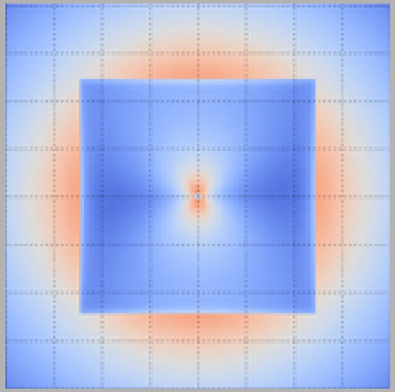
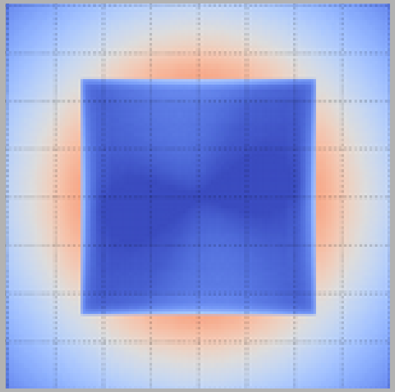
The Promise of Predictive Control:
A Factor of 10+





Jared Males (UofA)
Olivier Guyon (UofA, Subaru)

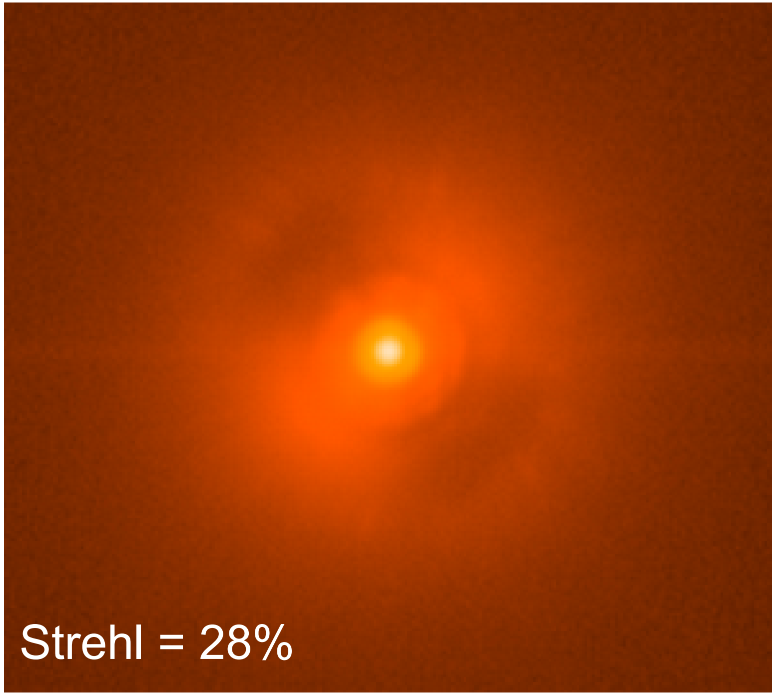
KISS HCI Workshop
2016-12-05

Outline

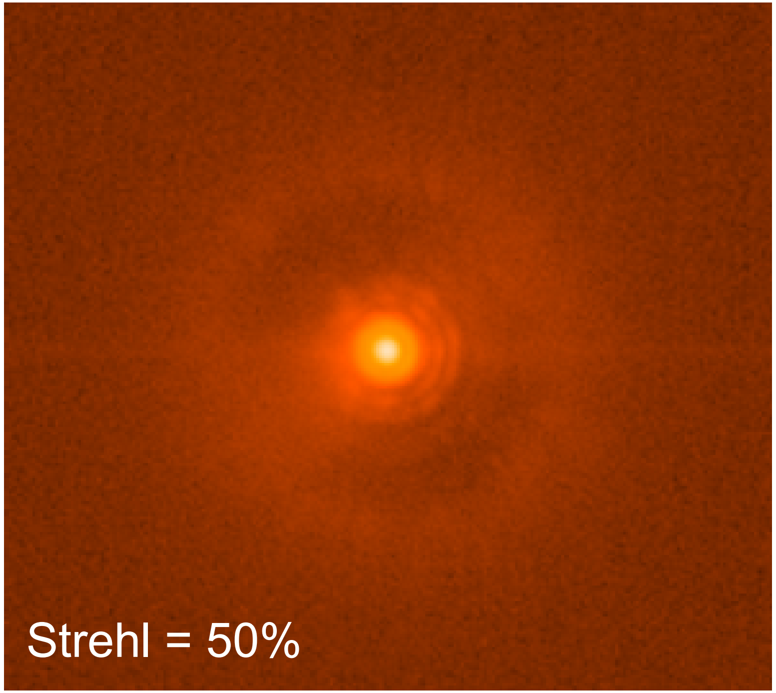
- MagAO update
- Wavefront variance and temporal PSDs
- Basic AO Control
- Predictive Control
  - LQG with frozen flow (PFC)
  - Linear Predictor
- Ideal performance with LP control.

MagAO-2K Results

Target: Fomalhaut
Seeing: 0.6” to 0.7” (LCO median is 0.63”)
Wind: 20-23 mph on ground
Conditions: variable clouds
Filter: z’ (0.9 microns)
Correcting 350 modes
with 3 λ/D modulation

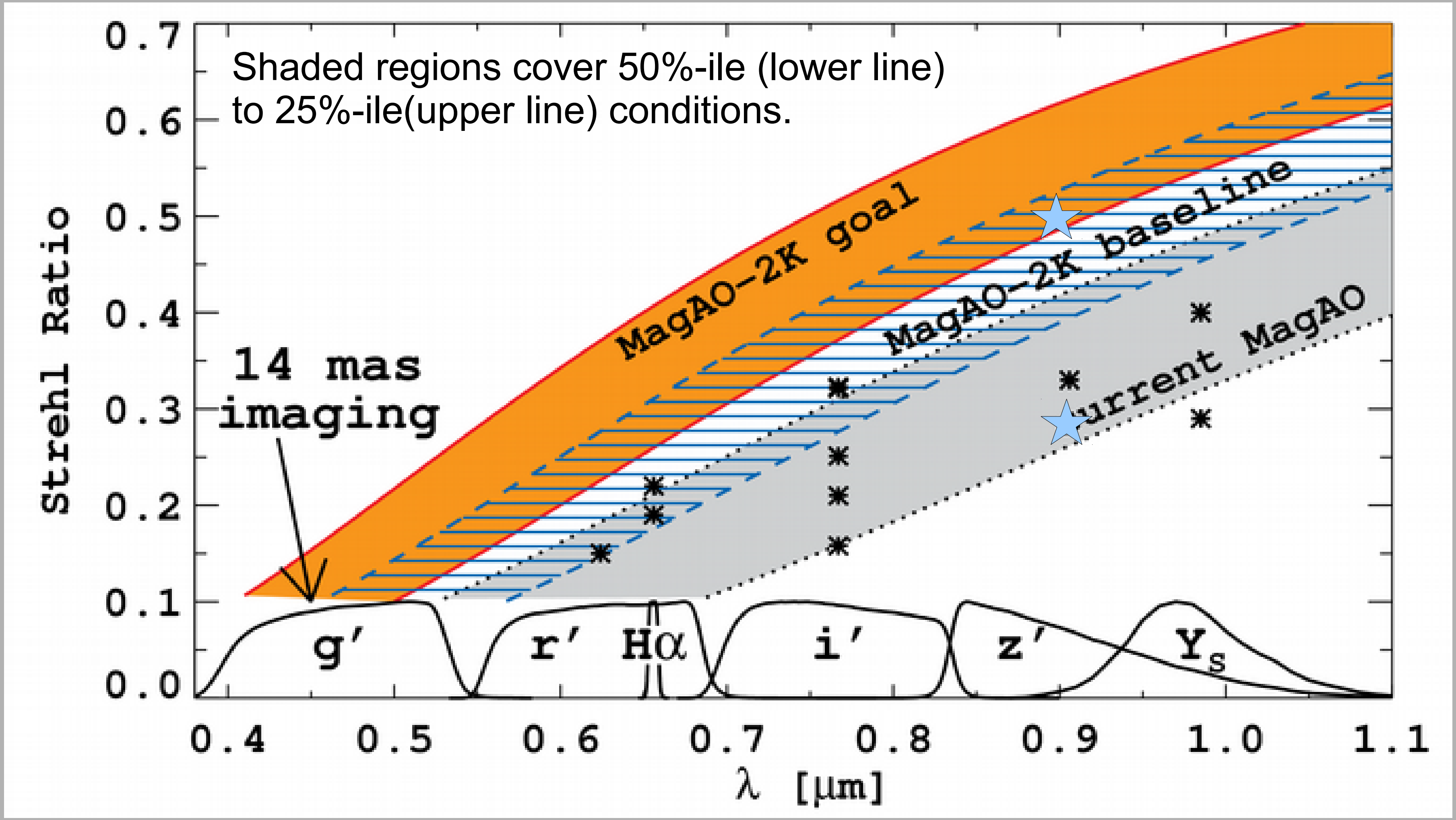


1000 Hz
FWHM = 4.7 pix (37 mas)

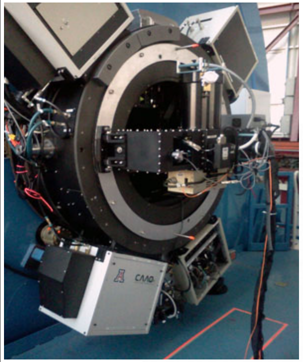


2000 Hz
FWHM = 4.3 pix (34 mas)
Less flux due to clouds
ADC residual evident in Airy rings

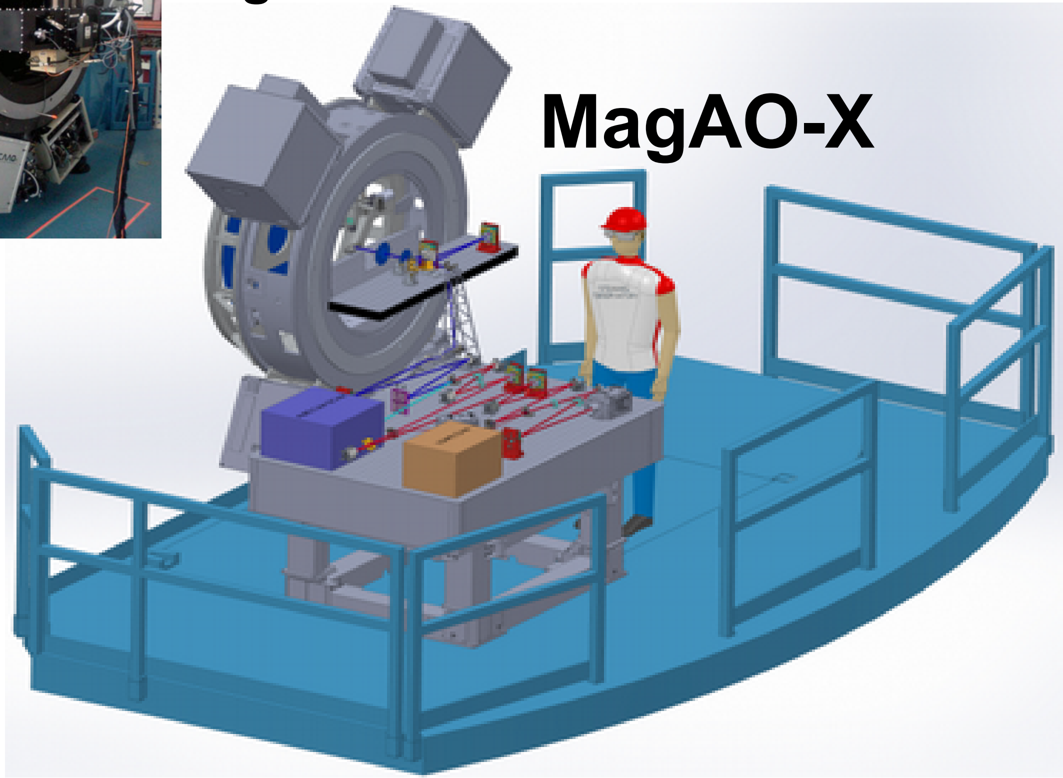
MagAO-2K Predictions



MagAO-X



Existing MagAO



NSF MRI funded ExAO Coronagraph

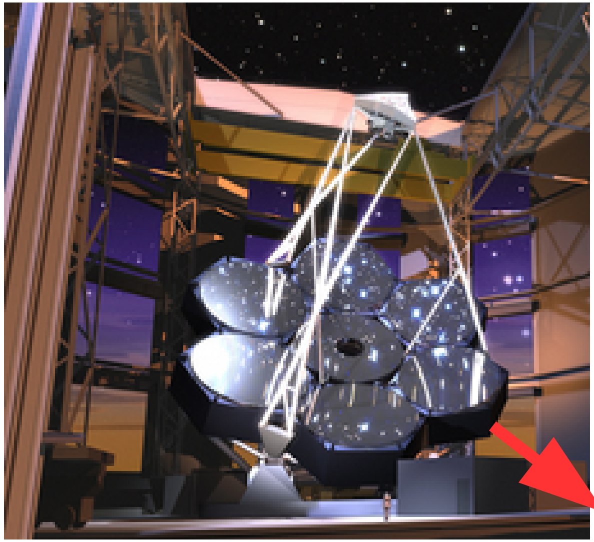
SCExAO-like 2nd stage at 6.5 m Magellan Telescope behind MagAO

BMC 2K MEMS DM

OCAM 2K 3.7 kHz PyWFS

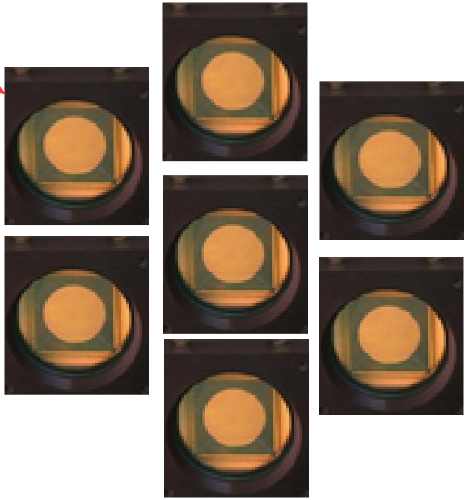
1 λ/D coronagraphy in the visible

MagAO-X to the GMT



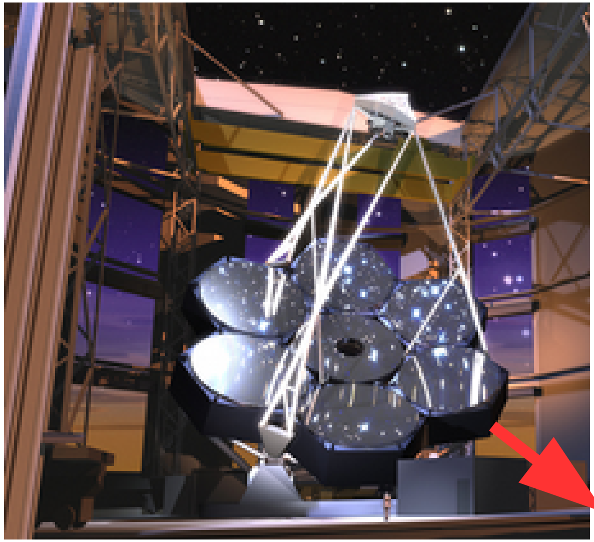
Can we get high-Strehl on an ELT?

