# NGS modes in laser tomography AO systems

Carlos M. Correia

July 21, 2015

**Abstract**

This note reviews the NGS modes modelling for laser tomography AO systems. Three models are presented:
*i)* tilt-tomography using a combination of tilt and high-altitude quadratic modes that produce pure tilt through cone projected ray-tracing through the wave-front profiles;
*ii)* a spatio-angular MMSE tilt estimation anywhere in the field that is more general.
*iii)* the (straightforward) generalisation to dynamic controllers using near-Markovian time-progression models from [Correia et al., 2015].

These controllers will be used for the HARMONI NGS modes.

# Contents

# 1 Fundamentals of the LQG control approach

The linear-quadratic (LQ) and linear-quadratic-Gaussian (LQG) approaches are two standard design tools in optimal control theory [Bar-Shalom and Tse, 1974, Anderson and Moore, 1995a] that stand on a state-space formalism.

This formalism consists in mathematically modelling a physical system as a set of inputs, outputs and state variables related by first order difference equations (discrete case). States are considered as column vectors and the algebraic and difference equations are written in matrix form. With respect to the classical single-input single-output (SISO) systems for which frequency analysis is often preferred (in the form of transfer-functions) using state-spaces is a very convenient manner to represent multiple-input multiple-output (MIMO) systems with real-time computational delays. In an AO context, phase (and its temporal dynamics) along with the WFS measurements and DM commands can be explicitly included in the model.

The initial LQG setup relies on three essential requirements: 1) the system is linear, 2) the criterion is quadratic and 3) the disturbances are zero-mean, white and Gaussian-distributed. Consider the following general discrete-time linear system of the order $n$, $\forall k \geq 0$,

$$\begin{cases} \mathbf{x}_{k+1} &= \mathcal{A}_\mathrm{d}\mathbf{x}_k + \mathcal{B}_\mathrm{d}\mathbf{u}_k + \mathbf{\Gamma} v_k \\ \mathbf{s}_k &= \mathcal{C}_\mathrm{d}\mathbf{x}_k + \mathcal{D}_\mathrm{d}\mathbf{u}_k + w_k \end{cases}, \tag{1}$$

The variables $v$ and $w$ are assumed to be Gaussian white noises, *i.e.* sequences of independent and identically distributed vector-valued Gaussian variables, with distributions $v_k \sim \mathcal{N}(0, \Sigma_v)$ and $w_k \sim \mathcal{N}(0, \Sigma_w)$.

In Eq. (1) the notion of state gathers both the disturbance and mirror states. The general system of Eq. (1) defines

1. a state evolution equation, where the state temporal dynamics is characterised by a state transition matrix $\mathcal{A}_\mathrm{d}$ and an input vector $\mathcal{B}_\mathrm{d}$ that brings about the effect of the control decisions $\mathbf{u}_k$ at instants $t = kT_s$ and state noise $v_k$ a zero-mean white noise with covariance matrix $\Sigma_v$. This latter component will be used to account for the random behaviour of disturbance;

2. a measurement equation indicating which information is available at $t = kT_s$ for computing the control $\mathbf{u}_k$. The measurement $z_k$ is assumed to be the sum of a linear function of the internal state $\mathbf{x}_k$ corrupted with additive zero-mean Gaussian white noise $w_k \sim \mathcal{N}(0, \Sigma_w)$ independent of $v_k$ (and hence of $\mathbf{x}_k$).

## 1.1 Optimality criteria

In what follows, the minimisation of the pupil-integrated mean-square residual phase after AO correction is considered. Minimising the variance of $\boldsymbol{\psi}^{\mathsf{res}}$ results in the maximisation of the Strehl-ratio (SR) [Herrmann, 1992] leading to the continuous-time criterion

$$J^{\mathrm{c}}(u) \triangleq \lim_{\tau \to +\infty} \frac{1}{\tau} \int_0^{\tau} \left\| \boldsymbol{\psi}^{\mathsf{res}}(t) \right\|^2 dt = \lim_{\tau \to +\infty} \frac{1}{\tau} \int_0^{\tau} \left\| \boldsymbol{\psi}^{\mathsf{tur}}(t) - \boldsymbol{\psi}^{\mathsf{cor}}(t) \right\|^2 dt, \tag{2}$$

where the residual phase is the difference between the turbulent and correction phases, $\boldsymbol{\psi}^{\mathsf{res}} = \boldsymbol{\psi}^{\mathsf{tur}} - \boldsymbol{\psi}^{\mathsf{cor}}$.

Define $J^{\mathrm{c}}(u)_k$ as the average value of the residual phase variance between sampling instants $t = kT_s$ and $t = (k+1)T_s$,

$$J^{\mathrm{c}}(u)_k \triangleq \frac{1}{T_s} \int_{kT_s}^{(k+1)T_s} \left\| \boldsymbol{\psi}^{\mathsf{res}}(t) \right\|^2 dt = \frac{1}{T_s} \int_{kT_s}^{(k+1)T_s} \left\| \boldsymbol{\psi}^{\mathsf{tur}}(t) - \boldsymbol{\psi}^{\mathsf{cor}}(t) \right\|^2 dt. \tag{3}$$

With this notation, the continuous-time criterion of Eq.(2) can be rewritten as an infinite-horizon average over all temporal frames of duration $T_s$

$$J^{\mathrm{c}}(u) = \lim_{M \to +\infty} \frac{1}{M} \sum_{k=0}^{M} J^{\mathrm{c}}(u)_k. \tag{4}$$

Note also that by construction $J^{\mathrm{c}}(u)_k$ depends only on control decisions made up to time $t = kT_s$, in other words on the discrete sequence $\mathbf{u}_0, \ldots, \mathbf{u}_k$.

Under standard assumptions, the criterion $J^{\mathrm{c}}(u)$ in Eq.(2) is almost surely (a.s.) equal to the mathematical expectation of $J^{\mathrm{c}}(u)_k$

$$J^{\mathrm{c}}(u) \overset{a.s.}{=} \mathrm{E}\left(J^{\mathrm{c}}(u)_k\right) \tag{5}$$

## 1.2 Step-by-step solution

Within a LQ/LQG framework, the problem at hand can be phrased as the determination of the control decisions $\mathbf{u}_k$ $\forall k > 0$ that minimise a quadratic criterion in the form

$$J^{\mathrm{d}}(u) = \lim_{M \to +\infty} \frac{1}{M} \sum_{k=0}^{M-1} \left( \mathbf{x}_k^{\mathsf{T}} Q \mathbf{x}_k + \mathbf{u}_k^{\mathsf{T}} R \mathbf{u}_k + 2\mathbf{x}_k^{\mathsf{T}} S \mathbf{u}_k \right), \tag{6}$$

where $Q$, $S$ and$R$ are weighting matrices to be obtained from the development of the AO MV criterion. To define a well-posed MV problem, it is necessary that Eq. (6) corresponds a a proper quadratic form, *i.e.*

$$P \triangleq \begin{pmatrix} Q & S \\ S^{\mathsf{T}} & R \end{pmatrix} \geq 0. \tag{7}$$

This positiveness condition is in turn equivalent to $Q = Q^{\mathrm{T}}$, $Q - SR^{-1}S^{\mathsf{T}} \geq 0$ $S$ and $R = R^{\mathsf{T}} > 0$.

Suppose that the state $\mathbf{x}_k$ $\forall k > 0$ is perfectly known. The solution to this *complete state information* (CSI) problem can be expressed in state feedback form as

$$\mathbf{u}_k = -\mathcal{K}_\infty \mathbf{x}_k, \tag{8}$$

where $\mathcal{K}_\infty$ is found from the solution of a control Riccati equation that involves the matrices $Q$, $R$ and $S$ in Eq. (6).

In the general and more realistic *incomplete state information* (ISI) case the *separation principle* [Bar-Shalom and Tse, 1974, Anderson and Moore, 1995a] applies: the optimal control has the same form as Eq. (8), where the state is replaced by its conditional expectation, denoted $\widehat{\mathbf{x}}$. Such state is estimated from the noisy and possibly delayed measurements.

In other words, the optimal control decisions are retrieved from

$$\mathbf{u}_k = -\mathcal{K}_\infty \widehat{\mathbf{x}}_{k|k}, \tag{9}$$

using this estimate as if it were the true state in the state-feedback control. The subscript in the form $(k|k-1)$ means the conditional expectation of $\mathbf{x}_k$ with respect to statistical priors and measurements $\mathcal{Y}_k \triangleq \{y_0, \ldots, y_k\}$.

