Performance limits of adaptive-optics/high-contrast imagers with pyramid wavefront sensors

Carlos M. Correia, Olivier Fauvarque, Charlotte Bond, Vincent Chambouleyron, Jean-François Sauvage, Thierry Fusco

Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **495,** 4380–4391 (2020) doi:10.1093/mnras/staa843

Performance limits of adaptive-optics/high-contrast imagers with pyramid wavefront sensors

Carlos M. Correia,[1,2]★ Olivier Fauvarque,[2] Charlotte Z. Bond,[1] Vincent Chambouleyron,[2] Jean-François Sauvage[2,3] and Thierry Fusco[2,3]

[1]*WM Keck Observatory, 65-1120 Mamalahoa Hwy, Kamuela, HI 96743, USA*
[2]*Aix Marseille Université, CNRS, CNES, LAM, Marseille, France*
[3]*ONERA, 29 avenue de la division Leclerc, 92322 Chatillon, France*

Accepted 2020 March 14. Received 2020 March 14; in original form 2020 January 10

ABSTRACT

Advanced adaptive-optics (AO) systems will likely utilize pyramid wavefront sensors (PWFSs) over the traditional Shack–Hartmann sensor in the quest for increased sensitivity, peak performance and ultimate contrast. Here, we explain and quantify the PWFS theoretical limits as a means to highlight its properties and applications. We explore forward models for the PWFS in the spatial-frequency domain: these prove useful because (i) they emanate directly from physical-optics (Fourier) diffraction theory; (ii) they provide a straightforward path to meaningful error breakdowns; (iii) they allow for reconstruction algorithms with $O(n\ \log(n))$ complexity for large-scale systems; and (iv) they tie in seamlessly with decoupled (distributed) optimal predictive dynamic control for performance and contrast optimization. All these aspects are dealt with here. We focus on recent analytical PWFS developments and demonstrate the performance using both analytic and end-to-end simulations. We anchor our estimates on observed on-sky contrast on existing systems, and then show very good agreement between analytical and Monte Carlo performance estimates on AO systems featuring the PWFS. For a potential upgrade of existing high-contrast imagers on 10-m-class telescopes with visible or near-infrared PWFSs, we show, under median conditions at Paranal, a contrast improvement (limited by chromatic and scintillation effects) of 2×–5× when just replacing the wavefront sensor at large separations close to the AO control radius where aliasing dominates, and of factors in excess of 10× by coupling distributed control with the PWFS over most of the AO control region, from small separations starting with an inner working angle of typically 1–2 λ/D to the AO correction edge (here 20 λ/D).

Key words: atmospheric effects – instrumentation: adaptive optics – methods: analytical – techniques: high angular resolution.

1 THE QUEST FOR PERFORMANCE AND CONTRAST

The first generation of high-contrast imagers on 10-m-class telescopes has been in operation for the last 5 years or so, producing amongst others remarkable images of scattered light from discs in circumstellar environments and exquisite observations of solar system objects (Mawet et al. 2016; Guyon 2018; Macintosh et al. 2018; Mouillet et al. 2018; Xuan et al. 2018; Beuzit et al. 2019). However, the discovery of new planets has been quite disappointing, with very few confirmed detections.

There are of course no culprits to blame, but limiting contrast has been raised as a limitation to be lifted in order to populate the long-awaited list of new discoveries (Mawet et al. 2014; Cantalloube et al. 2019). In this respect, it has been recognized that the inner working angle (IWA) of coronagraphs should be decreased to as close as possible to λ/D in the attempt to observe close-in new planets. For such an endeavour, novel coronagraph concepts abound (Mawet et al. 2012; Guyon 2018; Snik et al. 2018).

On a par, adaptive-optics-related (AO-related) residuals can be further reduced in the hopes of improving the contrast across the AO correction band (typically up to a separation of few tens of λ/D, depending on the number of deformable mirror actuators across the pupil; Correia et al. 2017). As pointed out and quantified in Guyon (2005), the effects that limit the performance of wavefront correction are as follows:

(i) noise on the wavefront sensor (WFS) (photon – fundamental, read-out – technological), requiring subelectron noise detectors;

★ E-mail: carlos.correia@lam.fr

(ii) aliasing arising from the discrete nature of the WFS measurement, limited with custom optical filtering;

(iii) servo-lag error owing to the dynamic rejection of residuals in a negative feedback loop, calling for faster/cleverer algorithms;

(iv) actuator fitting, demanding higher-density deformable mirrors.

To these adds chromatic optical path length difference (OPD) and amplitude errors between the WFS wavelength $\lambda_{\rm WFS}$ and the imaging wavelength $\lambda_{\rm im}$ (Hardy 1998; Guyon 2005; Fusco et al. 2006), which we revisit and fully take into account in what follows.

In this paper, we provide AO-limited performance and limiting contrast estimates when a perfect coronagraph is employed. We show the expected improvement from the use of a pyramid WFS (PWFS) in both the near-infrared (NIR) and visible (VIS) wavelengths with a realistic 2D physical-optics model from Fauvarque et al. (2019), capable of mimicking effects such as modulation, partial AO correction causing point spread function (PSF) broadening with subsequent drop in sensitivity, and extended sources. In addition, we investigate the usefulness of predictive control through the application of Kalman filters and distributed control in the spatial-frequency domain (Correia et al. 2017).

Throughout the paper, we use models in the spatial-frequency domain (Fourier for short) for a number of good reasons, each addressed in a dedicated section:

(i) optical transfer functions are naturally described in the Fourier plane – Section 2;

(ii) statistically independent error terms can be readily evaluated from the residuals – Section 3;

(iii) wavefront reconstruction can be seamlessly performed using standard filter operations, giving rise to the use of matrix-free operations with $O(n\ \log(n))$ complexity algorithms for large-scale systems – Section 4;

(iv) decoupled (distributed) optimal controllers for performance and contrast optimization are easily synthesized – Section 5.

We assume that non-common path errors are properly corrected for and thus do not enter the AO-centric error budget developed here, whereas in practice they limit the PWFS linear range and must be addressed operationally.

2 OPTICAL MODELS OF THE PYRAMID WAVEFRONT SENSOR USING DIFFRACTION THEORY

The behaviour of the pyramid in the spatial domain has been extensively studied by Vérinaud (2004), Vérinaud et al. (2005), Chew, Clare & Lane (2006), Korkiakoski et al. (2007), LeDue et al. (2009), Quirós-Pacheco, Correia & Esposito (2009), Wang et al. (2010), Shatokhina et al. (2013) and Fauvarque et al. (2015, 2017), following the seminal work of Ragazzoni (1996), who in turn built on Linfoot (1948) Foucault knife-edge diffraction model.

Here we stick to the original four-facet pyramid concept, although generalizations to any number of facets exist, using a coherent or incoherent recombination of light past the pyramid optic (Fauvarque et al. 2015, 2017). The latter can be designed to optimize contrast at certain separations, but the lack of a general design compelled us, with some loss of generality, to consider only the original pyramid WFS concept.

The intensity pattern on each of the four re-imaged pupils at the detector plane $i_q(\boldsymbol{x}, t)$, $q \in \{1, \cdots, 4\}$, indexed by a bi-dimensional coordinate $\boldsymbol{x} = (x, y)$ and time t is conveniently formulated using Fourier masking:

$$i_q(\boldsymbol{x}', t) = \int_{t-T_s}^{t} \left| \mathcal{F}^{-1} \left\{ H_q(\boldsymbol{\kappa}) \mathcal{F} \left\{ A(\boldsymbol{r}) \mathrm{e}^{i(\psi(\boldsymbol{r}) + \theta(\mathbf{r},t))} \right\} \right\} \right|^2 \mathrm{d}t. \tag{1}$$

Here, $A\mathrm{e}^{i\psi}(\boldsymbol{r})$ is the electric field in the pupil $\mathcal{A}$ (for aperture) indexed by $\boldsymbol{r} = (r_x, r_y)$, where $A(\boldsymbol{r})$ is the amplitude and $\psi(\boldsymbol{r})$ is the phase, which is Fraunhofer-propagated to the focal plane using a 2D Fourier transform $\mathcal{F}$. In this equation, $\theta(\boldsymbol{r}, t)$ is an additional time-dependent *modulation* signal, introduced as a phase increment to the aberrated wavefront over the integration time T_s in the pupil-plane. A customarily used signal is a time-varying tilt that shifts the focal-plane electric field and makes it wander across the four pyramid facets. A guide probe other than a point-source would merely alter the modulation because each point of the object would then add to the phasor $\mathrm{e}^{\,i\,\theta(\boldsymbol{r},\,t)}$ and would be summed incoherently to the output intensity. For simplicity we preferred not to overload the equation with multiple integrals over spatial coordinates. Next in line, H_q is a masking function (or transparency mask) placed at the focal plane and indexed by $\boldsymbol{\kappa} = (\kappa_x, \kappa_y)$ for each qth quadrant of the form

$$H_q(\boldsymbol{\kappa}) = \mathcal{H}_{\pm\kappa_x} \mathcal{H}_{\pm\kappa_y} \mathrm{e}^{-i\boldsymbol{\alpha}_q(\pm\kappa_x)\cdot(\pm\kappa_y)}, \tag{2}$$

where $\mathcal{H}_{\pm\kappa_x}$ is the Heaviside function for either positive or negative spatial frequencies, and $\alpha_q \in \mathbb{R}$is a real-valued variable that sets the output angle of the re-imaged pupils with respect to the chief ray. In practice, on a computer, we replace the integral by a sum on temporally incoherent intensity patterns each for a tilt value (from the modulation from the object). The number of sums is calculated based on the sampling of the PSF at the WFS detector focal plane, although we could opt to replace this regular sample by irregular sampling, namely finer across the pyramid edges and coarser on top of the facets, with consistent results (Fauvarque 2017).

2.1 Impulse response of a PWFS

Traditionally, the PWFS signals are extracted from the four re-imaged pupils using a slope-like formulation that stems from the original Foucault knife-edge test. This formulation provides a notional first-derivative measurement of the wavefront,

$$\begin{aligned} \boldsymbol{s}_x &= g^x_{\rm opt} \frac{i_1 + i_2 - i_3 - i_4}{\sum_q i_q} - s^0_x \\ \boldsymbol{s}_y &= g^y_{\rm opt} \frac{i_1 - i_2 + i_3 - i_4}{\sum_q i_q} - s^0_y, \end{aligned} \tag{3}$$

with $g^{x,y}_{\rm opt}$ the optical gain (Bond et al. 2017; Esposito et al. 2015) and $s^0_{x,y}$ the null-phase reference measurement. Henceforth, this definition is referred to as the *slopes-map* model.