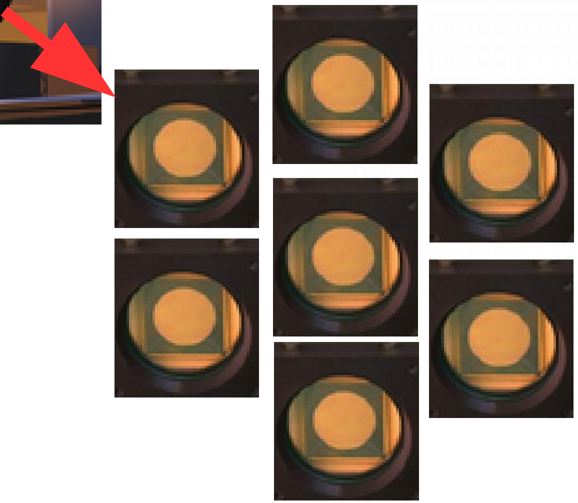
Scaling MagAO-X to the GMT: 21000 actuators

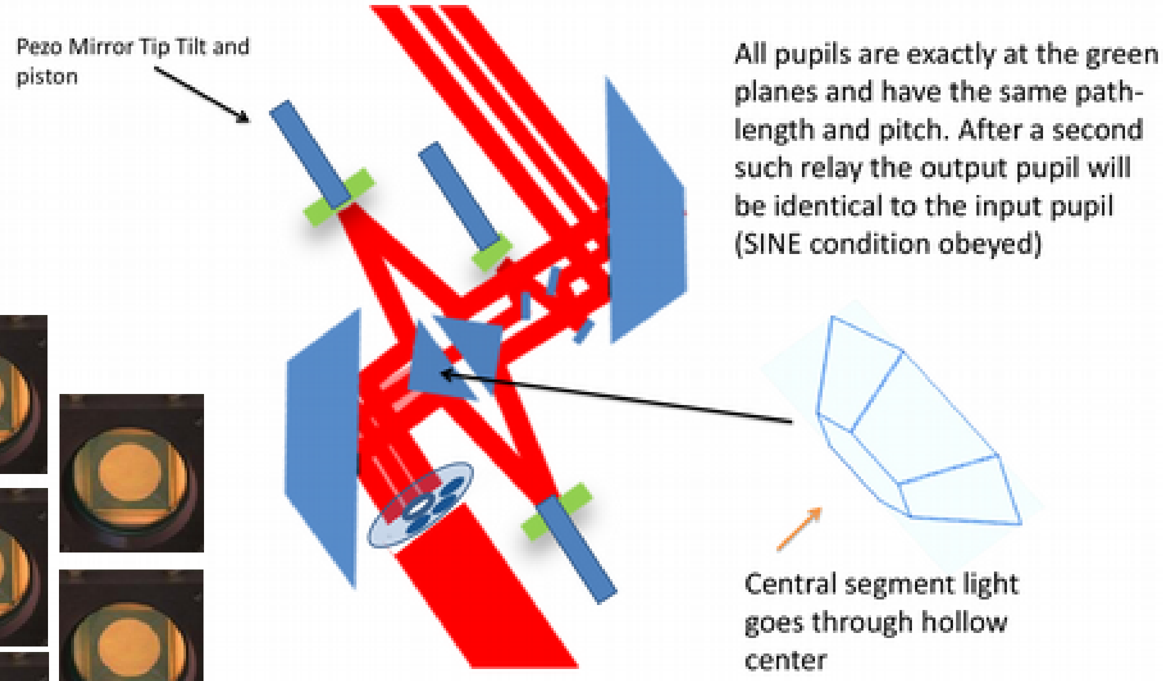


7x 3000 BMC MEMS

21,000 Actuators







OptoMech Powerpoint by L. Close

Computations

- How many FLOPS do we need?
  - GMagAO-X: M = 21,000 actuators
  - 4-Q PyWFS: N = 42,000 slopes
  - Matrix-vector multiply floating point ops:

2(N-1)M * 3.6kHz = 6.4 x 10^12 FLOPS

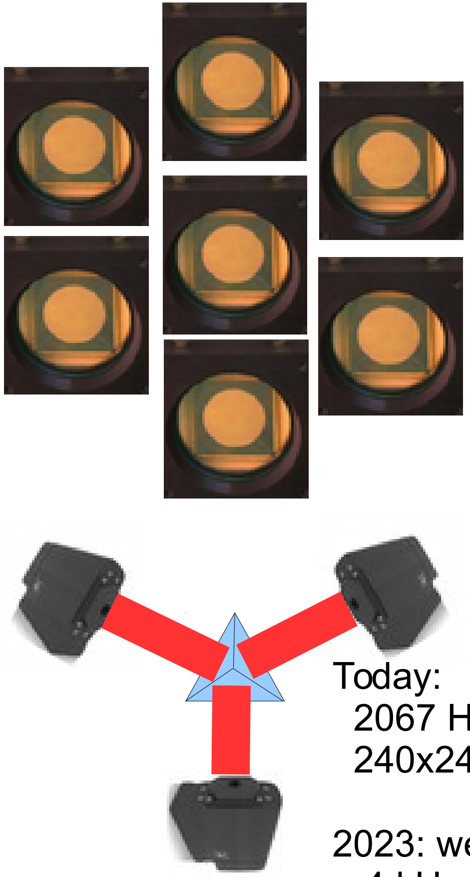
- Thanks video games.
  - NVIDIA Titan-X Pascal = 11 TeraFLOPS
  - Next generation (Volta) out in 2018

GMT: GMagAO-X

- What would MagAO-X look like on GMT?
  - 7 *(8.4/6.5)^2 times more DOF
  - 7x BMC 3000 actuator DMs *(can buy today)*
  - ~480x480 (0 noise) detector at > 3.6 kHz
    - 2 kHz already achievable with 3 or 4 OCAMs
      - 3 sided PyWFS is theoretically equiv. to 4 sided
    - Can expect improvment in OCAM-like EMCCDs
    - *MKIDS? MCT-APD arrays?*
  - PIAACMC design for GMT pupil

We can build a 21,000 actuator, > 2 kHz ExAO Coronagraph on GMT **TODAY**

- *Caveats:*
  - *We need a telescope*
  - *This is narrow FOV (see Laird's design)*
  - *This is not-optimized for > 2 um.*

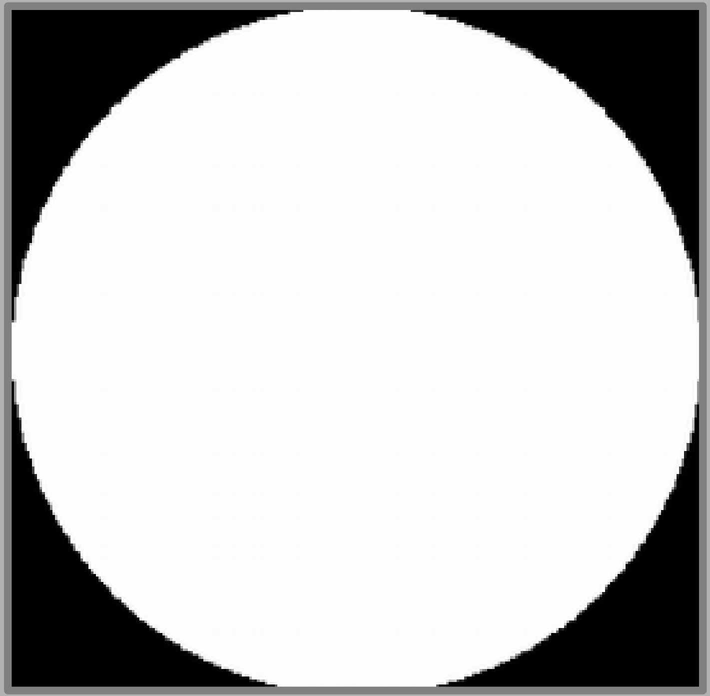


Today:
2067 Hz with
240x240 pupils

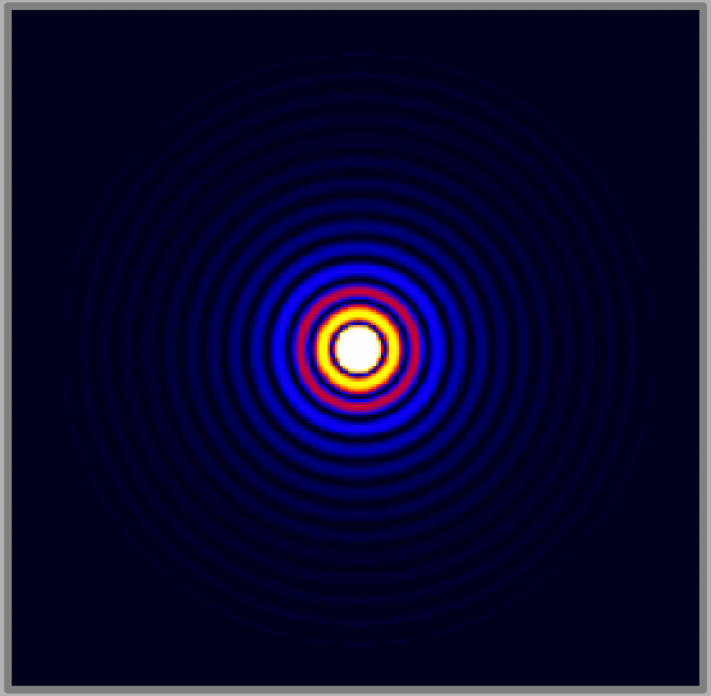
2023: we expect
~4 kHz available

Fourier Modes and Speckles

Pupil Plane



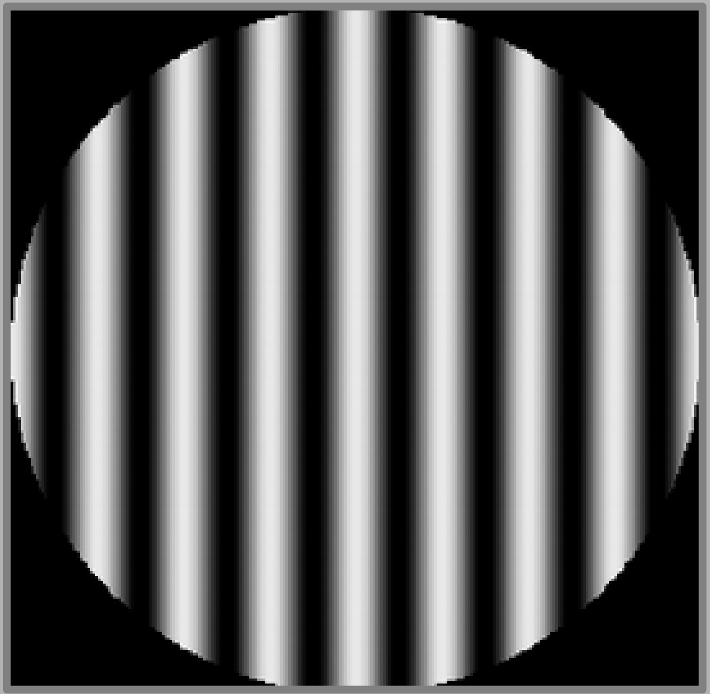
Focal Plane

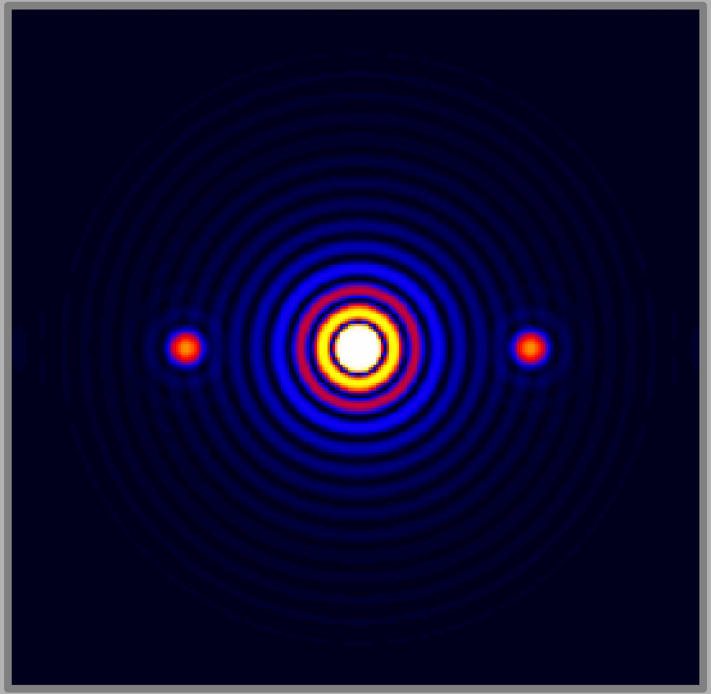


$$\mathrm{PSF}(\vec{r}) = \left|\mathcal{F}[\mathcal{A}(\vec{q})]\right|^2$$

Fourier Modes and Speckles

$$\phi(\vec{q}) \propto \cos(2\pi f u)$$

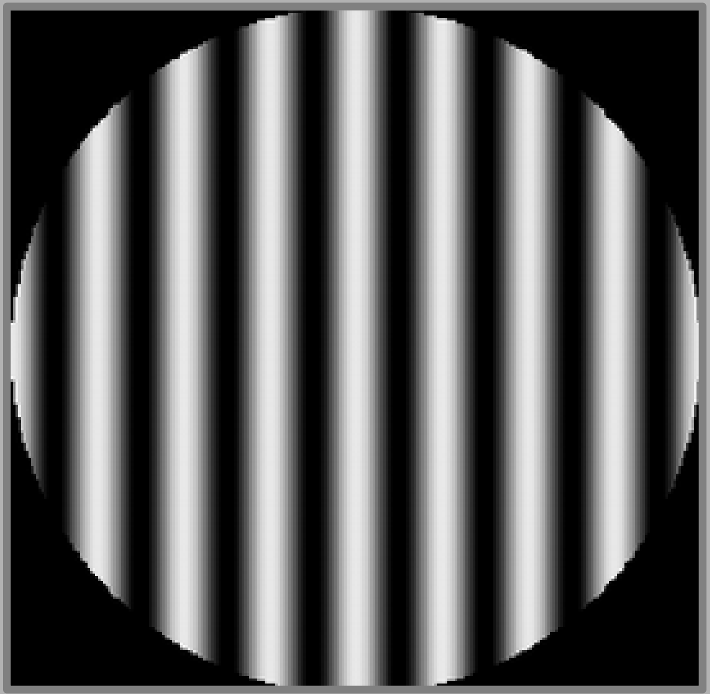


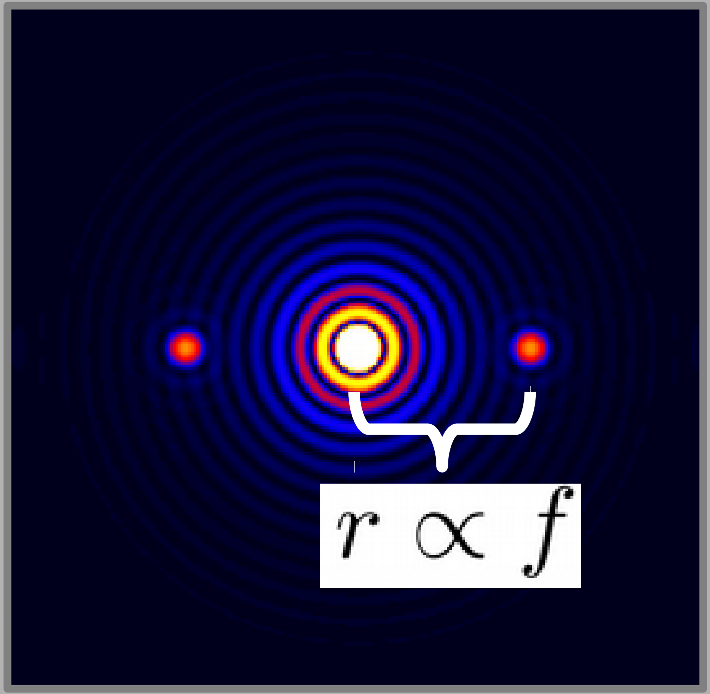


$$\mathrm{PSF}(\vec{r}) = \left| \mathcal{F}[\mathcal{A}(\vec{q})e^{i\phi(\vec{q})}] \right|^2$$

