

Off-line and on-line modal optimization of the MMT-AO system

B. Girardet^a, F. Wildi^{a,b} and G. Brusa^b

^aWest Switzerland University of Applied Science, EIVD, 1400 Yverdon, Switzerland;

^bSteward Observatory/Univ. of Arizona, 933 Cherry Av., Tucson AZ, 85721, USA

ABSTRACT

The MMT adaptive optics system has been in operation for over 2 years.^{1,2} Besides being a technological demonstration, it has achieved remarkable success.^{3,4} However, the system is presently limited by a few factors, one of which is the lack of an optimised controller. In this paper, we review the optimised modal integrator as introduced in Ref⁵ and evaluate its potential for MMT-AO. We find that it can indeed increase the system sensitivity by one magnitude but that a careful analysis of wave front sensor data must be performed to remove artefacts that can severely bias the outcome of the optimization.

Keywords: adaptive optics, modal control, MMT, optimal control

1. INTRODUCTION

Scarcity of bright stellar sources leads to the need of using AO with faint guide stars. In this photon-starved regime the performance of the AO system degrades rapidly due to the propagation of the photon noise to the wavefront error tracking loop. As the signal to noise ration (SNR) degrades on the sensor, the AO controller has to use information that is more and more noisy until the effect of the noise dominates the useful wavefront correction at which point the AO system starts introducing more aberrations than it corrects. Figure 1 shows typical Strehl ratio achieved versus guide star magnitude.

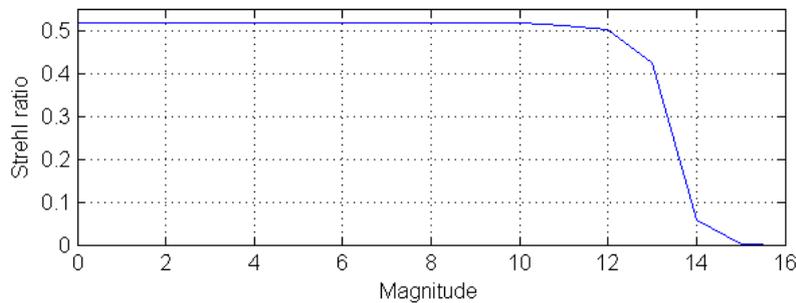


Figure 1. Strehl ratio v.s. guide star magnitude without modal optimization.

We have implemented the solution proposed by Gendron et al. 1994. This is a solution to delay the collapse of the performance in order to use AO with faint guide stars. The idea is to tune the bandwidth of the controller in order not to propagate more noise than signal through the system. This optimization is performed separately for each mode controlled by the AO. It is a modal optimization.

Further author information: (Send correspondence to B.G.)

B.G.: E-mail: benjamin.girardet@alumni-heig-vd.ch

2. MODAL OPTIMIZATION

2.1. Framework

We present here a summary of the modal optimization method, developed in Gendron et al. 1994, that we have chosen to implement. This method has been implemented in the Adonis system (ESO 3.6m telescope, Chile) in 1993, on the PUEO system (Cfht, Hawaii) in 1996, and recently in the NAOS system (ESO, VLT).

Let us consider the block diagram of a traditional AO system on Fig. 2. The block of the forward branch embeds the controller (integrator + delay) and the deformable mirror (DM). The feed-back branch is made of the wavefront sensor (WFS) camera and read-out electronics. The goal of the modal optimization is to tune the controller parameters to minimize the residual wavefront phase error, given certain observation conditions. In particular, we are going to derive the best integrator gain g_i for each mirror mode i . To be able to optimize the behaviour of each mode separately, we make the assumption that the modal base is orthogonal.

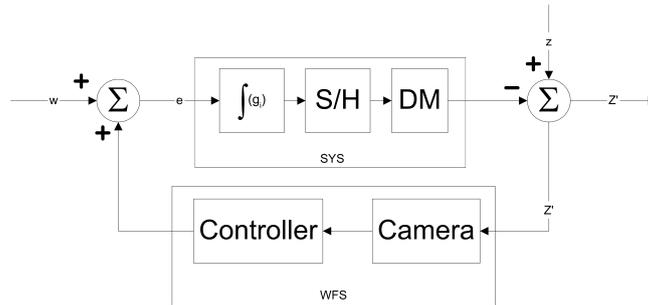


Figure 2. Traditional AO block diagram.

Considering each mode separately, we write the modal open-loop transfer function:

$$h_{ol} = h_{wfs} \cdot h_{sys} \quad (1)$$

And the corresponding (modal) correction transfer function and error rejection transfer function are respectively:

$$h_{cor} = \frac{1}{1 + h_{ol} \cdot g_i} \quad (2)$$

$$h_n = \frac{h_{sys} \cdot g_i}{1 + h_{ol} \cdot g_i} \quad (3)$$

Of course, the h_{cor} and h_n are functions of the modal gain g_i . Assuming a linear plant, we can write:

$$\langle z_i^2 \rangle = \int H_{cor}(f, g_i) \cdot \|z_i(f)\|^2 df + \int H_n(f, g_i) \cdot \|m_i(f)\|^2 df \quad (4)$$

Where H_{cor} and H_n are the squared amplitudes of the corresponding transfer functions h_{cor} and h_n . $\|z_i(f)\|^2$ and $\|m_i(f)\|^2$ are respectively the PSD of the uncorrected modal amplitude and the modal noise PSD. Note that the aliasing effect due to the spatial sampling of the WFS is not taken into account.