Provided one has a linear model describing the state's dynamics, $\widehat{\mathbf{x}}_{k|k-1}$ is obtained recursively from the output of a Kalman filter with gain $\mathcal{L}^{\mathrm{opt}}$ [Anderson and Moore, 1995a]. The state estimate is obtained from the recursive equation

$$\widehat{\mathbf{x}}_{k|k-1} = \mathcal{A}_{\mathrm{d}} \widehat{\mathbf{x}}_{k-1|k-2} + \mathcal{B}_{\mathrm{d}} \mathbf{u}_{k-1} + \mathcal{L}_\infty \left( \mathbf{s}_{k-1} - \widehat{\mathbf{s}}_{k-1|k-2} \right), \tag{10}$$

where the gain

$$\mathcal{L}_\infty = \mathcal{A}_{\mathrm{d}} \mathbf{\Sigma} \mathcal{C}_{\mathrm{d}}^{\mathsf{T}} \left( \mathcal{C}_{\mathrm{d}} \mathbf{\Sigma} \mathcal{C}_{\mathrm{d}}^{\mathsf{T}} + \mathbf{\Sigma}_w \right)^{-1} \tag{11}$$

is obtained from

$$\mathbf{\Sigma} = \mathcal{A}_{\mathrm{d}} \mathbf{\Sigma} \mathcal{A}_{\mathrm{d}}^{\mathsf{T}} + \mathbf{\Sigma}_v - \mathcal{A}_{\mathrm{d}} \mathbf{\Sigma} \mathcal{C}_{\mathrm{d}}^{\mathsf{T}} \left( \mathcal{C}_{\mathrm{d}} \mathbf{\Sigma} \mathcal{C}_{\mathrm{d}}^{\mathsf{T}} + \mathbf{\Sigma}_w \right)^{-1} \mathcal{C}_{\mathrm{d}} \mathbf{\Sigma} \mathcal{A}_{\mathrm{d}}^{\mathsf{T}} \tag{12}$$

In summary, the full LQG solution is found in two steps: first solve a deterministic control problem establishing the optimal regulator in the CSI case, and then solve a MV estimation problem to find $\widehat{\mathbf{x}}_{k|k-1}$ in the ISI case .

# 2 LQG controller for AO systems with infinitely fast mirrors

## 2.1 Solution in the *complete-state information* case

When the DM's transient response is negligible compared with the sampling period $T_s$ so that its dynamics can be reduced to its DC gain matrix $\mathbf{N} \in \Re^{n\times m}$ (the so-called "DM influence matrix") then

$$\boldsymbol{\psi}^{\mathsf{cor}}(t) = \mathbf{N}\mathbf{u}_k, \quad \forall \;\; kT_s \leq t < (k+1)\,T_s. \tag{13}$$

As shown in [Kulcsár et al., 2006], the globally optimal MV control - that is, the discrete control minimising the continuous-time criterion $J^{\mathrm{c}}(u)$ - is obtained by minimising the degenerate (because independent from the phase temporal evolution model, just the inter-sample phase mean value is used) equivalent discrete-time MV criterion[Le Roux et al., 2004, Petit et al., 2004], denoted by $J^{\mathrm{d}}_{\mathrm{ifm}}$, with subscript standing for *infinitely fast mirror*,

$$J^{\mathrm{d}}_{\mathrm{ifm}}(u) = \lim_{M\to+\infty} \frac{1}{M} \sum_{k=0}^{M-1} \|\overline{\boldsymbol{\psi}}^{\mathsf{tur}}_{k+1} - \mathbf{N}\mathbf{u}_k\|^2, \tag{14}$$

where $J^{\mathrm{c}}(u)$ relates to $J^{\mathrm{d}}_{\mathrm{ifm}}(u)$ by

$$J^{\mathrm{c}}(u) = J^{\mathrm{d}}_{\mathrm{ifm}}(u) + \delta J(T_s), \tag{15}$$

with $\delta J(T_s)$ the insurmountable error due to the use of averaged variables instead of continuous ones [Petit, 2006]

$$\delta J(T_s) = \lim_{M\to+\infty} \frac{1}{M} \sum_{k=0}^{M-1} \left( \frac{1}{T_s} \int_{kT_s}^{(k+1)\;T_s} \left\| \boldsymbol{\psi}^{\mathsf{tur}}(t) - \overline{\boldsymbol{\psi}}^{\mathsf{tur}}_{k+1} \right\|^2 dt \right). \tag{16}$$

Symbol $\bar{\cdot}$ represents time average, *i.e.*

$$\overline{\boldsymbol{\psi}}_k \triangleq \int_{(k-1)T_s}^{kT_s} \boldsymbol{\psi}(\tau)\mathrm{d}\tau \tag{17}$$

For the commonly used frame-rates of several hundred Hertz, $\delta J(T_s)$ is small and normally accounted for in the total error budget. The absolute error $\delta J(T_s)$ can be computed from

$$\delta J(T_s) = \int_{-\infty}^{\infty} \left(1 - |sinc(\kappa T_s)|^2\right) PSD_{TT}(\kappa)\mathrm{d}f. \tag{18}$$

for three combinations of atmospheric and wind conditions. At a nominal sampling frequency of 1500Hz, in bad conditions the insurmountable error is below 0.01 mas$^2$. This is below 1% the achieved performance (if taken to be 4 mas rms in general) and it is the improvement a fractional-delay controller is to bring about.

For the sake of simplicity, it shall be assumed that the DM influence matrix $\mathbf{N}$ is full column rank, in other words that $\mathbf{N}^\mathsf{T}\mathbf{N}$ is invertible. Since $J^{\mathrm{d}}_{\mathrm{ifm}}(u)$ relates to $J^{\mathrm{c}}(u)$ by a term independent of the control decision $\mathbf{u}$ in Eq. (15), it follows that $\arg\min_{\mathcal{U}} J^{\mathrm{c}}(u) = \arg\min_{\mathcal{U}} J^{\mathrm{d}}_{\mathrm{ifm}}(u)$, with $\mathcal{U}$ the set of admissible controls, which establishes the equivalence of both criteria.

The optimal control decisions $\mathbf{u}_k$ in the CSI case are hence given by

$$\mathbf{u}_k = \Theta\overline{\boldsymbol{\psi}}_{k+1} = \left(\mathbf{N}^\mathsf{T}\mathbf{N}\right)^{-1}\mathbf{N}^\mathsf{T}\overline{\boldsymbol{\psi}}^{\mathsf{tur}}_{k+1}, \tag{19}$$

with $\Theta \in \Re^{m\times n}$ the orthogonal projector of phase onto the DM's space.

Two interesting features of the CSI control are worth noting: 1) it does not depend on any particular assumption on the turbulent phase and 2) it effectively redefines *complete state information* as the advance knowledge not of the whole turbulent phase trajectory over the next sampling interval, but rather of its average value $\overline{\boldsymbol{\psi}}^{\mathsf{tur}}_{k+1}$. The latter is specially interesting for the ISI case covered next.

## 2.2 Solution in the *incomplete state information* case

In the *incomplete state information* (ISI) case the state is estimated from WFS measurements. These measurements can also be expressed as functions of the average phase $\overline{\boldsymbol{\psi}}^{\mathsf{tur}}$ since light flux emanating from a guide star is commonly integrated during the detector exposure time. Assuming a linear response, their output can be modeled to a first approximation as

$$\mathbf{s}_k = \mathbf{G}\overline{\boldsymbol{\psi}}^{\mathsf{tur}}_k - \mathbf{G}\mathbf{N}\mathbf{u}_{k-1} + w_k \tag{20}$$

where $w_k$ is a zero-mean, Gaussian white noise $w_k \sim \mathcal{N}(0, \boldsymbol{\Sigma}_w)$. The assumptions on the WFS read-out and processing delays are not limiting - the approach presented here could easily be adapted to cope with a different temporal diagram. In Eq. (20), the linear operator $\mathbf{G} \in \Re^{s\times n}$ is a device-specific phase-to-WFS influence matrix and $N$ is defined in Eq. (13).