Fourier Modes and Speckles

$$\phi(\vec{q}) \propto \cos(2\pi f u)$$





$$\mathrm{PSF}(\vec{r}) = \left| \mathcal{F}[\mathcal{A}(\vec{q}) e^{i\phi(\vec{q})}] \right|^2$$

Relating Variance To PSF Contrast

- Long exposure post-coronagraph irradiance due to 1 mode:

$$\langle I_{mn}(\vec{r}) \rangle = \sigma_{mn}^2 \left[\mathrm{PSF}(\vec{r} - \vec{k}_{mn}\lambda) + \mathrm{PSF}(\vec{r} + \vec{k}_{mn}\lambda) \right]$$

Separation from star in focal plane

Temporal variance of the amplitude of the Fourier mode, after AO control

Spatial frequency defining the Fourier mode

Relating Variance To PSF Contrast

- Long exposure post-coronagraph irradiance due to 1 mode:

$$\langle I_{mn}(\vec{r})\rangle = \sigma_{mn}^2 \left[\mathrm{PSF}(\vec{r} - \vec{k}_{mn}\lambda) + \mathrm{PSF}(\vec{r} + \vec{k}_{mn}\lambda)\right]$$

- Due to all modes:

$$\langle I_{\Phi}(\vec{r})\rangle = \sum_{mn} \langle I_{mn}(\vec{r})\rangle$$

Relating Variance To PSF Contrast

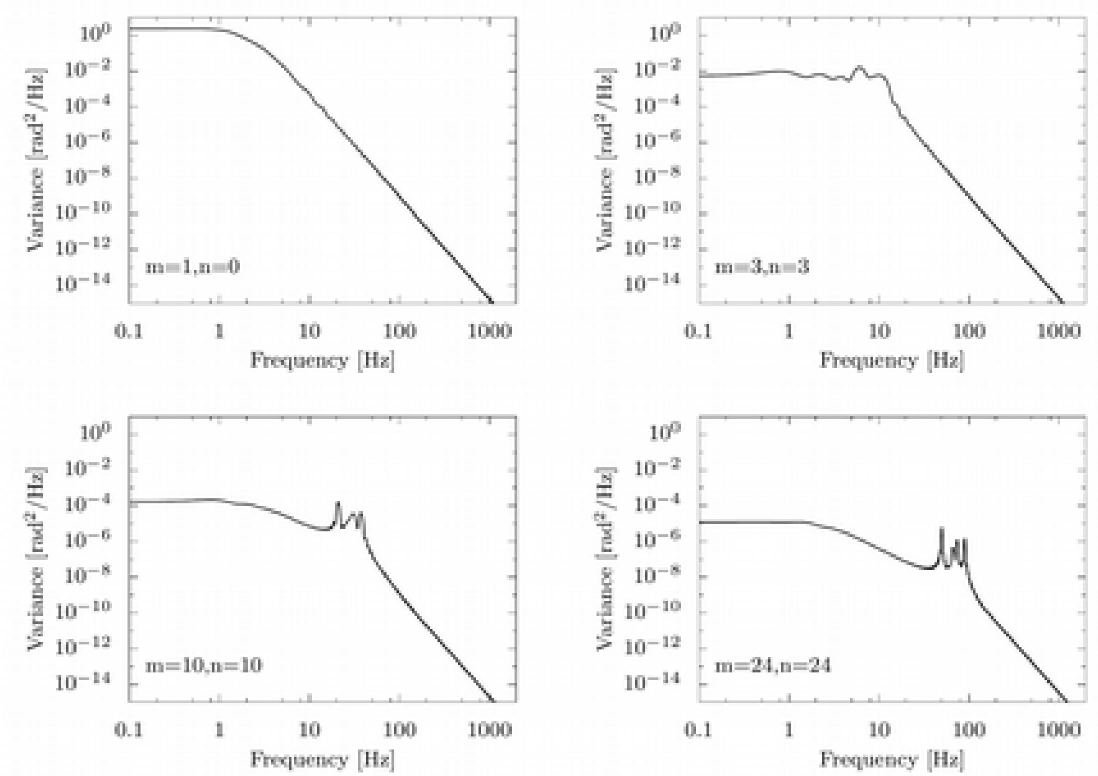
- Long exposure post-coronagraph irradiance due to 1 mode:

$$\langle I_{mn}(\vec{r})\rangle = \sigma_{mn}^2 \left[\mathrm{PSF}(\vec{r} - \vec{k}_{mn}\lambda) + \mathrm{PSF}(\vec{r} + \vec{k}_{mn}\lambda)\right]$$

- Post-coronagraph raw-contrast:

$$\langle I_\Phi(\vec{r})\rangle = \sum_{mn} \langle I_{mn}(\vec{r})\rangle \longrightarrow \langle C(\vec{r})\rangle = \frac{\langle I_\Phi(\vec{r})\rangle + \langle I_A(\vec{r})\rangle}{\langle S\rangle \ \mathrm{PSF}(0)}$$

- AO control as a time-domain problem problem is described by the temporal PSD



Theoretical input atmospheric turbulence PSDs

$$\sigma^2_{mn} = \int \mathrm{PSD}_{mn}(f)df$$

$$\sigma^2 = \sum_{mn} \sigma^2_{mn}$$

$$S = e^{-\sigma^2}$$

AO Control

- Basic control scheme: the Pure Integrator

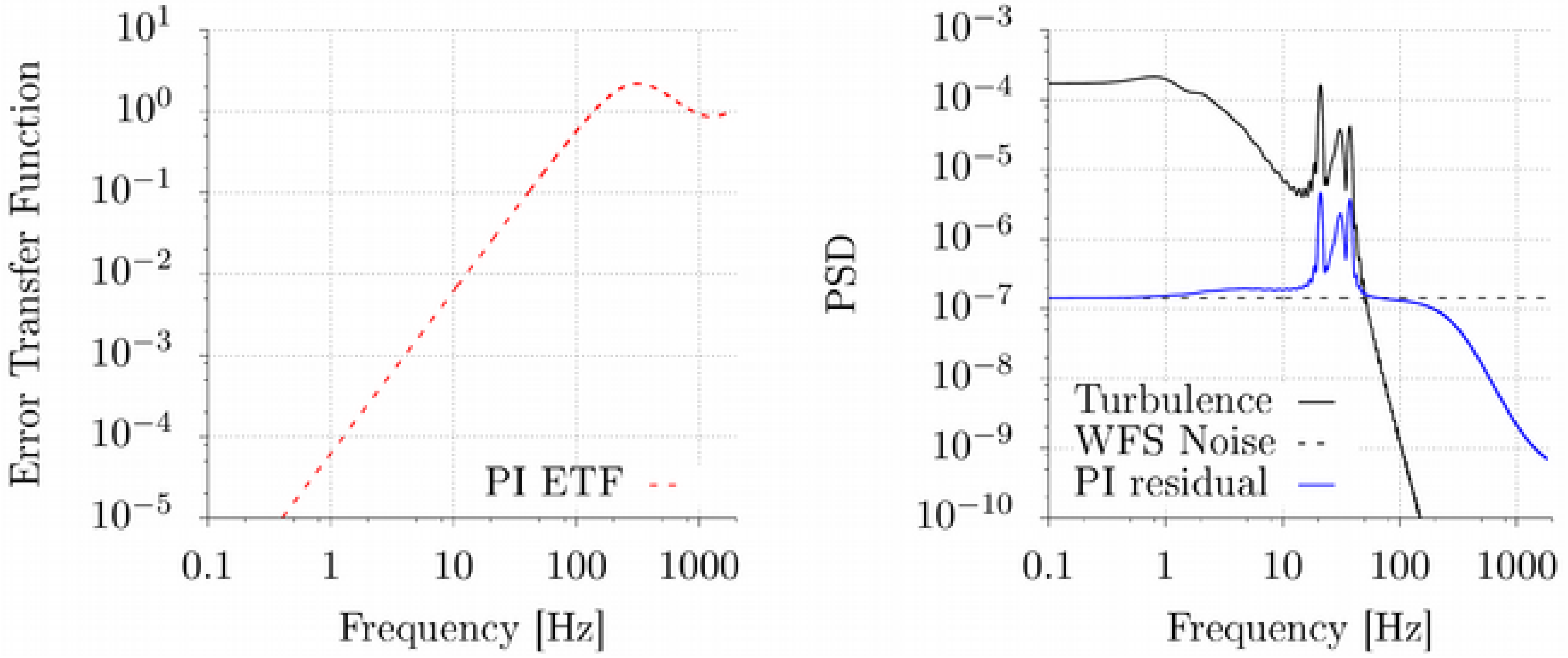
$$h_i = h_{i-1} + g\Delta h_i$$

- This takes a finite amount of time:
  - Integration time
  - Detector readout and data processing
  - Corrector motion
- Gain: what should g be?

Error Transfer Function

- The ETF quantifies the effects of sampling, delays, and the control law (e.g. the Pure Integrator).
- The ETF describes the relationship between the input PSD and the output PSD

$$\mathrm{PSD}_{cl,mn}(f) = \mathrm{PSD}_{ol,mn}(f)\left|\mathrm{ETF}_{mn}(f;g_{mn})\right|^2$$

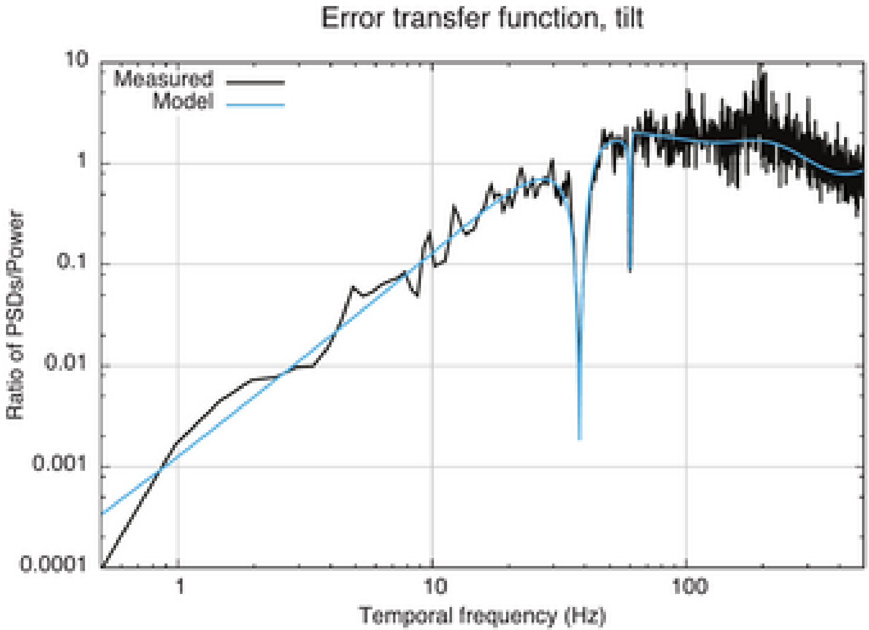


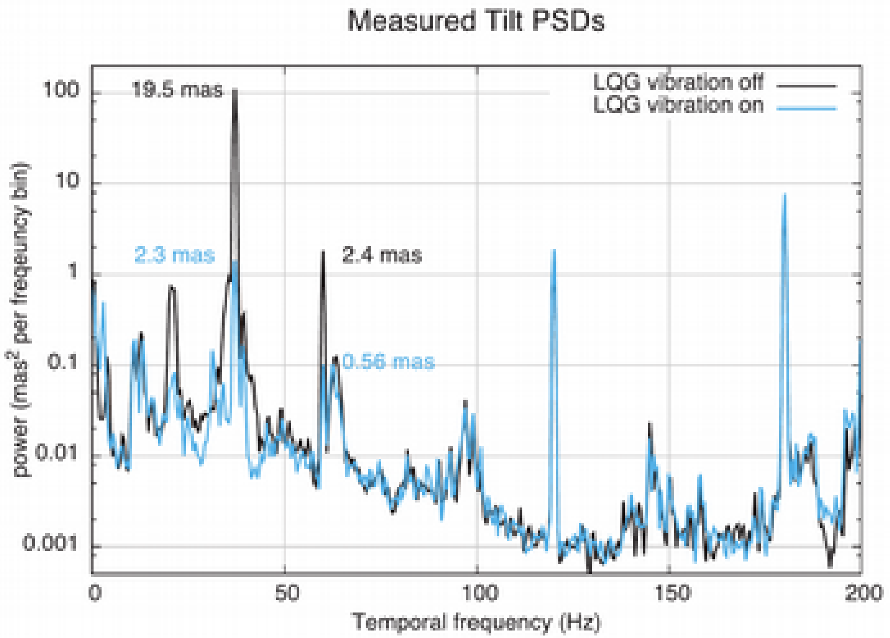
$$\mathrm{PSD}_{cl,mn}(f) = \mathrm{PSD}_{ol,mn}(f)\,|\mathrm{ETF}_{mn}(f; g_{mn})|^2$$

Current Status

- Pure Integrator controller is currently the main strategy in use
  - GPI: PI on Fourier modes
    - Gains optimized online, continuously
  - SPHERE: PI on Karhunen-Loeve modes
    - Gains optimized online, continously
  - LBTAO/MagAO: PI on KL modes
    - Gains optimized online, non-continously
- Noteworthy exception: LQG for vibrations

