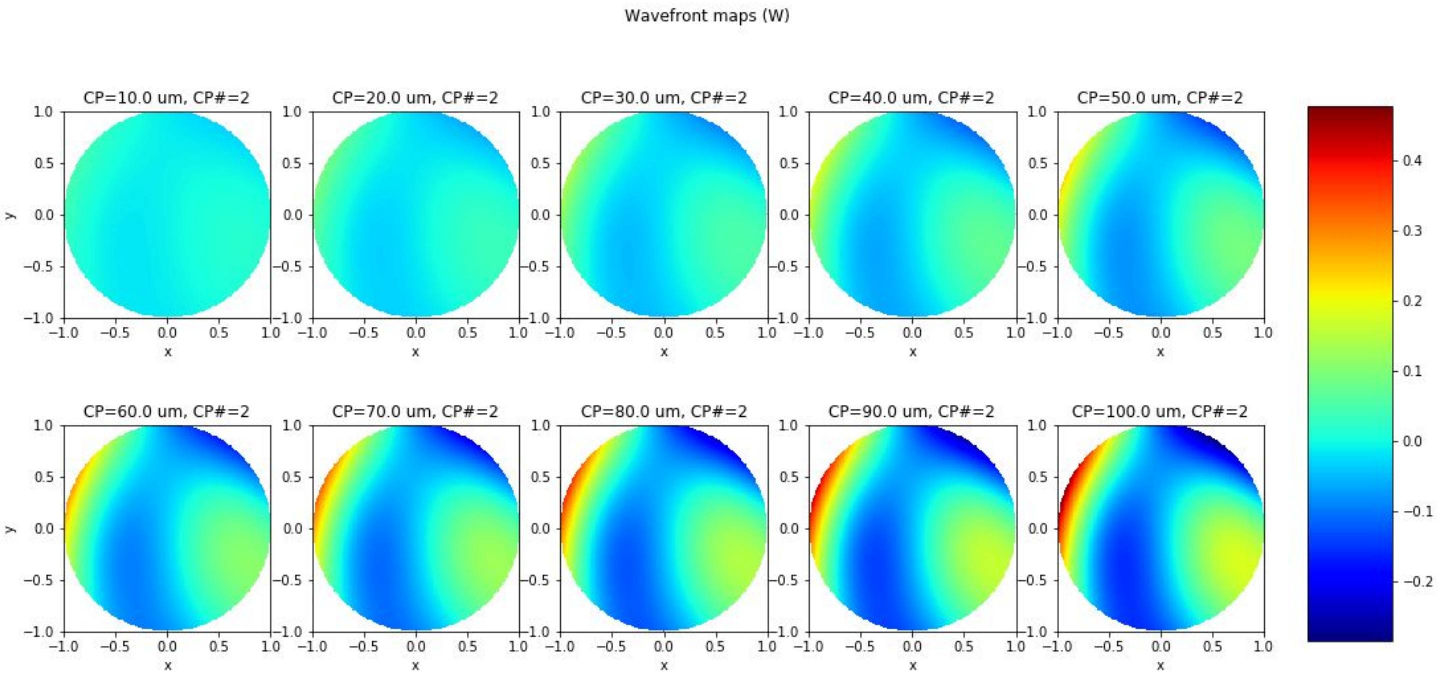
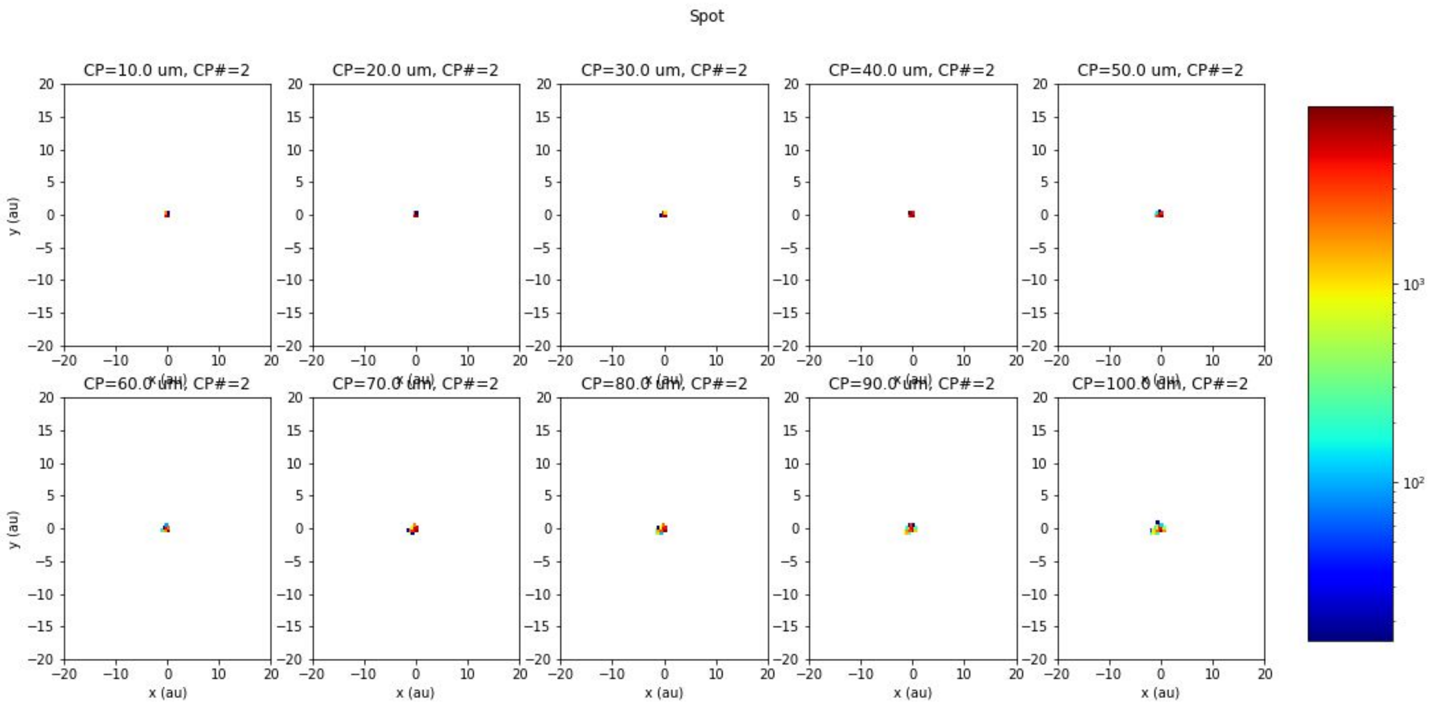
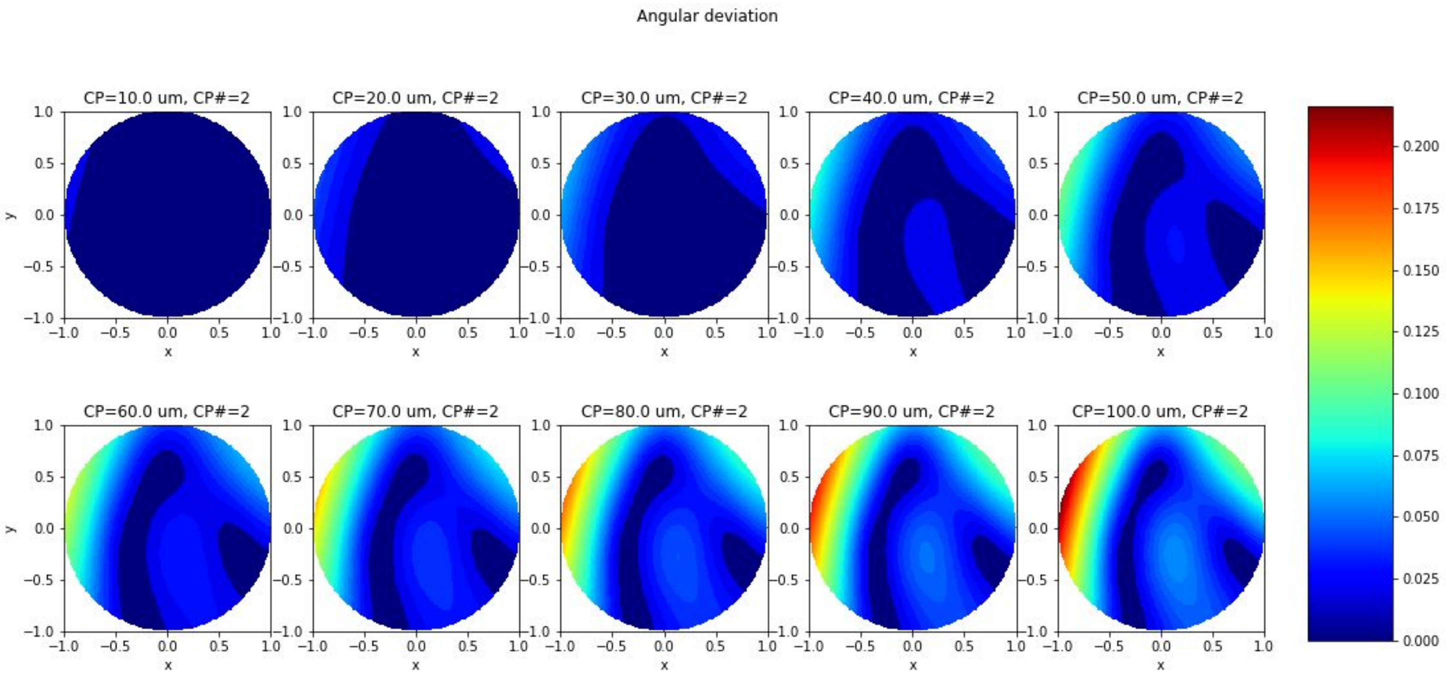


Image Spots for CP2.



Angular Deviation for CP2.



Wavefronts for CP3.