Special care must be paid to the denominator of (3), whether to normalize each value by the sum of the four corresponding intensities or by replacing $\sum_q i_q$ by a scalar value representing the total integrated flux, namely $\int_\Omega i_q \mathrm{d}\Omega$, with Ω the domain set by the valid pixels (Vérinaud 2004). The latter is considered a more robust option (Bond et al. 2016).

Here we focus on meaningful, yet practical, linear physical-optics forward model of the PWFS that can, under minimal simplifications, represent the bulk of its operation, yielding a convolution of the input

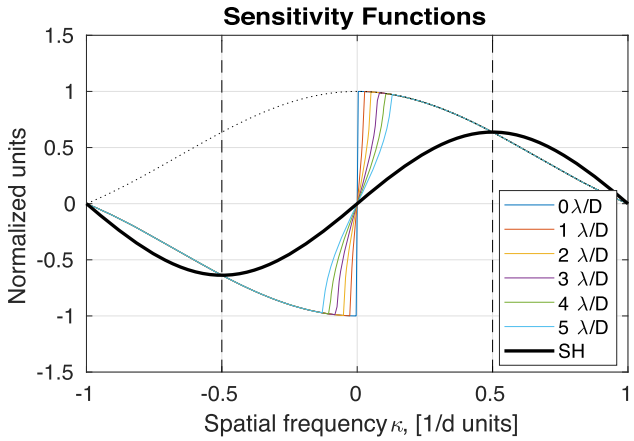

Figure 1. Pyramid 1D sensitivity plots for modulations $\{0, \cdots, 5\}\lambda/D$ from (8) overlaid with the Shack–Hartmann sensitivity when the discrete nature of the measurement is explicitly taken into account for a fair comparison between the PWFS and SH-WFS. Although the PWFS exhibits slope-like and phase-like measurement regimes, this misconception is clarified in the text.

phase by the sensor's impulse response (IR):

$$\boldsymbol{s}_x = \mathbf{IR}_{\mathbf{s}_x} \star \boldsymbol{\varphi}, \quad \boldsymbol{s}_y = \mathbf{IR}_{\mathbf{s}_y} \star \boldsymbol{\varphi}. \tag{4}$$

Both Conan (2003) and Shatokhina et al. (2013) show that the pyramid slopes-map can be asymptotically approximated as

$$\begin{aligned} \boldsymbol{s}_x &= -\frac{J_0(\alpha x)}{\pi x} \star \Pi_{\mathrm{p}} \star \boldsymbol{\varphi}(x, y), \\ \boldsymbol{s}_y &= -\frac{J_0(\alpha y)}{\pi y} \star \Pi_{\mathrm{p}} \star \boldsymbol{\varphi}(x, y) \end{aligned} \tag{5}$$

when the telescope aperture tends to infinity and the phase aberrations $\boldsymbol{\varphi} \ll 1$ rad. Conan went on to develop the effect of cross-terms (from adjacent and opposite quadrants of the PWFS), but we refrain from using it because the formulation that follows (Section 2.4) proved more practical by including realistic features of the PWFS in a condensed formulation. For completeness, $J_0(\cdot)$ is a zero-order Bessel function of the first kind, and α is a real-valued scalar representing the modulation in units of λ/D; we have added the function Π_{p} to Conan's and Shatokhina's to represent the pixel response – and likewise, a user-defined pixel binning for poor signal-to-noise ratio (SNR) regimes with dimmer stars – conveniently modelled as a door function (Oppenheim & Schafer 1999). This term carries the smearing of the sensitivity curves observed experimentally for spatial frequencies closer to the system's control radius. Unlike Vérinaud (2004), we consider the finite nature of the measurement to be an integral part of the sensing chain, leading to a different intuition into the nature of the measurements provided by both the PWFS and the Shack–Hartmann WFS (SH-WFS), as shown in Fig. 1. The 1D curves are insightful for understanding the PWFS behaviour, but rather limited because the 2D sensitivity is far from being radially symmetric, as shown in Fig. 2.

We note that this model can, as it happens, be formulated as a linear combination of intensity terms, each following a more general definition covering cases with coherent and incoherent recombination of light past the pyramid optic, in the form

$$i_{\mathrm{linear}} = \mathbf{IR} \star \boldsymbol{\varphi}, \tag{6}$$

which admits a closed-form expression. We follow Fauvarque et al. (2017) to refer to this as the *meta-intensity* model. These authors

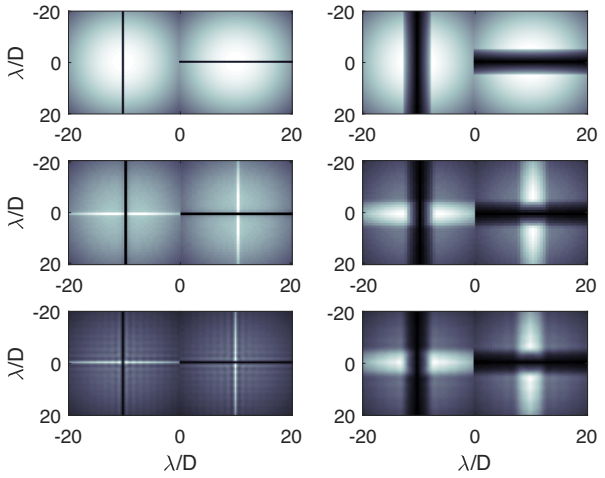

Figure 2. Comparison of measured with theoretical filters. (Top) model from equation (8); (middle) model from equation (9); (bottom) OOMAO-provided measurements (implementing (1)). (Left) modulation $1\lambda/D$; (right) $5\lambda/D$.

show that (1) can be linearized using a Taylor expansion series and the Cauchy product of two complex exponential functions to circumvent the squared modulus. This is particularly insightful, as they show that recombining the quadrant intensities as is customary (see equation 3) improves the PWFS linearity range because the even-power intensity terms cancel out, pushing the non-linearity further away to higher-order terms. This is the case for perfectly aligned systems, but in practice this assertion may not fully hold (Deo et al. 2018), which is a clear indication that it would be better to use the intensity signals directly instead of the slopes, at the expense of a reduced linearity range.

The models provided in Fauvarque et al. (2017) that we adopt in this study – as generalizations, which they are – allow for different transparency masks with a variable number of facets, extended guide-stars, the effect of the telescope pupil, and the presence of residual errors after AO partial compensation. Moreover, Fauvarque et al. (2019) developed the IR in equation (6) further, deriving an analytic formulation suitable for estimating the optical gains from the power-spectral density of AO residuals.

Equation (6) is very appealing for performing wavefront reconstruction, in that i_{linear} is a closer match to the PWFS physical-optics model than the model in equation (5) implies (as shown on Fig. 2). We chose explicitly to work with improved 'slopes-maps' models for ease of understanding and comparison with SH-WFS.

For analytic performance evaluation (in the absence of practical elements, such as calibration, optical defects, saturation ,etc.), the use of linear models in the form of either equation (4) or equation (6) leads to the same results. For real-time wavefront reconstruction, using detector intensities directly may lead to computational savings and to a more appropriate setting; however, we leave this issue open and do not consider it here.

The next section recasts the formulations seen so far in the Fourier domain, where the required mathematical operations admit simplifications and are soundly and effectively accomplished.

2.2 Transfer function of a PWFS

We now turn our attention to the physical-optics model in the spatial-frequency domain. This formulation is especially useful

because measurements are obtained as the convolution of the phase by the PWFS impulse-response, or, equivalently, as a point-wise multiplication in the Fourier domain. From Correia & Teixeira (2014) and Correia et al. (2017), let the following general-purpose linear measurement model,

$$\widetilde{\mathbf{s}}(\boldsymbol{\kappa}) = \widetilde{\mathcal{W}}\widetilde{\boldsymbol{\varphi}}_{||}(\boldsymbol{\kappa}) + \widetilde{\boldsymbol{\rho}}(\boldsymbol{\kappa}) + \widetilde{\boldsymbol{\eta}}(\boldsymbol{\kappa}), \tag{7}$$

where $\widetilde{\mathcal{W}} = \{\widetilde{\mathcal{W}}_x; \widetilde{\mathcal{W}}_y\}$ is a linear filter obtained by Fourier-transforming the impulse-response in equation (5), namely the PWFS optical transfer function (OTF), relating the in-band wavefront $\widetilde{\boldsymbol{\varphi}}_{||}$ to the measurements $\widetilde{\mathbf{s}}$; $\widetilde{\boldsymbol{\rho}}(\boldsymbol{\kappa})$ is the aliasing term acting as a generalized (coloured) noise term; and $\widetilde{\boldsymbol{\eta}}(\boldsymbol{\kappa})$ is the additive noise representing the photon and detector read noise.

Shatokhina et al. (2013) provide a closed-form equation for the Fourier transform of (5), which is a generalization from the 1D linearly-modulated PWFS from Vérinaud (2004) to the 2D case with circular modulation, allowing for an expression of the filter as

$$\widetilde{\mathcal{W}}_x(\boldsymbol{\kappa}) = \begin{cases} i\,\mathrm{sgn}(\kappa_x)\mathrm{sinc}(bd\boldsymbol{\kappa}) & \text{if } |\boldsymbol{\kappa}| > \kappa_{\mathrm{mod}} \\ \frac{2i}{\pi}\arcsin(\kappa_x/\kappa_{\mathrm{mod}})\mathrm{sinc}(bd\boldsymbol{\kappa}) & \text{if } |\boldsymbol{\kappa}| < \kappa_{\mathrm{mod}}, \end{cases} \tag{8}$$

with $\boldsymbol{\kappa} = (\kappa_x; \kappa_y)$ a 2D spatial-frequency vector in units of m^{-1}, κ_{mod} the modulation α from equation (5) expressed in m^{-1}, d the sampling of the pupil plane in metres (commonly the subaperture size), and $\widetilde{\mathcal{W}}_y(\kappa_x, \kappa_y) = \widetilde{\mathcal{W}}_x(\kappa_y, \kappa_x)$, namely the transpose of the 'x' filter. The discrete sampling and pixel averaging averaging at the detector level causes an attenuated high-frequency response closer to the AO control radius, given by a multiplicative separable factor $\mathrm{sinc}(\boldsymbol{\kappa}) = \mathrm{sinc}(\kappa_x)\mathrm{sinc}(\kappa_y)$, with $\mathrm{sinc}(x) = \sin(x\pi)/(x\pi)$. This term represents also the user-defined, post-facto binning with $b \in \mathbb{N}$ an integer scalar. Fig. 1 depicts 1D slices of $\widetilde{\mathcal{W}}(\boldsymbol{\kappa})$ across the spatial-frequency variables, representing the sensitivity of the pyramid optic jointly with the detector, the latter acting as discrete spatial sampler.