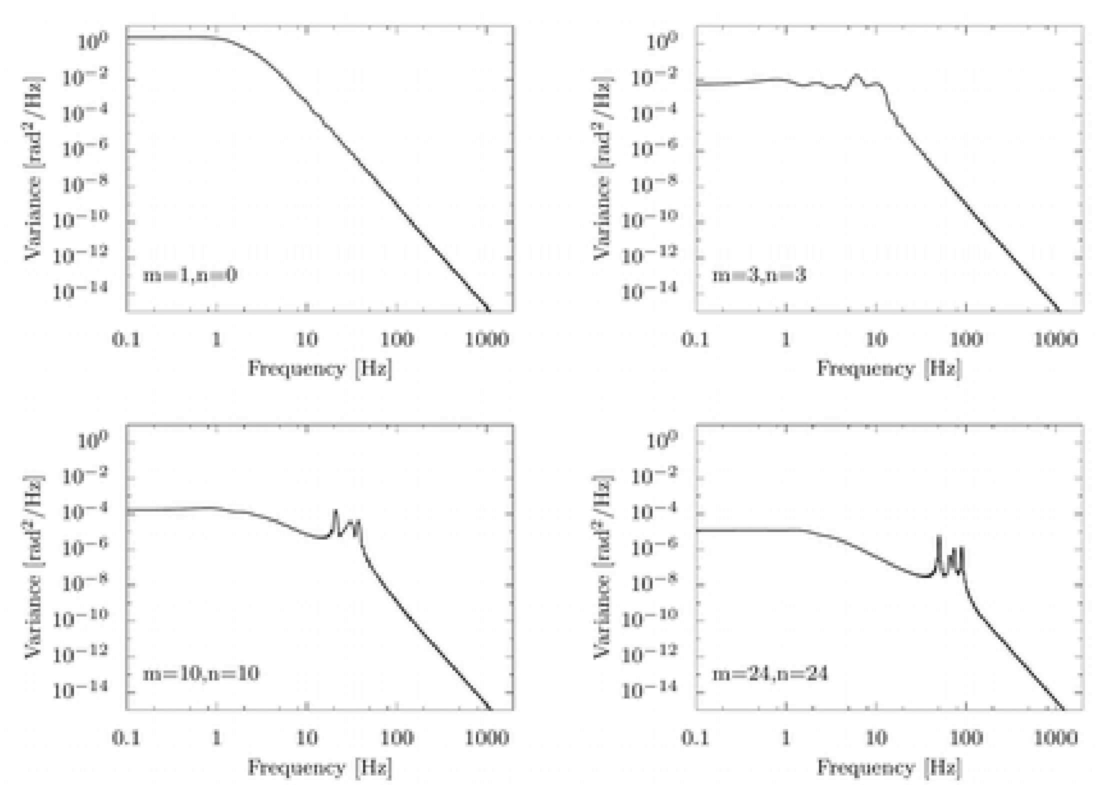
LQG Vibration Rejection





Results for GPI, from Poyneer et al 2016
See results from SPHERE, Petit et al 2014

Predictive Control

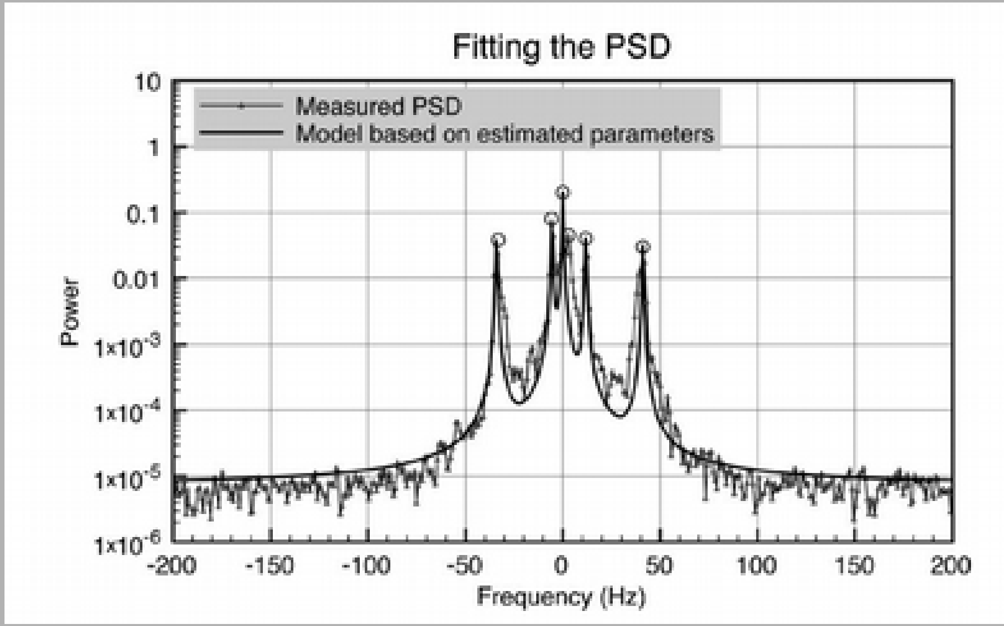


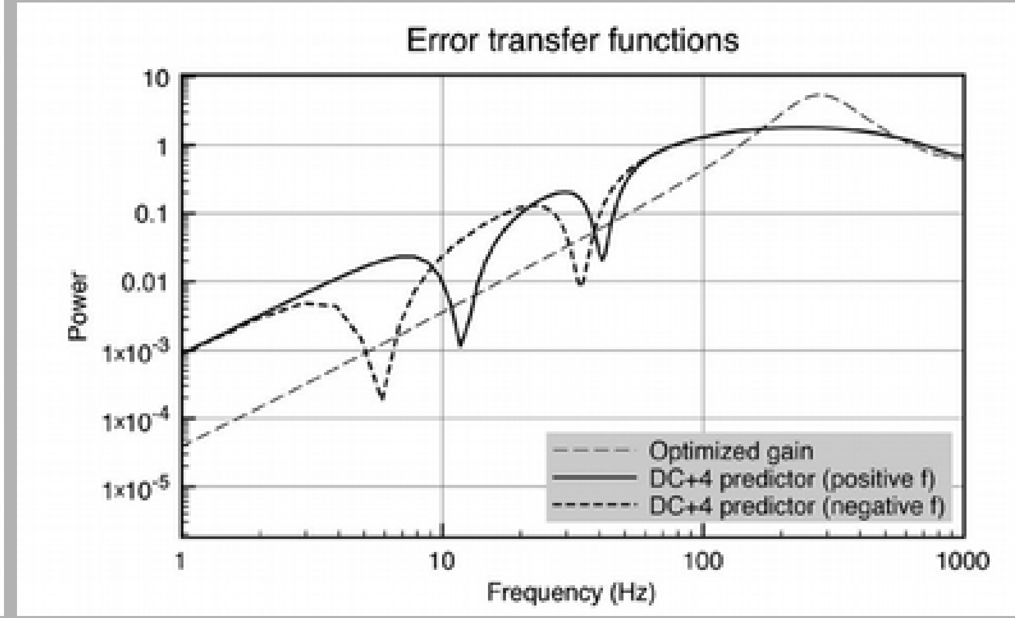
- The peaks from wind-layers are analagous to vibration "lines"
- Could be controlled with LQG
- This is "Predictive Fourier Control" (Poyneer, Macintosh, and Veran, 2007)

Predictive Control w/ PFC (LQG)

- Assumes frozen flow: discrete layers propagating at fixed velocities
    - These produce peaks at (V-dot-k) Hz
- Identify location and strength of wind-layer peaks
    - Build filter from those parameters