We are seeking to minimize the modal variance independently for each mirror mode. To do this, we must compute each term of Eq. (4). H_{cor} and H_n depend on the known dynamic behaviour of the AO system building blocks. Their model is assumed to be known. $z_i(f)$ and $m_i(f)$ are the modal spectra of the uncorrected mode i and the modal noise, respectively. These variables are not directly accessible because this only thing that can be measured is the noisy modal amplitude $\hat{z}_i(f)$. We know however that $\|\hat{z}_i(f)\|^2 = \|z_i(f)\|^2 + \|m_i(f)\|^2$, because the noise is decorrelated from the signal.

We use an approximation of Eq. (4) 1st suggested by E. Gendron,⁵ using only known quantities: The hypothesis is that the noise and signal are separable in the frequency domain. At low frequencies the signal (modal amplitude) dominates the noise while the opposite is true at high frequencies. Using the white noise hypothesis, one can write:

$$\langle z_i^2 \rangle = \int H_{cor}(f, g_i) \cdot \left(\|\widehat{z}_i(f)\|^2 - p(i) \cdot \sigma_n^2 \right) df + p(i) \cdot \sigma_n^2 \cdot \int H_n(f, g_i) df \quad (5)$$

Where $p(i)$ is the noise propagation coefficient between the noise measured at the WFS and the modal noise at reconstructor output. σ_n^2 corresponds to the phase error variance measured from the WFS slopes (the MMT uses a Shack-Hartmann WFS).

The propagation coefficients $p(i)$ depend on the reconstructor matrix. They express the way the slope noise is translated into modal noise, once the slope vector is multiplied by the reconstructor matrix D' . Their value is equal to the diagonal elements of $(D^t \cdot D)^{-1}$:

$$p(i) = (D^t \cdot D)_{ii}^{-1} \quad (6)$$

Where D is the interaction matrix, going from the slope space to the modal space. σ_n is the rms measurement phase noise in [rad] over each sub-aperture. The value σ_n used is the average of the slope noise over all sub-apertures. One can also measure the factor $p(i) \cdot \sigma_n^2$ by evaluating the modal noise after reconstruction, removing the need to compute the propagation coefficients.

2.2. Noise estimation by auto-correlation

There are several ways of computing $p(i) \cdot \sigma_n^2$. Each one requires extracting the noise from an unknown signal.

1. The noise variance can be measured on the slope signals. σ_n is the rms measurement phase noise. After computation of the propagation coefficients, the modal noise is known.
2. One can measure the term $p(i) \cdot \sigma_n^2$ from the modal amplitudes in the mirror space.

Two ways to estimate the noise have been used: auto-correlation and averaging the high frequency part of the spectrum. In both cases, the result is better when the SNR is lower. The noise must be estimated on the open-loop signal. For the autocorrelation method, we assume that the noise is decorrelated from the signal and from itself. The autocorrelation of the signal is determined by the modal time-frequency spectrum via the Wiener-Kinchin theorem. One can measure the noise of the autocorrelation function by taking the quadrature difference between the auto-correlation computed at $t = 0$ and the value of the auto-correlation at $t = 0$ interpolated from the samples at $t < 0$ and $t > 0$. The second way of measuring the noise variance is the integration of the PSD. The method work in an ad hoc manner, provided the PSD levels off clearly beyond a certain frequency. Experience has shown that satisfactory estimations can be made in our case from the upper 100 frequency samples (245-275 [Hz]).

Other filters can also be used for this estimation like auto-regressive filters (see Ref⁶).

2.3. Implementation

The only signal required for the optimization is the atmospheric turbulence. We are going to reconstruct the input wavefront from the slope signal measured by the WFS. As mentioned already, this measurement must be performed in open AO loop. Actually, on the MMT-AO system, it is not possible to work in these conditions because the WFS will saturate for most of the time. Therefore, on the real system, we measure the residual slopes $\vec{s}(t)$ in closed-loop, we translate this error into modal error using the reconstructor and add the current mirror position to have the best estimate of the modal amplitude. The knowledge of the DM transfer function $H_{DM}(s)$ relieves us from the need to know the DM position. It can be derived from the convolution of $\vec{s}(t)$ and $h_{DM}(t)$. Because of the time-invariance of $h_{DM}(t)$, we do not consider that inferring the mirror position from the time-serie of its inputs rather than measuring its position will add an error term to our error budget.

2.3.1. The slope signal

When working in open loop the slope signal is available directly. When in closed-loop the slope signal is estimated from the residual slopes and the modal position, using the influence matrix. Currently the MMT-AO system has 108 active sub-apertures and the slope signal is a matrix of 216 lines and k columns, k being the number of time samples.

From the slope signal, we derive the σ_n^2 term of Eq. (5), which is the average variance per sub-aperture. Using the noise propagation coefficients computed using Eq. (6), we can now compute the terms $p(i) \cdot \sigma_n^2$.

Figure 3 shows the propagations coefficients of the theoretical reconstructor as a function of the mode number. It strikingly shows how certain modes are less visible by the WFS than others for equivalent modal power. This is a well-known fact for SH-based systems. In particular, the SH being sensitive to the local slopes, it responds better the modes of higher spatial frequencies. Generally, the noise propagates well on the first modes, but since the atmosphere has more power on those modes their SNR will still be better than higher modes where noise propagation is weaker. Notice there are peaks of noise propagation on certain high order modes. We expect to have to reduce the feedback on those noisy modes.