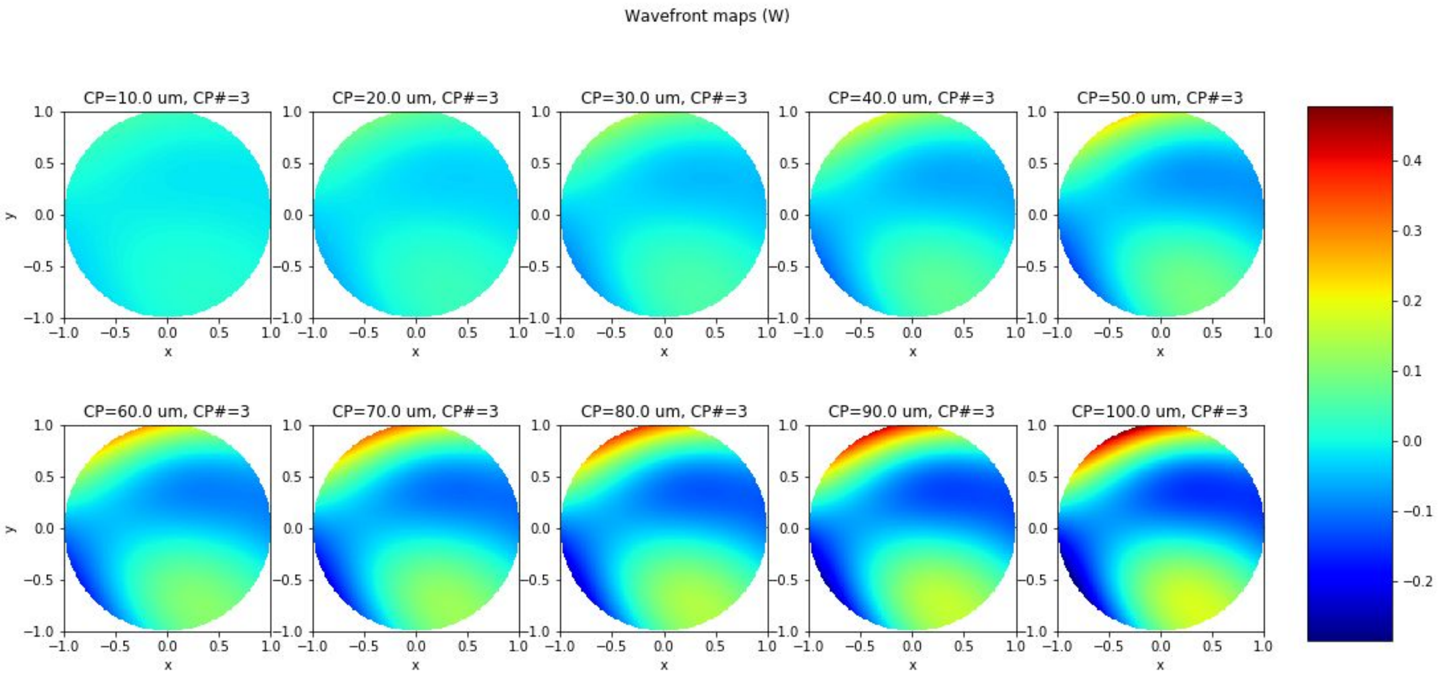
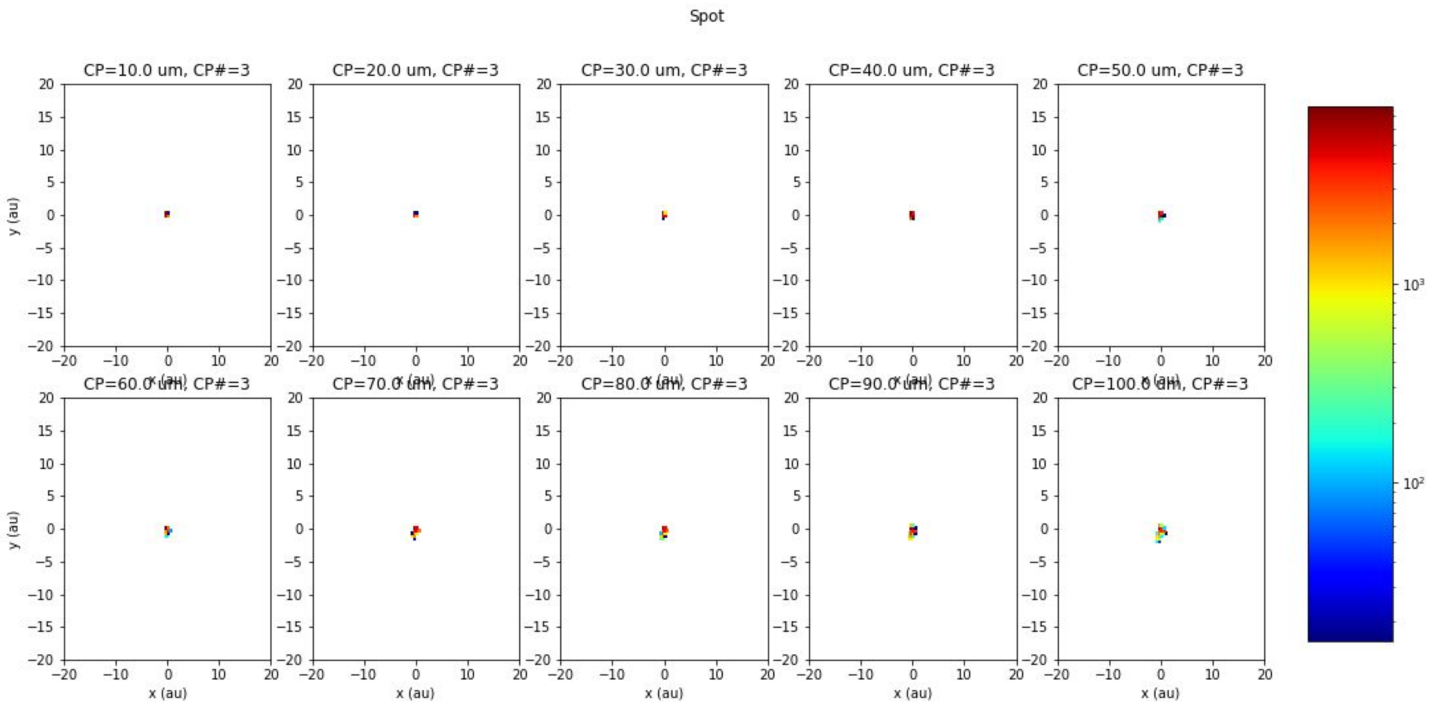
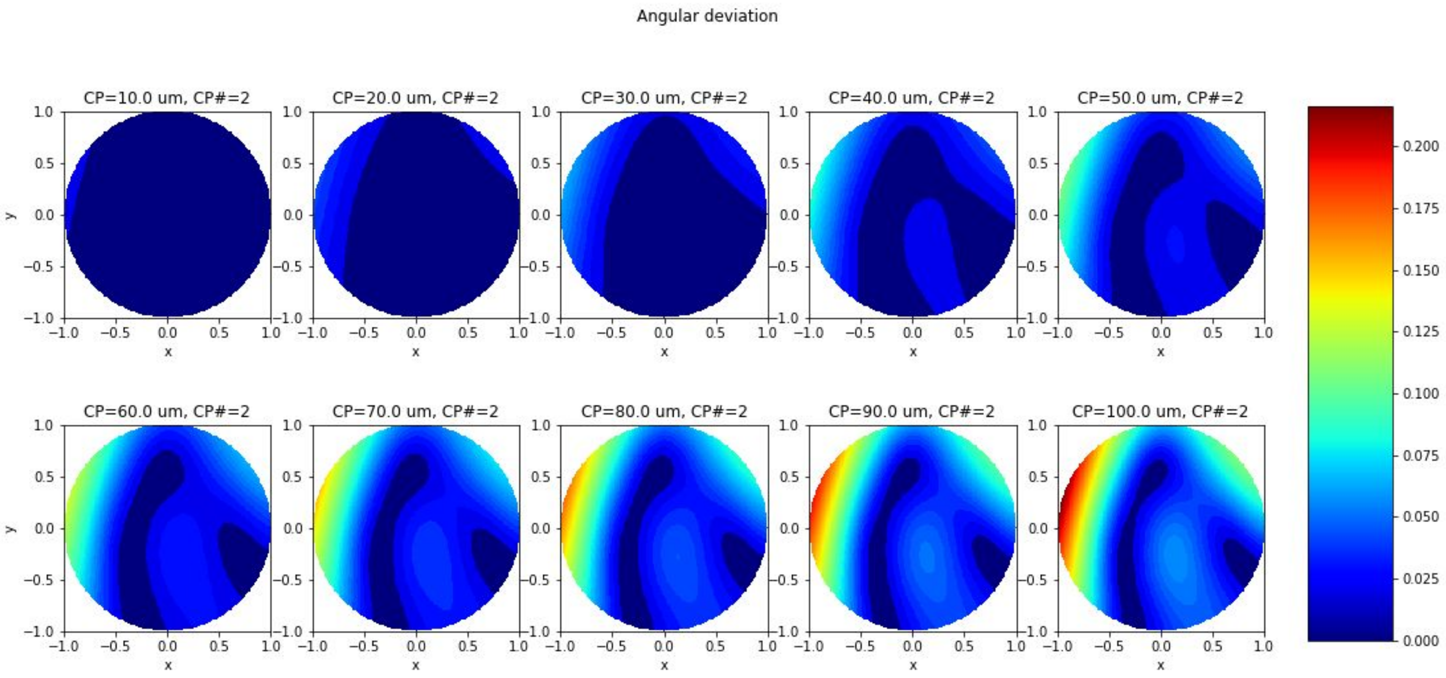


Image Spots for CP3.



Angular Deviation for CP3.



Wavefronts for CP4.

