Machine Learning for calibration of non-common path aberrations in ELT-HARMONI

Alvaro Menduina-Fernandez*, and Matthias Tecza. Department of Astrophysics, University of Oxford
*alvaro.menduinafernandez@physics.ox.ac.uk
AlvaroMenduina

AIA 2019
22-26 July
ESO Garching

Background

Non-common path aberrations (NCPA) inside astronomical instruments hinder optical performance. Novel calibration techniques are needed to estimate aberrations and compensate them. Improved performance means enhanced observations. Better data means new exciting science!
Goal: *Could we use ML to estimate the aberrations in our instrument?*

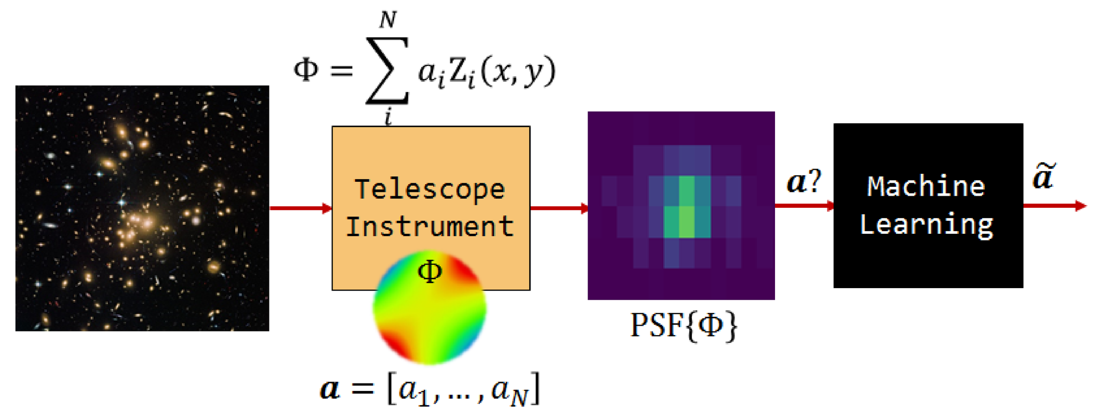

Figure 1 – Machine Learning in the context of NCPA calibration

Methodology

Single-Network approach: MLP regressor trained on PSF images to identify underlying aberrations $\bar{a} = [a_1, \ldots, a_N]$
Main caveat: *Curse of dimensionality.* Size of training set blows up with N number of aberrations we want to calibrate
Proposed solution: multi-network approach where we split the task among several networks, each trained to recognize a small subset of aberrations (keeps training sets manageable), see **Fig. 2**