Feeding the slope signal into the reconstructor, one obtains the reconstructed wavefront., expressed as a modal temporal evolution. A Fourier transform yields the term $\|\hat{z}_i(f)\|^2$.

2.3.2. Finding the minimum residual variance

The optimal controller bandwidth is the one that will minimise the residual modal phase variance. Using Eq. (5), we search a scalar gain g_i for which $\langle z_i^2 \rangle$ is minimal. The lower bound of $g(i)$ is 0 and its larger bound is the value that correspond to the stability limit of the system which we define by the phase margin of 45 degrees. His corresponds for our model to a gain $g(i)$ of 0.47. We have implemented a dichotomic search a the minimum using a relative gain change $< 10\%$ as an output criterion. The Fig. 4 below shows the optimal gains obtained on a 11th magnitude star. The gains are clipped to the system stability limit we have set.

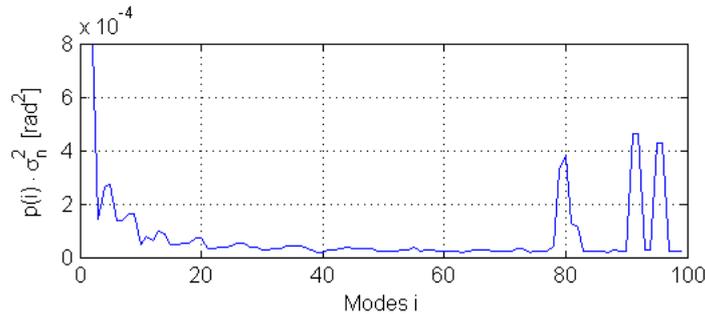


Figure 3. Propagation coefficients of the standard MMT-AO reconstructor.

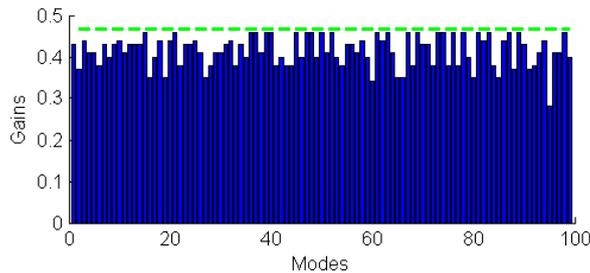


Figure 4. Modal gains obtained on a 12th mag guide star. The gains are clipped to the system stability limit.

3. SIMULATION

We are using a station-temporal simulation of our AO system which we use to quantify the performance of the modal optimization against a case of "flat gain" controller where all modal gains are set to the same value. The main parameters of the simulation are defined to closely match the MMT-AO:

- Atmosphere: Single frozen turbulence layer at 25m/s. Outer scale 100m, inner scale 9cm (limited to prevent spatial aliasing by the simulation). The atmosphere is the same for all simulation to remove statistical effects.
- WFS: 12x12 Shack-Hartmann, using quad-cells and 8 e- read-out noise.
- Modal base: first 100 Karhunen-Loeve modes.
- Controller: Pure integrator. In the flat gain case, it correspond to the stability limit (45deg phase margin) and is equal to 0.47. Sampling time is 1.8 ms.
- The sky brightness is 18 mag/(square ").
- Simulated duration: 3s per run.

The performance criterion is the Strehl ratio. Figure 5 shows the Strehl ratio versus guide star magnitude in the flat gain case and with modal optimization. As soon as the performance starts to be limited by photon noise, modal optimization can improve performance. Indeed, the optimization criterion will force the noisy modes (those with a big $p(i)$) to have a low noise integral, which means a low gain in the higher part of the spectrum. In reality, the gain of all modes will be reduced because they become all noisy. Modal optimization allows to find the most suitable gain for each of them.

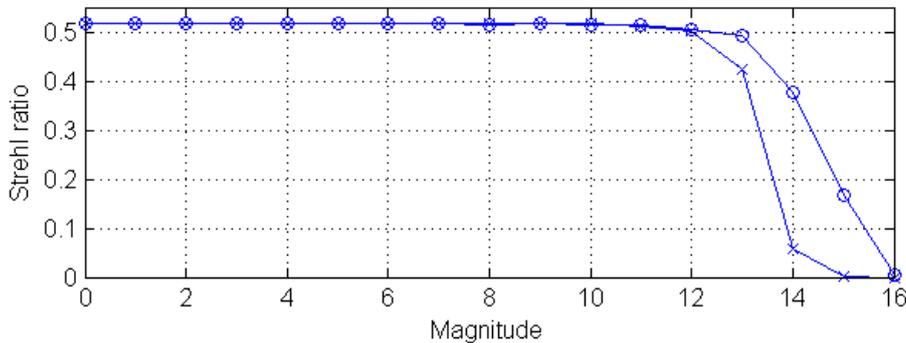


Figure 5. Strehl ration v.s. magnitude with (o) and without (x) modal optimization. The flat gain is set to 0.47, which corresponds to a 45 degree phase margin.

4. MEASUREMENTS ON THE SKY

In the preceding section the system has been optimized using open loop simulations. The actual MMT system saturates under these conditions and the optimization algorithms must be extended to be able to use data from the (un-optimized) AO working in closed-loop. This is a first step towards on-line implementation of this optimization.

4.1. Atmospheric Input turbulence

We have used a set of data measured on the sky during MMT-AO runs. We have computed the atmospheric turbulence from the closed-loop data as described in Sect. 2.3. We have chosen this method for the convenience of not having to synchronise 2 data sets (slope error and DM position).