As a consequence, *any* stochastic model of $\overline{\boldsymbol{\psi}}^{\mathsf{tur}}$ will enable to construct an "exhaustive control-oriented" model for the AO loop - *i.e.*, one from which the globally optimal MV control can be derived using standard LQG procedures.

Assuming that $\overline{\boldsymbol{\psi}}^{\mathsf{tur}}$ is a wide sense stationary random process, the *separation principle* applies; the MV control then becomes

$$\mathbf{u}_k = \Theta \widehat{\overline{\boldsymbol{\psi}}}^{\mathsf{tur}}_{k+1|\mathcal{Y}_k}, \tag{21}$$

where $\widehat{\overline{\boldsymbol{\psi}}}^{\mathsf{tur}}_{k+1|\mathcal{Y}_k}$ denotes the conditional expectation of $\overline{\boldsymbol{\psi}}_{k+1}$ with respect to the sequence of all measurements available at $t = kT_s$, namely $\mathcal{Y}_k = \{y_0, \dots, y_k\}$.

Any state-space model meant to describe these measurements should accommodate $\overline{\boldsymbol{\psi}}^{\mathsf{tur}}_{k+1}$ to compute the optimal control decisions in Eq. (19) and further $(\overline{\boldsymbol{\psi}}^{\mathsf{tur}}_k, \mathbf{u}_{k-1})$ in Eq. (20). Consider the average phase temporal evolution if given by a first order auto-regressive model

$$\overline{\boldsymbol{\psi}}^{\mathsf{tur}}_{k+1} = \mathcal{A}_{\mathsf{tur}} \overline{\boldsymbol{\psi}}^{\mathsf{tur}}_k + v_k. \tag{22}$$

Though these simplistic models cannot thoroughly represent Taylor's frozen flow hypothesis, they are accurate enough to grasp the spatial (inter-modal) statistics and the short term correlations.

All gathering in a complete state, the full discrete-time state-space becomes

$$\begin{pmatrix} \overline{\boldsymbol{\psi}}^{\mathsf{tur}}_{k+2} \\ \overline{\boldsymbol{\psi}}^{\mathsf{tur}}_{k+1} \\ \mathbf{u}_k \end{pmatrix} = \begin{pmatrix} \mathcal{A}_{\mathsf{tur}} & 0 & 0 \\ \mathbf{I} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \overline{\boldsymbol{\psi}}^{\mathsf{tur}}_{k+1} \\ \overline{\boldsymbol{\psi}}^{\mathsf{tur}}_{k} \\ \mathbf{u}_{k-1} \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ \mathbf{I} \end{pmatrix} \mathbf{u}_k + \begin{pmatrix} \mathbf{I} \\ 0 \\ 0 \end{pmatrix} v_k, \tag{23}$$

with

$$\mathbf{s}_k = \mathbf{G} \begin{pmatrix} 0 & \mathbf{I} & -\mathbf{N} \end{pmatrix} \mathbf{x}_k + w_k. \tag{24}$$

The control decisions are found from

$$\mathbf{u}_k = \Theta \begin{pmatrix} \mathbf{I} & 0 & 0 \end{pmatrix} \widehat{\mathbf{x}}_{k|k}, \tag{25}$$

where the estimation version of the controller uses information up to $\mathbf{s}_k$. This ensures that the loop delays are correctly taken into account in the complete model of Eqs. (23-25).

Expanding Eq. (14) and comparing it to Eq. (6) the LQ weighting matrices are readily

$$Q = T^{\mathsf{T}} T \geq 0, \quad S = -\mathbf{N} T^{\mathsf{T}}, \quad R = \mathbf{N}^{\mathsf{T}} \mathbf{N} > 0, \tag{26}$$

and $T = \begin{pmatrix} \mathbf{I} & 0 & 0 \end{pmatrix}$. This yields an equivalent discrete-time criterion of the form

$$J^{\mathrm{d}}_{\mathrm{ifm}}(u) = \lim_{M \to +\infty} \frac{1}{M} \sum_{k=0}^{M-1} \left( \begin{pmatrix} \overline{\boldsymbol{\psi}}^{\mathsf{tur}}_{k+1} \\ \mathbf{u}_k \end{pmatrix}^{\mathsf{T}} \begin{pmatrix} \mathbf{I} & -\mathbf{N} \\ -\mathbf{N}^{\mathsf{T}} & \mathbf{N}^{\mathsf{T}}\mathbf{N} \end{pmatrix} \begin{pmatrix} \overline{\boldsymbol{\psi}}^{\mathsf{tur}}_{k+1} \\ \mathbf{u}_k \end{pmatrix} \right). \tag{27}$$

In this case, the optimal control gain, as shown above, can be computed directly as an orthogonal projection. The optimal observer gain (Kalman gain), on the other hand, is computed from an estimation Riccati equation.

## 2.3 Extension to vibration suppression

Let the standard second-order differential equation

$$\ddot{\theta} + 2\xi\omega_0^2\dot{\theta} + \omega_0^2\theta = \nu \tag{28}$$

where $\theta$ is the vibration signal, $\xi$ is a damping factor (defines the width and overshoot of the vibration peak) and $\omega_0$ is the natural oscillatory frequency in rad/s. $\nu$ is a continuous excitation signal, chosen here to be a zero-mean Gaussian white noise.

The solution to the ordinary differential-equation is an oscillatory response in the form

$$\theta = \frac{e^{-\xi\omega_0 t}}{\sqrt{1-\xi^2}} sin\left(2\pi\sqrt{1-\xi^2}\omega_0 t + \Delta\varphi\right) \tag{29}$$

where $\Delta\varphi$ is a phase shift found from the initial conditions set to Eq. (28).

[Petit et al., 2008] and [Meimon et al., 2010] used a discretised version of the vibration given by (28) that led to the $\mathcal{Z}-$domain model

$$H(z) = \frac{\varepsilon_k}{1 - 2e^{-\xi\omega_0^2 T_s} cos\left(\omega_0 T_s\sqrt{1-\xi^2}\right) z^{-1} + e^{-2\xi\omega_0^2 T} z^{-2}} \tag{30}$$

where $T_s$ is the sampling interval.

The correspondence with a second-order auto-regressive formulation is straightforward (using standard $\mathcal{Z}-$transform properties, [Oppenheim and Willsky, 1997])

$$\theta_{k+1} = \alpha_1\theta_k + \alpha_2\theta_{k-1} + \varepsilon_k \tag{31a}$$

$$\alpha_1 = 2e^{-\xi\omega_0^2 T_s} cos\left(\omega_0 T\sqrt{1-\xi^2}\right) \tag{31b}$$

$$\alpha_2 = e^{-2\xi\omega_0^2 T_s} \tag{31c}$$

where $\epsilon_k$ is the process excitation noise, considered zero-mean, Gaussian-distributed and temporally white.

# 3 NGS modes in laser-tomography AO: LTAO and MCAO

The NGS modes in laser-tomography AO are defined as the null modes of the high-order LGS measurement space, i.e., modes that produce average slope $\neq 0$, but that due to the LGS tilt indetermination, cannot be measured by the latter. In other words, the null space can be thought as the combination of all the modes that have non-null projection onto the angle-of-arrival ( = not just Zernike tip and tilt but also higher order Zernike modes).

## 3.1 Measurement model with time-averaged variables

Assume the following measurement model

$$\boldsymbol{s}_k = \int_{(k-1)T_s}^{kT_s} \mathbf{G} \left(\boldsymbol{\psi}(\tau)^{\mathsf{tur}} - \boldsymbol{\psi}^{\,\mathsf{cor}}(\tau)\right) \mathrm{d}\tau + \boldsymbol{\eta}_k \tag{32a}$$

$$= \mathbf{G}\left(\overline{\boldsymbol{\psi}_k^{\mathsf{tur}}} - \overline{\boldsymbol{\psi}_k^{\mathsf{cor}}}\right) + \boldsymbol{\eta}_k \tag{32b}$$

$$= \mathbf{G}\overline{\boldsymbol{\psi}_k^{\mathsf{res}}} + \boldsymbol{\eta}_k \tag{32c}$$

where $\boldsymbol{s}_k \in \Re^{12\times 1}$ are the $T_s$-averaged slopes over each OIWFS sub-aperture and over the integration time $T_s$, $\mathbf{G} \in \mathbb{R}^{12\times 9}$ is the wave-front-to-measurements matrix, 'dm' and 'res' stand respectively for correction and residual phase; $\boldsymbol{\eta}_k$ is a zero-mean Gaussian-distributed spectrally white noise vector with known covariance matrix $\boldsymbol{\Sigma}_\eta$ – $\boldsymbol{\eta} \sim \mathcal{N}(0, \boldsymbol{\Sigma}_\eta)$.