If instead we use the developments by Fauvarque et al. (2019), then the PWFS IR can be formulated as

$$\widetilde{\mathcal{W}}_x(\boldsymbol{\kappa}) = \sqrt{|\widetilde{\mathcal{T}}_x(\boldsymbol{\kappa}) \star \widetilde{\Pi}_{\mathrm{p}}|^2}, \tag{9}$$

where the function $\widetilde{\mathcal{T}}_x(\boldsymbol{\kappa})$ can be expanded as

$$\widetilde{\mathcal{T}}_x(\boldsymbol{\kappa}) = 2i\,(H_3 \star H_2\omega - H_2 \star H_3\omega + H_1 \star H_4\omega - H_4 \star H_1\omega) \tag{10}$$

with H_n Heaviside functions as before and ω a weighting function that characterizes the modulation signal; that is, it encodes the normalized time spent on the modulation phase $\boldsymbol{\theta}(\boldsymbol{r}, t)$ over one integration frame. Provided it is expanded as a linear series of n modes, $\boldsymbol{\theta}(\boldsymbol{r}, t) = \sum_{i=1}^{n} a_i(t)\mathcal{M}(\boldsymbol{r})$ with $\mathcal{M}(\boldsymbol{r})$ an orthonormal basis set, ω becomes

$$\omega = \widetilde{\mathcal{A}} \int_{t-T_s}^{T_s} a(t)\mathrm{d}t, \tag{11}$$

where $\tilde{\mathcal{A}}$ is the Fourier-transformed aperture function.

The most common choice is tilt modulation, in which case we have $\boldsymbol{\theta}(\boldsymbol{r}, t) = a_1(t)x + a_2(t)y$. If the modulation describes a perfect radially symmetric ring, $a_1 = a_2$, and consequently equation (11) becomes (Baddour 2011)

$$\omega = \mathbf{TF}^{-1}\{\mathcal{A}(\boldsymbol{\rho}) \times J_0(\alpha\boldsymbol{\rho})\}, \tag{12}$$

where $J_0(\alpha\boldsymbol{\rho})$ is a Bessel function resulting from the Fourier transform of a (modulation) circle in the focal plane. The relationship between modulation and tilt amplitude is $\alpha = \pi/4a$ in units of λ/D. This is the formulation that we will use in the remainder of this paper, unless otherwise specified.

We note from equation (9) that the OTF is in fact the modulation transfer function (MTF), which provides the magnitude response of the optical system to harmonic functions of different spatial frequencies. The PWFS phase transfer function (PTF) is therefore null, a consequence of using intensity signals to measure a complex field.

A remarkable feature of this model is that the PWFS instantaneous response can now be understood and potentially used to estimate instantaneous optical gains through the use of a complex-valued $\tilde{\mathcal{A}}$ function to (i) optimize the run-time AO performance and (ii) estimate (and remove) quasi-static (pinned) speckles that limit the contrast achievable with high-contrast imagers.

2.3 A note on the nature of the PWFS signals

The dual behaviour of the modulated pyramid sensor, acting as a *slope-like sensor* for low spatial frequencies and as a *phase-like sensor* for high spatial frequencies, is now well established. But it corresponds to a misconception. It stems from an erroneous analogy between the sensitivity of the pyramid and the nature of the measured signal initially stated in Vérinaud (2004) and represented in Fig. 1. Although for spatial frequencies above the modulation (often in multiples of λ/D units) the sensitivity is that of a phase sensor, the PWFS still provides a signal akin to the first spatial derivative of the phase (in the form of a Hilbert transform), with a frequency-dependent scaling at the origin of the misconception. Dubbing the PWFS measurements as 'slopes' is, in this light, justified. Paradoxically, it is commonplace in the AO community.

In theory, for a large modulation the sensor will act more fully as a gradient sensor (with the correct frequency-dependent gains), and it may be possible to reconstruct from its measurements using previously derived Shack–Hartmann filters by Correia & Teixeira (2014). One such successful albeit suboptimal attempt can be found in Quirós-Pacheco et al. (2009).

2.4 Modelled versus measured PWFS filter functions

Fig. 2 shows the measured PWFS OTF obtained from implementing equation (1) in the full end-to-end physical optics model in the Object-oriented Matlab Adaptive-Optics Simulator (OOMAO) featured in Conan & Correia (2014). The procedure is reminiscent of the 'poke-matrix', in that we record the PWFS response to the complete set of complex-exponential functions in our basis set. It is compared with the model in equation (8) (which does not take into account the cross-terms for ease of presentation, although they are formulated in Conan 2003 and Wang et al. 2010) and that in equation (9), for which one can clearly see the correct fit to the low-, high- and cross-term frequencies.

2.5 PWFS measurement noise model

Feeney (2001) established that the effect of photon noise (yet not limited to) on the WFS measurements is such that

$$\sigma^2_{s_x,\mathrm{photon}} = \sum_q \sigma^2_{i_q} \left(\frac{\partial s_x}{\partial i_q}\right)^2, \tag{13}$$

where $\sigma^2_{i_q}$ is the signal variance on the qth quadrant, and

$$s_x = f(i_1, i_2, i_3, i_4) \tag{14}$$

from equation (3). The PWFS diffracted field gives rise to local intensity variations in the re-imaged pupil planes, leading to $\sigma^2_{i_q} =$

$\langle i_q \rangle$ under Poisson statistics, where $\langle \cdot \rangle$ denotes ensemble-averaging. Furthermore, light falls outside the valid re-imaged pupils most prominently for low-modulation cases, leading to a lower SNR. Although one noise model taking into account these features can be obtained straightforwardly, it lacks practicality. We will assume for the sake of simplicity that the number of incident photons on each and every pixel is the same, yielding

$$\sigma^2_{i_q} \approx n_{\rm ph}/4, \tag{15}$$

where $n_{\rm ph}$ is the average number of photon detections on the PWFS.

Assuming the raw measurement

$$\boldsymbol{r}_x = \frac{i_q + i_{\rm n}}{i_q + i_{\rm d}}, \tag{16}$$

with $i_{\rm d}$ and $i_{\rm n}$ shorthand for the other-than-q quadrant intensities in the denominator and numerator, respectively, of equation (3), the measurement partial derivatives are readily found:

$$\frac{\partial \boldsymbol{s}_x}{\partial i_q} = \frac{\partial \boldsymbol{r}_x}{\partial i_q}\frac{\partial \boldsymbol{s}_x}{\partial \boldsymbol{r}_x} = \frac{i_q + i_{\rm d} + i_q + i_{\rm n}}{(i_q + i_{\rm d})^2}\frac{\partial \boldsymbol{s}_x}{\partial \boldsymbol{r}_x}, \tag{17}$$

with $\partial \boldsymbol{s}_x/\partial \boldsymbol{r}_x$ playing the role of the pixel-dependent optical gain $g_{\rm opt}$. For brevity and practical reasons, that as we shall see below lead to very a good match, we assume it to be a scalar value. Using equation (15),

$$\frac{\partial \boldsymbol{r}_x}{\partial i_q} = \frac{4n_{\rm ph}/4}{n^2_{\rm ph}} = \frac{1}{n_{\rm ph}}. \tag{18}$$

Substituting (18) into (17) and then into (13), one finally obtains

$$\sigma^2_{s_x,\rm photon} = \left[4\frac{n_{\rm ph}}{4}\left(\frac{1}{n_{\rm ph}}\right)^2\left(\frac{\partial \boldsymbol{s}_x}{\partial \boldsymbol{r}_x}\right)^2\right] = \frac{g^2_{\rm opt}}{n_{\rm ph}}. \tag{19}$$

For the read-out noise, following the same assumptions,

$$\sigma^2_{s_x,\rm ron} = \sum_q \left(\frac{\partial \boldsymbol{s}_x}{\partial i_q}\right)^2, \tag{20}$$

yielding

$$\sigma^2_{s_x,\rm ron} = 4g^2_{\rm opt}\left(\frac{ron}{n_{\rm ph}}\right)^2, \tag{21}$$

where *ron* is the average read-out-noise in photo-electrons per frame and per pixel.

Fig. 3 compares the various models with physical-optics simulations, showing the high accuracy of the photon and read noise models, with variances represented respectively by σ^2_η and $\sigma^2_{\rm ron}$. The multiple markers correspond to different modulations because the photo-electron count varies accordingly, with different factor of light diffracted outside of the pupils, albeit only slightly.

2.5.1 *Comparison with the SH-WFS*

Because the PWFS has often been (wrongly) likened to the SH-WFS in quad-cell mode, we provide general expressions for the latter to enable the comparison offered in Fig. 3.

From Thomas et al. (2006),

$$\sigma^2 = A\frac{1}{\bar{n}_{\rm ph}} + B\left(\frac{ron}{\bar{n}_{\rm ph}}\right)^2, \tag{22}$$

with A and B a function of the algorithm used. $\bar{n}_{\rm ph}$ is the number of photo-electrons/sub-aperture/frame, and *ron* is the effective read-out noise in photo-electrons rms.

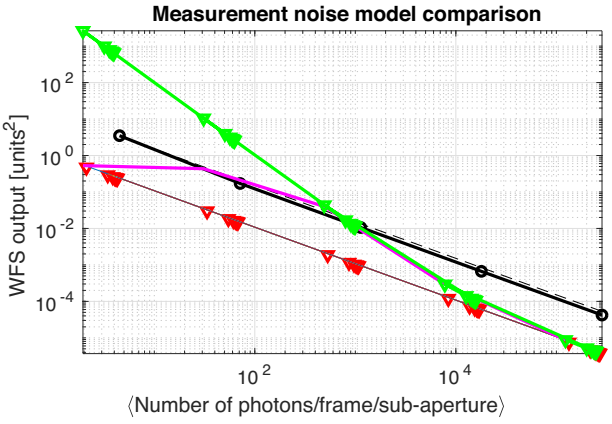

Figure 3. Illustrative example of noise models for the PWFS. Curves with markers are from theoretical expressions in this section; red, photon-noise only; green, photon+read noise (50e-), highly exaggerated for illustration purposes only. Curves with no markers are from physical-optics simulation models embedded in OOMAO realizing equation (1); multiple colours overlapping under the red triangles: photon-noise only; magenta, photon+read noise (50e-). The saturation at low photon-counts is the manifestation of the non-linear measurement regime (intensities are always positive). SH-WFS curves are in black.