The slope vector \vec{s} measured by the WFS yields the instantaneous modal error. From the AO loop model we know the transfer function between this error and the mirror position and therefore we can derive the position from the error. Since we have an integrator in the loop, the position is only known to a constant but this does not matter since we assume 0-average modes. Even with the flat gain $g(i)$ known, the model has a parameter that must be estimated and this is the gain of the wavefront sensor. For a Shack-Hartmann working in quad-cell mode, this gain depends on the seeing.^{7,8} A real-time estimation of this gain will be designed for the MMT-AO.

4.2. Sample duration

We have also estimated the number of time sampled required to have an effective optimization. This number will influence the output of the optimization in several ways: A longer sample will reduce the statistical variation of the modal spectrum, will increase the spectral resolution and will also reduce the noise on the auto-correlation function used to estimate the noise. Rather than trying to derive an analytical expression for the minimal number of samples, we have observed the evolution of the Strehl ratio with respect to the sample duration used for modal optimization (see Fig. 6).

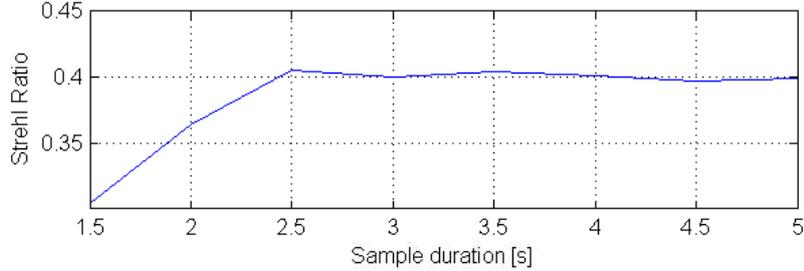


Figure 6. This figure shows the Strehl ratio obtained after a modal optimization performed from an open-loop time sequence of the duration indicated on the abscissa. From this figure, it is apparent that an observation time beyond 2.5s is not required in our case. (we anticipate that lower wind speeds will require longer integration times.).

5. OPTIMAL GAINS DERIVED FROM MMT DATA SETS

We investigate the modal optimization from one data set acquired in closed-loop on a 5.4 magnitude guide star. At this light level, the signal is essentially noiseless. Knowing the noise propagation coefficient, we are progressively adding more Poisson noise to it and observe the evolution of optimum gains. Figure 7 illustrate the effect of adding more noise. The solid curve is turbulence spectrum measured on the 5.4 mag guide-star, while the dotted line is the same spectrum after noise has been added on the slopes and propagated to the mode. Note that the SNR is not calibrated in equivalent guide star magnitude.

5.1. Results

5.1.1. Important parameters

Figure 8 shows the evolution of optimized modal gains in noisier and noisier conditions.

The optimal loop gain influenced mostly by 3 parameters :

- The modal temporal spectrum
- The transfer function overshoot

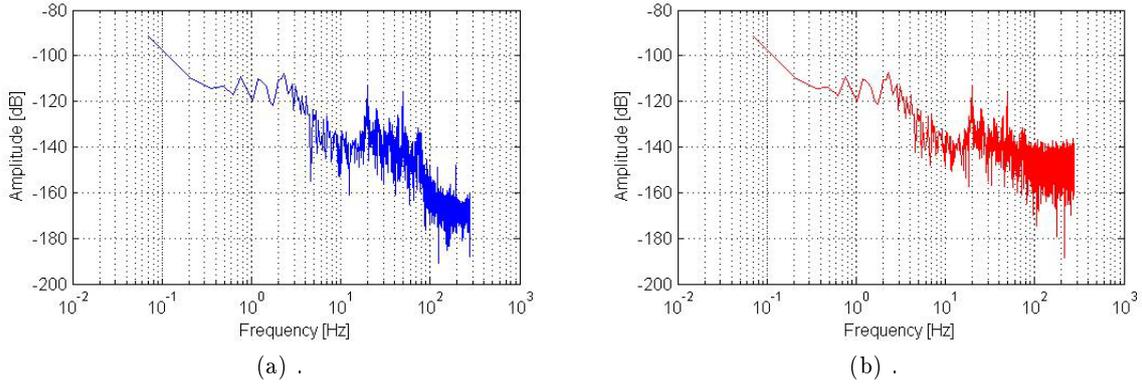


Figure 7. Effect of adding white noise Poisson noise on the WFS signal. Mode #1 (quasi-tip) of a 56-modes MMT modal base, based on mirror eigen modes.