The modal matrix $\mathbf{G}$ translates modal coefficients of TT, TT and TTFA modes into average slopes over the illuminated sub-region of each sub-aperture

$$
\begin{aligned}
\mathbf{G} \triangleq \begin{pmatrix} \mathbf{G}_{\mathsf{TT1}} \\ \mathbf{G}_{\mathsf{TT2}} \\ \mathbf{G}_{\mathsf{TTFA}} \end{pmatrix}
&= \begin{pmatrix}
\gamma_{\mathsf{T}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & \gamma_{\mathsf{T}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & \gamma_{\mathsf{T}} & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & \gamma_{\mathsf{T}} & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & \gamma_{\mathsf{T}} & 0 & \gamma_{\mathsf{F}} & \gamma_{\mathsf{A}} & \gamma_{\mathsf{A}} \\
0 & 0 & 0 & 0 & 0 & \gamma_{\mathsf{T}} & \gamma_{\mathsf{F}} & \gamma_{\mathsf{A}} & -\gamma_{\mathsf{A}} \\
0 & 0 & 0 & 0 & \gamma_{\mathsf{T}} & 0 & -\gamma_{\mathsf{F}} & \gamma_{\mathsf{A}} & -\gamma_{\mathsf{A}} \\
0 & 0 & 0 & 0 & 0 & \gamma_{\mathsf{T}} & \gamma_{\mathsf{F}} & -\gamma_{\mathsf{A}} & -\gamma_{\mathsf{A}} \\
0 & 0 & 0 & 0 & \gamma_{\mathsf{T}} & 0 & -\gamma_{\mathsf{F}} & -\gamma_{\mathsf{A}} & -\gamma_{\mathsf{A}} \\
0 & 0 & 0 & 0 & 0 & \gamma_{\mathsf{T}} & -\gamma_{\mathsf{F}} & -\gamma_{\mathsf{A}} & \gamma_{\mathsf{A}} \\
0 & 0 & 0 & 0 & \gamma_{\mathsf{T}} & 0 & \gamma_{\mathsf{F}} & -\gamma_{\mathsf{A}} & \gamma_{\mathsf{A}} \\
0 & 0 & 0 & 0 & 0 & \gamma_{\mathsf{T}} & -\gamma_{\mathsf{F}} & \gamma_{\mathsf{A}} & \gamma_{\mathsf{A}}
\end{pmatrix},
\end{aligned}
\tag{33}
$$

For the TTF OIWFS, the average slope produced by the TT on the quarter of the aperture $S_{\frac{1}{4}} = \frac{\pi}{4}$ (aperture units are normalised by the aperture radius) is given by

$$
\begin{aligned}
\gamma_{\mathsf{T}} &= \frac{1}{S_{\frac{1}{4}}} \int_0^1 \int_0^{\sqrt{1-y^2}} \frac{\partial}{\partial x} Z_{2,3}(x,y) \partial x \partial y \\
&= 2
\end{aligned}
\tag{34}
$$

The average slope produced by the focus mode is given by

$$
\begin{aligned}
\gamma_{\mathsf{F}} &= \frac{1}{S_{\frac{1}{4}}} \int_0^1 \int_0^{\sqrt{1-y^2}} \frac{\partial}{\partial x} Z_4(x,y) \partial x \partial y \\
&= \frac{16\sqrt{3}}{3\pi}
\end{aligned}
\tag{35}
$$

whereas for the astigmatisms

$$
\begin{aligned}
\gamma_{\mathsf{A}} &= \frac{1}{S_{\frac{1}{4}}} \int_0^1 \int_0^{\sqrt{1-y^2}} \frac{\partial}{\partial x} Z_{5\cdots6}(x,y) \partial x \partial y \\
&= \frac{8\sqrt{6}}{3\pi}
\end{aligned}
\tag{36}
$$

The signal $\pm$ attached to $\gamma_{\mathsf{F}}$ and $\gamma_{\mathsf{A}}$ in Eq. (33) is a function of the exact quadrant where each sub-aperture is located.

## 3.2 Noise model

### 3.2.1 Diffraction-limited case

In the following, the noise model detailed in [Clare et al., 2006] is used. It is assumed that spots are diffraction limited. Therefore, these equations apply for a Nyquist-sampled spot, *i.e.* with $2 \times 2$ pixels, by other words a quadrant detector.

The noise added to each sub-aperture measurement is given by (in angle rms units)

$$\sigma_\eta = \frac{\theta_b}{\mathsf{SNR}}, [rad] \tag{37}$$

where $\theta_b$ is the effective spot size of the sub-aperture, and SNR is the signal-to-noise ratio of a single sub-aperture. For a quadrant detector, the SNR is given by

$$\mathsf{SNR} = \frac{N_p}{\sqrt{N_p + 4N_b + 4\sigma_e^2}}, \tag{38}$$

where $N_p$ is the number of photo-detection events per sub-aperture, $N_b$ is the number of background photo-detection events per sub-aperture, and $\sigma_e$ is the rms detector read noise per pixel.

In the IR (H band), the NGS images are assumed to contain a diffraction-limited core, for which case the effective spot size is given by [Hardy Eq. 5.13]

$$\theta_b = \frac{3\pi\lambda\sqrt{N_{sa}}}{16D_0}, \tag{39}$$

where $N_{sa}$ is the total number of sub-apertures for the NGS WFS. The $2 \times 2$ NGS WFS is therefore noisier than any of the two single sub-aperture NGS WFS. Note $\theta_b$ is twice that of the latter, since $N_{sa}$ is 4 instead of 1 and that the number of photo-detections per sub-aperture is also cut by a factor of 4 – providing a $\sim 2$ times smaller SNR. One thus ends up with a factor $\sim$ 4 noisier measurement. The same reasoning applied to a general SH WFS leads to standard estimations of the full-aperture gain.

To convert to mas rms, multiply $\sigma_\eta[mas] = 180/\pi \times 3600 \times 1000 \times \sigma_\eta[rad]$

### 3.2.2 Seeing-limited case

For the seeing-limited case, use instead for the photon noise

$$\sigma_\eta^2 = \frac{\pi^2}{2ln(2)} \frac{1}{n_{ph}} \left(\frac{N_T}{N_D}\right)^2 [rad^2], \tag{40}$$

where $N_T$ and $N_D$ are the FWHM of image and sub-aperture respectively. With $N_T = \lambda/r_0$ and $N_D = \lambda/D$ one finds

$$\sigma_\eta^2 = \frac{\pi^2}{2ln(2)}\frac{1}{n_{ph}}\left(\frac{D}{r_0}\right)^2 [rad^2], \tag{41}$$

For the read-out noise we use

$$\sigma_\eta^2 = \frac{\pi^2}{3}\frac{\sigma_e^2}{n_{ph}^2}\left(\frac{N_S^2}{N_D}\right)^2 [rad^2], \tag{42}$$

where $\sigma_e$ is the rms number of photoelectron event per pixel and per frame, $N_S^2$ is the total number of pixels in the CoG computation.

# 4 Static reconstruction

## 4.1 Isoplanatic tilt correction

A first approach consists in averaging the tilt measurements obtained across the field

$$\mathbf{E} = \frac{1}{nGs}\sum_{i=1}^{nGs}\mathbf{G}^\dagger \mathbf{s}_{\boldsymbol{\alpha},i} \tag{43}$$

and/or eventually weigh the partial contributors by $1/\sigma_{\eta,i}^2$, i.e. the inverse of the noise variance level on each direction.