For a quad-cell

$$A = \pi^2\kappa, \tag{23}$$

$$B = 4\pi^2\kappa^2, \tag{24}$$

with $\kappa = 1$ for a diffraction-limited spot.

We observe that using the quad-cell noise model from the SH-WFS applied to the PWFS leads to results different from those developed in equation (19) and equation (21).

In the Appendix, we show the necessary steps to calculate the noise propagation expressed on an orthogonal basis of modes. Although for the SH-WFS the output variance can be assimilated to an OPD at the edges of the subapertures, we caution that the same cannot be achieved with the PWFS on account of the nature of its measurement (see Section 2.3).

2.6 Considerations about VIS versus NIR WF sensing

2.6.1 *Photometric argument*

Let the accuracy of the measurement be proportional to the diffraction λ/D; that is, take the optical gain in equation (17) to be inversely proportional to the angular size of an hypothetical PSF (its full width at half-maximum) that would be recorded at the vertex of the PWFS. A back-of-the-envelope calculation tells us that it is more beneficial to sense at near-infrared (NIR) instead of visible (VIS) wavelengths should the number of photons $n_{\rm ph,\,NIR} > (\lambda_{\rm NIR}/\lambda_{\rm VIS})^2 n_{\rm ph,\,VIS}$. If this is the case, a factor of 4 more photons is required in the NIR than in the VIS. Taking the case of SPHERE (Fusco et al. 2016), the photometric budget for VIS and NIR detectors (CCD220 and Saphira respectively) coupled with the throughput of the whole instrument, we obtain 2×10^9 ph s^{-1} m^{-2} at the central $\lambda_{\rm VIS} = 800$ nm, whereas at $\lambda_{\rm NIR} = 1300$ nm we obtain 8×10^9 ph s^{-1} m^{-2} for a G5 star. This seems to indicate that there is no huge gain in performing NIR sensing, avoiding further operational overheads of operating in the IR. The validity of this

argument depends on the type of star used and the scientific potential of NIR targets. Such discussion is beyond this paper, however.

2.6.2 *Morphological argument*

As far as the PWFS is concerned, because we are not measuring the position of a PSF (as is the case for the SH-WFS), the previous argument is flawed, at least to the extent that the PWFS in its normal operating regime features mixed slope-like and phase-like sensitivities (Section 2.3). Inasmuch as the relationship of the morphology of the PSF and its optical gain is non-linear, operating in the NIR is advantageous because the residuals at $\lambda_{\rm NIR}$ are lower, allowing the PWFS to work closer to its linear regime. Conversely, in the VIS, the wavefront distortions are higher, causing the PWFS to work in a 'less linear' regime and thereby causing optical gain variations that are not fully compensated by an increase in photon collection from those sources (Bond et al. 2018).

3 ANALYTICAL ERROR BUDGET EVALUATION

3.1 AO-induced OPD effects

The results in this section closely follow those in Correia et al. (2017). There we made a comprehensive presentation of how calculations of servo-lag (or aniso-servo-lag in case the guide-star and the science target are not the same), aliasing, measurement noise and fitting error can be conveniently evaluated using power-spectral densities in the spatial-frequency domain under temporally filtered, closed-loop control. Here we provide a shortened presentation.

Throughout this paper, the parameters in Table 1 are used as default.

Our goal is to evaluate the total residual (piston-removed) phase variance, defined by

$$\sigma^2_{\rm Tot} \triangleq \int \widetilde{\mathcal{P}} \left\langle |\widetilde{\boldsymbol{\varphi}}(\boldsymbol{\kappa}) - \widetilde{\boldsymbol{\varphi}}^{\rm cor}(\boldsymbol{\kappa})|^2 \right\rangle \partial\boldsymbol{\kappa}, \tag{25}$$

which is a function of $\{d, D, r_0, L_0, \sigma^2_\eta\}$, the actuator pitch, the telescope diameter, the atmosphere coherence length, the outer scale, and the measurement-noise variance. The piston-removal function is given by $\widetilde{\mathcal{P}} = [1 - |2J_1(\pi\kappa D)/\pi\kappa D|^2]$, with the term within the modulus the Fourier transform of a circular pupil function of diameter D.

In the remainder of the paper, we suppose that the DM corrects entirely for the reconstructed phase, namely $\widetilde{\boldsymbol{\varphi}}^{\rm cor}(\boldsymbol{\kappa}) = \widehat{\widetilde{\boldsymbol{\varphi}}}(\boldsymbol{\kappa}) = \widetilde{\mathcal{R}}\widetilde{\mathbf{s}}(\boldsymbol{\kappa})$, when the anti-folding filter is applied (Correia & Teixeira 2014).

Equation (25) is expanded using equation (7), yielding

$$\begin{aligned}\widetilde{\mathcal{P}}\left\langle |\widetilde{\boldsymbol{\varphi}}(\boldsymbol{\kappa}) - \widehat{\widetilde{\boldsymbol{\varphi}}}(\boldsymbol{\kappa})|^2 \right\rangle &= \widetilde{\mathcal{P}} \left\langle |\tilde{\varepsilon}_\perp(\boldsymbol{\kappa})|^2 \right\rangle + \widetilde{\mathcal{P}} \left\langle |\tilde{\varepsilon}_{||}(\boldsymbol{\kappa})|^2 \right\rangle \\ &= \left\langle |\widetilde{\boldsymbol{\varphi}}_\perp|^2 \right\rangle + |1 - \widetilde{\mathcal{R}}\widetilde{\mathcal{W}}|^2 \widetilde{\mathcal{P}} \left\langle \widetilde{\boldsymbol{\varphi}}(\boldsymbol{\kappa})\widetilde{\boldsymbol{\varphi}}(\boldsymbol{\kappa})^* \right\rangle \\ &\quad + \mathbf{W}_{\rm RA} + \left\langle \widetilde{\mathcal{P}} |\widetilde{\mathcal{R}}\widetilde{\boldsymbol{\eta}}|^2 \right\rangle \end{aligned} \tag{26}$$

with $\langle|\widetilde{\boldsymbol{\varphi}}_\perp|^2\rangle$ the Power Spectral Desnsity (PSD) of the fitting error (where we approximate $\widetilde{\mathcal{P}}(\boldsymbol{\kappa}) = 1$ for $|\boldsymbol{\kappa}| > 1/(2d)$). The term

$$|1 - \widetilde{\mathcal{R}}\widetilde{\mathcal{W}}|^2 \widetilde{\mathcal{P}} \langle \widetilde{\boldsymbol{\varphi}}(\boldsymbol{\kappa})\widetilde{\boldsymbol{\varphi}}(\boldsymbol{\kappa})^* \rangle = |1 - \widetilde{\mathcal{R}}\widetilde{\mathcal{W}}|^2 \mathbf{W}'_\varphi(\boldsymbol{\kappa}) \tag{27}$$

is the PSD of the open-loop phase reconstruction error, and

$$\mathbf{W}_{\rm RA} = \widetilde{\mathcal{P}} \sum_{\boldsymbol{m}\neq 0} |\widetilde{\mathcal{R}}(\boldsymbol{\kappa})\widetilde{\mathcal{W}}(\boldsymbol{\kappa} + \boldsymbol{m}/d)|^2 \mathbf{W}_\varphi(\boldsymbol{\kappa} + \boldsymbol{m}/d) \tag{28}$$

Table 1. Default simulation parameters. The turbulence model represents median Paranal conditions. Bold refers to the default parameters used in the analysis.

Telescope	
D	8.0 m
Throughput	50%
Guide-star	
Zenith angle	**0**°–60°
Magnitude	0–12
Atmosphere	
r_0	15 cm
L_0	25 m
Fractional r_0	[53.28;1.45;3.5;9.57;10.83; 4.37;6.58;3.71;6.71]/100
Altitudes	[0.042;0.140;0.281;0.562;1.125; 2.25;4.5;9;18] km
Wind speeds	[15;13;13;9;9; 15;25;40;21] m/s
Wind direction	[38;34;54;42;57; 48;−102;−83;−77]π/180°
Wavefront sensor	
Order	40 × 40
RON	1 e$^-$
$n_{\rm pix}$	4
$f_{\rm sample} = 1/T_s$	0.1–**1**–5 kHz
Modulation m	0–**2**–6 λ/D
$\lambda_{\rm WFS}$	**0.64**–1.65–2.2 μm
Centroiding algorithm	Thresholded CoG
DM	
Order	41 × 41
AO loop	
Pure delay	τ_{lag} =1 ms
Loop gain	$g = \{0.01, \cdots, \mathbf{0.3}, ..., 0.5\}$
Imaging wavelength	
$\lambda_{\rm im}$	0.75–**1.65**–2.2 μm

is the PSD of the reconstructed aliasing error. Finally,

$$\mathbf{W}_\eta = \langle \widetilde{\mathcal{P}} |\widetilde{\mathcal{R}}\widetilde{\boldsymbol{\eta}}|^2 \rangle \tag{29}$$

is the PSD of the propagated noise. This model can be (and was) further generalized to the closed-loop regime by Correia et al. (2017) when factoring in spatio-temporal functions characteristic of the loop filtering into equation (26).

3.1.1 *Aliasing rejection*

Fig. 4 shows the propagated aliasing after least-squares wavefront reconstruction (Correia et al. 2017) (i.e. with no temporal loop filtering). When comparing it with fig. 9 in Vérinaud (2004), we note the general agreement. However, as a result of the 2D reconstruction and the way x- and y-frequencies are mixed in the reconstructor's denominator the propagated aliasing is now more structured. As expected, it reaches still the level of the fitting error at the edge of the control radius.

In either case, the SH-WFS term is provided in black curves for comparison. Its amplitude is always greater than that of the PWFS. The face-on patterns provide further insight into the propagated aliasing and its spatial distribution.