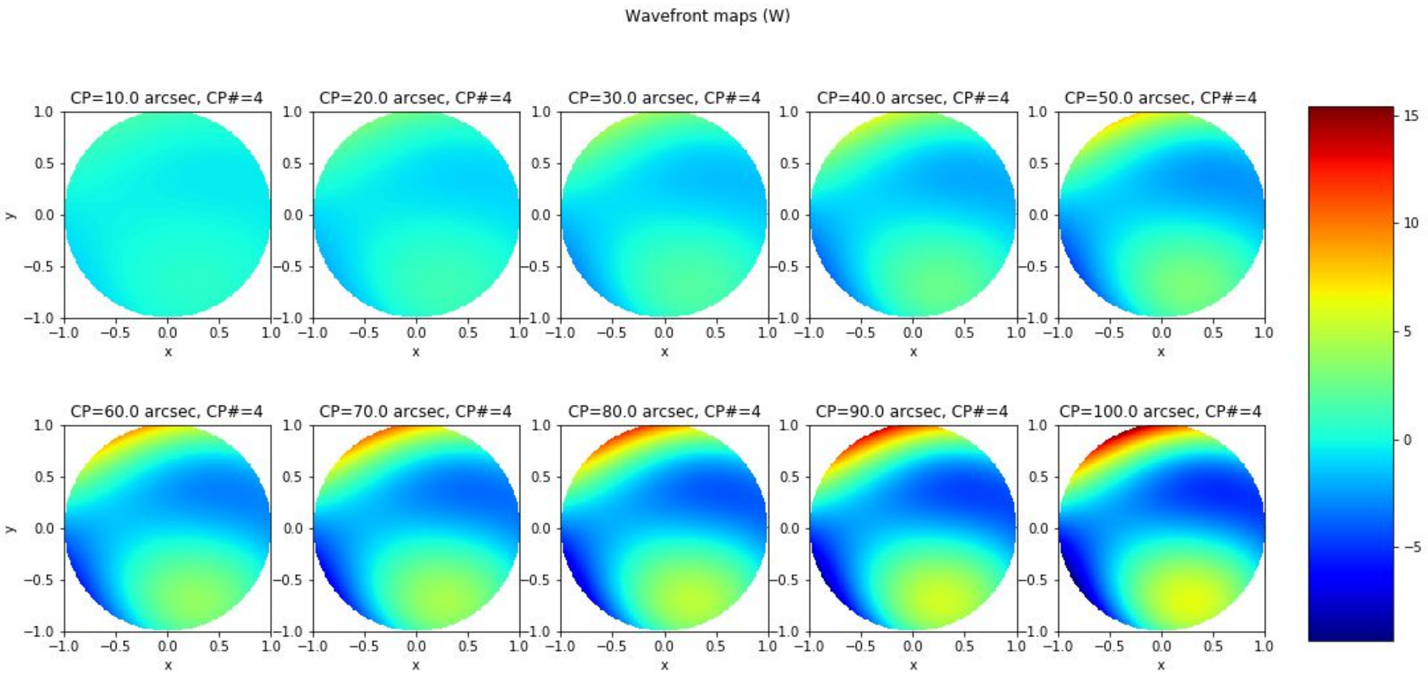
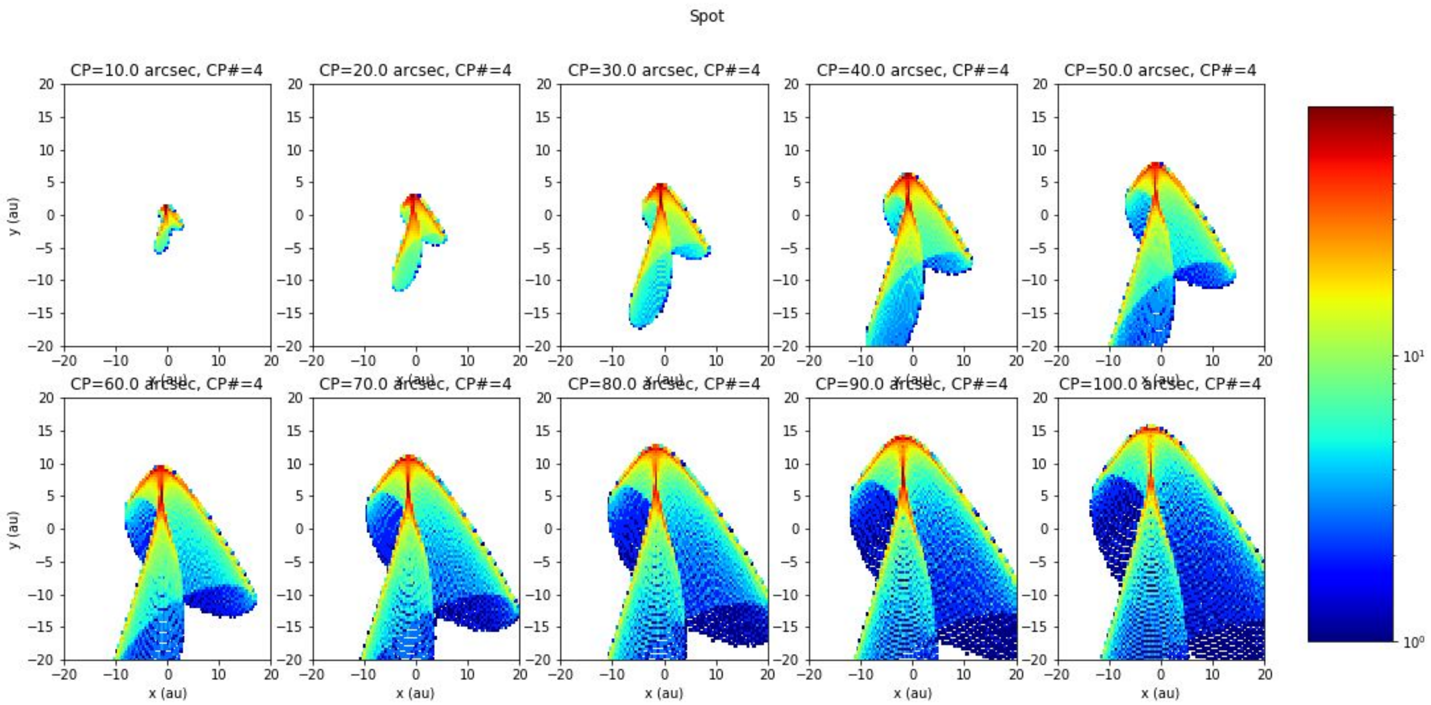
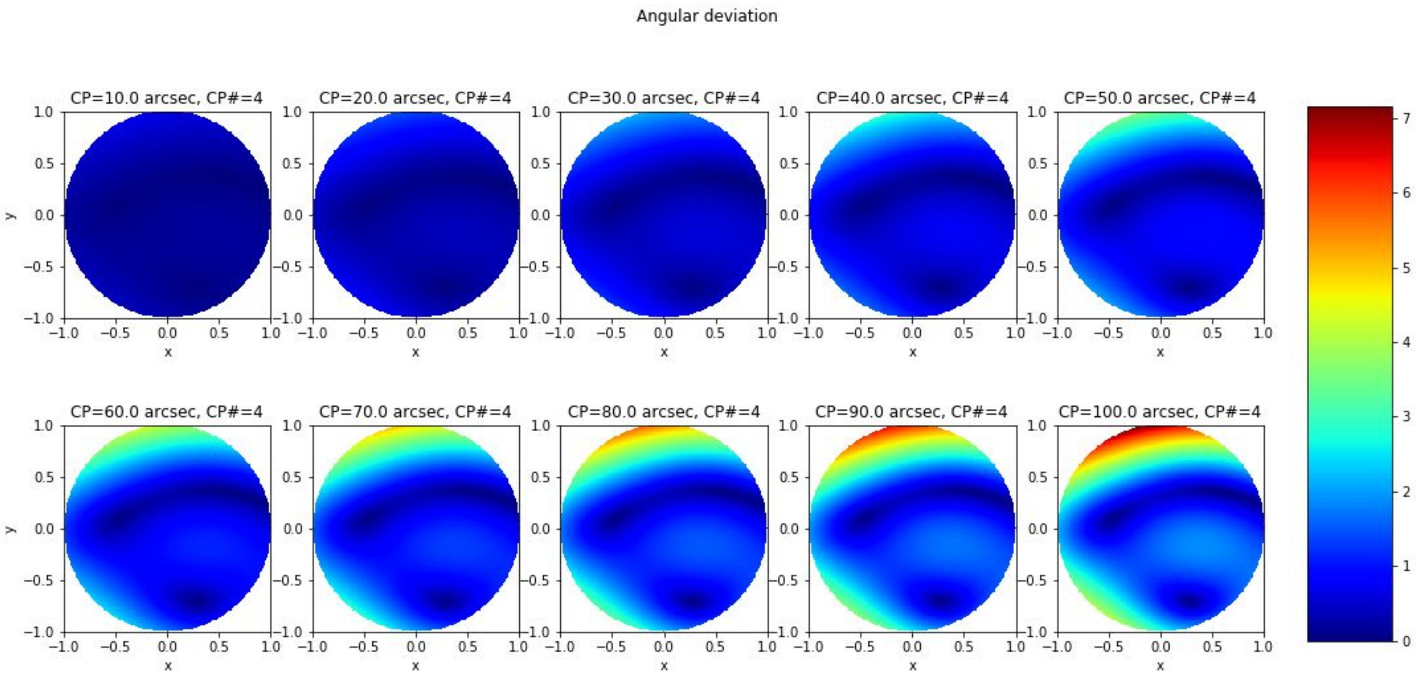


Image Spots for CP4.



Angular Deviation for CP4.



Wavefronts for CP5.

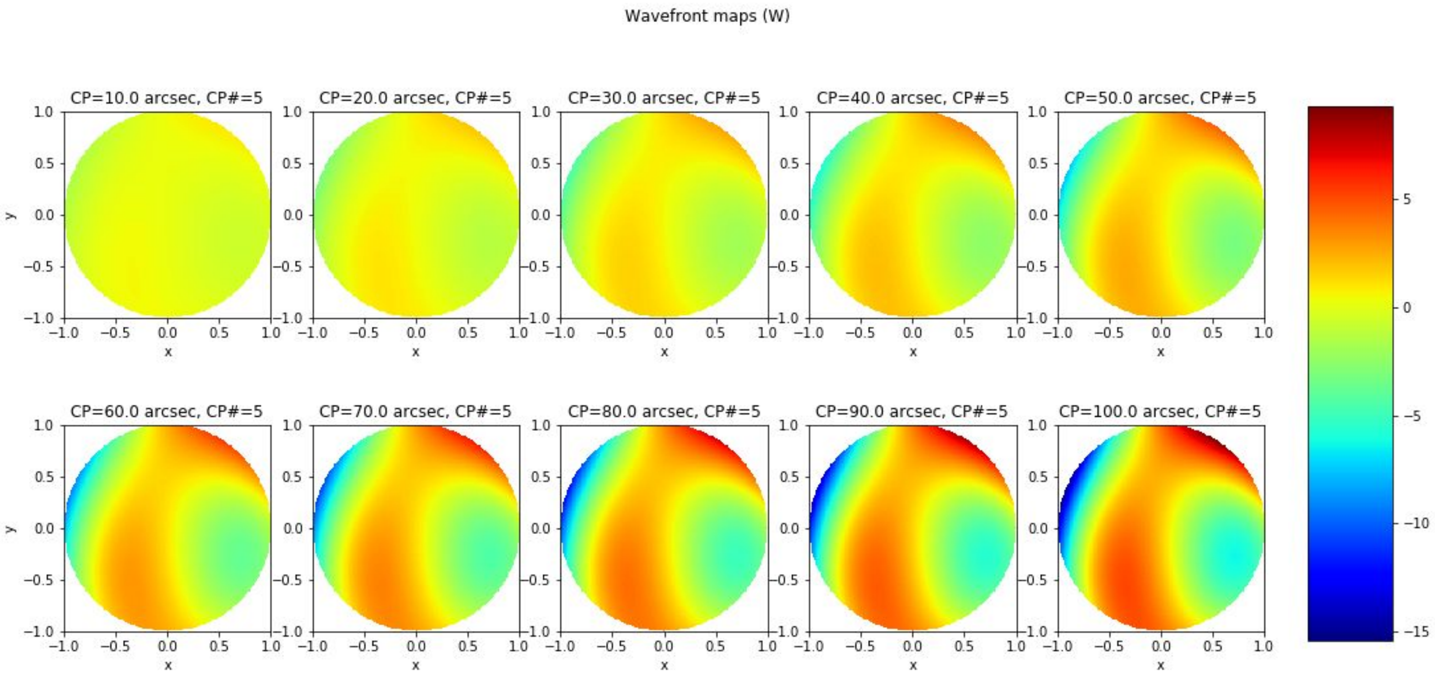
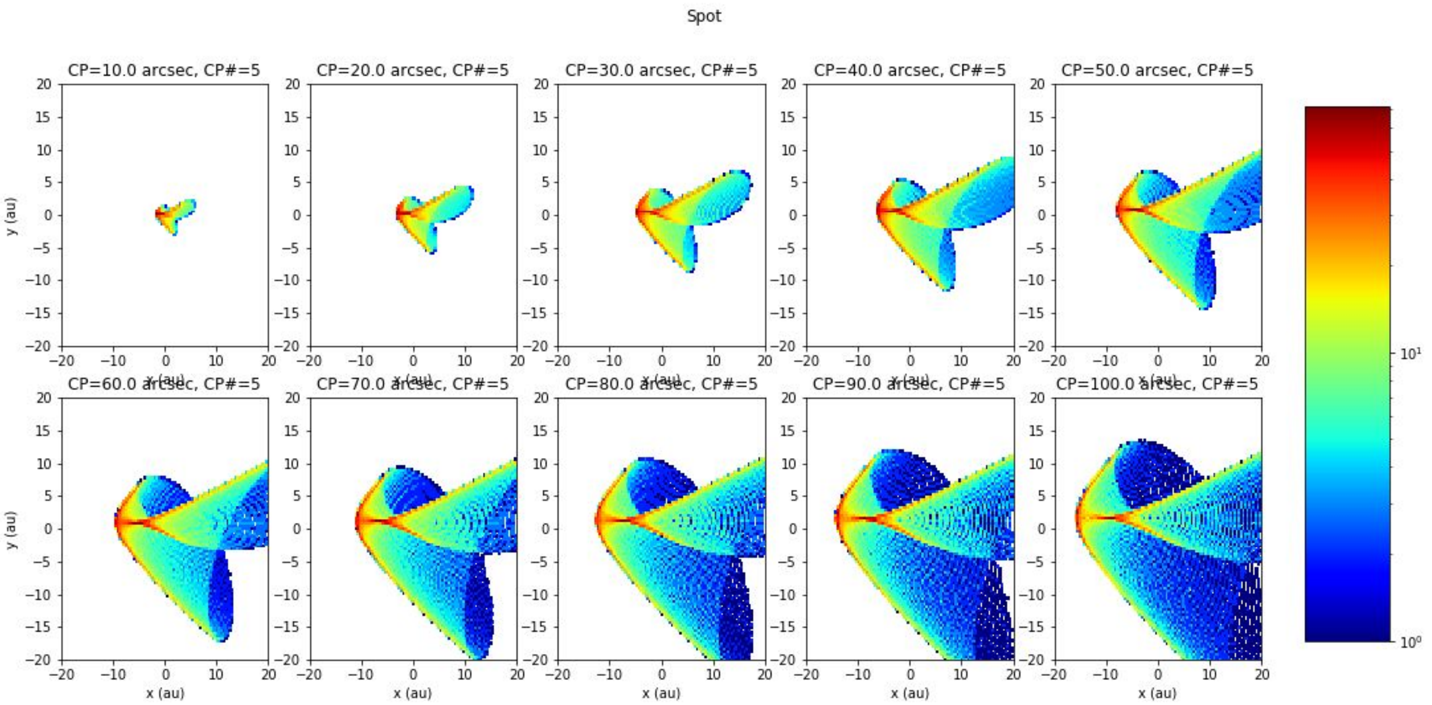
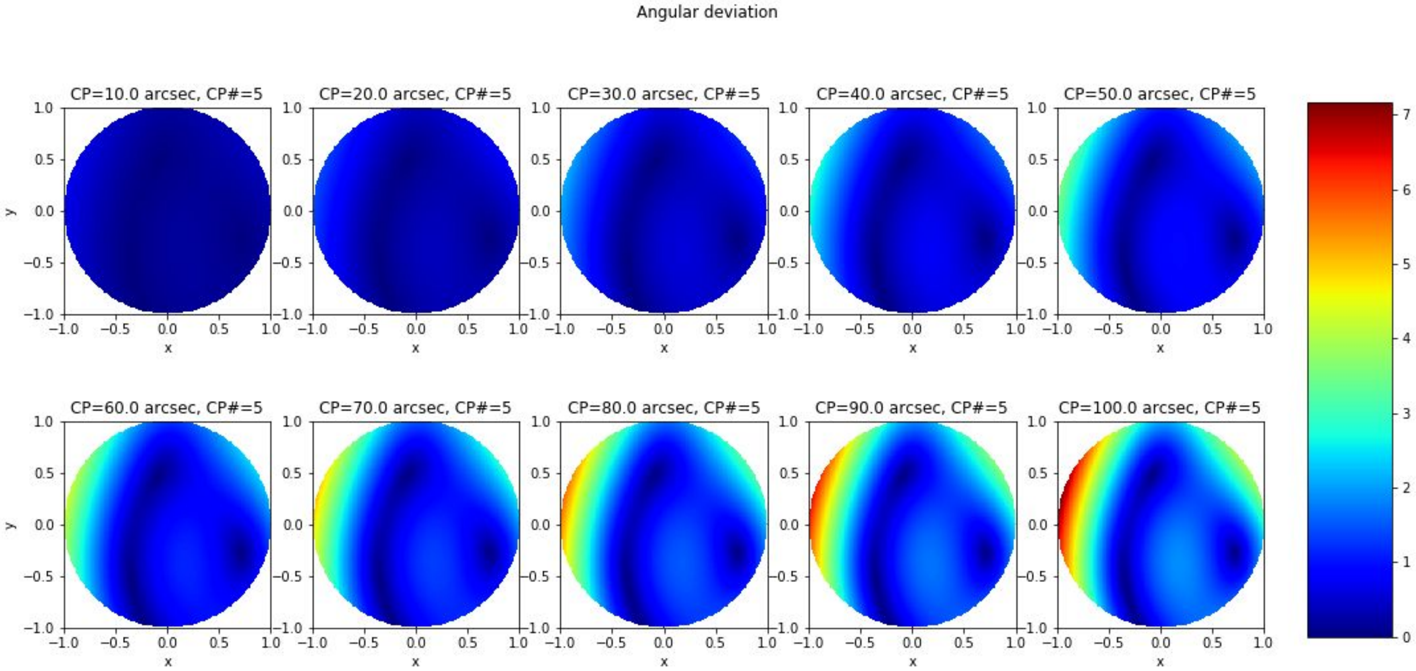