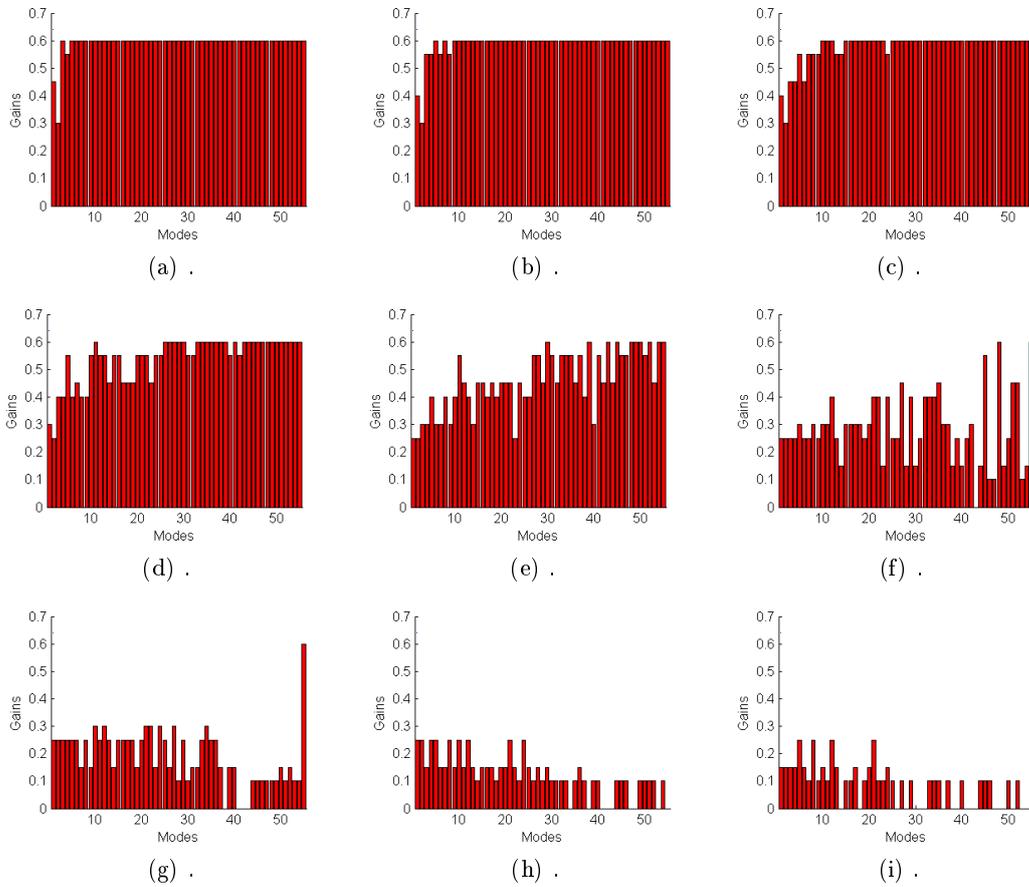


Figure 8. Evolution of optimal gain with increasing (by artificial noise added) GS magnitude. (a) using the original 5.4 mag. GS and going up to (i) approximately 9 mag GS. Here, the stability limit has moved up to 0.6 because the temporal model is a little bit different from the real data from the MMT.

- The noise level

Let us analyze quantitatively these parameters to understand the modal gains shown in Fig. 8. First let us note that to first order the DM dynamics is mode-independent, and only the temporal spectrum and the noise level are different between the modes.

In the case of a bright GS, the noise becomes negligible and Eq. (5) simplifies to:

$$\langle z_l^2 \rangle = \int H_{cor}(f, g_i) \cdot \|\hat{z}_i(f)\|^2 df \quad (7)$$

Where modes differ only by their modal spectra. When noise is added, the right part of Eq. (5) starts weighting in the total modal variance. We are using the MMT-AO 56 modes base where modes are ordered by decreasing singular value as shown in Fig. 9. Modes #1 and #2 are almost equal to the Zernike modes *tip* and *tilt*. The noise propagation coefficients are therefore monotonically decreasing too. From this perspective only, one would expect the modal gain to increase with the mode number.

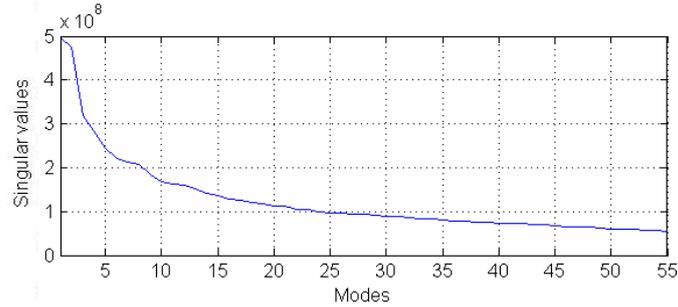


Figure 9. In the MMT 56 modes case, the modes are ordered by decreasing singular value

5.1.2. Effect of Spectral power distribution

Figure 10 shows the frequency distribution of mode #1 (quasi-tip), where 2 peaks are clearly present:

1. One peak at 20 Hz due to mechanical vibrations on the secondary mirror support hub.⁷ These rotary oscillations of the hub produce essentially tip and tilt on the WFS. One solution being implemented to remove this effect is to correct it in open loop, using a set of accelerometers as sensors and adding their tip-tilt measurement to the AO controller position output.
2. One peak at 49 Hz which origin is not yet identified. Its bandwidth is quite narrow (<0.5 Hz) and it is present on modes (#1, #2 and #4). Some system considerations lead to thinking that it is not linked to mechanical vibrations and might of electrical origin. We are going to consider it as "added" noise not linked to the atmospheric turbulence.

Other modes also contain spurious peaks of power. Some around 30 Hz. Since they come and go, it is difficult to identify their origin, And we do not have enough data sets to make statistically significant conclusions. However, we believe that these peaks do not correspond to atmospheric turbulence. We must nonetheless determine their effect on the final image quality. For instance, the 20 Hz certainly degrades the image (we were able to take it out with stack and add) while the 49 Hz likely degrades only the measurement.

Figure 11 compares the frequency distribution of modes #3 and #30 in noise conditions similar to Fig. 8(d). In the 20-30 Hz band, mode #30 has more power than mode #3. The band is located before the overshoot of the rejection transfer function, therefore increasing the gain is going to reduce the power in this band. However, this mode also has more power in the spectral band where the rejection transfer function overshoots and this is going to drive the optimal gain lower. Our algorithm is going to look for the optimal considering the complete spectrum. (In this particular case the gain is larger for mode #30 than for mode #3).