## 4.2 Tilt tomography using virtual DMs

The classical approach to derive tilt anisoplanatism is to consider that field-dependent tilt is a linear combination of aperture-plane tilt with pupil and high-altitude quadratic modes of opposite signs whose resultant through ray-racing produces pure tilt (the quadratic terms vanish) [Flicker and Rigaut, 2002, Ellerbroek and Rigaut, 2001]. A similar approach can still be used in laser tomography provided the quadratic terms are scaled by the cone shrinking factor – this is the general case shown next. At least 3 independent measurements of tilt are required in the field (i.e. 6 measurements) from which 2 tip/tilt and 3 quadratic modes (differential focus and astigmatisms) are estimated. In real systems, since the LGS cannot measure focus neither due to Na-range fluctuations, one of the NGS WFSensors is actually a 2x2 WFS providing focus and astigmatism measurements. We thus end up with a total of 12 measurements to estimate

6 modes – which is somehow sub-optimal since we're not making use of the full information contained in the measurements – as was done more recently in [Gilles et al., 2011].

In the following we assume a split-tomography framework [Gilles and Ellerbroek, 2008] and follow on the footsteps and notation of [Correia et al., 2013].

For natural tomography AO (with tilt-removed high-order WFS measurements) is can be computed from what follows with ease.

Following the above considerations the resulting aperture-plane wave-front (WF) is conveniently expanded onto a truncated orthonormal Zernike polynomial's basis defined over 2 layers. For the NGS modes model, only modes $Z_{2\cdots6}$ are used that correspond to the TT and the quadratic modes of Eq. (44).

The total aperture-plane wave-front error induced by the TT/TA modes at time instant $t$, in direction $\boldsymbol{\theta}$ is given by

$$W(\boldsymbol{\rho},\boldsymbol{\theta},t) = \sum_{i=2}^{6} \alpha_i(t) Z_i\left(\frac{\boldsymbol{\rho}}{R_0}\right) + \sum_{i=2}^{6} \beta_i(t) Z_i\left(\frac{\boldsymbol{\rho}+\boldsymbol{\theta} h_{\text{DM}}}{R_h}\right) \tag{44}$$

where $\boldsymbol{\rho}$ is a two-dimensional space coordinate vector and $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are vectors containing the Zernike coefficients (following the ordering of [Noll, 1976]) defined over the lower and upper DM-conjugate planes of radius $R_0 = D_0/2$ and $R_h = D_h/2$ Note $i \in \{2,\cdots,6\}$ which effectively represents 5 modes; $i = 1$ refers to piston and is disregarded as this mode has no impact on AO performance. Global TT have the corresponding Zernike modes applied to the ground DM only, *i.e.* $\beta_2 = \beta_3 = 0$.

The coefficients $\beta_i$ can be worked out such that $W(\boldsymbol{\rho},\boldsymbol{\theta},t)$ only produces TT in the LGS WFSs (but not on the NGS WFSs). This constraint provides $\beta_i = -r_l^{-2}\alpha_i$, with $r_l$ given by the ratio of the cone-intersected pupil and meta-pupil in the DM conjugate range is

$$r_l = r_n \underbrace{\left(1 - \frac{h_{\text{DM}}}{h_{\text{Na}}}\right)}_{r_c}, \tag{45}$$

where $r_c \triangleq 1 - h_{\text{DM}}/h_{\text{Na}}$ is the shrinking factor of the cone-intersected meta-pupil diameter with respect to the cylinder-intersected meta-pupil diameter at $h_{\text{DM}}$ km, translating the cone effect for a DM conjugated to range $h_{\text{DM}}$ and an LGS at range $h_{\text{Na}}$ and

$$r_n \triangleq \frac{D_0}{D_0 + FoV \times h_{\text{DM}} \times 1000}, \tag{46}$$

normalizes the upper modal coefficient to the particular choice of underlying meta-pupils over which the modes are defined, with $FoV$ the field-of-view in radians, $h_{\text{DM}}$ the conjugate altitude of the upper DM in km.

Setting $\beta_i = -r_l^{-2}\alpha_i$ in Eq. (44) and solving for the aperture-plane field-dependent TT

$$\mathsf{TT}(\boldsymbol{\rho}, \boldsymbol{\theta}, t) = \sum_{i=2}^{3} \xi_i(\theta, t) Z_i\left(\frac{\boldsymbol{\rho}}{R_0}\right) \tag{47}$$

with the field-dependent TT coefficients $\boldsymbol{\xi}(\boldsymbol{\theta}, t)$ [Flicker et al., 2003]

$$\boldsymbol{\xi}(\boldsymbol{\theta}, t) = \begin{pmatrix} 1 & 0 & -r_c^{-2}\frac{2\sqrt{3}h\theta_x}{R_0} & -r_c^{-2}\frac{\sqrt{6}h\theta_y}{R_0} & -r_c^{-2}\frac{\sqrt{6}h\theta_x}{R_0} & 0 \\ 0 & 1 & -r_c^{-2}\frac{2\sqrt{3}h\theta_y}{R_0} & -r_c^{-2}\frac{\sqrt{6}h\theta_x}{R_0} & r_c^{-2}\frac{\sqrt{6}h\theta_y}{R_0} & 0 \end{pmatrix} \boldsymbol{\chi}(t) \tag{48}$$

In Eq. (48) $\boldsymbol{\chi}(t) \in \Re^{6\times 1}$ is a vector with the coefficient of the NGS modes (numerically equal to $\boldsymbol{\alpha}(t)$ in Eq. (44) but not to be confounded with it) plus a sixth coefficient to model pure focus induced by $N_a$-layer that does not produce TT.

Expressing the Zernike polynomials using cartesian coordinates one gets

$$\mathbf{x}_k^{\mathsf{tip}} = 2(x_0, 0); \tag{49a}$$
$$\mathbf{x}_k^{\mathsf{tilt}} = 2(y_0, 0); \tag{49b}$$
$$\mathbf{x}_k^{\Delta F} = \sqrt{3}([x_0^2 + y_0^2]; -[x_c^2 + y_c^2]/r_l^2); \tag{49c}$$
$$\mathbf{x}_k^{\Delta A_0} = \sqrt{6}([x_0^2 - y_0^2]; -[x_c^2 - y_c^2]/r_l^2); \tag{49d}$$
$$\mathbf{x}_k^{\Delta A_{45}} = \sqrt{6}([x_0 y_0]; -[x_c y_c]/r_l^2); \tag{49e}$$

where $x_0$, $y_0$ and $x_c$, $y_c$ are the actuator coordinates on the ground and upper DMs, normalised by the meta-pupil radius at 0 and $h_{\text{DM}}$.

The coefficients $\boldsymbol{\chi}(t)$ can be found from least-squares fitting the tilt and quadratic modes to the atmosphere (in passing, this is by no means different from the DM fitting step to an estimated tomographic phase – in the latter the DM influence functions are used instead of the quadratic modes defined previously).

$$\boldsymbol{\chi} = \arg\min_{\boldsymbol{\chi}} \left\langle |\mathbf{H}_\alpha \boldsymbol{\varphi} - \mathbf{H}_\alpha^\chi \mathbf{H}_{\mathsf{PS2Q}} \boldsymbol{\chi}|^2 \right\rangle_{\text{FoV}} \tag{50}$$

where

$$\boldsymbol{\psi} = \mathbf{H}_\alpha \boldsymbol{\varphi} \tag{51}$$

is the pupil-plane integrated wave-front profile in direction $\boldsymbol{\alpha}$ and

$$\mathsf{TT}_{\boldsymbol{\alpha}} = \mathbf{H}_\alpha^\chi \mathbf{H}_{\mathsf{PS2Q}} \boldsymbol{\chi} \tag{52}$$

is the tip/tilt produced by the TT/TTA modes on those directions.

In addition,

$$\mathbf{H}_{\mathsf{PS2Q}} \triangleq \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1/r_l^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1/r_l^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1/r_l^2 & 0 \end{pmatrix}, \tag{53}$$

providing the 2-layer Zernike modal decomposition of the WF from the NGS modal coefficients.

By setting $\boldsymbol{M} \triangleq \mathbf{H}_\alpha^\chi \mathbf{H}_{\mathsf{PS2Q}}$ one then gets

$$\boldsymbol{\chi} = \left\langle \boldsymbol{M}^\intercal \boldsymbol{M} \right\rangle_{\text{FoV}}^{-1} \left\langle \boldsymbol{M}^\intercal \mathbf{H}_{\boldsymbol{\alpha}} \right\rangle_{\text{FoV}} \tag{54}$$

In practice this is done over a few directions in the FoV with weights applied on them (e.g. Simpson weights).

Although the NGS modes produce only field-dependent TT not seen by the LGS WFS, for a NGS looking upwards through a cylinder and not a cone, focus and astigmatisms are indeed probed.