Image Spots for CP5.



Angular Deviation for CP5.



Wavefronts for CP6.

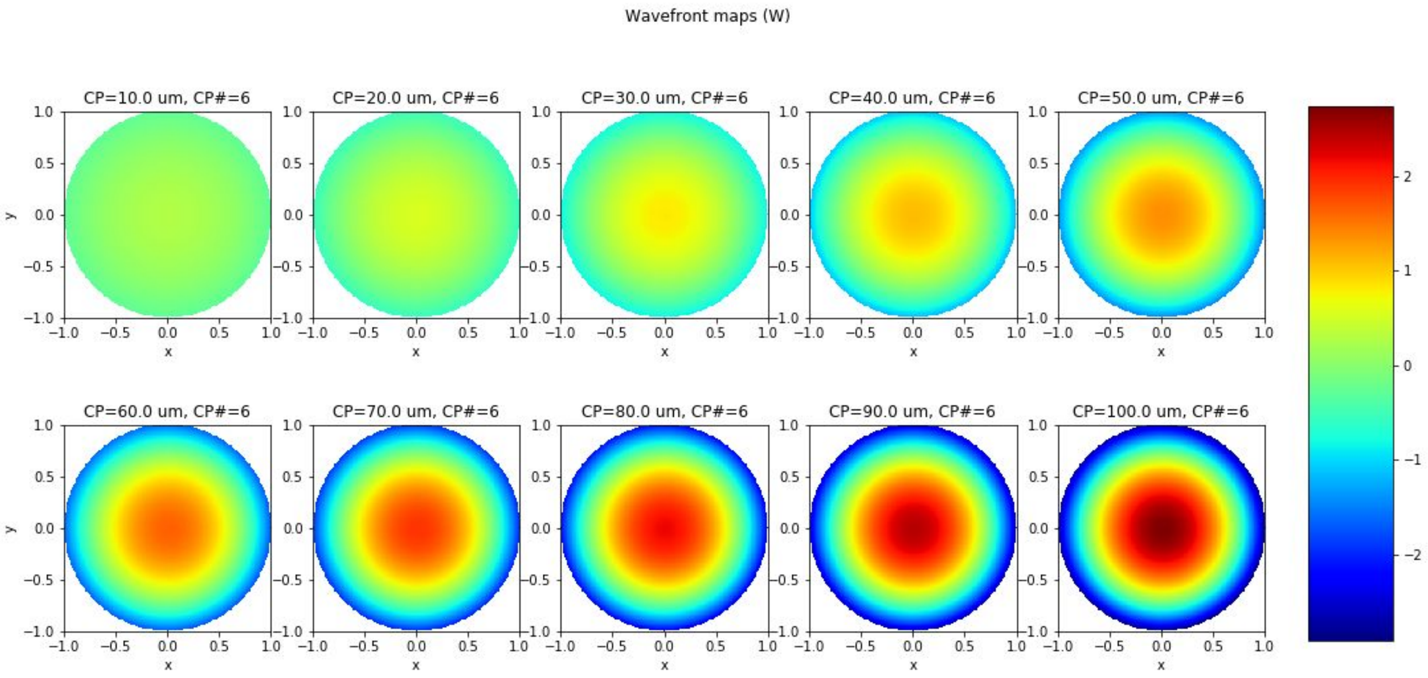
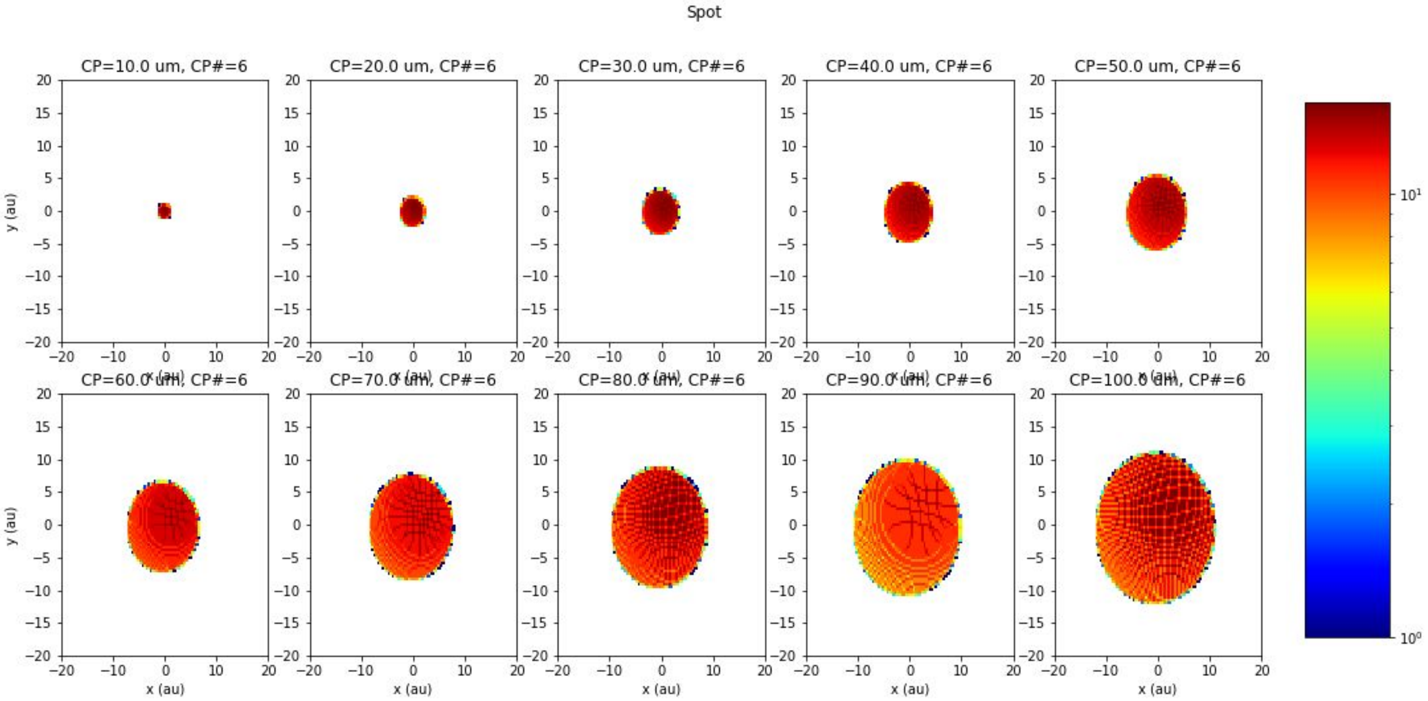
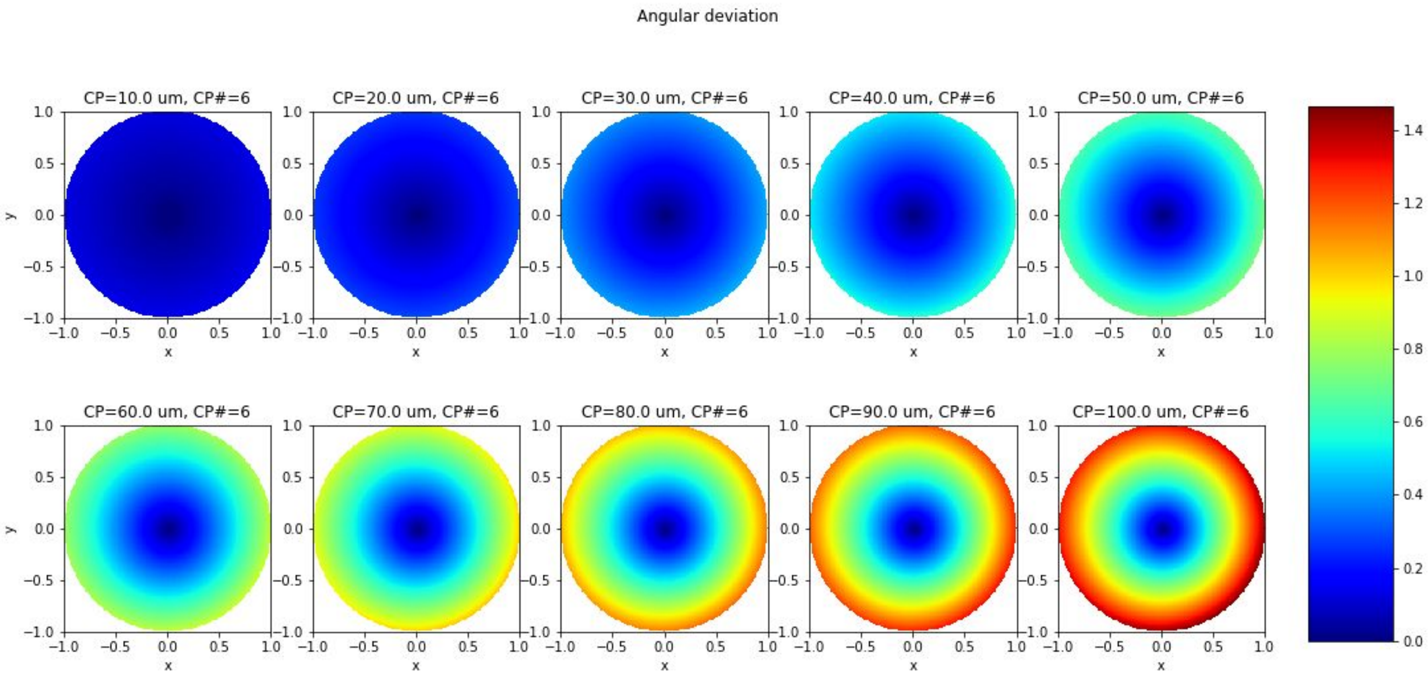