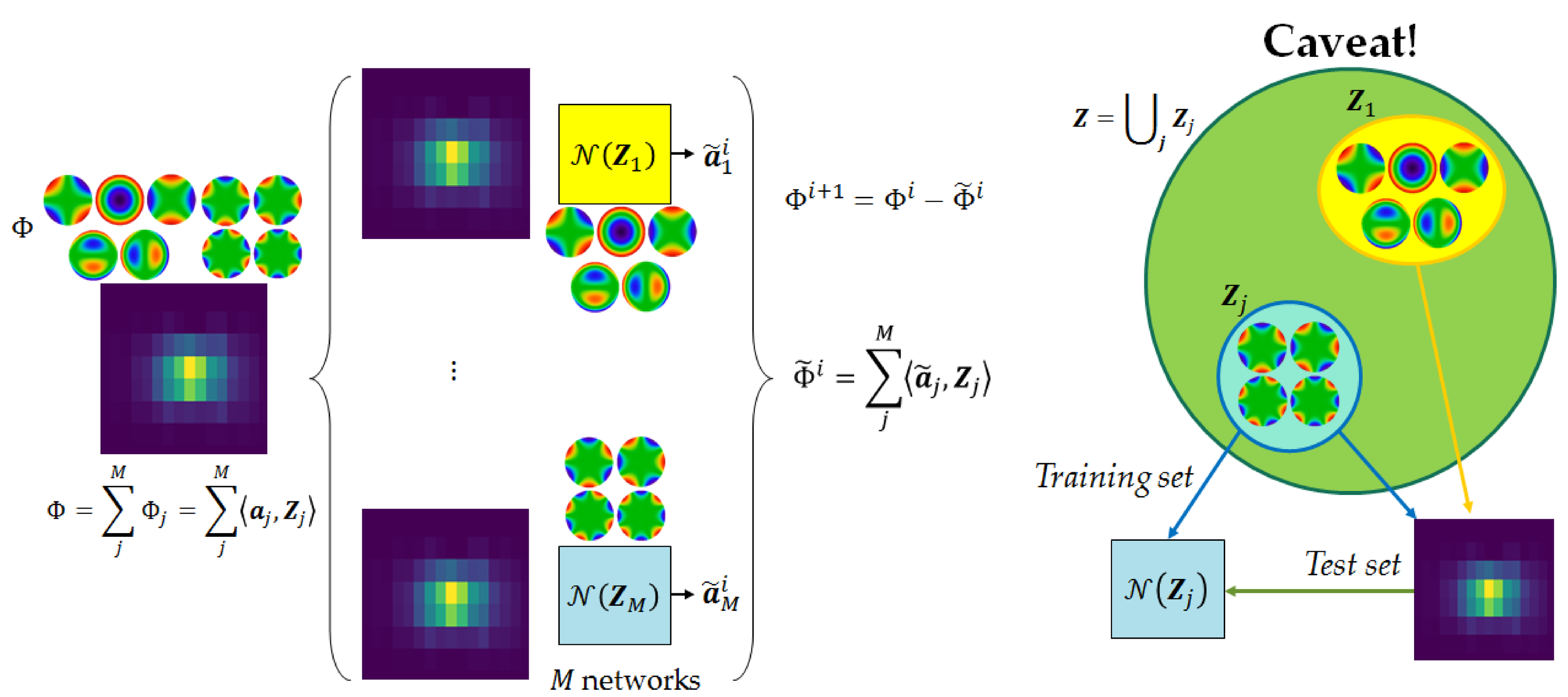

Figure 2 – **Multi-Network approach**. Each network is trained to recognize a small set of aberrations. Problem: distribution mismatch training-testing!

Problem: each network sees images with features from aberrations it doesn't know about! **Feature contamination** degrades performance
What can we do to clean the images?

Denoising AutoEncoders [1]: trained to minimize reconstruction error between corrupted input and clean output. We teach them how to remove 'noise' from other aberrations.

Our calibration architecture: parallel DAE + MLP

DAE remove the spurious features to help the MLP do their task. But how do we use the data? Two options: reconstructed: train MLP on **clean** images $\mathbf{r} = g(f(\tilde{\mathbf{x}}))$; vs. encoded: train MLP on encoded data $\mathbf{h} = f(\tilde{\mathbf{x}})$.