## 4.3 Tilt tomography with spatio-angular reconstruction in NTAO and LTAO

In the NTAO[1]/LTAO case, the tilt compensation is greatly simplified since an explicit tomographic estimate of tilt is not mandatory. We thus fall in the spatio-angular case whereby the computation of field-dependent covariances is done off-line and applied directly on-line.

We now develop the *minimum mean square error* (MMSE) solution with a simplified measurement model involving the pupil-plane turbulence only

$$\mathbf{s}_{\boldsymbol{\alpha}}(t) = \mathbf{G}\boldsymbol{\psi}_{\boldsymbol{\alpha}}(t) + \boldsymbol{\eta}(t) \tag{55}$$

Assuming $\mathbf{s}$ and $\boldsymbol{\psi}$ are zero-mean and jointly Gaussian, direct application of the MMSE solution to estimate the aperture-plane phase in the $N_\beta$-science

[1] Assume tilt-removed measurements.

directions of interest yields [Anderson and Moore, 1995b]

$$\mathcal{E}\{\boldsymbol{\psi}_{\boldsymbol{\beta}}|\mathbf{s}_{\boldsymbol{\alpha}}\} \triangleq \boldsymbol{\Sigma}_{(\boldsymbol{\psi}_{\boldsymbol{\beta}},\mathbf{s}_{\boldsymbol{\alpha}})}\boldsymbol{\Sigma}_{\mathbf{s}_{\boldsymbol{\alpha}}}^{-1}\mathbf{s}_{\boldsymbol{\alpha}} = \widehat{\boldsymbol{\psi}}_{\boldsymbol{\beta}} \tag{56}$$

where $\mathcal{E}\{X|Y\}$ stands for mathematical expectation of $X$ conditioned to $Y$. Since in general $\boldsymbol{\beta} \neq \boldsymbol{\alpha}$, Eq. (56) follows from $\mathcal{E}\{\boldsymbol{\psi}_{\boldsymbol{\beta}}|\mathbf{s}_{\boldsymbol{\alpha}}\} = \mathcal{E}\{\boldsymbol{\psi}_{\boldsymbol{\beta}}|\mathcal{E}\{\boldsymbol{\psi}_{\boldsymbol{\alpha}}|\mathbf{s}_{\boldsymbol{\alpha}}\}\}$. Matrices $\boldsymbol{\Sigma}_{(\boldsymbol{\psi}_{\boldsymbol{\beta}},\mathbf{s}_{\boldsymbol{\alpha}})}$ and $\boldsymbol{\Sigma}_{\mathbf{s}_{\boldsymbol{\alpha}}}$ are spatio-angular covariance matrices that relate the tilt on direction $\theta$ to that on direction $\beta$. These matrices are computed using formulae in [Whiteley et al., 1998b] from the simulator OOMAO [Conan and Correia, 2014].

Under the spatio-angular approach, it is also possible to compute $\boldsymbol{\psi}(\boldsymbol{\rho}, \boldsymbol{\beta}, t + \Delta)$, i.e., temporally predict the TT ahead in time by adjusting the angles over which the correlations are computed [Correia et al., 2015].

Given that the conditioning relates only to the very last available measurement (as opposed to present and previous measurements), these reconstructors are labelled as *static*. Developing terms in Eq. (56) using Eq. (55) the reconstructor becomes

$$\widehat{\boldsymbol{\psi}}_{\boldsymbol{\beta}} \triangleq \left\langle \boldsymbol{\psi}_{\boldsymbol{\beta}}\boldsymbol{\psi}_{\boldsymbol{\alpha}}^{\mathsf{T}} \right\rangle \mathbf{G}^{\mathsf{T}} \left(\mathbf{G} \left\langle \boldsymbol{\psi}_{\boldsymbol{\alpha}}\boldsymbol{\psi}_{\boldsymbol{\alpha}}^{\mathsf{T}} \right\rangle \mathbf{G}^{\mathsf{T}} + \left\langle \boldsymbol{\eta}\boldsymbol{\eta}^{\mathsf{T}} \right\rangle\right)^{-1} \mathbf{s}_{\boldsymbol{\alpha}} \tag{57}$$

This MMSE reconstructor is dubbed *spatio-angular* (SA) reconstructor on account of the nature of the covariance matrices involved in its definition. It can be seen as a generalisation of the work of Whitely *et al* [Whiteley et al., 1998b] in seeking the optimal anisoplanatic reconstuctor in classical AO to the tomographic, multiple sensor case. It has several convenient features: is much faster to compute off-line than the *explicit tomography* reconstructor for large modal sets and circumvents the truncated expansion on a modal basis.

When phase is expanded onto the Zernike polynomials the spatio-angular cross-covariance functions can be analytically computed for the infinite outer-scale case of turbulence [Valley and Wandzura, 1979, Chassat, 1989]

$$\langle \phi_i(0)\phi_j(\xi)\rangle = 3.895 \left(\frac{D}{r_0}\right)^{\frac{5}{3}} \frac{\int_0^{h_{\max}} C_n^2(h) I_{ij}\left(\frac{\xi h}{R}\right) \mathrm{d}h}{\int_0^{\infty} C_n^2(h)\mathrm{d}h} \tag{58}$$

with $D = 2R$ the telescope diameter, $r_0$ the Fried parameter, $h$ the altitude $\xi$ the angle between the pupils over which the Zernike polynomials are defined, $C_n^2$ the atmospheric vertical profile and $I_{ij}(x)$ a term involving 1-D numerical integration. Equation (58) has been extended for the finite outer-scale case in [Takato and Yamaguchi, 1995] and later extensively used and generalised in [Whiteley et al., 1998a, Whiteley et al., 1998b]. The layered spatial covariance matrix $\left\langle \boldsymbol{\varphi}\boldsymbol{\varphi}^{\mathsf{T}} \right\rangle$ is block-diagonal (layers are independent) and can be found in [Noll, 1976] for the infinite outer scale case and in [Winker, 1991] for the finite case.

## 4.4 Review of reconstructors

- MMSE

$$\mathbf{E} = \boldsymbol{\Sigma}_{(\psi_{\boldsymbol{\beta}},\mathbf{s}_{\boldsymbol{\alpha}})}\boldsymbol{\Sigma}_{\mathbf{s}_{\boldsymbol{\alpha}}}^{-1}\mathbf{s}_{\boldsymbol{\alpha}} \tag{59}$$

- GLAO-like

$$\mathbf{E} = \frac{1}{nGs}\sum_{i=1}^{nGs}\mathbf{G}^{\dagger}\mathbf{s}_{\boldsymbol{\alpha},i} \tag{60}$$

- virtual DM on two layers

$$\mathbf{E} = \mathbf{G}\mathbf{H}_{\boldsymbol{\beta}}\left(\mathbf{H}_{\boldsymbol{\alpha}}^{\intercal}\mathbf{G}^{\intercal}\boldsymbol{\Sigma}_{\eta}^{-1}\mathbf{G}\mathbf{H}_{\boldsymbol{\alpha}}\right)^{-1}\mathbf{H}_{\boldsymbol{\alpha}}^{\intercal}\mathbf{G}^{\intercal}\boldsymbol{\Sigma}_{\eta}^{-1} \tag{61}$$

This reconstructor is a noise-weighted reconstructor. In case of equally noisy measurements it boils down to the simple averaging (GLAO-like) case.

## 4.5 Performance assessment

For the MMSE case, the standard performance assessment formulae apply.