Image Spots for CP6.



Angular Deviation for CP6.



Wavefronts for CP7.

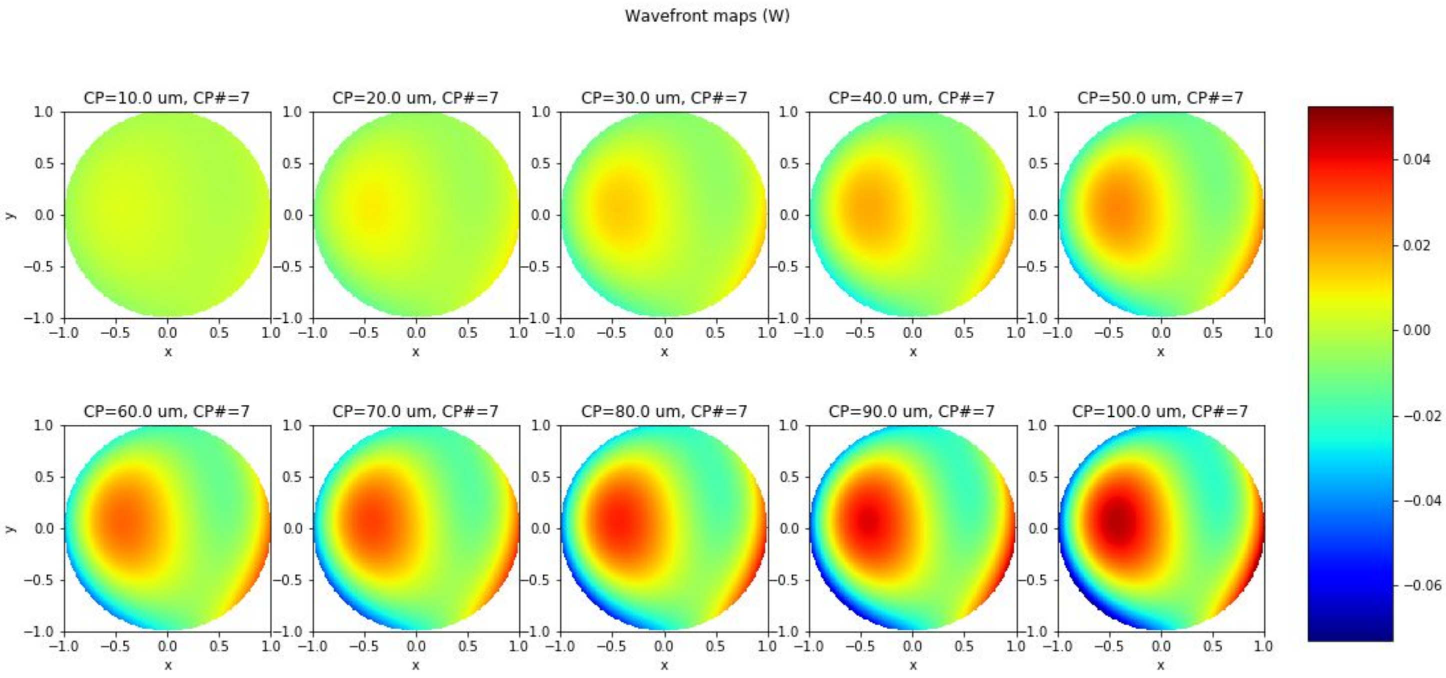
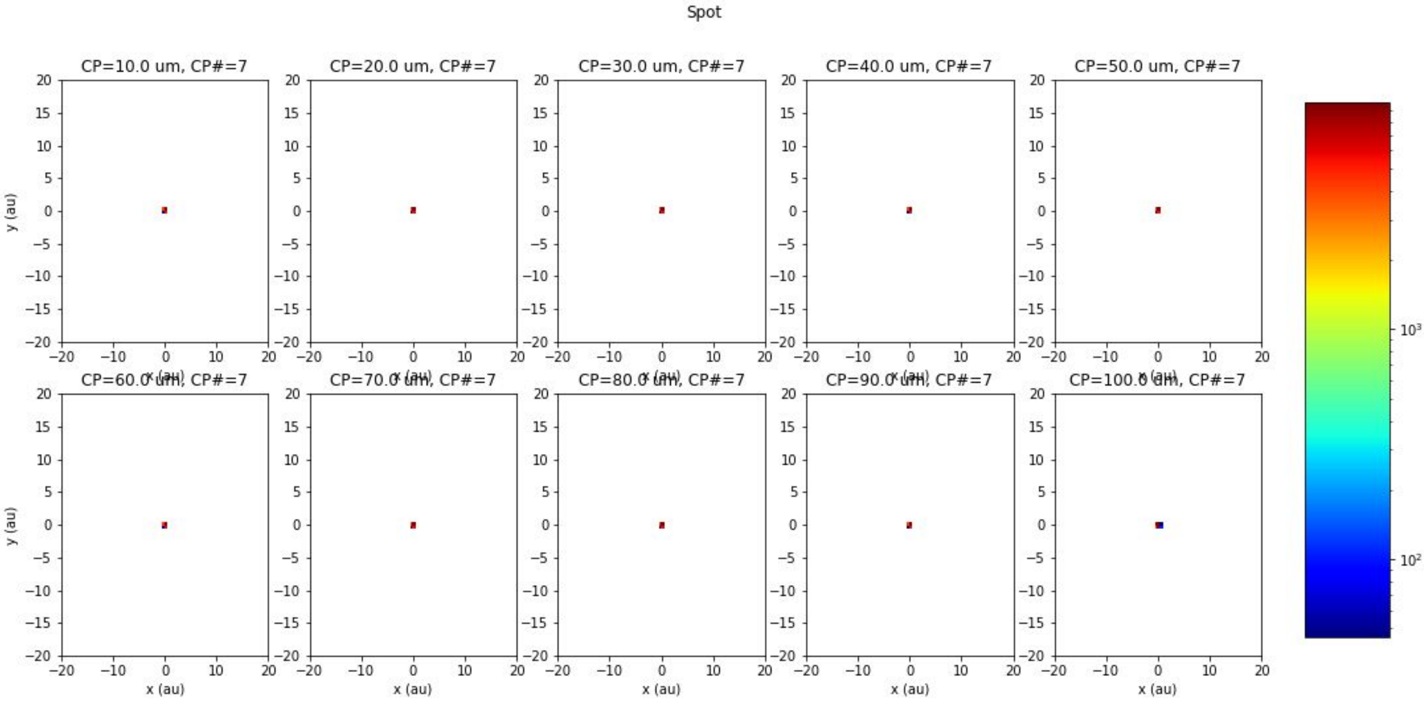
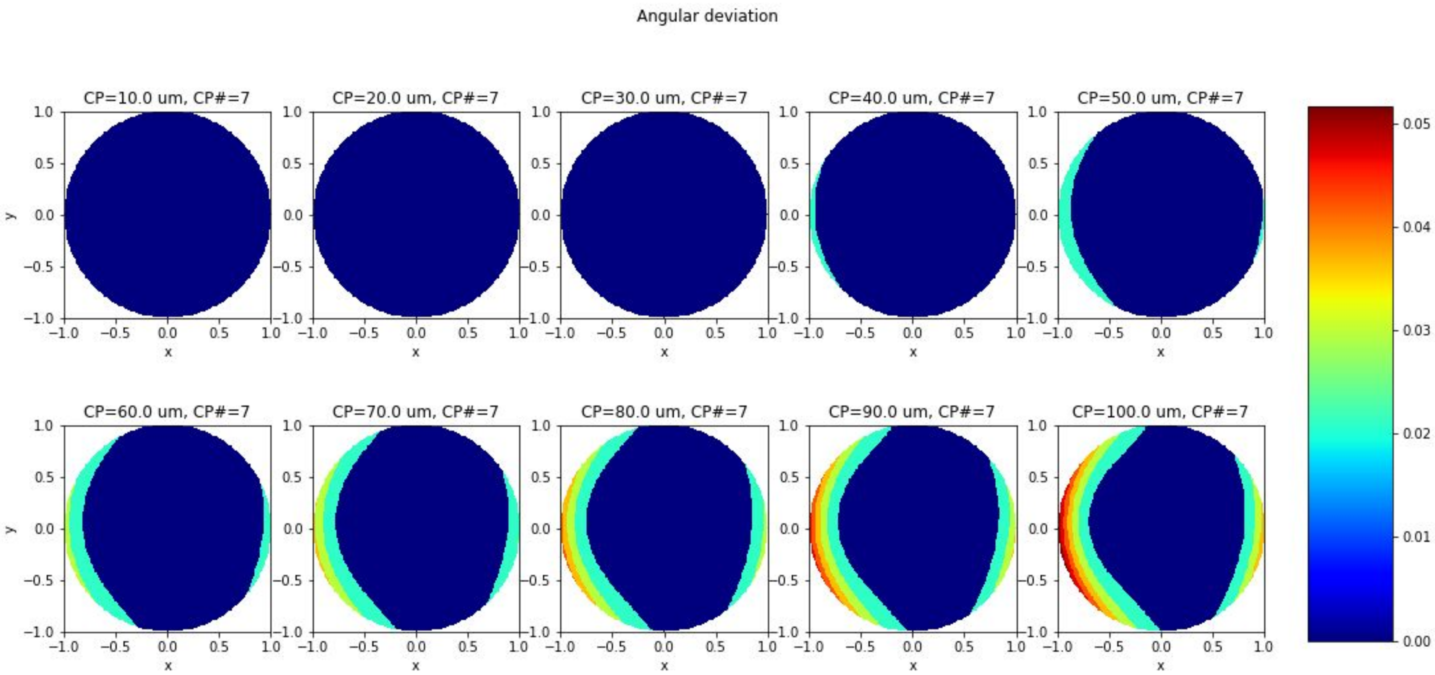


Image Spots for CP7.