From Fig. 1, there are hints that the amount of aliasing affecting either the PWFS or the SH-WFS is about the same. However, it is the

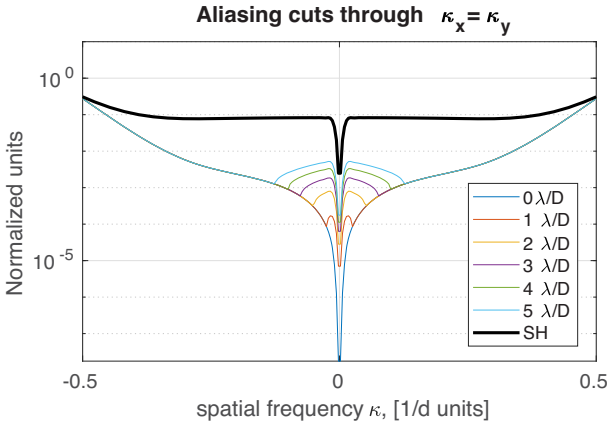

Figure 4. Propagated aliasing through a LS reconstructor. Slabs $\kappa_y = \kappa_x$.

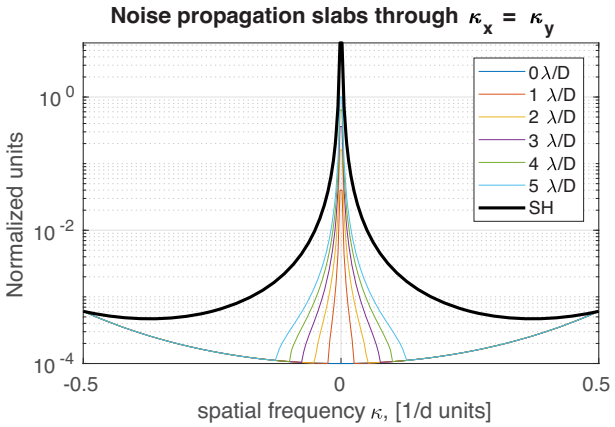

Figure 5. Propagated noise through a LS reconstructor: slab $\kappa_y = \kappa_x$.

propagation through the reconstructor that proves more beneficial with the PWFS.

3.1.2 *Noise propagation*

The noise propagated through a LS reconstructor is shown in Fig. 5 (cf. Vérinaud 2004, fig. 7). Here we depict a diagonal slab through the 2D noise propagation filter, yet the same behaviour applies to $\kappa_x = 0$ or $\kappa_y = 0$ slabs. As with the aliasing, the noise propagation is always lower for the pyramid WFS across all the spatial frequencies within the control radius.

3.2 Chromatic effects and scintillation

In addition to AO-induced residual OPD effects, we now consider additional limits to contrast resulting from chromatic effects and scintillation, which generate both amplitude and OPD variations. Guyon (2005) and Fusco et al. (2006) provide quantifications of such effects. We revisit those calculations and provide what we hope is a more comprehensive taxonomy.

A wavelength-dependent index of refraction gives rise to three distinct errors (Hardy 1998):

(i) angular dispersion due to the angle of incidence and refraction as the field propagates through the medium, which can be seen as a cumulative version of Snell's law over the vertical path;

(ii) differential refraction error: chromatic path-length and amplitude errors for different wavelengths traversing the same path because they travel at different velocities, creating a chromatic effect (termed correction chromatism by Fusco et al. (2006); term C6 in Guyon (2005));

(iii) dispersion displacement error caused by the differential bending of wavefronts at different wavelengths, causing rays to probe slightly different patches of turbulence, resulting in an angular anisoplanatism-like error

To these can be added scintillation plus OPD errors resulting from Fresnel propagation, even in the absence of chromatic refraction (Guyon 2005) and despite the weak-turbulence regime (Roddier 1981).

The first effect listed – angular dispersion – has no impact on contrast, but solely on the angular positions of point sources on the focal plane; it is therefore disregarded in the following.

Now, let the index of refraction fluctuations for standard pressure and temperature from Edlén (1966) (later slightly adjusted by Owens 1967)

$$\Delta n(\lambda) = 8.34213 \times 10^{-5} + \frac{0.0240603}{130 - \lambda^{-2}} + \frac{0.00015997}{38.9 - \lambda^{-2}} \quad (30)$$

here taken to coincide with the refractivity index, namely $n = 1 + \Delta n$.

We assume that the differential refraction error is proportional to the ratio of fluctuations $\Delta n(\lambda_1)/\Delta n(\lambda_0)$, as suggested by Fusco et al. (2006), and not to the ratio of indices of refraction $n(\lambda_1)/n(\lambda_0)$, as was considered by Guyon (2005).

For the dispersion displacement error we follow Fusco et al. (2006) to compute an error PSD (both OPD and amplitude), considering anisoplanatic imaging with an angle

$$\theta = (\Delta n(\lambda_1) - \Delta n(\lambda_0)) \tan(ZA), \quad (31)$$

where ZA is the zenith angle in radians.

Using standard expressions for Fresnel propagation (real and imaginary components) and differential chromaticity from Guyon (2005), we produced Fig. 6 with visible WFS and NIR imaging, and Fig. 7 with both NIR WFS and imaging. As expected, amplitude effects are more pronounced when using very different wavefront sensing and imaging wavelengths.

We note that the contrast estimates are in good agreement with the results obtained with SPHERE on the VLT (Cantalloube et al. 2019; Vigan et al. 2019) and Keck (Xuan et al. 2018).

In our implementation, we do not consider any coupling between servo-lag OPD errors and scintillation, although it gives rise to tangible effects in post-coronagraphic images, in particular asymmetric haloes in the AO correction region (commonly known as the butterfly because of its shape) caused by a correlation of temporal delay and wind velocity (Cantalloube et al. 2018).

We further notice (as seen elsewhere by Guyon 2005, Fusco et al. 2006 and Guyon 2018) that the AO residuals – namely the servo-lag error – is by far the limiting factor. This provides a compelling motivation for the investigation of predictive control approaches; see, for instance, Correia et al. (2017), Males (2018) and Massioni, Gilles & Ellerbroek (2015).

4 WAVEFRONT RECONSTRUCTION IN THE SPATIAL-FREQUENCY DOMAIN

Having developed PWFS formulations and evaluated the AO-centric error budget from functions in the continuous spatial-frequency domain in previous sections, we devote this section to

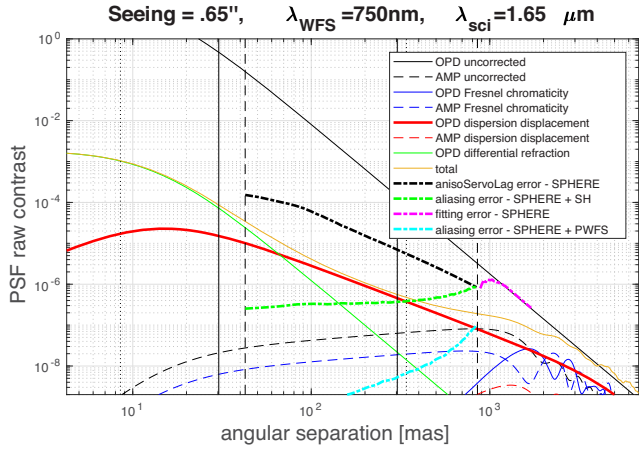

Figure 6. Radially-averaged post-coronagraphic contrast curves including atmospheric dispersion effects and Fresnel propagation on both amplitude (AMP) and optical path length (OPD) for a WFS in the VIS and a camera in the *H* band for a 30deg off-zenith observation. The maximum improvement is comprised between the black dashed curve (the aniso-servo-lag error) and the orange curve, indicating the total of the non-AO limiting contrast terms. In the range 200–1000 mas, the gap has an upper bound of a factor of 10. Vertical black lines indicate the correction bands of Keck's AO system (solid), the Very Large telescope's SPHERE instrument (dashed) and a first generation European Extremely Large Telescope (ELT) AO system(dotted). We caution that the radial averages shown do not represent in full the 2D structure of the post-coronagrahic contrast structure.

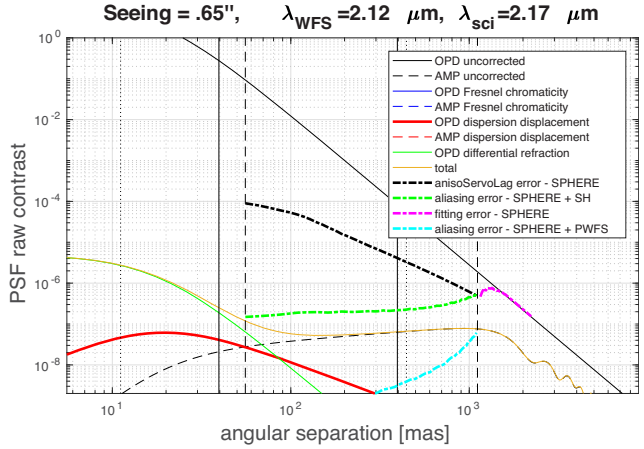

Figure 7. Radially-averaged post-coronagraphic contrast curves including atmospheric dispersion effects and Fresnel propagation on both amplitude (AMP) and phase (optical path length – OPD) for both *K*-band wavefront sensing and imaging. The maximum improvement is comprised between the black dashed curve (the anisoServoLag error) and the orange curve, indicating the total of the non-AO limiting contrast terms. In the range 50–1000 mas, the gap reaches two orders of magnitude. The vertical black lines are as in Fig. 6. Some of the terms are below the 10^{-8} level and thus not shown in the plot.

the real-time wavefront reconstruction from pyramid signals using discrete deconvolution-based processing as a natural extension of the preceding results.

The extensive development of Fourier reconstruction methods has focused on the Shack–Hartmann sensor, with compelling results – the most prominent being the Gemini Planet Imager Fourier Domain Reconstructor (Poyneer & Véran 2005).

Advanced systems such as upgrades to existing telescopes (e.g. VLT's SPHERE, Gemini's GPI, Subaru's ScExAO and Keck's KPIC) and AO systems for future GSMTs are likely to utilize the pyramid wavefront sensor over the more commonly used Shack–Hartmann sensor.