Starting from

$$\sigma^2(\boldsymbol{\beta}) = \left\langle \left\| \boldsymbol{\psi}(\boldsymbol{\beta}) - \widehat{\boldsymbol{\psi}}(\boldsymbol{\beta}) \right\|^2 \right\rangle \tag{62}$$

with $\widehat{\boldsymbol{\psi}}(\boldsymbol{\beta}) = \mathbf{R}\mathbf{s}(\boldsymbol{\alpha})$ we get

$$\sigma^2(\boldsymbol{\beta}) = \mathrm{trace}\{\boldsymbol{\Sigma}_{\boldsymbol{\beta}}\} + \mathrm{trace}\{\mathbf{R}\mathbf{G}\boldsymbol{\Sigma}_{\boldsymbol{\alpha}}\mathbf{G}^{\intercal}\mathbf{R}^{\intercal}\} \tag{63}$$

$$- \mathrm{trace}\{\mathbf{R}\mathbf{G}\boldsymbol{\Sigma}_{\boldsymbol{\beta},\boldsymbol{\alpha}}\} - \mathrm{trace}\{\boldsymbol{\Sigma}_{\boldsymbol{\alpha},\boldsymbol{\beta}}\mathbf{G}^{\intercal}\mathbf{R}^{\intercal}\} \tag{64}$$

$$+ \mathrm{trace}\{\mathbf{R}\boldsymbol{\Sigma}_{\boldsymbol{\eta}}\mathbf{R}^{\intercal}\} \tag{65}$$

We can as well account for temporal errors by estimating instead the following

$$\sigma^2(\boldsymbol{\beta}, T_s) = \left\langle \|\boldsymbol{\psi}(\boldsymbol{\beta}, t + T_s) - \boldsymbol{\psi}(\boldsymbol{\beta}, t)\|^2 \right\rangle \tag{66}$$

leading to similar expressions as before but where a lag equal to $T_s$ is considered when computing the integrals with a displacement between the meta-pupils that increases by $vT_s$ with $\mathbf{v}$ the wind velocity vector ($v = |\mathbf{v}|$).

A more sound way to compute the residuals is by considering still Eq. (68) when the reconstructor has built-in prediction

$$\mathbf{R} = \boldsymbol{\Sigma}_{(\psi_{\boldsymbol{\beta},k+1},\mathbf{s}_{\boldsymbol{\alpha},k})}\boldsymbol{\Sigma}_{\mathbf{s}_{\boldsymbol{\alpha},k}}^{-1} \tag{67}$$

and computing

$$\sigma^2(\boldsymbol{\beta}) = \text{trace}\{\boldsymbol{\Sigma}_{\boldsymbol{\beta}}\} + \text{trace}\{\mathbf{R}\mathbf{G}\boldsymbol{\Sigma}_{\boldsymbol{\alpha}}\mathbf{G}^\intercal\mathbf{R}^\intercal\} \tag{68}$$

$$- \text{trace}\{\mathbf{R}\mathbf{G}\boldsymbol{\Sigma}_{(\boldsymbol{\beta},k+1),(\boldsymbol{\alpha},k)}\} - \text{trace}\{\boldsymbol{\Sigma}_{(\boldsymbol{\alpha},k),(\boldsymbol{\beta},k+1)}\mathbf{G}^\intercal\mathbf{R}^\intercal\} \tag{69}$$

$$+ \text{trace}\{\mathbf{R}\boldsymbol{\Sigma}_{\boldsymbol{\eta}}\mathbf{R}^\intercal\} \tag{70}$$

where I omit the temporal dependence whenever possible due to stationarity.

# 5 Dynamic controller of NGS modes

We'll use the approach laid out in [Correia et al., 2015] in which the states are directly the pupil-integrated TT modes in the GS directions coupled to a state-transition model using a near-Markovian time-progression model. We can thus include information from a large number of layers but never explicitly estimate those layered modal contributions to the final NGS modes in the pupil in the GS directions.

The angular cross-covariance matrices give rise to order-1 Markovian models with non-null temporal cross-spectra (iff more that TT are considered). A comparison to decoupled high-order models (AR for instance) for the NGS-only modes must be conducted. A temporal TF analysis is most welcome – if this problem is tractable.

## 5.1 near-Markovian time-progression model of the 1st-order

The use of auto-regressive models has been given wide attention using the Zernike polynomials' expansion basis set [Petit et al., 2009, Sivo et al., 2014] as well as with Fourier modes for which a complex model encodes perfectly spatial shifts and thus frozen-flow [Poyneer and Véran, 2008].

We
restrict our attention to the near-Markovian first-order time-evolution model [Gavel and Wiberg, 2003] which has delivered best overall performance albeit with increased computational complexity [Piatrou and Roggemann, 2007, Correia et al., 2014, Jackson et al., 2015] (spatial dependence omitted)

$$\boldsymbol{\psi}_{k+\Delta}(\boldsymbol{\alpha}) = \mathbf{A}\boldsymbol{\psi}_k(\boldsymbol{\alpha}) + \boldsymbol{\nu}_k(\boldsymbol{\alpha}) \tag{71}$$

with $\Delta = \tau/T_s$ is a delay in units of sample step $T_s$, where the transition matrix minimizes the quadratic criterion [Correia et al., 2014]

$$\mathbf{A} = \arg\min_{\mathbf{A}'} \left\langle \left\| \boldsymbol{\psi}_{k+\Delta}(\boldsymbol{\alpha}) - \mathbf{A}'\boldsymbol{\psi}_k(\boldsymbol{\alpha}) \right\|^2_{L_2(\Omega)} \right\rangle, \tag{72}$$

The solution to Eq. (72) is

$$\mathbf{A} = \left\langle \boldsymbol{\psi}_{k+\Delta}\boldsymbol{\psi}_k^\intercal \right\rangle \left\langle \boldsymbol{\psi}_k\boldsymbol{\psi}_k^\intercal \right\rangle^{-1}(\boldsymbol{\alpha}) \tag{73}$$

Note that we do not remove piston in this equation although we could following formulae in [Wallner, 1983].

Assuming stationarity, the state excitation noise covariance matrix is found for a first order time-evolution model from the covariance equality (implicit indices dropped out)

$$\left\langle \boldsymbol{\psi}\boldsymbol{\psi}^\intercal \right\rangle = \mathbf{A}\left\langle \boldsymbol{\psi}\boldsymbol{\psi}^\intercal \right\rangle \mathbf{A}^\intercal + \left\langle \boldsymbol{\nu}\boldsymbol{\nu}^\intercal \right\rangle, \tag{74}$$

since $\left\langle \boldsymbol{\psi}_{k+1}\boldsymbol{\psi}_{k+1}^\intercal \right\rangle = \left\langle \boldsymbol{\psi}_k\boldsymbol{\psi}_k^\intercal \right\rangle = \left\langle \boldsymbol{\psi}\boldsymbol{\psi}^\intercal \right\rangle$. The excitation noise covariance matrix is therefore

$$\left\langle \boldsymbol{\nu}\boldsymbol{\nu}^\intercal \right\rangle = \left\langle \boldsymbol{\psi}\boldsymbol{\psi}^\intercal \right\rangle - \mathbf{A}\left\langle \boldsymbol{\psi}\boldsymbol{\psi}^\intercal \right\rangle \mathbf{A}^\intercal \tag{75}$$

The model driving noise covariance matrix $\left\langle \boldsymbol{\nu}\boldsymbol{\nu}^\intercal \right\rangle$ is a key element of the KF design.

## 5.2 State-space representation

In the spatio-angular framework, we will make use of a state vector with the pupil-plane phase integrated over directions $\boldsymbol{\alpha} \in [1, \cdots, N_\alpha]$ at instant $k$ to provide the *Spatio-Angular* LQG formulation.

Selecting thus $\mathbf{x}_k \triangleq \boldsymbol{\psi}_k(\boldsymbol{\alpha})$

$$\boldsymbol{\psi}_k(\boldsymbol{\alpha}) = \begin{pmatrix} \boldsymbol{\psi}(\boldsymbol{\alpha}_1) \\ \cdots \\ \boldsymbol{\psi}(\boldsymbol{\alpha}_{N_\alpha}) \end{pmatrix}_k \tag{76}$$

where the state is a concatenation of phase vector in the $N_\alpha$ guide-star directions, one defines the state space terms for Eq. (71)

$$\left[\begin{array}{c|c|c} \mathcal{A}_{\mathsf{tur}} & \mathcal{B}_{\mathsf{tur}} & \boldsymbol{\Gamma} \\ \hline \mathcal{C}_{\mathsf{tur}} & \mathcal{D}_{\mathsf{tur}} & 0 \end{array}\right] = \left[\begin{array}{c|c|c} \mathbf{A} & 0 & \mathbf{I} \\ \hline \mathbf{G} & -\mathbf{G}z^{-d} & 0 \end{array}\right] \tag{77}$$

The implementation of the KF involves a real-time state update and prediction equations which in the SA case are

$$\widehat{\boldsymbol{\psi}}_{k|k}(\boldsymbol{\alpha}) = \widehat{\boldsymbol{\psi}}_{k|k-1}(\boldsymbol{\alpha}) + \mathcal{H}_\infty\left(\mathbf{s}_k(\boldsymbol{\alpha}) - \mathbf{G}\widehat{\boldsymbol{\psi}}_{k|k-1}(\boldsymbol{\alpha}) + \mathbf{G}\mathbf{u}_{k-d}\right) \tag{78a}$$

$$\widehat{\boldsymbol{\psi}}_{k+1|k}(\boldsymbol{\alpha}) = \mathbf{A}\widehat{\boldsymbol{\psi}}_{k|k}(\boldsymbol{\alpha}) \tag{78b}$$

where $\mathcal{H}_\infty$ is the asymptotic Kalman gain computed from the solution of an estimation Riccati equation [Correia et al., 2010]. The use of the asymptotic value is justified, like elsewhere [Petit et al., 2008] for one seeks long-exposure performance therefore employing the steady-state gain with no loss of performance. Whenever implicit we drop the notation $(\boldsymbol{\alpha})$.