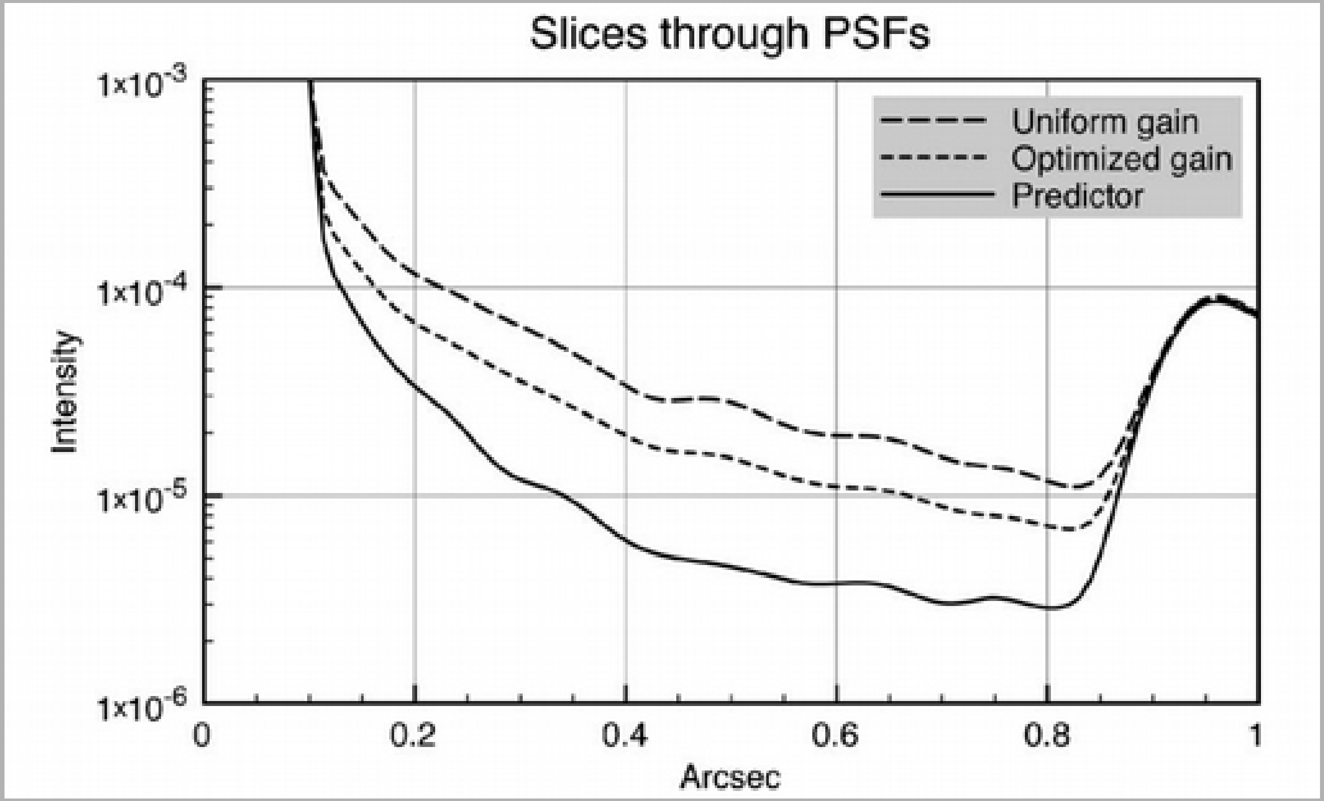
PFC In Action





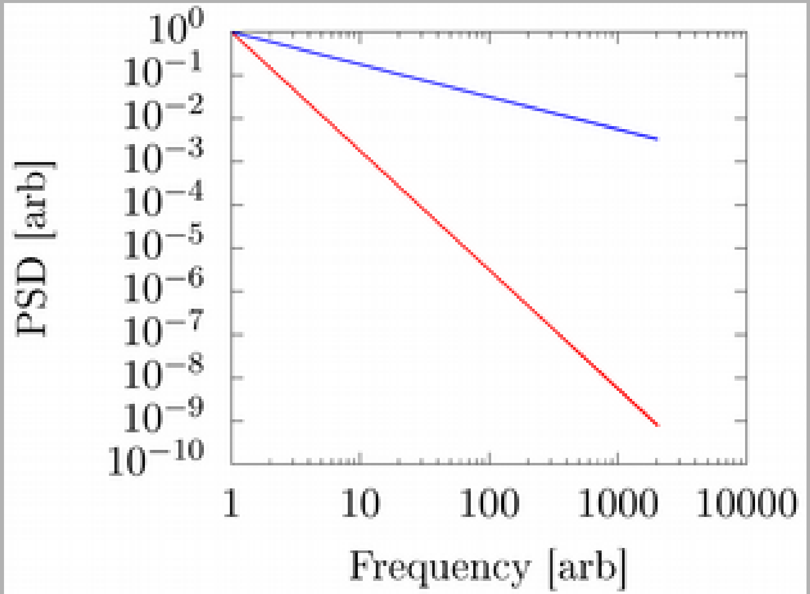
Poyneer et al, 2007

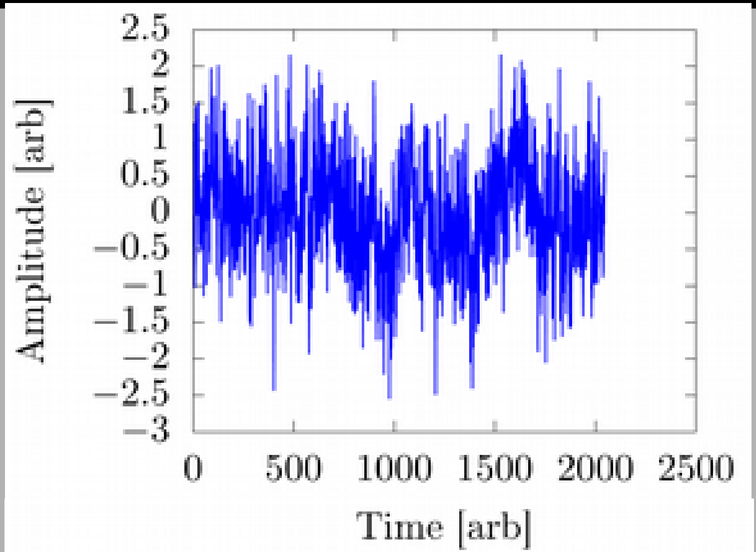
The Promise of PFC

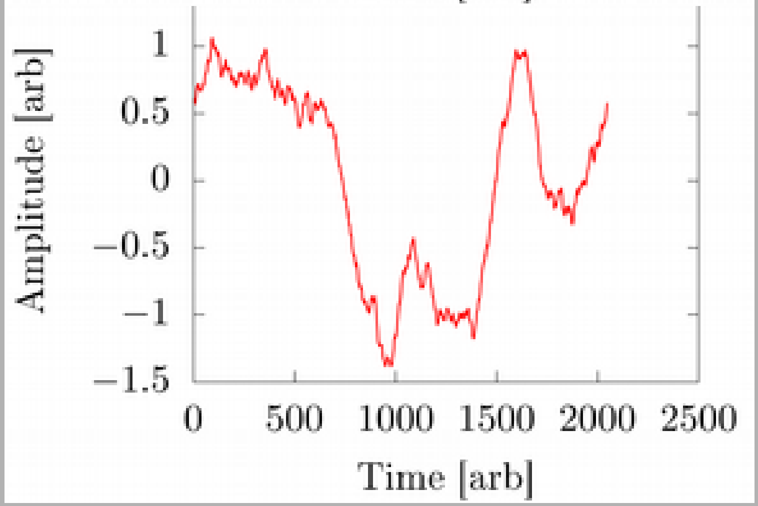


Note: PFC not yet implemented at GPI, see Alex Rudy et al, SPIE 2016

The Peaks Aren't Everything

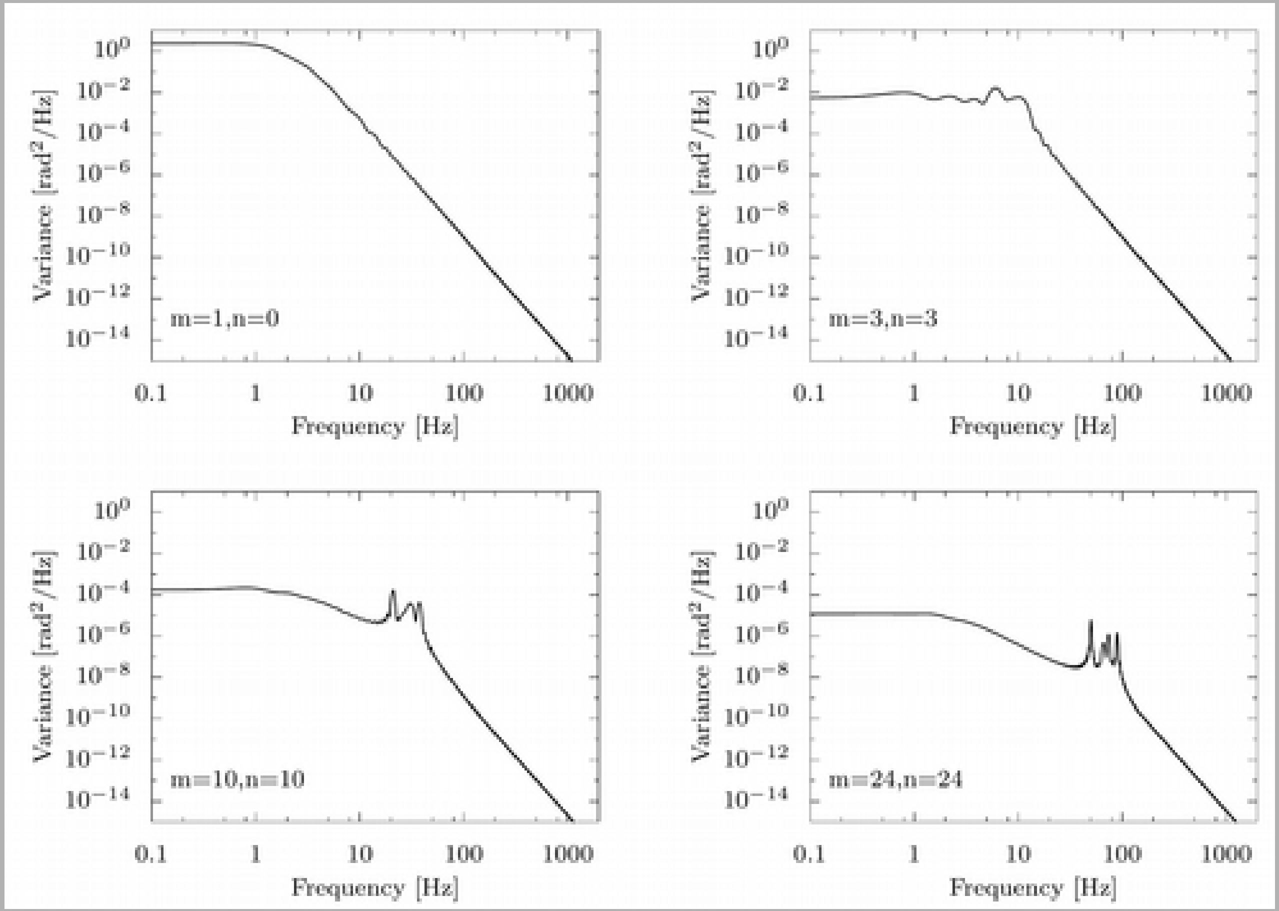






Correlations (i.e. predictability) exist in PSDs without peaks too . . .

More Than Peaks



Linear Prediction

- The general linear filter: $$h(t_i) = \sum_{j=1}^{J} a_j h(t_{i-j}) + \sum_{l=0}^{L} b_l \Delta h(t_{i-l})$$
- Problem: determine coefficients
  - PFC peak-identification is equivalent to picking a and b
- Dessenne et al (1998)
  - Used Regressive Least Squares on telemetry
  - Minimize LSE of h(t_i).
  - Computationally expensive...

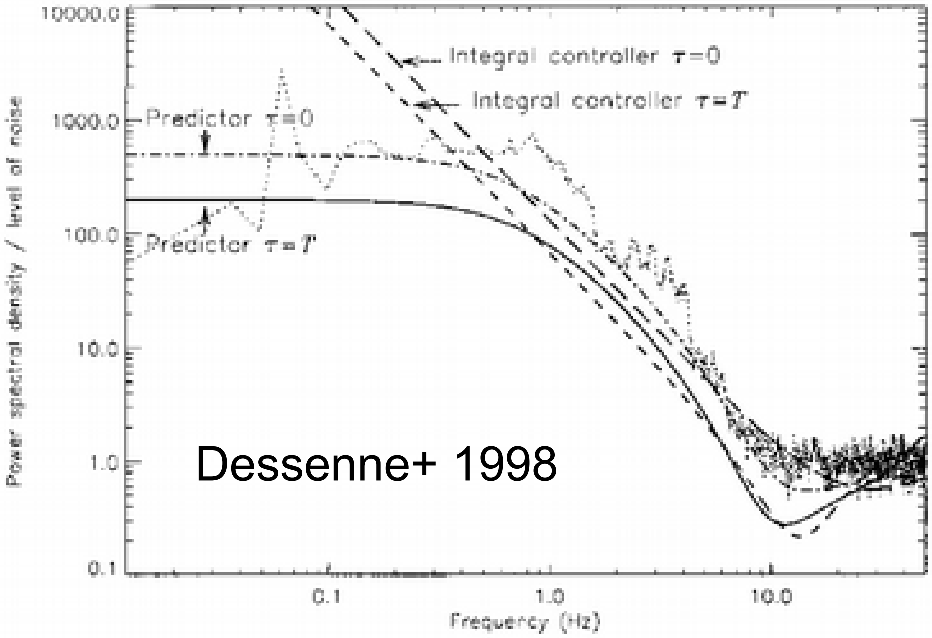
Role Of The Coefficients

- Form an "All-Pole" model of the input PSD

$$h(t_i) = \sum_{j=1}^{J} a_j h(t_{i-j}) + \sum_{l=0}^{L} b_l \Delta h(t_{i-l})$$

$$H_{con}(z) = \frac{\sum_{l=0}^{L} b_l z^{-l}}{1 + \sum_{j=1}^{J} a_l z^{-j}}$$

$$\widehat{\mathrm{PSD}(z)} \rightarrow |1 - H(z)|^2$$



Dessenne+ 1998

Linear Prediction

- Choose coefficients to minimize LSE
- This can be done with the "Yule-Walker" or "Normal" equations of Linear Prediction (see The Theory of Linear Prediction, P. P. Vaidyanathan, California Institute of Technology)

$$\mathbf{R}\vec{c} = -\vec{r}.$$

- It's all about the autocorrelation:

$$\mathbf{R} = \begin{bmatrix} \mathcal{R}_{mnp}(0) & \mathcal{R}_{mnp}(1) & \cdots & \mathcal{R}_{mnp}(N-1) \\ \mathcal{R}_{mnp}(1) & \mathcal{R}_{mnp}(0) & \vdots & \mathcal{R}_{mnp}(N-2) \\ \vdots & \cdots & \ddots & \vdots \\ \mathcal{R}_{mnp}(N-1) & \mathcal{R}_{mnp}(N-2) & \cdots & \mathcal{R}_{mnp}(0) \end{bmatrix}$$

$$\vec{r} = \begin{bmatrix} \mathcal{R}_{mnp}(1) \\ \mathcal{R}_{mnp}(2) \\ \vdots \\ \mathcal{R}_{mnp}(N) \end{bmatrix}$$

Linear Prediction

- Choose coefficients to minimize LSE
- This can be done with the "Yule-Walker" or "Normal" equations of Linear Prediction (see )

The Theory of Linear Prediction

P. P. Vaidyanathan

California Institute of Technology

$$\mathbf{R}\vec{c} = -\vec{r}.$$

- It's all about the autocorrelation:

$$\mathbf{R} = \begin{bmatrix} \mathcal{R}_{mnp}(0) & \mathcal{R}_{mnp}(1) & \cdots & \mathcal{R}_{mnp}(N-1) \\ \mathcal{R}_{mnp}(1) & \mathcal{R}_{mnp}(0) & \vdots & \mathcal{R}_{mnp}(N-2) \\ \vdots & \cdots & \ddots & \vdots \\ \mathcal{R}_{mnp}(N-1) & \mathcal{R}_{mnp}(N-2) & \cdots & \mathcal{R}_{mnp}(0) \end{bmatrix}$$

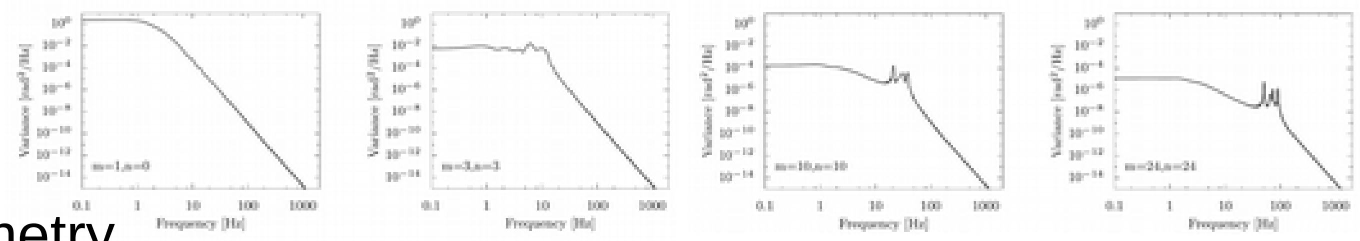
- Wiener-Khinchin:

$$\mathcal{R}(\tau) = \mathcal{F}\{\mathrm{PSD}(f)\}$$

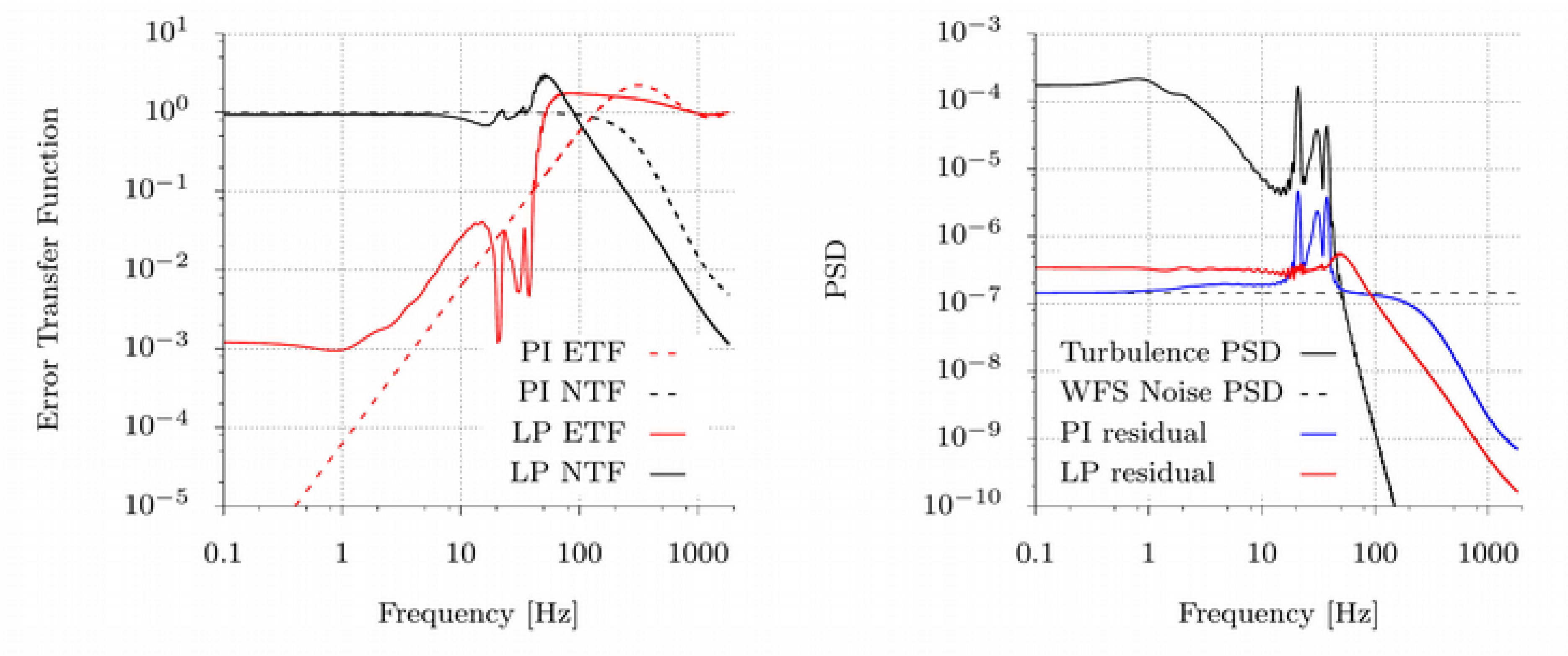
$$\vec{r} = \begin{bmatrix} \mathcal{R}_{mnp}(1) \\ \mathcal{R}_{mnp}(2) \\ \vdots \\ \mathcal{R}_{mnp}(N) \end{bmatrix}$$

A Recipe for LP

- Calculate PSD
  - Here: from theory
  - Online: from telemetry

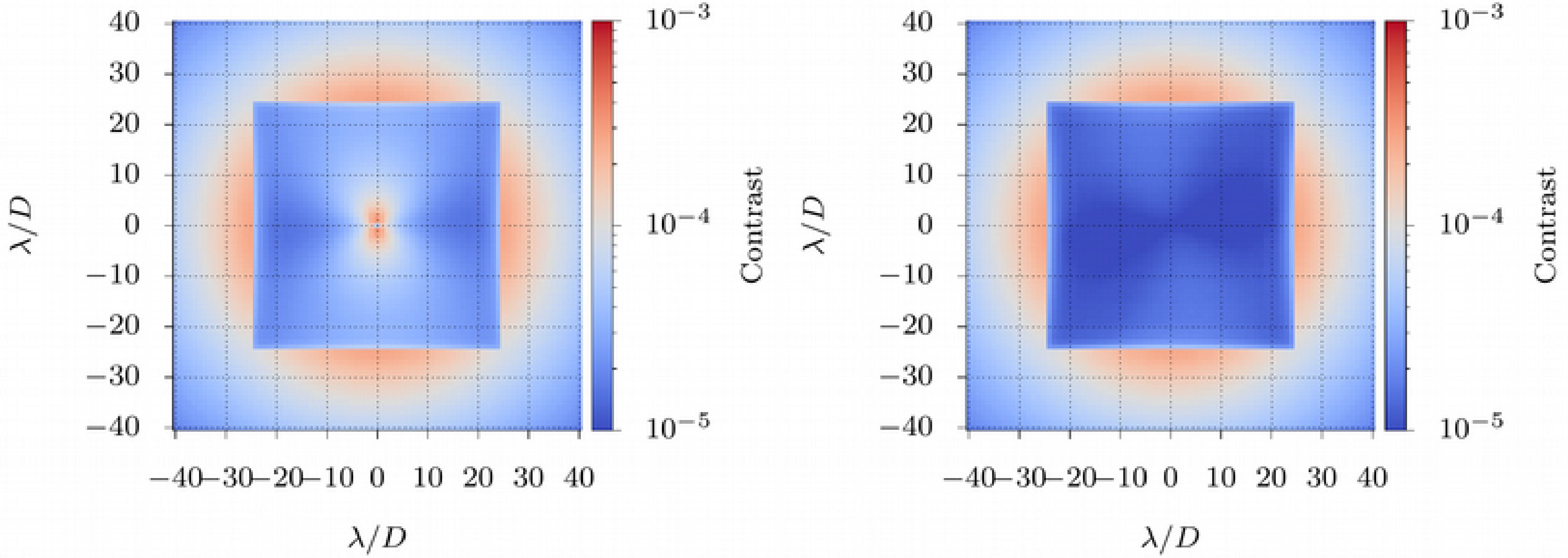


- Calculate autocorrelation: $\mathcal{R}(\tau) = \mathcal{F}\left\{\mathrm{PSD}(f)\right\}$
- Solve "Yule-Walker" equations by inverting R-matrix
  - The R matrix is "Toeplitz", use Levinson Recursion (very very fast)
- Use the c as a and b: $h_i = \sum_{j=0}^{N-1} c_j \left(h_{i-j-1} + \Delta h_{i-j}\right)$
- Don't forget to optimize gain (equivalent to adjusting the b)
  - And make sure it's stable . . .