Angular Deviation for CP7.



Wavefronts for CP8.

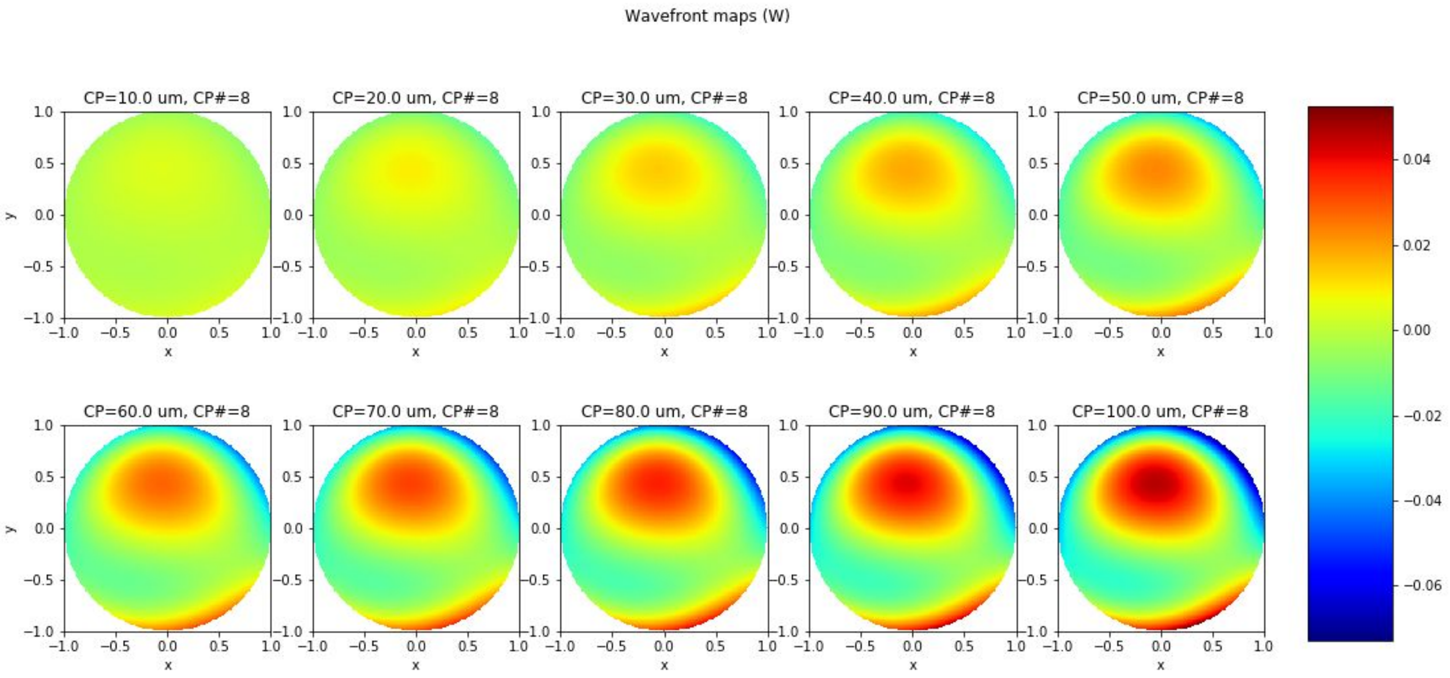
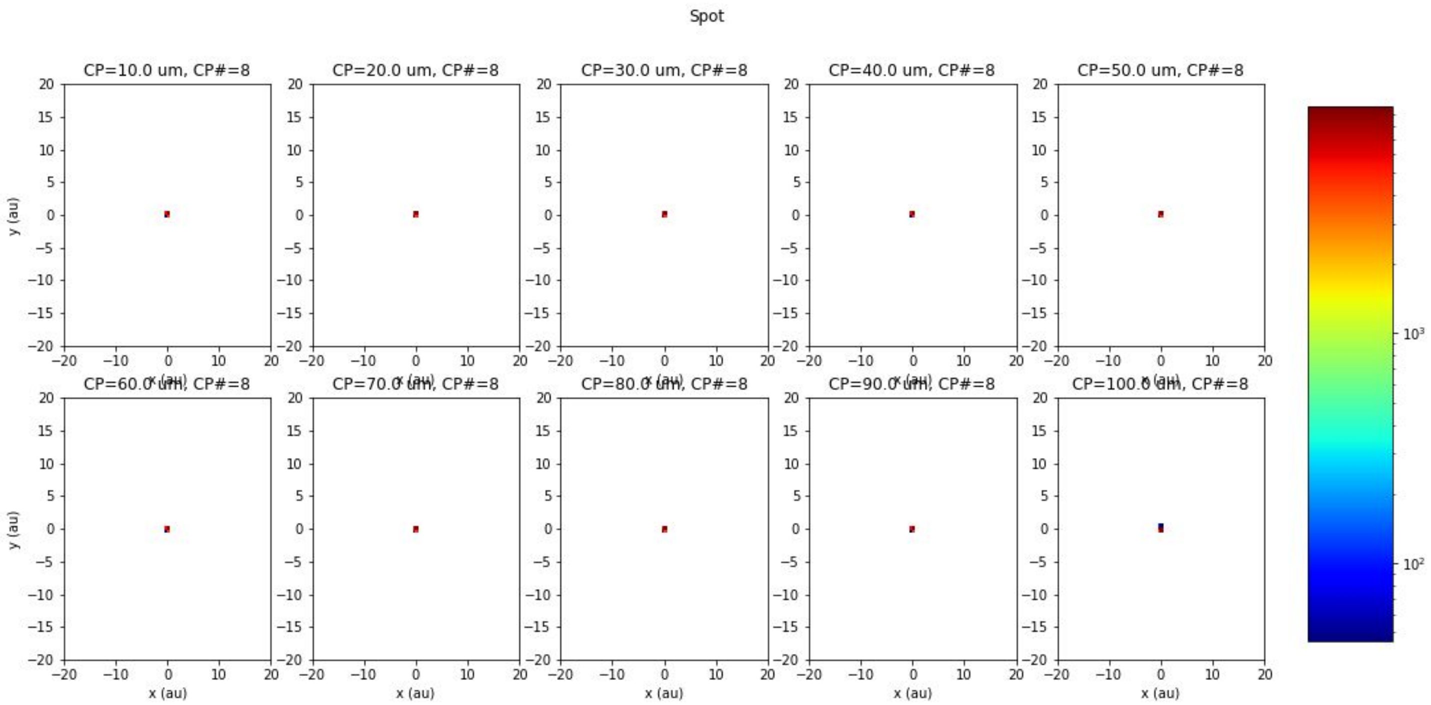
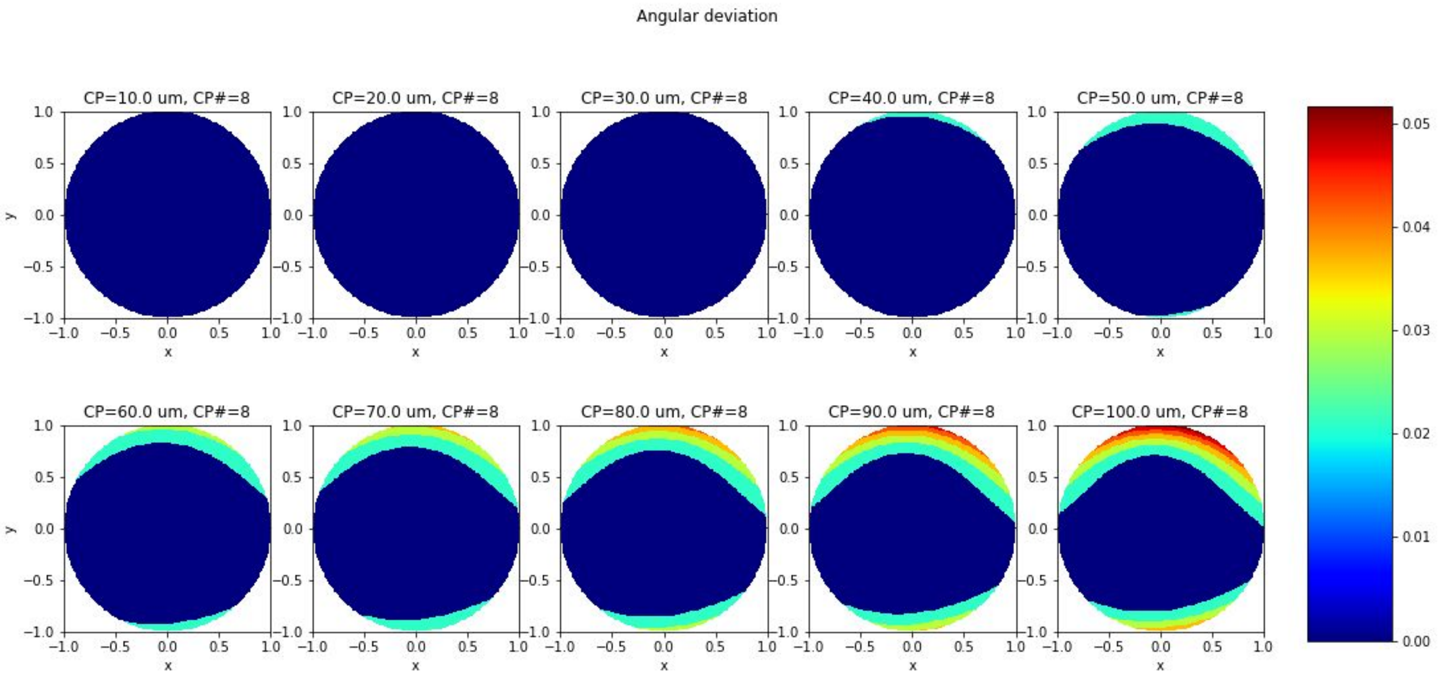


Image Spots for CP8.



Angular Deviation for CP8.



Wavefronts for CP9.