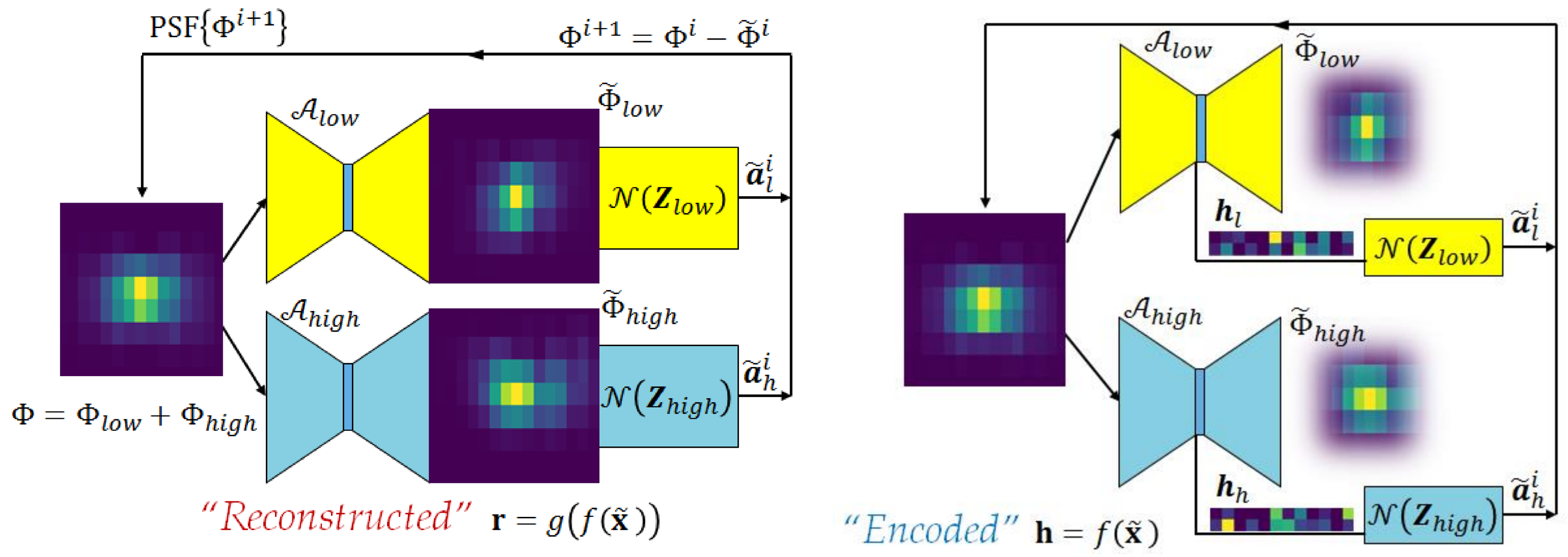

Figure 3 – Two architectures, depending on how we use the DAE data to train the MLP.

Clean is not enough. Taking denoising one step further

Encoded: denoising is **not** the goal; it is a training criterion for learning to extract useful features to guide the learning of other tasks **[2]**.

Learning to denoise = learning to represent the aberrations

Training MLP on encoded outperforms reconstructed approach (**Fig. 4**).
Advantages: faster training (encoded data has lower dimensionality), better performance "robust" features learned by the AE help calibration.