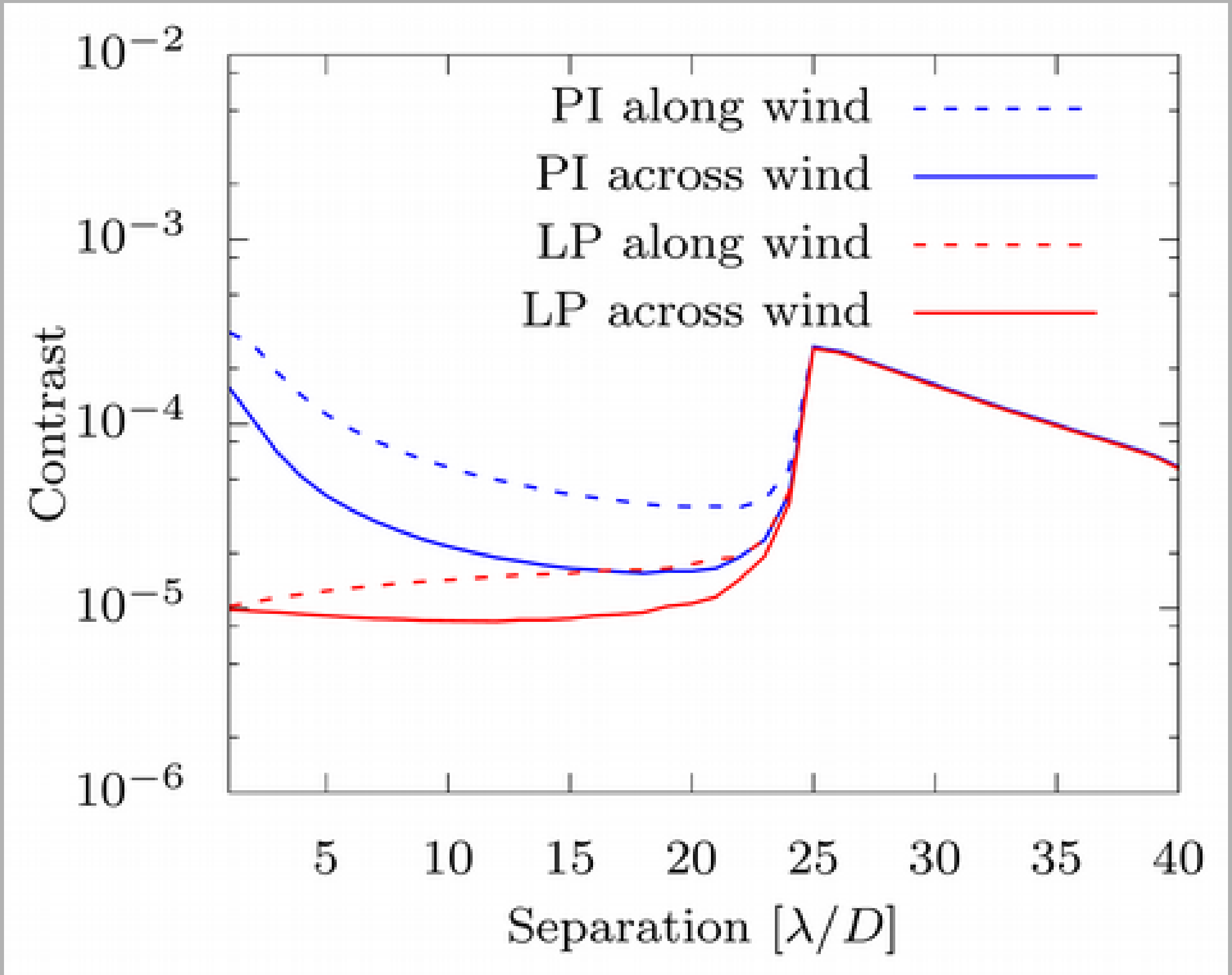


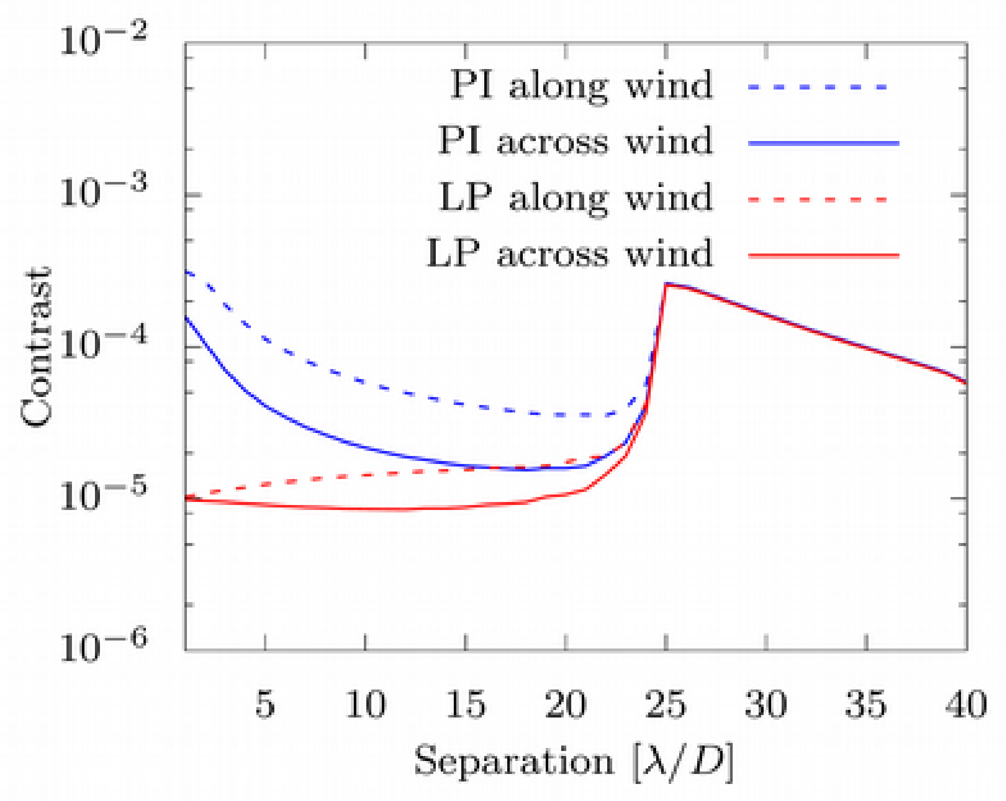
Error Transfer Function
10^1
10^0
10^-1
10^-2
10^-3
10^-4
10^-5
0.1
1
10
100
1000
Frequency [Hz]
PI ETF
PI NTF
LP ETF
LP NTF
PSD
10^-3
10^-4
10^-5
10^-6
10^-7
10^-8
10^-9
10^-10
Frequency [Hz]
Turbulence PSD
WFS Noise PSD
PI residual
LP residual