As the Fig. 8, Fig. 12 shows the evolution of optimized modal gains in noisier and noisier conditions. This second data set use an original 8.75 mag GS.

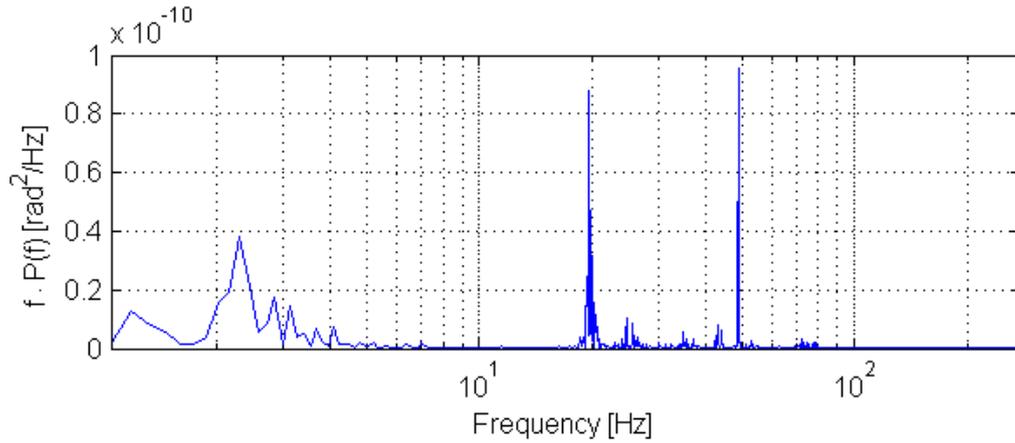


Figure 10. Frequency distribution of mode #1 with the 5.4 mag GS data set.

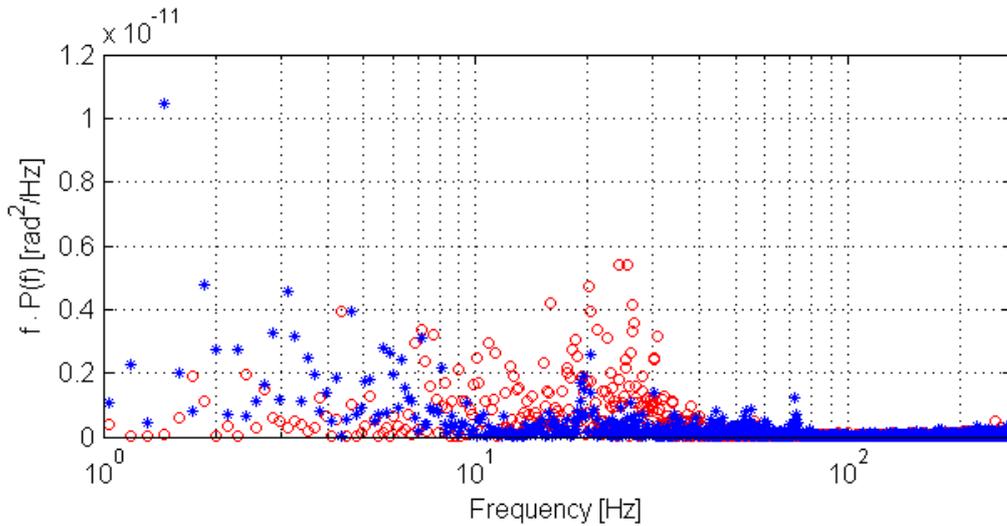


Figure 11. Frequency distribution of the modes #3 (blue stars) and #30 (red circles) with the 5.4 mag GS data set.

5.1.3. Behavior of the last mode

Figures 8 and 12 show that the highest mode has a particular behavior: The optimal gain for this mode tends to stay substantially higher than the one of the other high order modes until it drops to 0 as the SNR collapses. Figure 13 compares the frequency distribution of mode #55 and that of mode #54 for noise conditions corresponding to Fig. 8(g) (Fig. 13(a)) and 8(h) (Fig. 13(b)) respectively. Notice that the vertical scale is different between figure 14a and 14b. In atmospheric terms mode 55 has more power in the 8-30 Hz band but when the GS is dim enough, this effect is masked by the noise.

5.1.4. Behavior of the first mode

In the second data set (Fig. 12) the mode #1 also exhibits a particular behavior. In our modal base modes 1 and 2 are very close to the Zernike tip and tilt modes, therefore their behavior is very close. In this set however, their spectra are strikingly different as shown in Fig. 14. where the frequency distribution of these 2 modes are compared in conditions corresponding the ones in Fig. 12(g). Mode number one shows the 20 Hz peak already mentioned corresponding to the secondary hub vibration. During this acquisition, the vibration was affecting the tip axis but not the tilt axis. Since the oscillation frequency is within the control bandwidth, the optimization algorithm maximizes the gain to minimize this component.