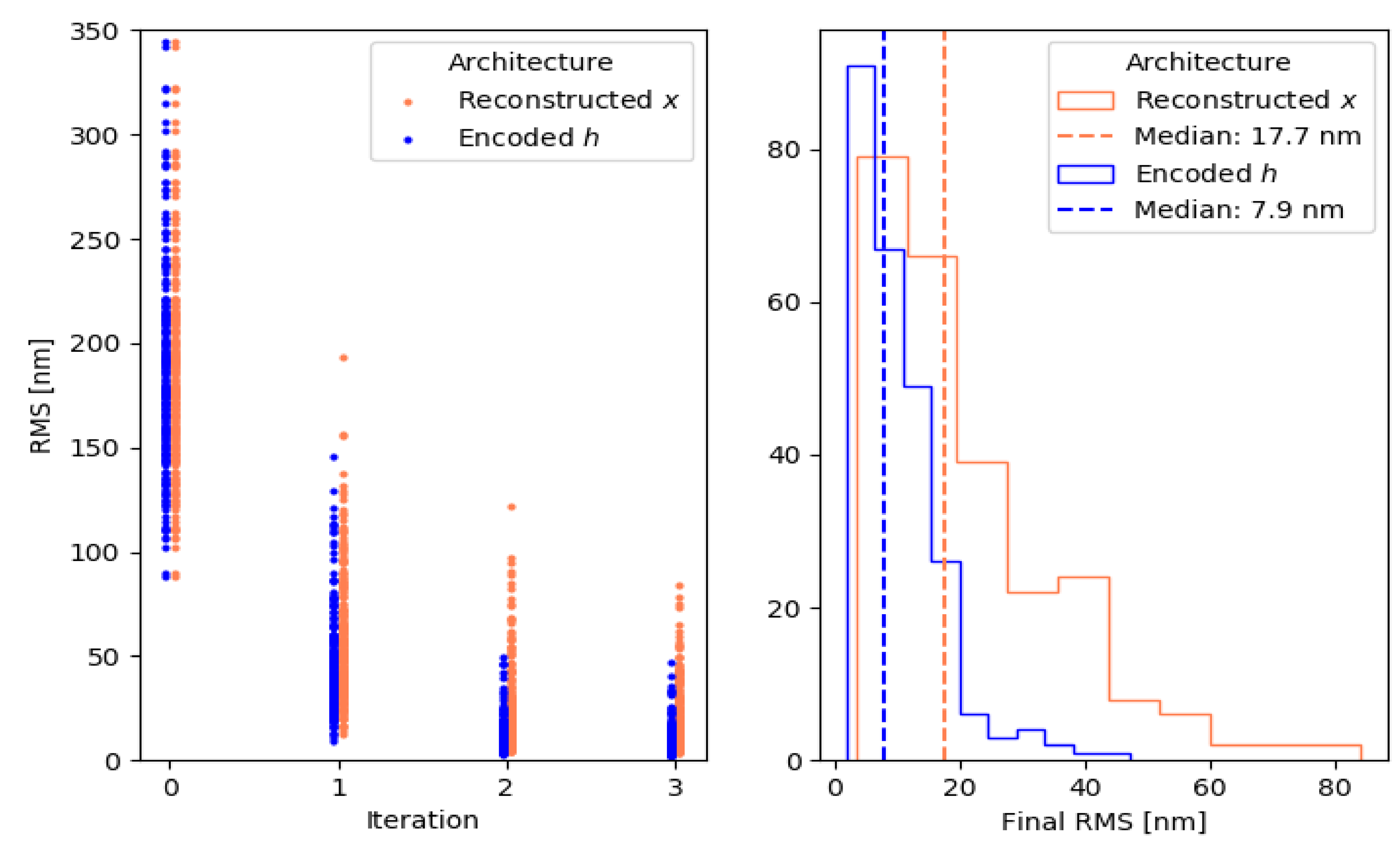

Figure 4 – Calibration results for the two architectures. MLP trained on encoded data leads to lower residual errors than when trained on reconstructed PSF images.

What is the autoencoder really *learning*?

For $\mathcal{A}_{high}$, $\mathbf{h}$ doesn't depend on the removed features. It learns to encode the information needed to recover $\mathbf{\Phi}_{high}$ and disregards $\mathbf{\Phi}_{low}$ (**Fig. 5**)