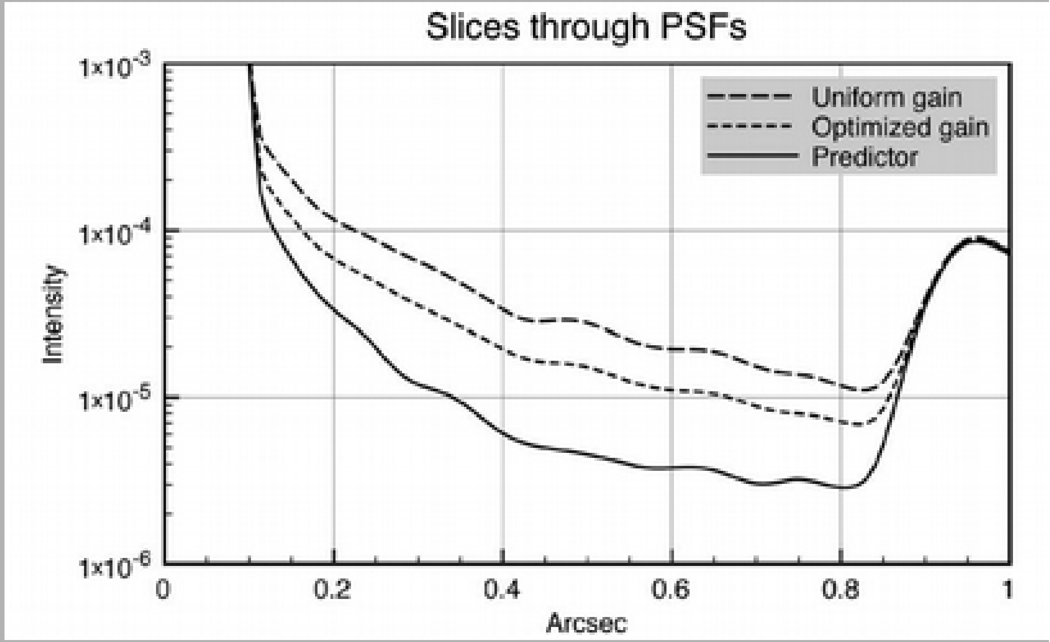
Linear Prediction on 5th mag Star



Linear Prediction on 5th mag Star

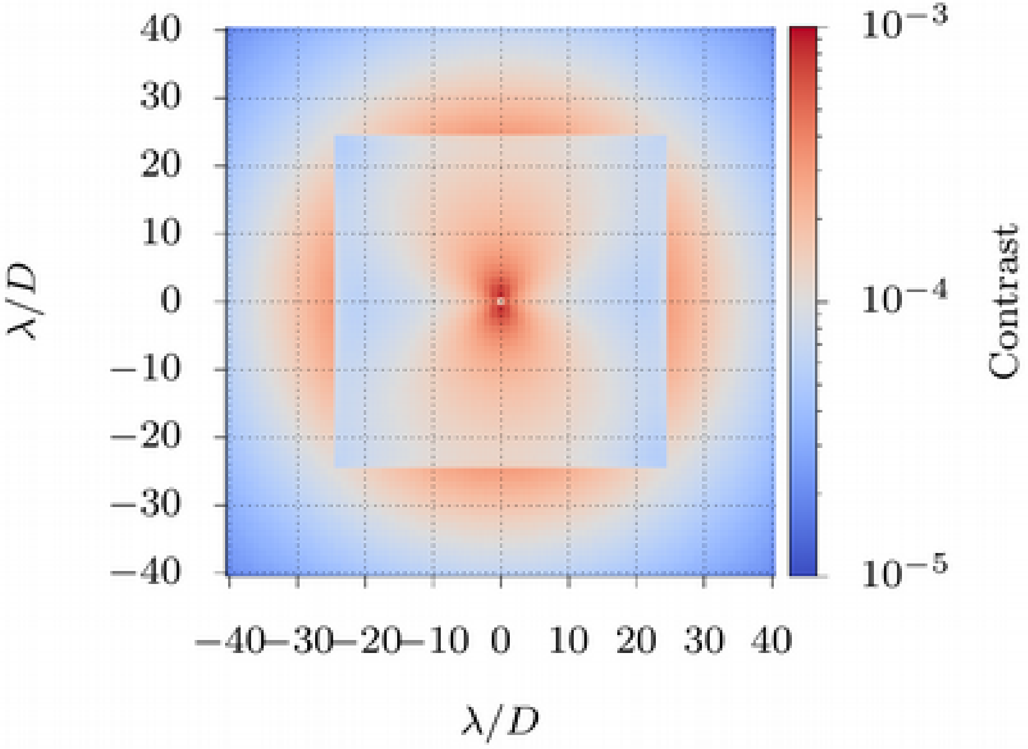


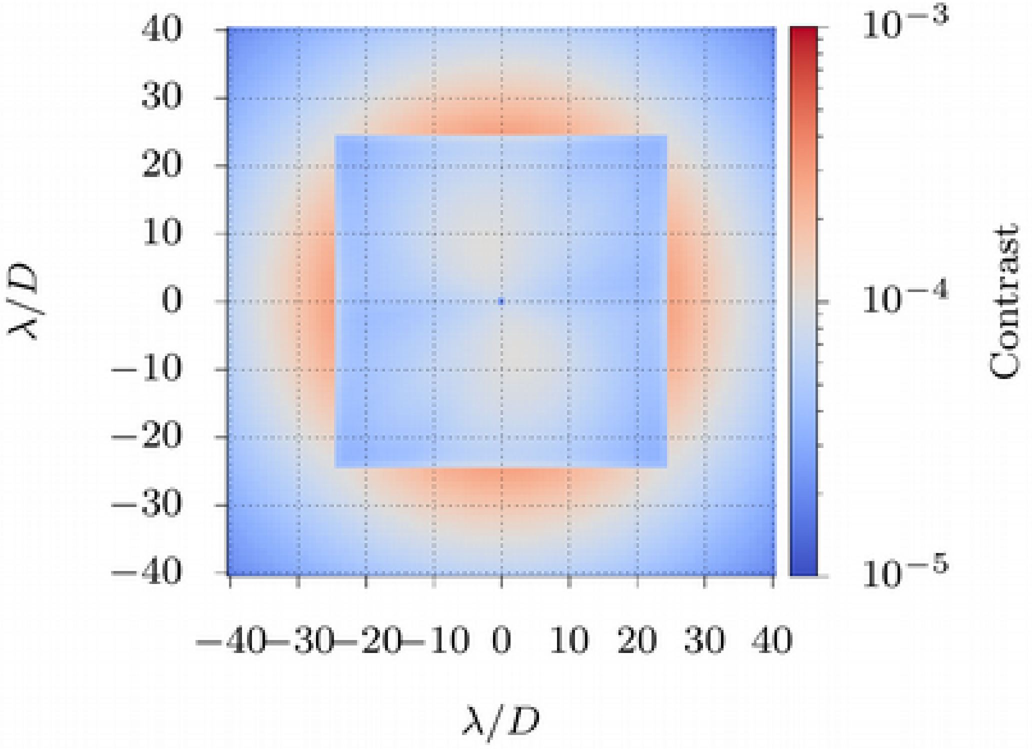




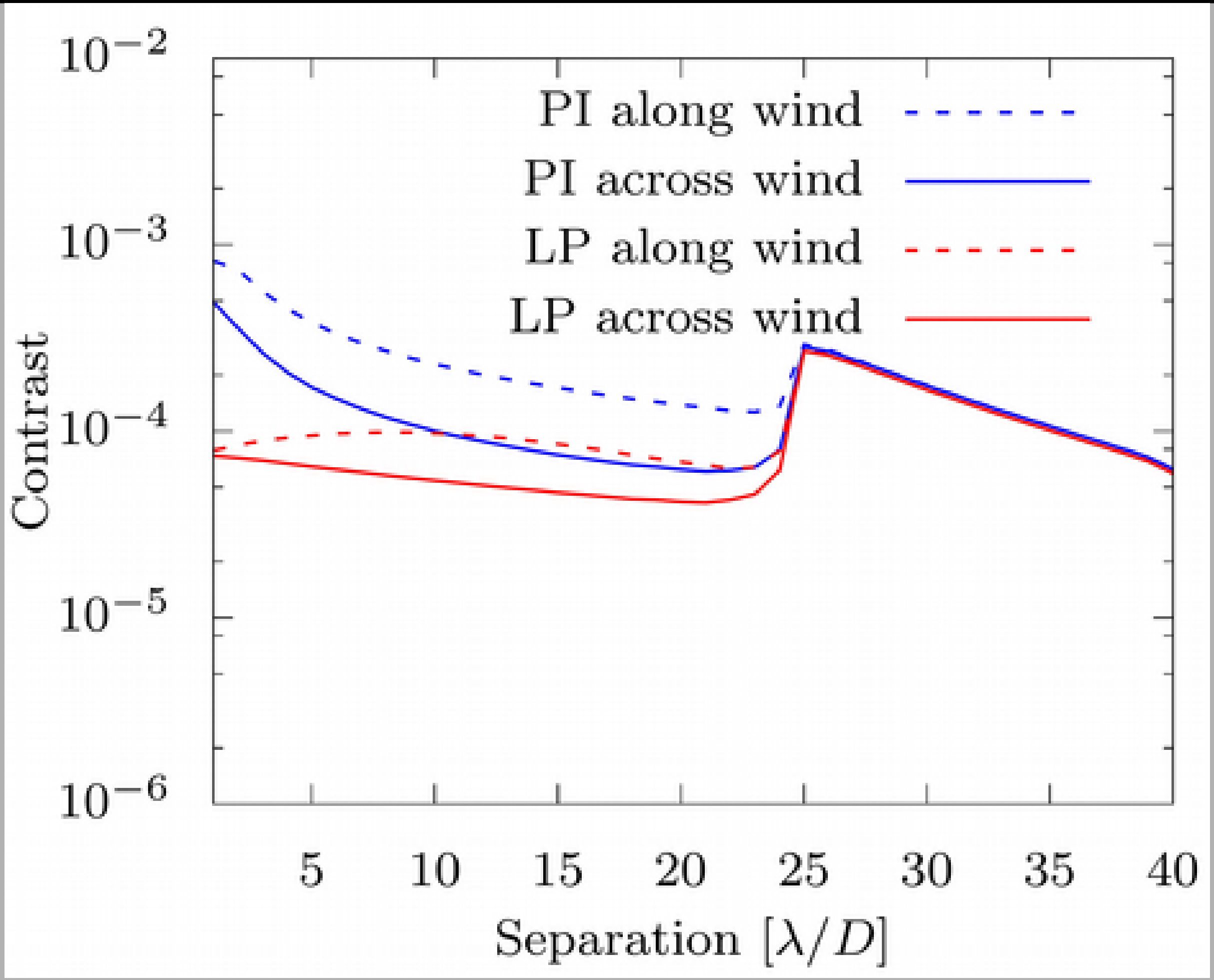
From Poyneer+ 2007

Linear Prediction on 8th mag Star

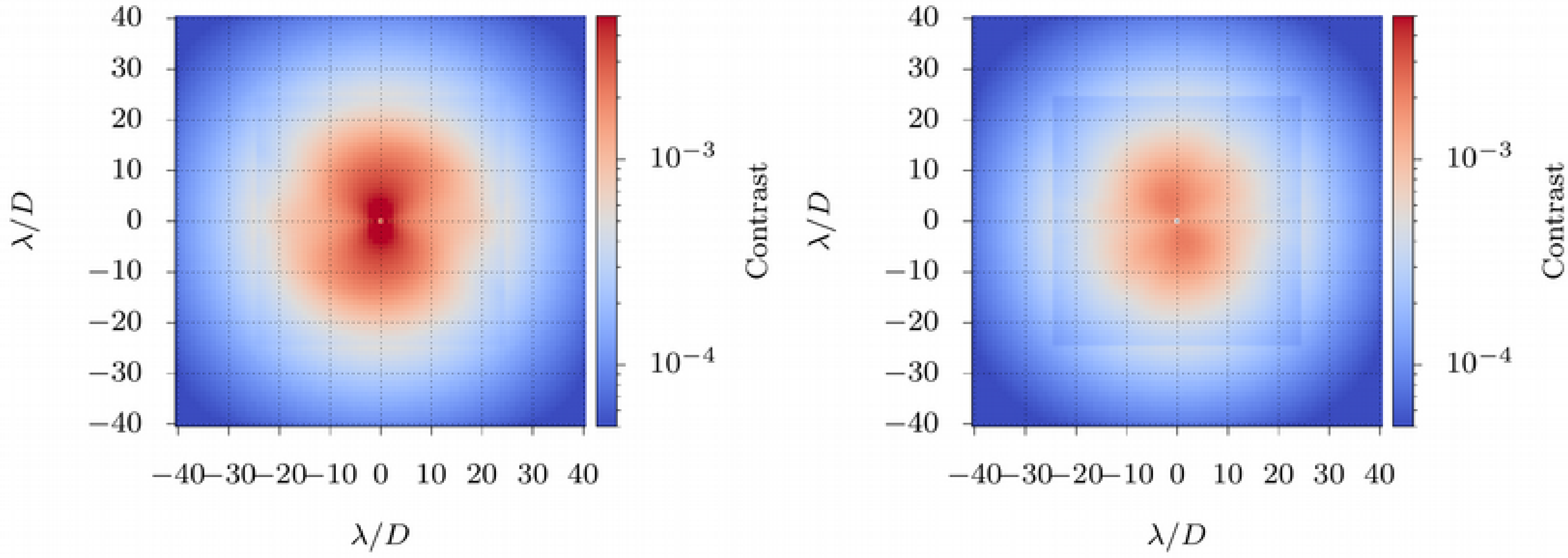




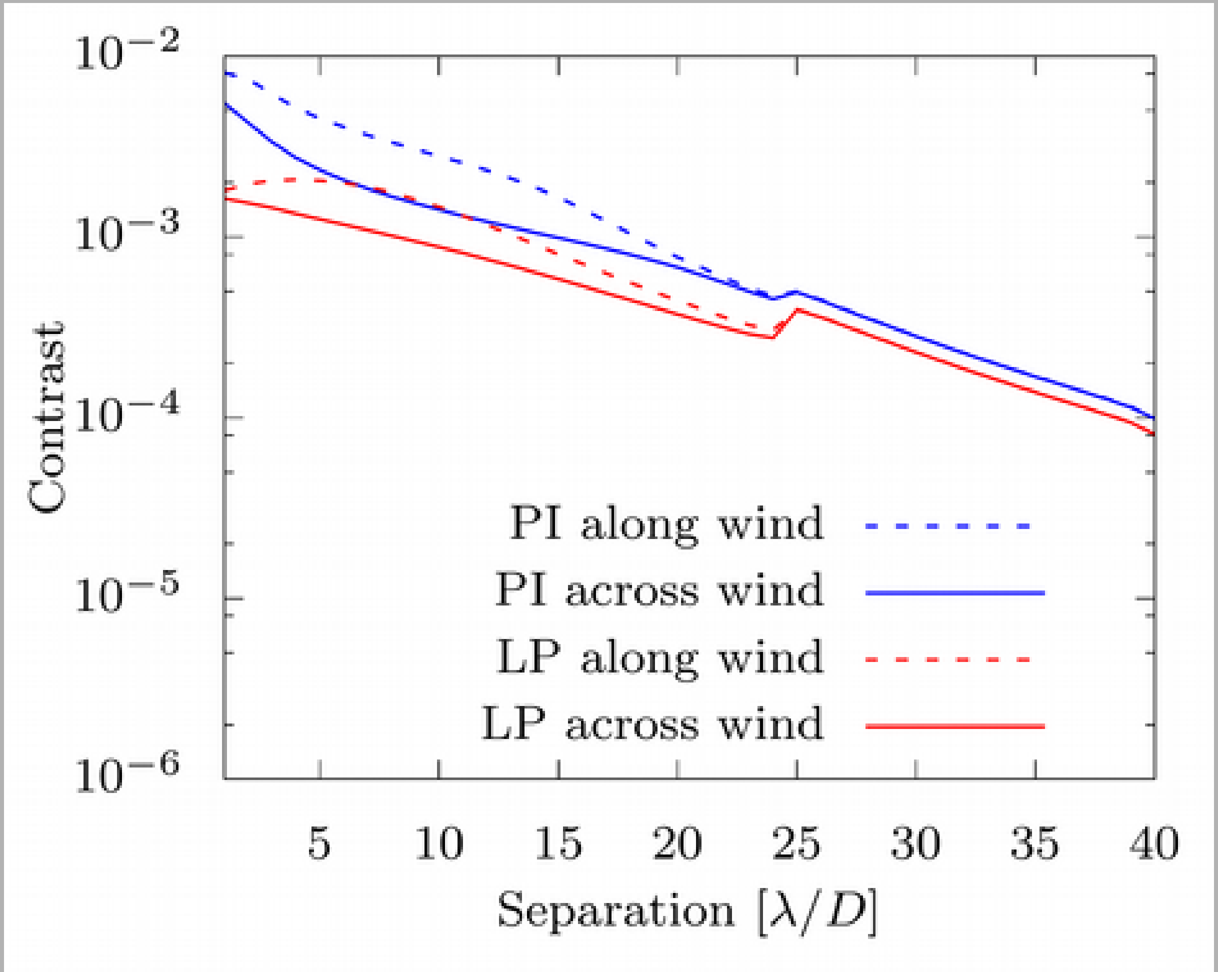
Linear Prediction on 8th mag Star



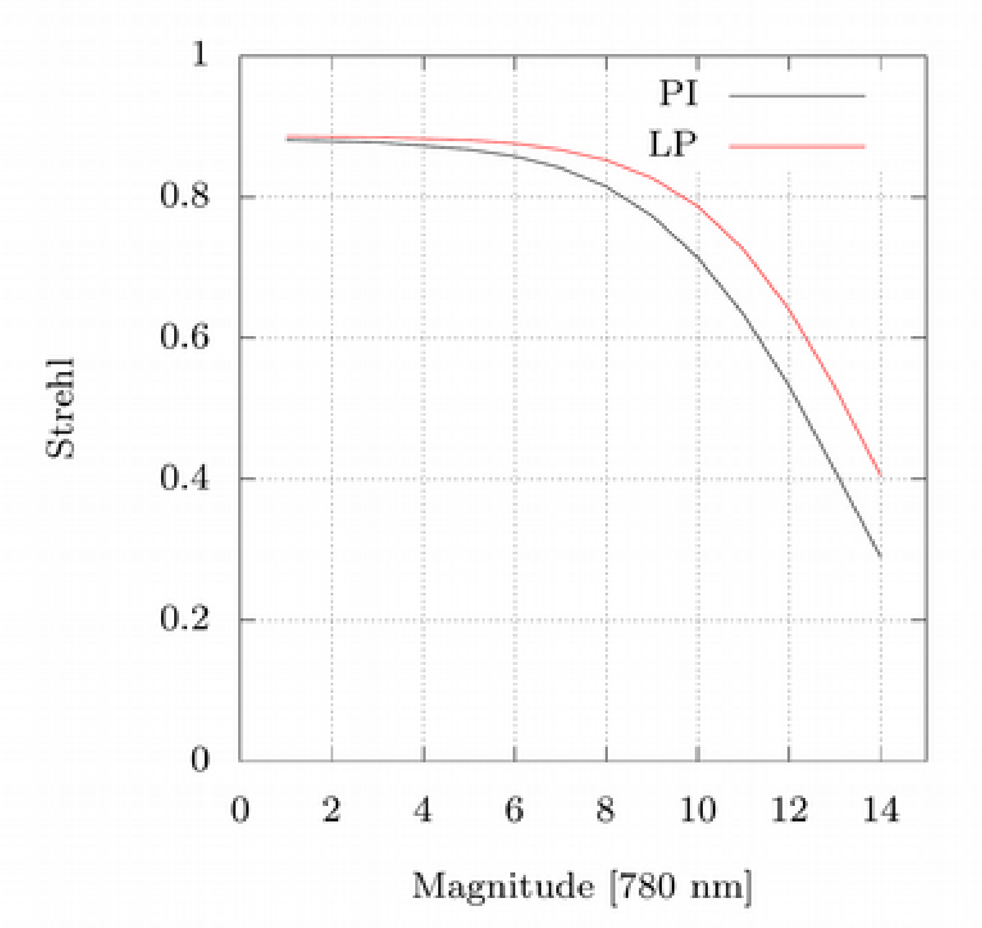
Linear Prediction on 12th mag Star



Linear Prediction on 12^{th} mag Star

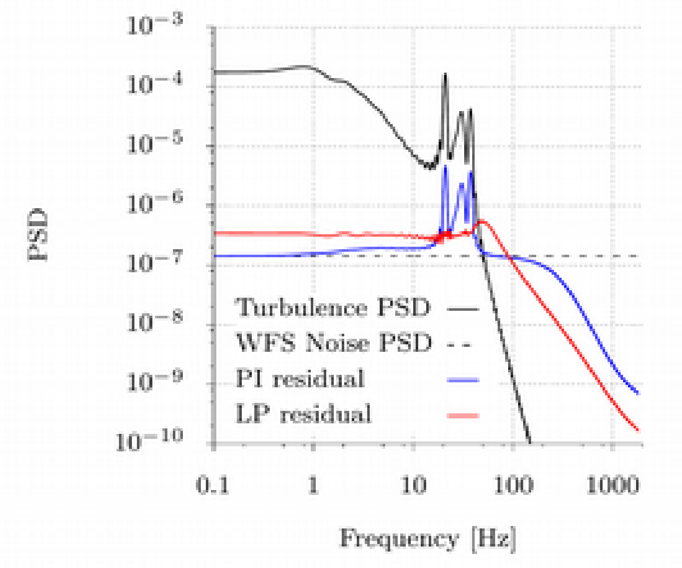


Impact on Strehl Ratio



Temporal Improvement Too

- Residual PSD whitening:



- To-do: perform this analysis without appealing to long exposures, and analyze **speckle lifetimes** (see Frazin's work)
  - Address limits of FPWFS
  - Address limits of P.P.