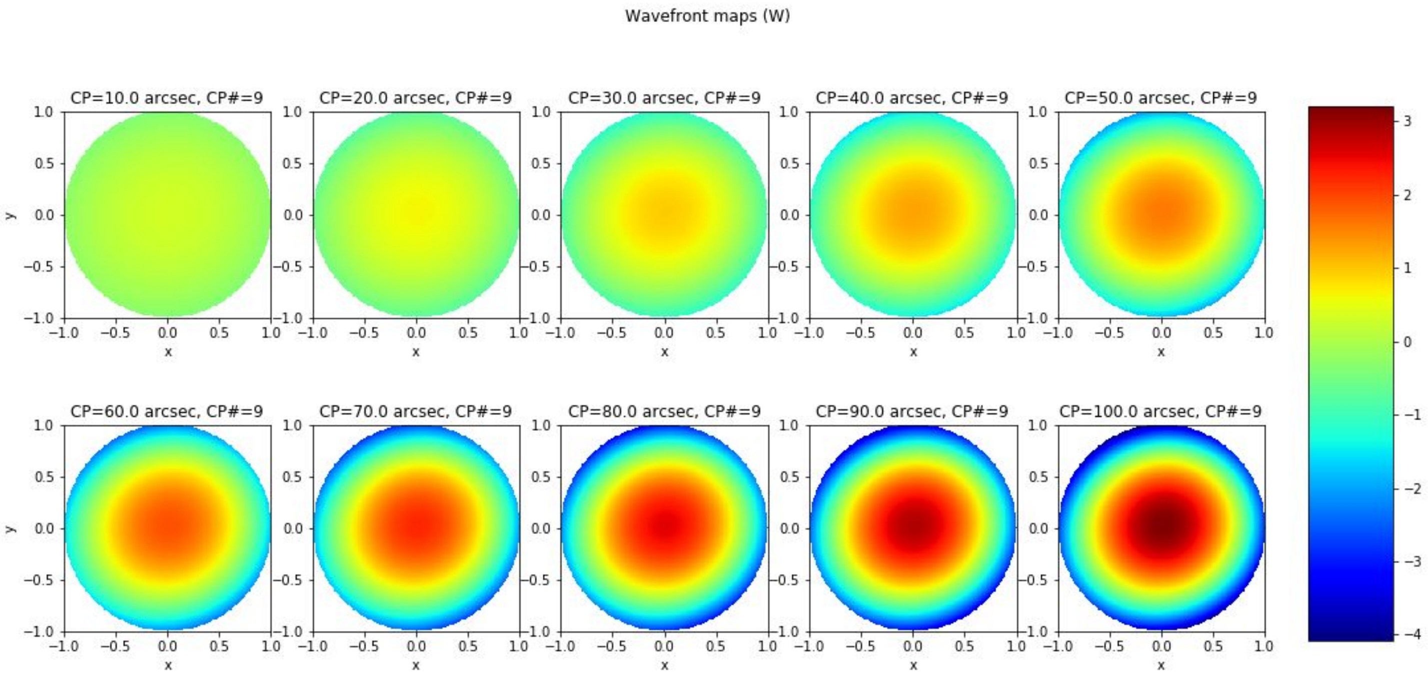
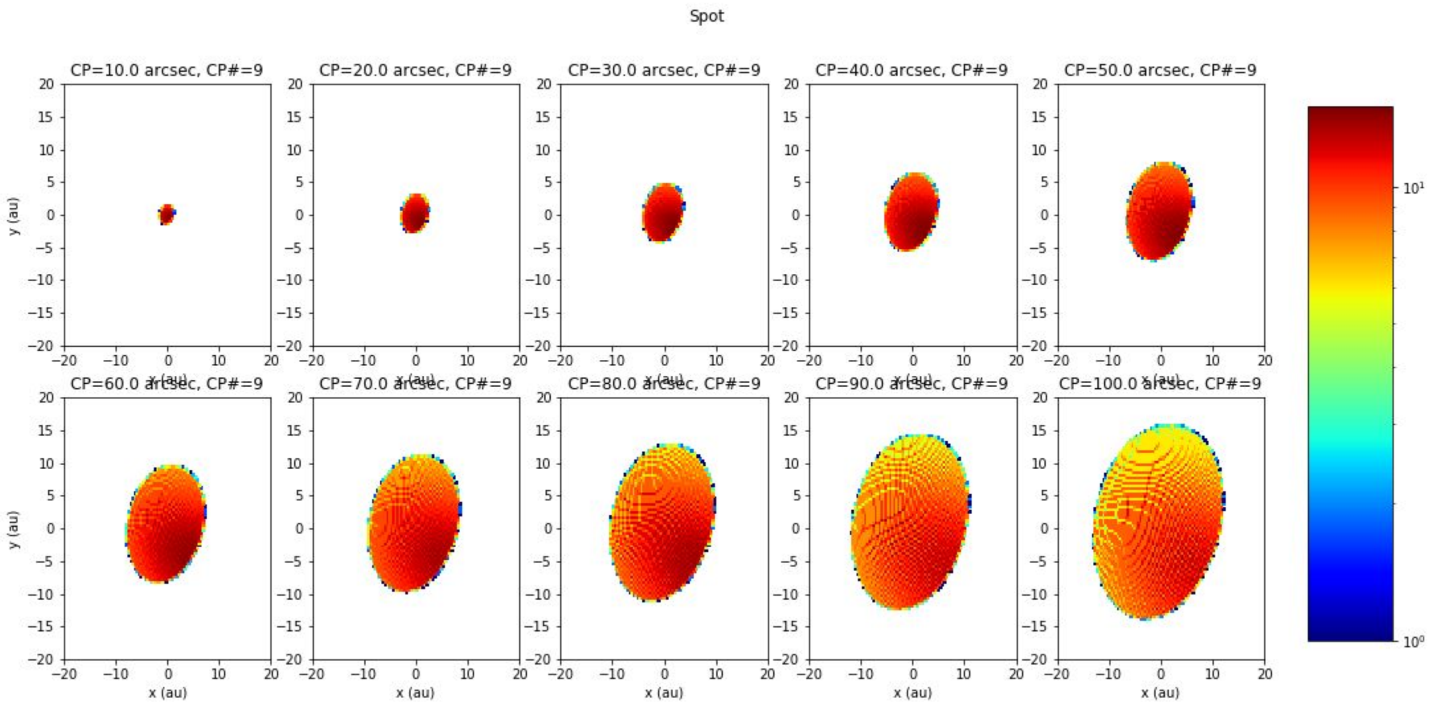
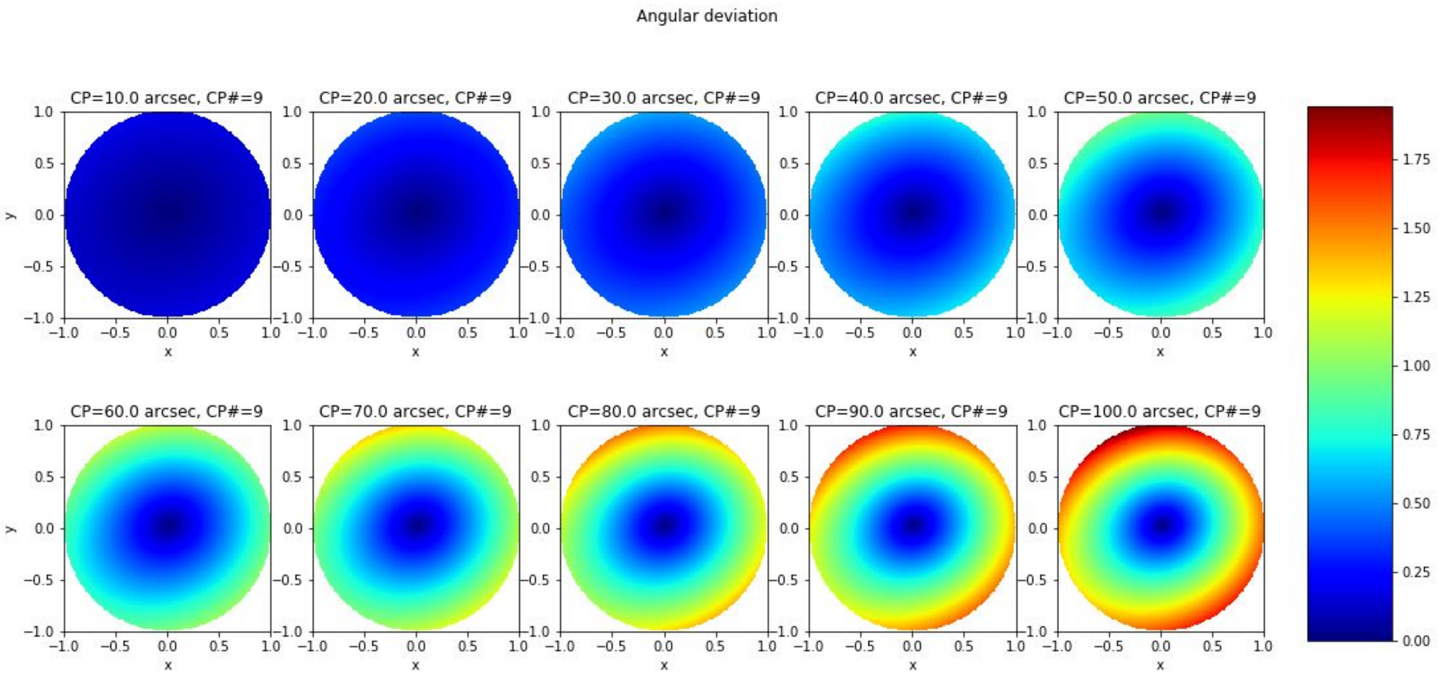


Image Spots for CP9.



Angular Deviation for CP9.



Wavefronts for CP10.