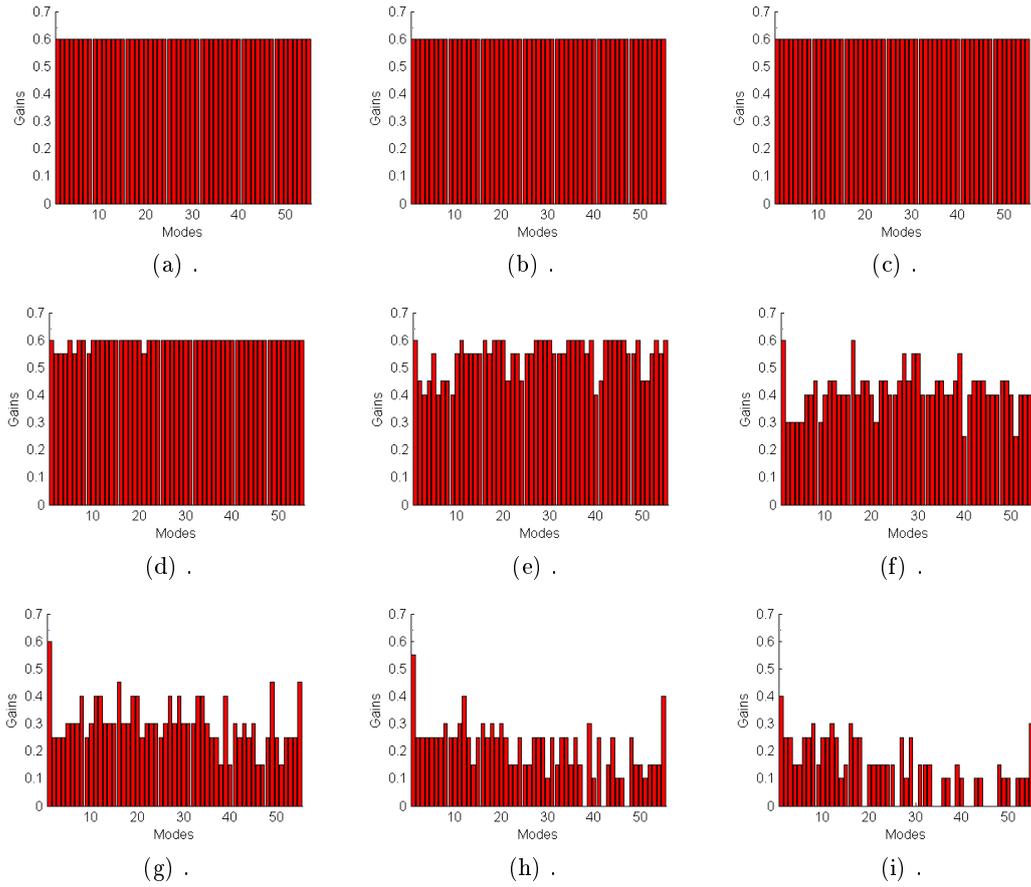


Figure 12. Evolution of optimal gain with increasing (by artificial noise added) GS magnitude. (a) using the original 8.75 mag. GS and going up to (i) approximatively 12.5 mag GS.

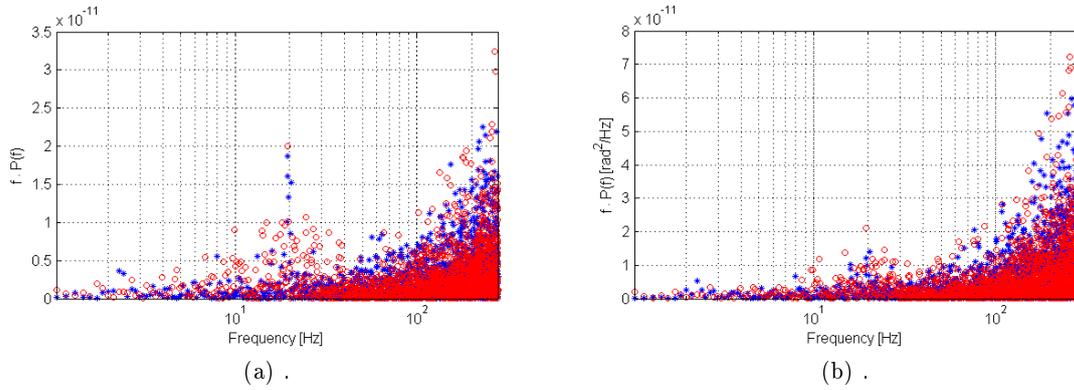


Figure 13. This is the comparison between mode #54 (blue stars) and #55 (red circles). Sub-fig (a) in the same conditions than Fig. 8(g) and (b) the same than Fig. 8(h).

5.1.5. Pre-processing

After examining the pattern of optimal gains, and the origin of certain specific patterns we conclude that pre-processing the data is required. Ideally, this processing should be embedded within the modal optimization

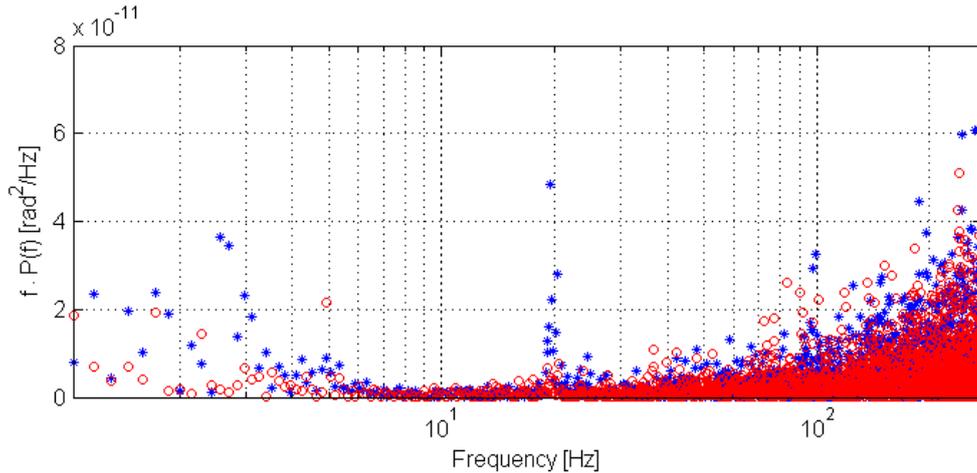


Figure 14. Frequency distribution of mode quasi-tip #1 (blue stars) and quasi-tilt #2 (red circles).

environnement. It cannot be based only on analytical reasoning but must include the diffuse but deep knowledge of the system we have used to support the interpretation we have made in this chapter.