Caveats

- This LP analysis is purely semi-analytic
  - No simulations
  - Think of it as the closed-loop version of Guyon, 2005
- Numerical stability issues
  - Large numbers of coefficients and numerical accuracy
  - See Poyneer, 2008
- Statistical stability
  - For how long does a PSD describe the process?
  - How long is the filter optimum?

Combining LP with Sensor Fusion

- Extend LP to include non-WFS measurements
  - e.g. Accelerometer measurements on top-end.

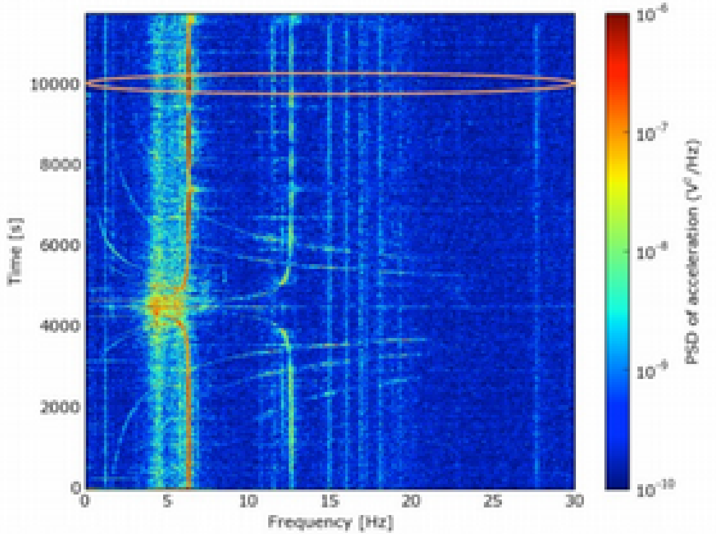


Fig. 12. Data from accelerometers attached to the top ring of the telescope. The data is only presented for the elevation axis. The orange ellipse indicates the period which the data was acquired with the photometer.

From SCExAO
See Lozi+ @SPIE 2016

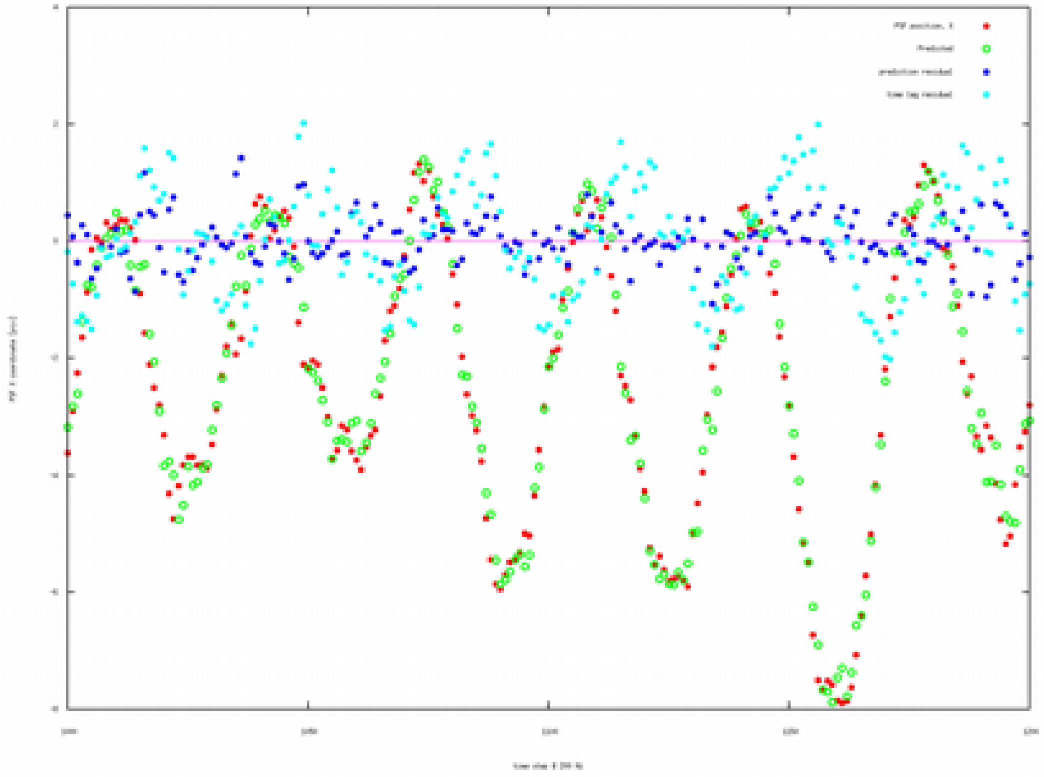
EOFs

- Empirical Orthogonal Functions
  - PCA in Space & Time
  - Space => any measurement of "space" you can come up with
    - WFS measurements (Pupil for Focal plane)
    - Accelerometers
    - The stock market
- Using regression over time history (similar to Dessenne)
  - Find coefficients of filter which minimize LSE
  - Includes a time-delay
- About to be tested (hopefully) on SCExAO (Guyon+ in prep.)

Algorithm

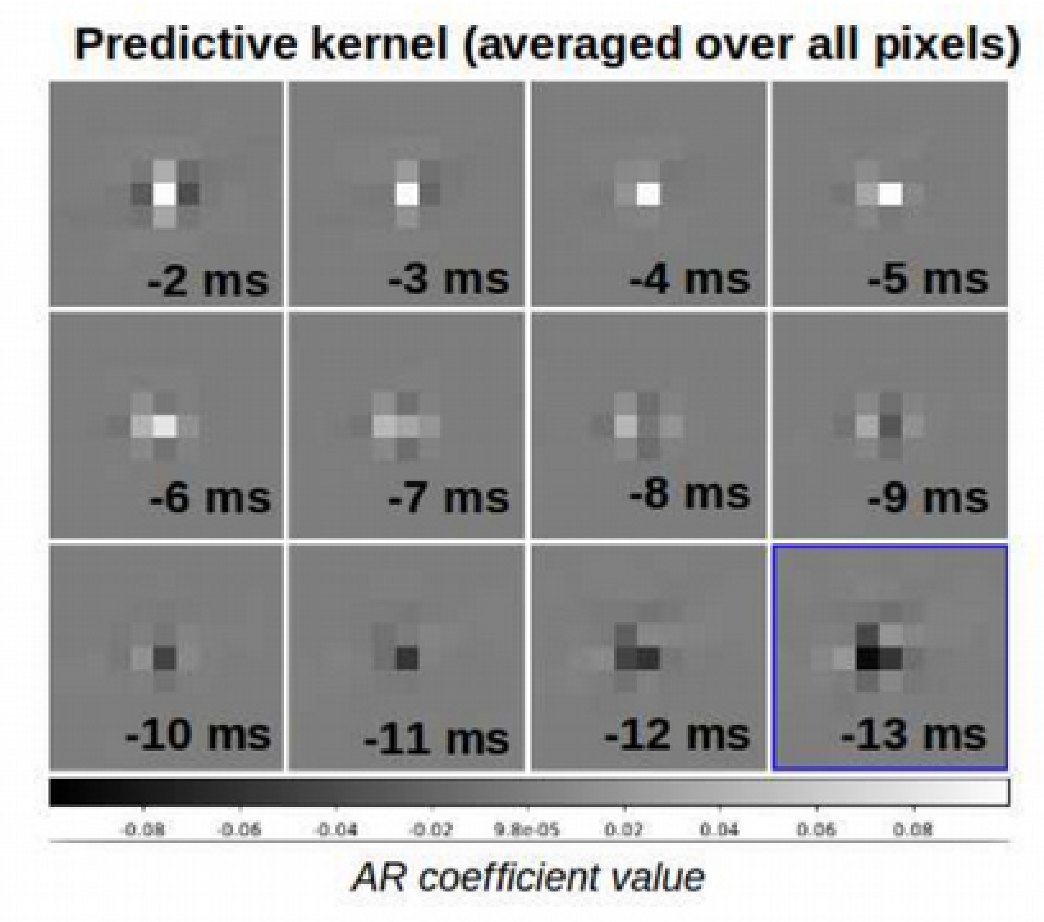
Steps:

- Record simultaneously accelerometer telemetry + PSF position (open loop)
- Construct optimal linear predictor for whole data set → extract Empirical Orthogonal Functions (EOFs)
- Project real-time accelerometer data on EOFs basis
- Extract predicted PSF position from EOFs projection
- Apply real-time correction to AO188 TT mount

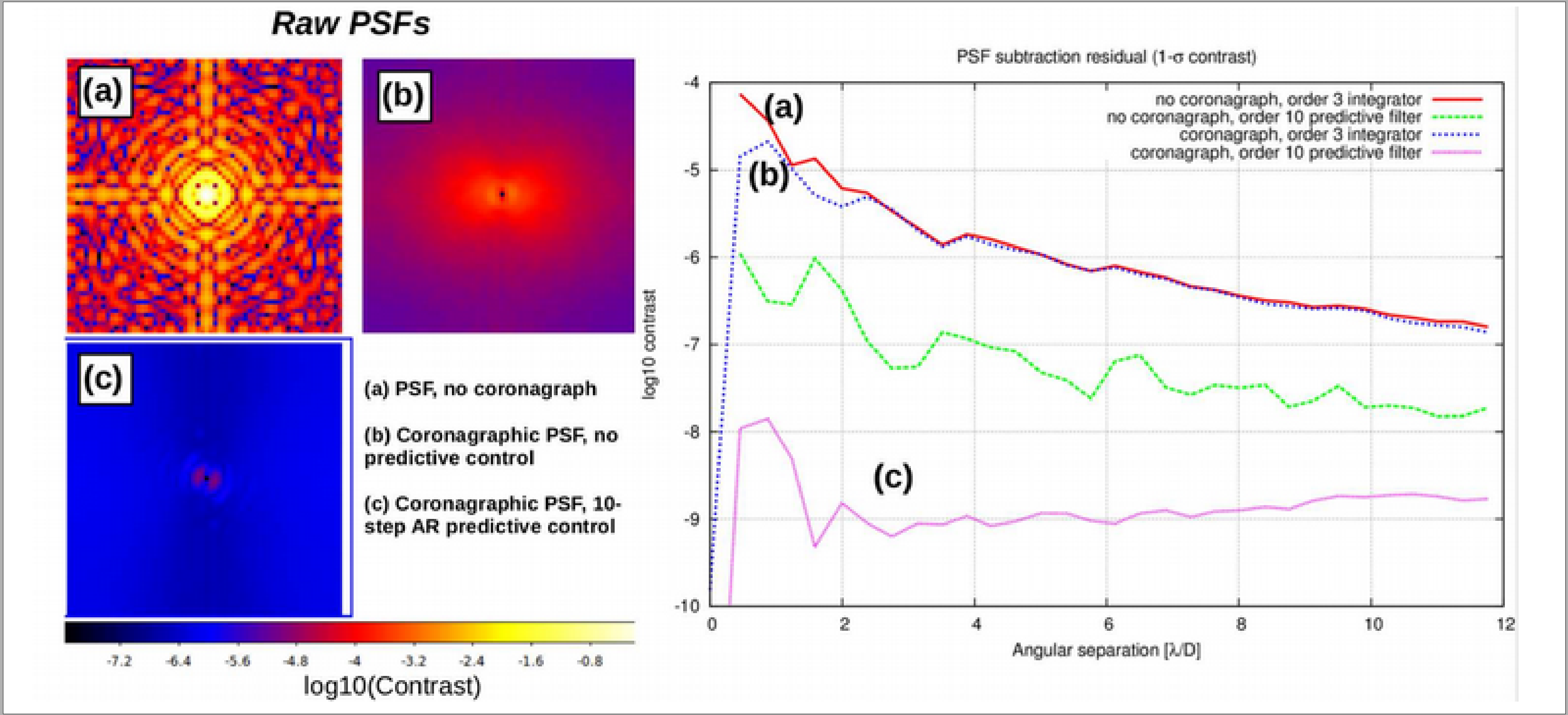


Slide from Olivier Guyon
Guyon et al, in prep.

Spatio-Temporal Correlation

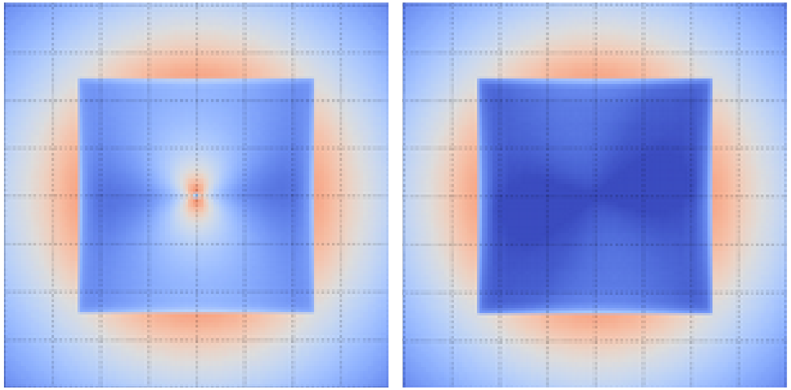


10-100 Gain After P.P.



Why Aren't We Doing This?

- We've been talking about this for > 2 decades
  - Yet neither SPHERE nor GPI are doing P.C. on all modes (just LQG on T/T/F)
- Not a huge benefit in Strehl, esp. on bright stars
  - Might talk ourselves out of expending resources on it
- Are we not there yet?
  - I.e., are we too limited by other issues?
  - We aren't working at the IWA where it matters?
  - Do we think computers aren't up to it?
- Does it just not work?
  - Statistical stability?
  - Numerical robustness?



Selected References

- Poyneer et al
  - 2005 (Optimal Fourier Control) http://adsabs.harvard.edu/abs/2005JOSAA..22.1515P
  - 2007 (Predictive Fourier Control) http://adsabs.harvard.edu/abs/2007JOSAA..24.2645P
  - 2009 (Frozen Flow Verification) http://adsabs.harvard.edu/abs/2009JOSAA..26..833P
  - 2016 (GPI Performance) http://adsabs.harvard.edu/abs/2016ApOpt..55..323P
- Dessenne et al
  - 1997 (Predictive Control) http://adsabs.harvard.edu/abs/1997OptL...22.1535D
  - 1998 (Predictive Control) http://adsabs.harvard.edu/abs/1998ApOpt..37.4623D
  - 1999 (On-sky demo) http://adsabs.harvard.edu/abs/1999OptL...24..339D
- LQG / Kalman
  - Le Roux et al (2004) http://adsabs.harvard.edu/abs/2004JOSAA..21.1261L
- Linear Prediction
  - Vaidyanathan (book) http://authors.library.caltech.edu/25063/1/S00086ED1V01Y200712SPR003.pdf
  - https://en.wikipedia.org/wiki/Linear_prediction