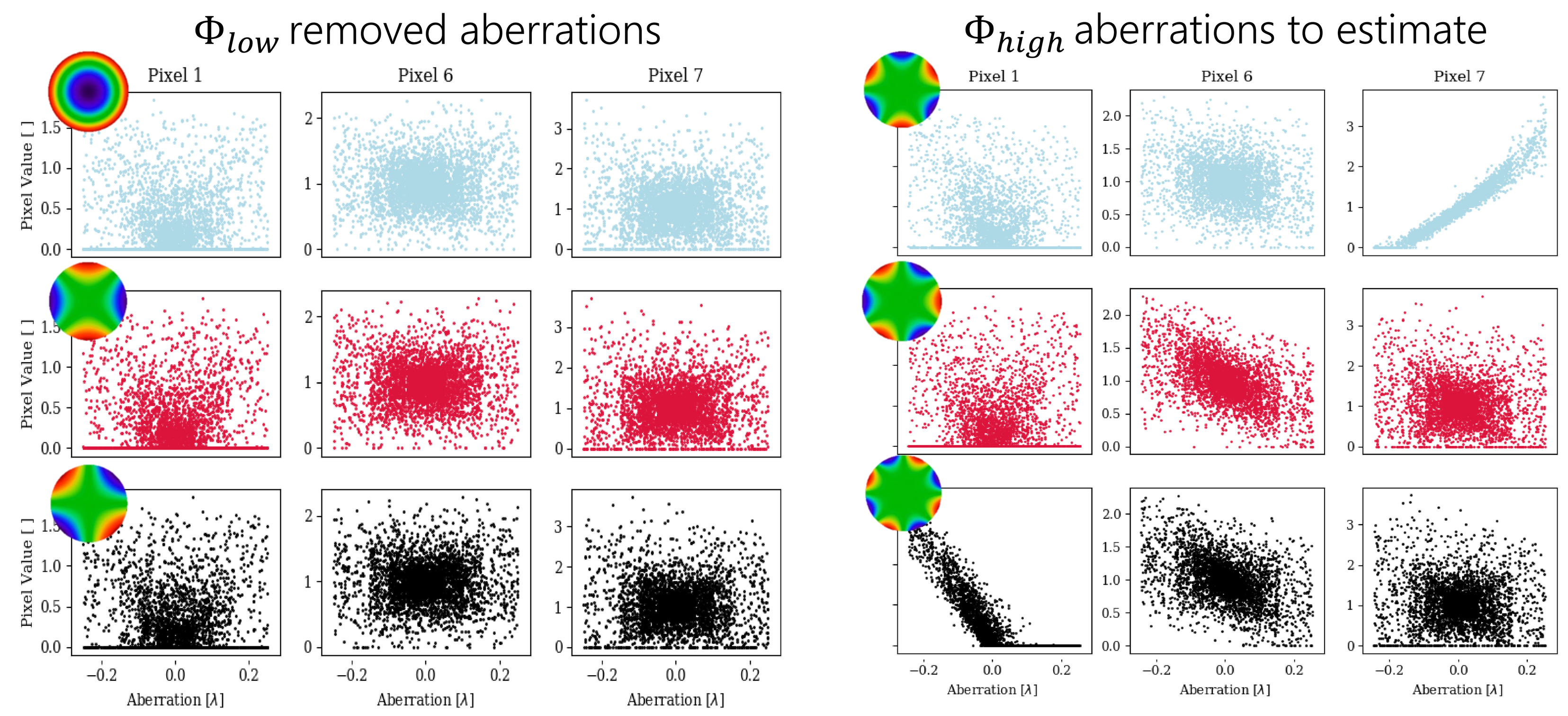

Figure 5 – $\mathcal{A}_{high}$(trained to remove Φ_{low}). The encoded features as a function of aberration intensity.

PCA analysis of encoded features (**Fig. 6**) reveals $\mathcal{A}_{high}$ learns a new data representation of the aberrations $\mathbf{\Phi}_{high}$. **Advantages**: sparse, robust against feature contamination, leads to better learning of MLP.

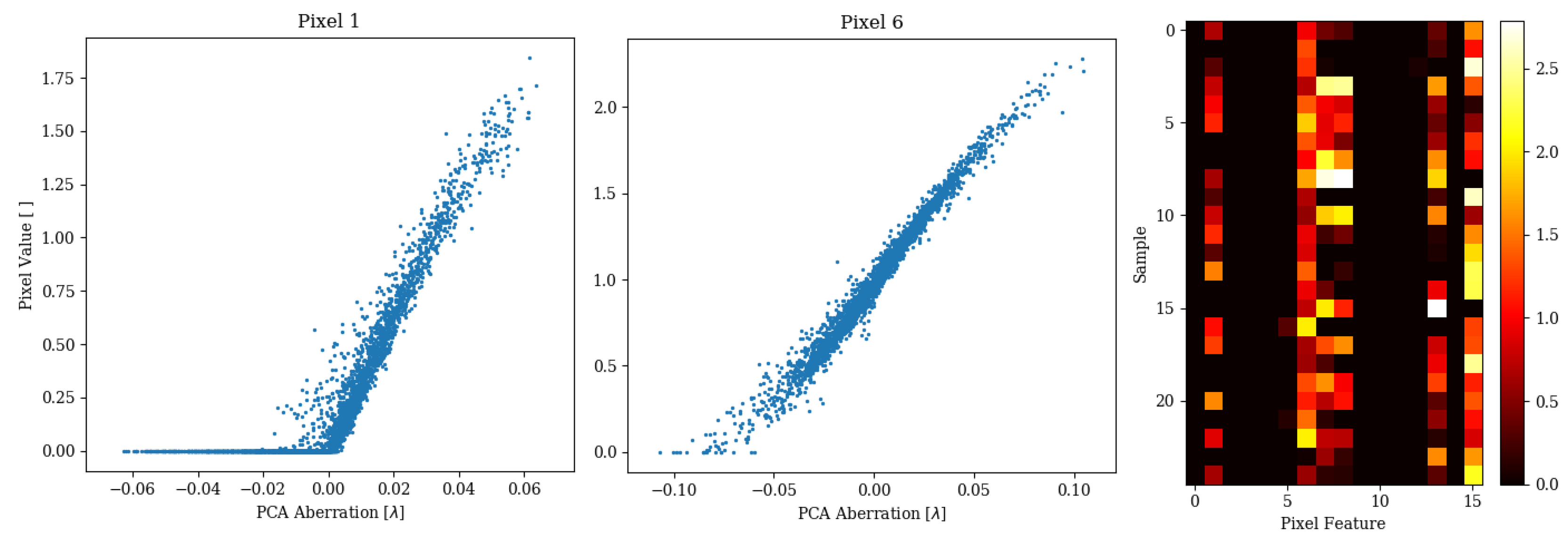

Figure 6 – Encoded data projected onto PCA $\mathbf{\Phi}_{high}$ aberration space

Conclusions

- **Novel approach** to NCPA calibration via ML
- Use of **Autoencoders** to mitigate **feature contamination**
- Denoising criterion guides **learning of new data representation**
- Using that data representation leads to **better performance** on calibration task.

References

[1] Vincent et al. 2008 - Extracting and Composing Robust Features with Denoising Autoencoders
[2] Vincent et al. 2010 – Staked Denoising Autoencoders. Learning useful representations...