Other authors have made similar conclusions. Notably Gendron,⁹ talks about the need of "incident detection" and of pre-processing the data.

For the optimization to work effectively, one must identify possible perturbations, track their origin and evaluate their consequences. If a perturbation does affect the system and degrades the science image, it must be included in the optimization of the residual phase variance. If on the other hand it only affects the measurement it must be eliminated before it can influence the output of the optimization.

6. CONCLUSION

The our case, modal optimization can only been considered when using closed loop slope data because of the WFS non-linearities. This is not a limitation since using close-loop data is the desired method when on-line modal optimization is the ultimate target. The model of our system has been reviewed to best match the elements of the AO system. Using this model, we are able to derive the actual position of the mirror and the atmospheric turbulence. (In this we are also helped by the fact that the adaptive secondary technology uses local position control and provides an accurate position measurement of each actuator several times per WFS integration).

Since the optimization is only valid for the current statistics of the perturbation, the optimization must be performed and applied rapidly, before the atmospheric statistics change. Steps have been made to shorten the processing time, including determining shortest integration times and implementing a fast dichotomic algorithm.

We were expecting to see the optimization converge towards a gain v.s. mode number curve that was generally monotonic. This behavior is not generally verified using data taken on the sky. We were able to understand the pattern of modal gains obtained after optimization and were able to link some odd behaviors to measurement artifacts and others to telescope vibrations than must be corrected for even if they have nothing to do with atmospheric turbulence. Pre-processing of the WFS data to identify and possibly eliminate artifacts must be implemented as an integral part of the optimization.

Presently, in the MMT-AO case, the modal optimization is not routinely available. The output of the study presented in this paper is a toolbox tailored to simplify as much as possible the optimization. It generates a vector of optimal gains off a regular laptop PC running MATLAB, using the time serie of slope vectors as an input. However, the results are encouraging enough and work in being done to implement the algorithms in the real-time AO software.

ACKNOWLEDGMENTS

B. Girardet wishes to thank the West Switzerland University of Applied Science for funding this study and Steward Observatory for hosting him within the MMT-AO group.

REFERENCES

1. G. Brusa, A. Riccardi, F. P. Wildi, M. Lloyd-Hart, H. M. Martin, R. Allen, D. L. Fisher, D. L. Miller, R. Biasi, D. Gallieni, and F. Zocchi, "Mmt adaptive secondary: first ao closed-loop results," *Astronomical Adaptive Optics Systems and Applications* **5169**(1), pp. 26–36, SPIE, 2003.
2. G. Brusa, A. Riccardi, P. Salinari, F. P. Wildi, M. Lloyd-Hart, H. M. Martin, R. Allen, D. Fisher, D. L. Miller, R. Biasi, D. Gallieni, and F. Zocchi, "Mmt adaptive secondary: performance evaluation and field testing," *Adaptive Optical System Technologies II* **4839**(1), pp. 691–702, SPIE, 2003.
3. L. M. Close, F. Wildi, M. Lloyd-Hart, G. Brusa, D. Fisher, D. Miller, A. Riccardi, P. Salinari, D. W. McCarthy, R. Angel, R. Allen, H. Martin, R. G. Sosa, M. Montoya, M. Rademacher, M. Rascon, D. Curley, N. Siegler, and W. J. Duschl, "High-resolution images of orbital motion in the trapezium cluster: First scientific results from the multiple mirror telescope deformable secondary mirror adaptive optics system," *ApJ* **599**, p. 537–547, 2004.
4. B. Biller, L. Close, A. Li, J. Biegging, W. Hoffmann, P. Hinz, D. Miller, G. Brusa, M. Lloyd-Hart, F. Wildi, D. Potter, and B. Oppenheimer, "High-resolution mid-infrared imaging of the asymptotic giant branch star ν bootis with the steward observatory adaptive optics system," *ApJ* **620**, p. 450–458, 2005.
5. E. Gendron and P. Léna, "Astronomical adaptive optics i. modal control optimization," *Astronomy and Astrophysics* **291**, pp. 337–347, 1994.
6. C. Dessenne, *Commande modale et prédictive en optique adaptative*. PhD thesis, ONERA, 1999.
7. G. Brusa, D. L. Miller, M. A. Kenworthy, D. L. Fisher, and A. Riccardi, "Mmt-ao: two years of operation with the first adaptive secondary," *Advancements in Adaptive Optics* **5490**(1), pp. 23–33, SPIE, 2004.
8. L. K. Saddlemyer, G. Herriot, J.-P. Véran, M. Smith, and J. Dunn, "Innovations within the altair real-time wave-front reconstructor," *Advancements in Adaptive Optics* **5490**(3), pp. 1384–1392, SPIE, 2004.
9. E. Gendron and P. Léna, "Astronomical adaptive optics ii. experimental results of an optimized modal control," *Astronomy and Astrophysics. Suppl. Ser.* **111**, pp. 153–167, 1995.