The use of Fourier reconstruction has the potential to significantly increase the reconstruction speed (or otherwise lessen the computational burden), particularly for high-order systems such as those on giant segmented mirror telescopes (GSMTs) (Poyneer & Véran 2005; Correia et al. 2007). On the other hand, and admitting that reconstruction speed in no longer of first priority owing to the huge progress in real-time architectures over the last decade, we note that the use of spatial frequencies extends the AO-correctable area by a factor of $\tilde{}\pi/4$ with respect to orthonormal modes defined on a circular pupil, which do not correct frequencies beyond $\kappa = \sqrt{\kappa_x^2 + \kappa_y^2} > 1/2d$ (although such an orthonormal basis could be formulated, to the authors knowledge it has not been used in the past).

A discrete version of the analytic model provided in equation (8) can be applied to the real-time wavefront reconstruction from pyramid slope data. An initial attempt was made in Quirós-Pacheco et al. (2009), assuming the PWFS sensitivity function to be that of a SH-WFS (which of course is only valid in the highly modulated case). An implementation customized to the PWFS is presented in Shatokhina & Ramlau (2017) using 1D reconstruction from PWFS signals in the x and y directions and then averaging.

Here instead we follow the approach by Bond et al. (2017) upgraded with the PWFS model in equation (9) and the Fourier-domain implementations in Correia et al. (2007); Correia et al. (2008) which use jointly the x and y measurement data and take special care of the finite aperture edge effects and boundary conditions using the circularity and divergence of the gradient field to ensure compatibility with the Fourier series. We find that this treatment is key to obtaining high levels of performance with minimal losses compared with the case that uses the full physical-optics PWFS model and SVD filtering to compute the reconstructor.

4.1 Discrete pyramid filter filters

We make use of a minimum mean square error (MMSE) minimization criterion to find the following filters:

$$\widehat{\tilde{\boldsymbol{\psi}}} = \mathbf{R}^x[k,l]\tilde{\boldsymbol{s}}_x + \mathbf{R}^y[k,l]\tilde{\boldsymbol{s}}_y, \tag{32}$$

$$\widehat{\boldsymbol{\psi}}[k,l] = \begin{cases} 0 & \text{if } k,l = 0; \\ \hat{\mathbf{R}}^x[k,l]\tilde{\boldsymbol{s}}_x[k,l] + \hat{\mathbf{R}}^y[k,l]\tilde{\boldsymbol{s}}_y[k,l] & \text{otherwise.} \end{cases} \tag{33}$$

It is straightforward to demonstrate (Correia & Teixeira 2014) that the MMSE solution can be expressed in the form of a Wiener filter as

$$\widehat{\boldsymbol{\psi}} = \frac{\mathbf{Q}_x^*\tilde{\boldsymbol{s}}_x + \mathbf{Q}_y^*\tilde{\boldsymbol{s}}_y}{|\mathbf{Q}_x|^2 + |\mathbf{Q}_y|^2 + \gamma\frac{\mathbf{W}_n}{\mathbf{W}_{\boldsymbol{\psi}}}}, \tag{34}$$

where $\mathbf{Q}$ is the discrete version of the continuous model in equation (8) with filters of the form

$$\hat{\mathbf{R}}^x[k,l] = \frac{\mathbf{Q}_x^*}{|\mathbf{Q}_x|^2 + |\mathbf{Q}_y|^2 + \gamma\frac{\mathbf{W}_n}{\mathbf{W}_{\boldsymbol{\psi}}}}, \tag{35}$$

$$\hat{\mathbf{R}}^y[k,l] = \frac{\mathbf{Q}_y^*}{|\mathbf{Q}_x|^2 + |\mathbf{Q}_y|^2 + \gamma\frac{\mathbf{W}_n}{\mathbf{W}_{\boldsymbol{\psi}}}}. \tag{36}$$

The least-squares solution is easily obtained by taking $\gamma = 0$.

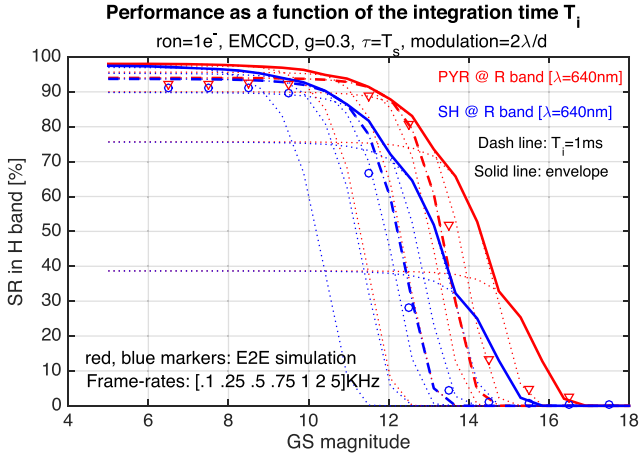

Figure 8. Performance comparison between the analytic and the Monte Carlo models. Dashed curves indicate the 1-kHz frame-rate case to be compared with the results (markers only) of the Monte Carlo simulations. In the latter, the PWFS optical gain is compensated for by adjusting a single scalar gain, which in our opinion is at the origin of the slight drop in performance at the bright-star end.

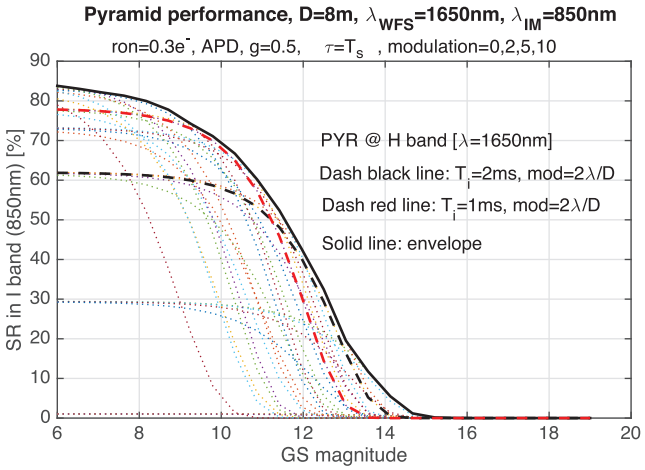

Figure 9. Performance expected at visible wavelengths by a PWFS system as a function of the GS magnitude for different loop frame-rates and PWFS modulation amplitudes.

The priors $\mathbf{W}_n$ and $\mathbf{W}_\psi = 0.49 r_0^{-5/3}\{(2\pi)^2[f_x^2 + f_y^2 + (1/L_0)^2]\}^{-11/6}$ are the spatial PSDs of the noise and the phase. The noise is assumed to be white and uncorrelated, and thus constant over all the frequencies; that is, $\mathbf{W}_n \propto k \in \Re$. An anti-aliasing Wiener filtering solution is developed in Correia & Teixeira (2014) by suitably modifying the whiteness of the noise. A further scalar factor γ is introduced to properly weigh the priors term to account for other unknown system parameters.

5 LIMITING PERFORMANCE AND CONTRAST

We are now in a position to apply the results from the preceding sections to representative cases of high-contrast imagers on existing 10-m-class AO systems; as a matter of choice application to giant segmented mirror telescopes is not shown, yet it can easily be achieved . In this section, we investigate (analytically and with Monte Carlo models) the performance for our simulated system as a function of exposure time, modulation and guide-star magnitude. In so doing we revisit the work of Vérinaud (2004) and extend it to the 2D case. Further observing conditions and system parameters can be found in Table 1.

5.1 λ_{im} = NIR, λ_{WFS} = VIS in imaging mode

As an example of the analytic error breakdown and reconstruction accuracy, Fig. 8 compares the limiting performance expected from a visible PWFS and SH-WFS on an 8-m-class telescope as a function of guide-star magnitude and AO loop sampling frequency. Using the developments in Section 4, we over-plot (circles and triangles) the results of Monte Carlo simulations performed with OOMAO.

It is interesting to note that for bright guide-stars there is only a slight advantage for the pyramid - certainly due to a smaller propagated aliasing. Looking instead at the full 2D PSFs would give further insight into the differences between these two wavefront sensors that the Strehl-ration (SR) is incapable of showing. The structure of the errors is actually key to coronagraphic applications.

At the faint-star end, the pyramid is at its best. The lower noise propagation due to its increased sensitivity, makes it push the limiting magnitude by roughly two stellar magnitudes. Whereas the SH drop-off knee is around magnitude 12, the pyramid ensures good performances down to magnitude 14. At magnitude 15, the pyramid still achieves 30% SR, which is a fairly high value.

5.2 λ_{im} = VIS, λ_{WFS} = NIR in imaging mode

We can likewise explore the performance at visible wavelengths of a pyramid-based high-contrast AO system. The motivation is two-fold: (i) to provide the performance in a parameter space complementary to that of space-borne and ground-based high-resolution spectrographs used for the indirect detection and characterization of extra-solar planets and discs; and (ii) to guide on NIR stars and brown dwarfs, the latter being relatively fainter at visible wavelengths but likely to host planetary systems as well.

Fig. 9 shows the performance in the I band ($\lambda_{im} = 850$ nm) when the sensing is carried out in the H band ($\lambda_{WFS} = 1650$ nm) as a function of guide-star magnitude, frame-rate and pyramid modulation. We have considered fast IR detectors with sub-electron noise, providing for fast reading. We can observe the huge impact of running at higher frame-rates, with the performance increasing from 60% at 500 Hz to 80+% at 5000 Hz. This is a strong indication that servo-lag error is the dominant factor.

5.3 Integral versus distributed control in coronagraphic mode

In an attempt to minimize residual AO errors after correction, the results that follow build on the predictive capabilities of distributed control with Kalman filters (DKF) following the formulation presented in Correia et al. (2017).

We first single-out the effect of using predictive control over scalar-gain integral control in Fig. 10, for a SH-based high-contrast imager (although this WFS is not the main focus of this paper). There, we plot the improvement/degradation as a ratio of achievable contrast (negative/positive values respectively, because we are using a log scale). The 2D nature of these plots adds to the limiting contrast curves in Figs 6 and 7, which are radially symmetric in nature; we can observe at the bright-star end that contrast improvements can surpass a factor of 10 at small separations (typically below 5 λ/D) and in certain wind-dependent directions. In appendix more combinations for the couple controller/WFS are shown.