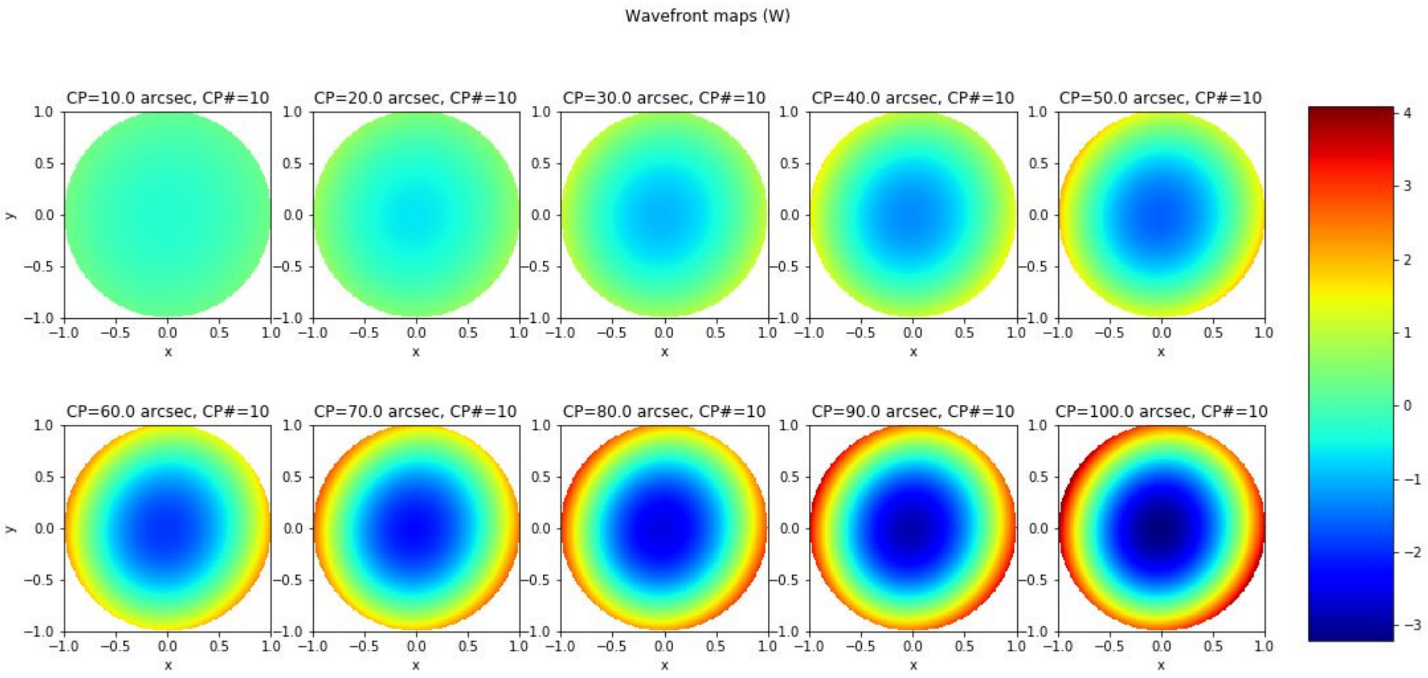
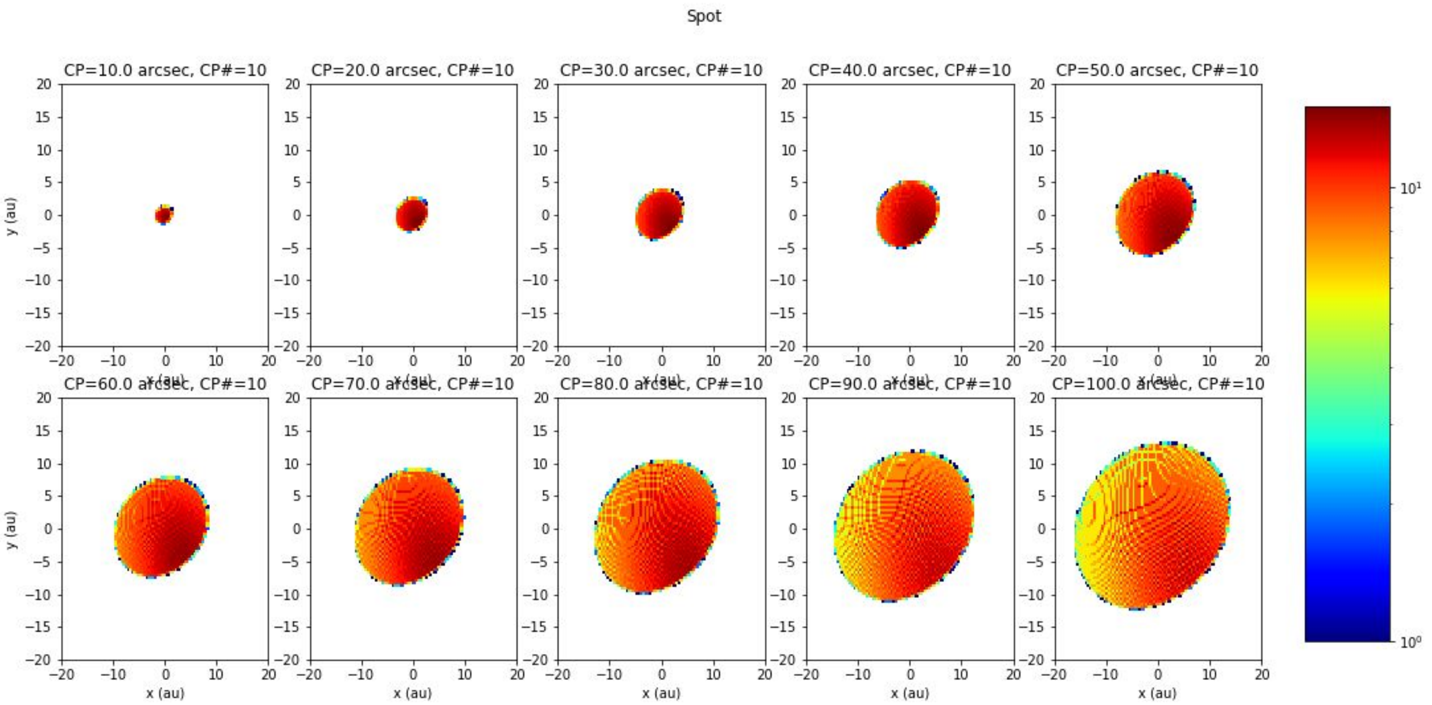
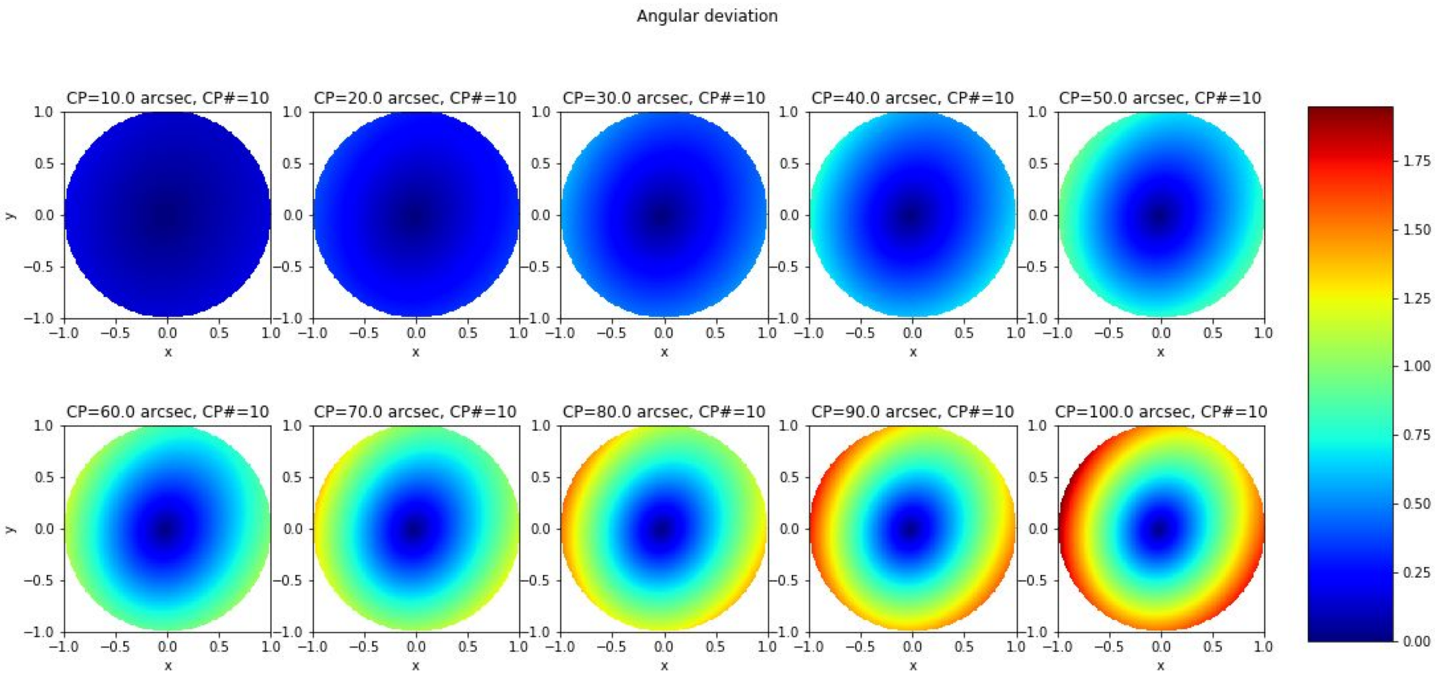


Image Spots for CP10.



Angular Deviation for CP10.



Shape of M1/M3 mirror.

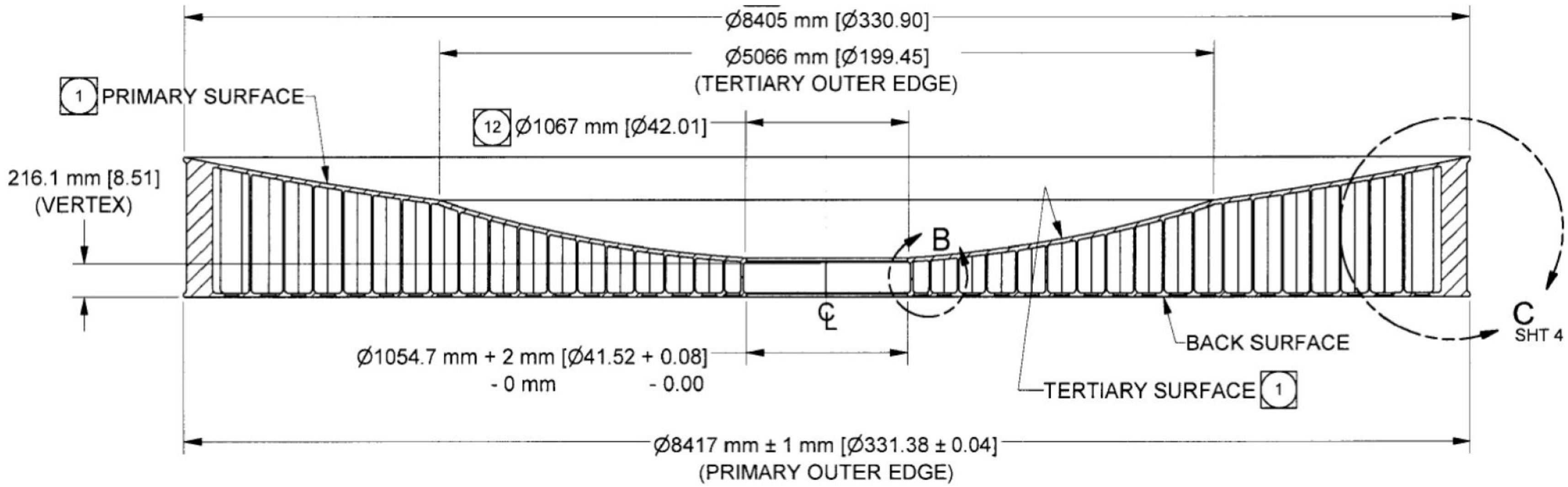


Figure 2.6: Design and dimensions of the primary and tertiary mirror, showing that the two are built out of a single mirror blank.

From LSST Science Book 2.

Control parameters of AOS

Table 3 Controlled degrees of freedom [**x**] in AOS

Optical element	Degrees of Freedom	Notes
M2	Piston x/y decenter x/y tilt	Realized by the M2 hexapod
Camera as a whole	Piston x/y decenter x/y tilt	Realized by the Camera hexapod/rotator
M1M3 substrate	Shape (#1-#15 bending modes of the M1M3 substrate)	The full set of (156-3) bending modes form a full basis set for all possible shapes the 156 figure control actuators can generate.
M2	Shape (#1-#15 bending modes)	The full set of (72-3) bending modes form a full basis set for all possible shapes the 72 figure control actuators can generate.

From: Real time wavefront control system for the Large Synoptic Survey Telescope (LSST).

More Details on Error Budget

Copied from Angeli et al. (Real Time Wavefront Control System for the Large Synoptic Survey Telescope (LSST), Section I):

- **Inherent design aberrations: low but non-zero (80 mas).**
- **Telescope errors constant in time: residual uncertainty of the LUT.**
  - (i) mirror and lens surface fabrication errors
  - (ii) other manufacturing and installation errors, including the focal plane
  - (iii) mirror coating thickness non-uniformity
  - (iv) fabricated relative positioning errors of M1 and M3.
- **Telescope errors that are changing faster than the temporal bandwidth of the wavefront control loop: cannot be corrected for by the AOS:**
  - (i) wind buffeting
  - (ii) tracking/guiding jitter
  - (iii) vibration
  - (iv) dome seeing
- **Wavefront correction errors (AOS errors: budget = 79 mas).**
- **Slowly varying errors: mostly corrected for by the AOS.**
- **Camera aberrations (69 mas)**
- **Camera errors constant in time (92 mas):**
  - initial position alignment errors
  - optical fabrication errors.

Recovery errors for bending modes

cp 10 = 0.015434016994144445

cp 11 = 0.015117857045320704

cp 12 = 0.018324356118510368

cp 13 = 0.010786114095912448

cp 14 = 0.008711288821202147

cp 15 = 0.008384204115433893

cp 16 = 0.008318321874112942

cp 17 = 0.00269014927401698

cp 18 = 0.0026417346471405818

cp 19 = 0.004093847010310072