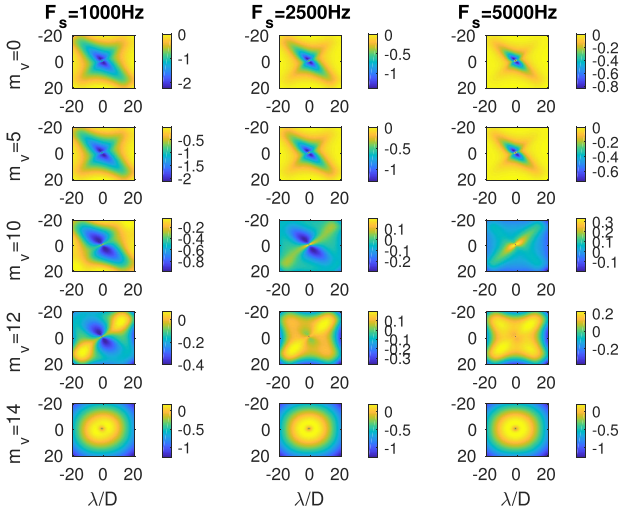

Figure 10. Contrast ratio improvement map (log scale) as a function of the frame-rate and stellar magnitude for an AO loop featuring a SH-WFS with a LS reconstructor and integral controller versus a DKF controller.

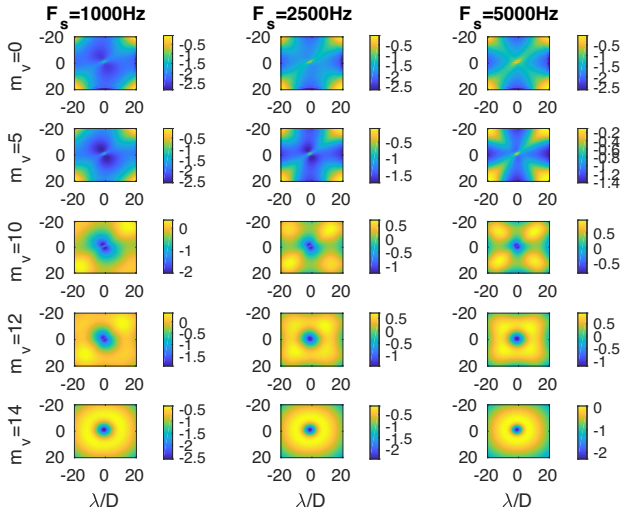

Figure 11. Contrast ratio improvement map (log scale) as a function of the frame-rate and stellar magnitude for a SH with a LS reconstructor and integral controller versus a PWFS with a DKF controller.

When a PWFS is employed, the potential contrast improvement increases, as shown in Fig. 11 in both depth and extent.

For a magnitude 0 star, a combination of PWFS + DKF can clean up the AO control region almost entirely, resulting in an extended zone with an improvement in the contrast ratio of up to two orders of magnitude. This happens on account of the almost complete removal of servo-lag (typically $\in \pm 5\lambda/D$) and aliasing (closer to the AO correction edges) terms. As we move towards higher frame-rates, the improvement brought by the DKF is lower because the servo-lag error is correspondingly smaller (as in Fig. 10).

In real observations, the presence of quasi-static speckles will limit the contrast figures provided. With the possibility offered by the models in Section 2 to estimate instantaneously the PWFS optical gains, we leave it for the interested reader to explore pushing the contrast further when using noiseless detectors (or very low read noise), from which high-quality, short-exposure images can be collected with custom post-processing techniques that include knowledge of the variability of these optical gains.

6 CONCLUSION

We have explored the performance limits of the PWFS for both imaging and high-contrast applications illustrated on sample AO systems on 10-m-class telescopes. In order to do this, we produced an AO-centric error breakdown using a practical, convolution-based PWFS model in the spatial-frequency domain (developed in Fauvarque et al. 2019) featuring some highly desirable properties:

(i) the meta-intensity linear model (from which the slopes-maps are computed as linear combinations) represents to a broader extent the diffractive nature of the PWFS optic.

(ii) this model can be generalized to finite pupils, coherent and incoherent recombinations of light resulting from overlapping and non-overlapping re-imaged pupils respectively, extended objects, and off-line optical gain retrieval

Our calculations back the generally accepted result whereby the PWFS extends the limiting WFS magnitude by up to two stellar magnitudes on 10-m-class telescopes, allowing for a greater sky coverage (not quantified here).

On existing high-contrast imagers equipping 10-m-class telescopes with a visible or near-infrared PWFS under the median Paranal turbulence conditions outlined in Table 1, we demonstrated a contrast improvement (limited by chromatic and scintillation effects) of 2× up to 10× when replacing the wavefront sensor only at large separations close to the AO control radius where aliasing dominates, and by factors in excess of 10× by coupling distributed control with the PWFS over most of the AO control region, from small separations starting with an IWA of typically 1–2 λ/D to the AO correction edge.

ACKNOWLEDGEMENTS

All the simulations and analysis were carried out using the object-oriented Matlab AO simulator (OOMAO). The class *spatialFrequencyAdaptiveOptics* implementing the analytics developed in this paper as well as the results herein is packed with the end-to-end library freely available from https://github.com/cmcorreia/oomao.

The research leading to these results received support from the A∗MIDEX project (no. ANR-11-IDEX-0001-02) funded by the 'Investissements d'Avenir' French Government program, managed by the French National Research Agency (ANR).

REFERENCES

Baddour N., 2011, Adv. Imag. Elect. Phys., 165, 1
Beuzit J. L. et al., 2019, A&A, 631, A155
Bond C. Z., El Hadi K., Sauvage J. F., Correia C., Fauvarque O., Rabaud D., Neichel B., Fusco T., 2016, Proc. Fourth AO4ELT Conference. eScholarship online at http://escholarship.org/uc/ao4elt4
Bond C. Z., Correia C. M., Sauvage J.-F., Neichel B., Fusco T., 2017, Opt. Express, 25, 11452
Bond C. Z. et al., 2018, in Close L. M., Schreiber L., Schmidt D., eds, Proc. SPIE Conf. Ser. Vol. 10703, Adaptive Optics Systems VI. SPIE, Bellingham, p. 107031Z
Cantalloube F. et al., 2018, A&A, 620, L10
Cantalloube F., Dohlen K., Milli J., Brandner W., Vigan A., 2019, Messenger, 176, 25
Chew T. Y., Clare R. M., Lane R. G., 2006, Opt. Comm., 268, 189
Conan R., 2003, Technical report, Fourier optics and distribution theory applied to pyramid wavefront sensors. LAOG/ONERA
Conan R., Correia C., 2014, in Marchetti E., Close L. M., Véran J.-P., eds, Proc. SPIE Conf. Ser. Vol. 9148, Adaptive Optics Systems IV. SPIE, Bellingham, p. 91486C

Correia C., Conan J. M., Kulcsár C., Raynaud H. F., Petit C., Fusco T., 2007, in Bouvier J., Chalabaev A., Charbonnel C., eds, SF2A-2007: Proc. Annu. Mtg French Soc. Astron. Astrophys.p. 25
Correia C., Kulcsar C., Conan J.-M., Raynaud H.-F., 2008, in Hubin N., Max C. E., Wizinowich P. L., eds, Proc. SPIE Conf. Ser. Vol. 7015, Adaptive Optics Systems. SPIE, Bellingham, p. 701551
Correia C. M., Teixeira J., 2014, J. Opt. Soc. Am. A., 31, 2763
Correia C. M., Bond C. Z., Sauvage J.-F., Fusco T., Conan R., Wizinowich P. L., 2017, J. Opt. Soc. Am. A, 34, 1877
Deo V., Gendron É., Rousset G., Vidal F., Buey T., 2018, in Close L. M., Schreiber L., Schmidt D., eds, Proc. SPIE Conf. Ser. Vol. 10703, Adaptive Optics Systems VI. SPIE, Bellingham, p. 1070320
Edlén B., 1966, Metrologia, 2, 71
Esposito S., Pinna E., Puglisi A., Agapito G., Veran J. P., Herriot G., 2015, in Proc. Fourth AO4ELT Conference. p. E36, eScholarship online at http://escholarship.org/uc/ao4elt4
Fauvarque O., 2017, PhD thesis. Aix-Marseille Univ.
Fauvarque O., Neichel B., Fusco T., Sauvage J.-F., 2015, Optics Lett., 40, 3528
Fauvarque O., Neichel B., Fusco T., Sauvage J.-F., Girault O., 2017, J. Astron. Telesc., Instrum. Syst., 3, 019001
Fauvarque O., Janin-Potiron P., Correia C., Brule Y., Neichel B., Chambouleyron V., Sauvage J.-F., Fusco T., 2019, J. Opt. Soc. Am. A, 36, 1241
Feeney O. A., 2001, Technical report. Theory and Laboratory Characterisation of Novel Wavefront Sensor for Adaptive Optics Systems. National Univ. Ireland, Galway
Fusco T. et al., 2006, Opt. Express, 14, 7515
Fusco T. et al., 2016, in Marchetti E., Close L. M., Véran J.-P., eds, Proc. SPIE Conf. Ser. Vol. 9909, Adaptive Optics Systems V. SPIE, Bellingham, p. 99090U
Guyon O., 2005, ApJ, 629, 592
Guyon O., 2018, ARA&A, 56, 315
Hardy J. W., 1998, Adaptive Optics for Astronomical Telescopes. Oxford Univ. Press, New York
Males J. R., Guyon O., 2018, J. Astron. Telesc., Instrum. Syst., 4, 4
Korkiakoski V., Vérinaud C., Louarn M. L., Conan R., 2007, Appl. Opt., 46, 6176
LeDue J., Jolissaint L., Véran J.-P., Bradley C., 2009, Opt. Express, 17, 7186
Linfoot E. H., 1948, MNRAS, 108, 428
Macintosh B. et al., 2018, in Close L. M., Schreiber L., Schmidt D., eds, Proc. SPIE Conf. Ser. Vol. 10703, Adaptive Optics Systems VI. SPIE, Bellingham, p. 107030K
Massioni P., Gilles L., Ellerbroek B., 2015, J. Opt. Soc. Am. A, 32, 2353
Mawet D. et al., 2012, in Clampin M. C., Fazio G. G., MacEwen H. A., Oschmann J. M., Jr., eds, Proc. SPIE Conf. Ser. Vol. 8442, Space Telescopes and Instrumentation 2012: Optical, Infrared, and Millimeter Wave. SPIE, Bellingham, p. 844204
Mawet D. et al., 2014, ApJ, 792, 97
Mawet D. et al., 2016, in Marchetti E., Close L. M., Véran J.-P., eds, Proc. SPIE Conf. Ser. Vol. 9909, Adaptive Optics Systems V. SPIE, Bellingham, p. 99090D
Mouillet D. et al., 2018, in Close L. M., Schreiber L., Schmidt D., eds, Proc. SPIE Conf. Ser. Vol. 10703, Adaptive Optics Systems VI, SPIE, Bellingham, p. 107031Q
Oppenheim A. V., Schafer R. W., 1999, Discrete-time signal processing, 2nd . Prentice-Hall, Inc
Owens J. C., 1967, Appl. Opt., 6, 51
Poyneer L. A., Véran J.-P., 2005, J. Opt. Soc. Am. A, 22, 1515
Quirós-Pacheco F., Correia C., Esposito S., 2009, in Clénet Y., Conan J.-M., Fusco T. F., Rousset G., eds, Proc. Adaptative Optics for Extremely Large Telescopes - 1st edition. EDP Sciences
Ragazzoni R., 1996, J. Modern Opt., 43, 289
Rigaut F., Gendron E., 1992, A&A, 261, 677
Roddier F., 1981, Prog. Opt., 19, 281
Shatokhina I., Ramlau R., 2017, Appl. Opt., 56, 6381
Shatokhina I., Obereder A., Rosensteiner M., Ramlau R., 2013, Appl. Opt., 52, 2640
Snik F. et al., 2018, in Navarro R., Geyl R., eds, Proc. SPIE Conf. Ser. Vol. 10706, Advances in Optical and Mechanical Technologies for Telescopes and Instrumentation III. SPIE, Bellingham, p. 107062L
Thomas S., Fusco T., Tokovinin A., Nicolle M., Michau V., Rousset G., 2006, MNRAS, 371, 323
Vérinaud C., 2004, Opt. Communications, 233, 27
Vérinaud C., Le Louarn M., Korkiakoski V., Carbillet M., 2005, MNRAS, 357, L26
Vigan A. et al., 2019, A&A, 629, A11
Wang J., Bai F., Ning Y., Huang L., Wang S., 2010, Opt. Express, 18, 27534
Xuan W. J. et al., 2018, AJ, 156, 156

APPENDIX A: NOISE PROPAGATION EXPRESSED ON AN ORTHONORMAL BASIS OF MODES

The results in this section follow closely the approach of Rigaut & Gendron (1992). The propagated noise covariance matrix on a predefined basis set of modes is defined as

$$\mathbf{\Sigma}_m = \langle \boldsymbol{m}\boldsymbol{m}^T \rangle, \tag{A1}$$

with the reconstructed modal coefficient vector $\boldsymbol{m}$ defined as

$$\boldsymbol{m} = \mathbf{G}^\dagger \boldsymbol{\eta}, \tag{A2}$$

where $\boldsymbol{\eta}$ is a vector of WFS measurement signals, $\mathbf{G}^\dagger = (\mathbf{G}\mathbf{G}^T)^{-1}\mathbf{G}^T$, and $\mathbf{G}$ is the system interaction matrix $\boldsymbol{s} = \mathbf{G}\boldsymbol{\psi}$ containing the responses of the WFS to each mode in $\boldsymbol{m}_i$ of $\boldsymbol{\psi}$. For the PWFS, this is the covariance function that needs be evaluated, namely

$$\mathbf{\Sigma}_m = \mathbf{G}^\dagger \langle \boldsymbol{\eta}\boldsymbol{\eta}^T \rangle \mathbf{G}^\dagger \tag{A3}$$

and $\langle \boldsymbol{\eta}\boldsymbol{\eta}^T \rangle$ provided in equation (19) and equation (21).

For linear WFSs, a simplification applies. If we consider now a measurement of pure noise, which is assumed to be of constant variance σ_η^2 across the pupil, then the noise propagated is

$$\mathbf{\Sigma}_m = \mathbf{G}^\dagger \langle \boldsymbol{s}\boldsymbol{s}^T \rangle \mathbf{G}^\dagger \tag{A4}$$

$$= (\mathbf{G}\mathbf{G}^T)^{-1}\sigma_\eta^2 \tag{A5}$$

provided that $\langle \boldsymbol{s}\boldsymbol{s}^T \rangle = \mathbf{I}\sigma_\eta^2$. Using singular value decomposition,

$$[\mathcal{U}\mathcal{S}\mathcal{V}^T] = \mathrm{svd}(\mathbf{G}), \tag{A6}$$

where $\mathcal{S} = \mathrm{diag}(\beta_{ii})$ is a diagonal matrix with the singular values in it.

One can likewise write

$$(\mathbf{G}\mathbf{G}^T)^{-1} = (\mathcal{V}\mathcal{S}^T\mathcal{U}^T\mathcal{U}\mathcal{S}\mathcal{V}^T)^{-1}. \tag{A7}$$

Using the equalities $\mathcal{U}^T = \mathcal{U}^{-1}$ and $\mathcal{V}^T = \mathcal{V}^{-1}$, one has

$$(\mathbf{G}\mathbf{G}^T)^{-1} = \mathcal{V}\mathcal{S}^{-T}\mathcal{S}^{-1}\mathcal{V}^T, \tag{A8}$$

from which the kth element of the diagonal can be written as

$$\sigma_k^2 = \sum_i \mathcal{V}_{k,i}\frac{1}{\beta_{ii}^2}\mathcal{V}_{ik}^T \tag{A9}$$

$$= \sum_i \mathcal{V}_{ki}\frac{1}{\beta_{ii}^2}\mathcal{V}_{ki} \tag{A10}$$

$$= \sum_i \frac{\mathcal{V}_{ki}^2}{\beta_{ii}^2}, \tag{A11}$$

which is a more straightforward demonstration but otherwise equivalent to that of Feeney (2001).

We caution the reader, however, that the PWFS noise propagation coefficients cannot be quoted in units of angle-on-sky, because its measurements, unlike the case for the SH-WFS, are not straight wavefront gradients (or slopes).

APPENDIX B: VARIATIONS AROUND WFS AND CONTROLLER CHOICES

In addition to the contrast plots provided in the main text, Figs B1, B2 and B3 depict the contrast ratios obtained as a function of the stellar magnitude and frame-rate for different combinations of WFS and controller (for a constant pure loop delay of 3 ms).

Fig. B1 shows the contrast improvements by increasing the frame rate for a SH-based system employing the DKF controller. Gains are observed in a butterfly-shaped region for brighter stars, but this vanishes as noise becomes the dominant factor for fainter stars. As expected, the relative gains of using predictive control (the DKF) lower as we use higher frame-rates.

If we now look into the contrast gains of just replacing the SH-WFS by a PWFS, but keeping a single-gain integral controller, we find the results in Fig. B2. Here the improvement achieved is less striking than before, but still up to a factor 10x in the outer regions of the AO control radius, because the key difference is PWFS aliasing propagation.

Finally, Fig. B3 shows the improvement from only increasing the frame-rate of AO loops with a SH-WFS and integral controller. Compared with Fig. 10, the contrast improvements are not nearly as spectacular, because the controller adds no predictive knowledge to the wavefront estimation in order to further improve the contrast at small separations.

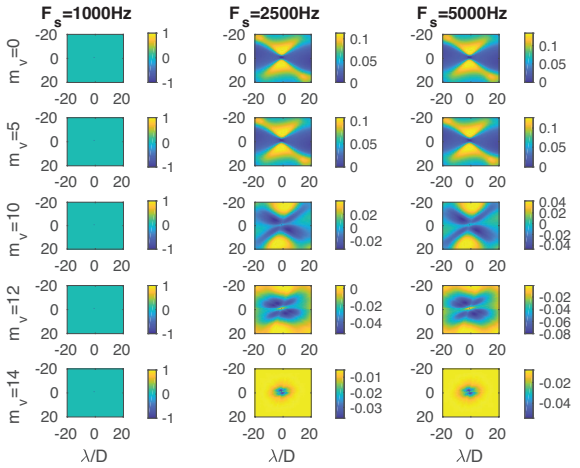

Figure B1. Contrast ratio improvement map (log scale) as a function of the frame-rate and stellar magnitude for an AO loop featuring a SH-WFS and a DKF controller.

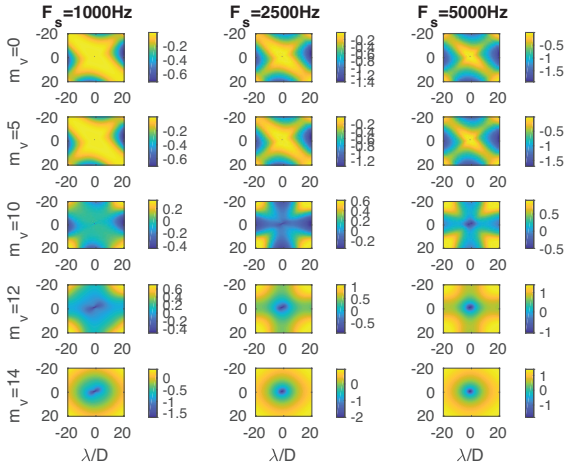

Figure B2. Contrast ratio improvement map (log scale) as a function of the frame-rate and stellar magnitude for PWFS-based versus a SH-based AO loop, both with a LS reconstructor and integral controller.

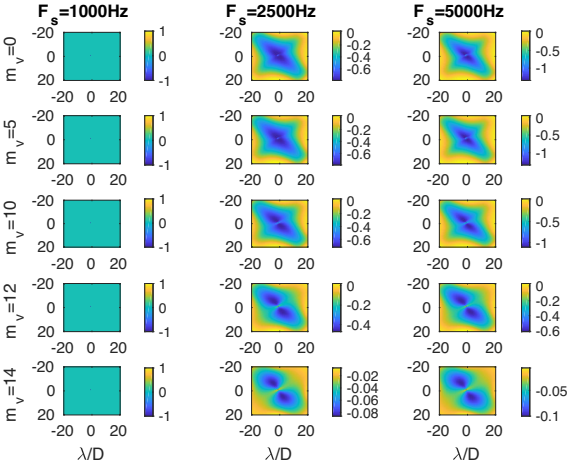

Figure B3. Contrast ratio improvement map (log scale) as a function of the frame-rate and stellar magnitude for an AO loop featuring a SH-WFS and a scalar-gain integrator.

This paper has been typeset from a TEX/LATEX file prepared by the author.