## 5.3 LQG + time-smoothing operation

For low frame-rates when guiding on dim stars replacing the central point estimate by the average over tha sampling interval may lead to further performance gains. There should be no added real-time complexity added, only off-line extra computation.

This up-sampling model may be required anyway one the (way less computationally intensive) TT vibration modes, following [Correia et al., 2012].

The command up-sampling could be easily evaluated with the SA algorithm, since one could estimate

$$\mathbf{u}_{k+\Delta} = \mathbf{F}\boldsymbol{\Sigma}_{\boldsymbol{\psi}_{\boldsymbol{\beta}},\boldsymbol{\psi}_{\boldsymbol{\alpha}}}\boldsymbol{\Sigma}^{-1}_{\boldsymbol{\psi}_{\boldsymbol{\alpha}},\boldsymbol{\psi}_{\boldsymbol{\alpha}}}\mathcal{A}_{\Delta}\widehat{\boldsymbol{\psi}}(\boldsymbol{\alpha})_{k+1|k} \tag{79}$$

$$\tag{80}$$

could be replaced by

$$\begin{aligned}\mathbf{u}_{k+\Delta} &= \mathbf{F}\mathcal{A}'_{\Delta}\boldsymbol{\Sigma}_{\boldsymbol{\psi}_{\boldsymbol{\beta}},\boldsymbol{\psi}_{\boldsymbol{\alpha}}}\boldsymbol{\Sigma}^{-1}_{\boldsymbol{\psi}_{\boldsymbol{\alpha}},\boldsymbol{\psi}_{\boldsymbol{\alpha}}}\widehat{\boldsymbol{\psi}}(\boldsymbol{\alpha})_{k+1|k} &(81)\\ &= \mathbf{T}'_{\Delta}\widehat{\boldsymbol{\psi}}(\boldsymbol{\alpha})_{k+1|k} &(82)\end{aligned}$$

i.e. first estimate the phase in the $\boldsymbol{\beta}$ direction and then make the temporal prediction.

However, this raises a question that has not yet been addressed: the phase estimate is an approximation of the average at the middle of the integration interval. The actual average phase could be computed (at least numerically) by doing

$$\mathbf{u}_{k+\Delta} = \mathbf{F}\overline{\mathcal{A}}'_{\Delta}\boldsymbol{\Sigma}_{\boldsymbol{\psi}_{\boldsymbol{\beta}},\boldsymbol{\psi}_{\boldsymbol{\alpha}}}\boldsymbol{\Sigma}^{-1}_{\boldsymbol{\psi}_{\boldsymbol{\alpha}},\boldsymbol{\psi}_{\boldsymbol{\alpha}}}\widehat{\boldsymbol{\psi}}(\boldsymbol{\alpha})_{k+1|k} \tag{83}$$

with

$$\begin{aligned}\overline{\mathcal{A}}'_{\Delta} &\triangleq \frac{1}{M}\sum_{i=1}^{M-1}\mathcal{A}_{\Delta+i/M} &(84)\\ &= \frac{1}{M}\sum_{i=0}^{M-1}\left\{\left\langle \phi_{k+\Delta}\left[\boldsymbol{\beta}-\mathbf{v}(-T_s/2+i/MT_s)\right]\phi_k^{\mathsf{T}}\right\rangle\right\}\left\langle\phi_k\phi_k^{\mathsf{T}}\right\rangle^{-1} &(85)\\ &= \frac{1}{M}\sum_{i=0}^{M-1}\left\{\left\langle \phi_{k}\left[\boldsymbol{\beta}-\mathbf{v}(\Delta/T_s-T_s/2+i/MT_s)\right]\phi_k^{\mathsf{T}}\right\rangle\right\}\left\langle\phi_k\phi_k^{\mathsf{T}}\right\rangle^{-1} &(86)\end{aligned}$$

which can probably be further simplified.

The time-averaged phase is

$$\overline{\boldsymbol{\psi}}_k = \frac{1}{T_s}\int_{-T_s/2}^{T_s/2} \boldsymbol{\psi}(kT_s + t)\mathrm{d}t = \overline{\mathcal{A}}'\boldsymbol{\psi}_k \tag{87}$$

where

$$\overline{\mathcal{A}}' = \overline{\left\langle \phi_{k+\Delta}\phi_k^{\mathsf{T}}\right\rangle}\left\langle \phi_k\phi_k^{\mathsf{T}}\right\rangle^{-1} \tag{88}$$

where $\overline{\left\langle \phi_{k+\Delta}\phi_k^{\mathsf{T}}\right\rangle}$ is computed in a single-step from $\overline{C}(\boldsymbol{\rho})$ using the averaged covariance function

$$\overline{C}(\boldsymbol{\rho}, T_s) = \sum_{l=1}^{L}\omega_l\frac{1}{T_s}\int_{-T_s/2}^{T_s/2} C(\boldsymbol{\rho} - \mathbf{v}t)\mathrm{d}t \tag{89}$$

or by averaging multiple instances of the covariance matrices as done in Eq. (84).

With this smoothed model one can instead have, still with $\mathbf{x}_k \triangleq \boldsymbol{\psi}_{\boldsymbol{\alpha},k}$

$$\left[\begin{array}{c|c|c} \mathcal{A}_{\mathsf{tur}} & \mathcal{B}_{\mathsf{tur}} & \boldsymbol{\Gamma} \\ \hline \mathcal{C}_{\mathsf{tur}} & \mathcal{D}_{\mathsf{tur}} & 0 \end{array}\right] = \left[\begin{array}{c|c|c} \mathcal{A}_{\boldsymbol{\alpha}} & 0 & \mathbf{I} \\ \hline \mathbf{G}\overline{\mathcal{A}}' & -\mathbf{G}z^{-d} & 0 \end{array}\right] \tag{90}$$

with the smoother commands given in Eq. (83) and $\mathbf{G}\overline{\mathcal{A}}'$ a matrix that produces the WFS gradients from the averaged wave-front over the integration period $T_s$. Matrix $\overline{\mathcal{A}}'$ approaches identity as the integration time decreases.

### 5.4 LQG + multi-rate

It has been show (unpublished though) that the multi-rate case cannot be exploited in the TT measurement out of 2 TT and 1 TTFA SH-WFS for there are only 6 independent vectors to estimate 5 unknowns (2 TT and 3 plate-scale modes).

Multi-rate could be potentially applied if one of the measurements would provide more modes than just linear and quadratic. This could eventually be the case for TT measured close to on-axis in NIR with partial correction by the LGS loop. In that case LIFT would operate and the multi-rate controller can potentially be further explored.

## 6 Split tomography v. integrated tomography

To be written later.

We need here the block diagrams for the spit and integrated tomography

# 7 Design of HARMONI (E-IFU) tilt anisoplanatism controller

To be written later.

## 7.1 Isoplanatic tilt

Wind buffeting on the telescope

## 7.2 Anisoplanatic tilt

Atmospheric tilt

## 7.3 Non-stationary tilt from M4 off-loading

The famous "schlonk" every 5 minutes or so.

# A Average-slope removal from measurements

For E2E simulations , the average slope removal is done in gradient space by projecting off the average x and y slopes.

$$\begin{aligned} s_{ol,TTR} &= s_{ol} - MM_\eta^\dagger s_{ol} \qquad (91)\\ &= (\mathbf{I} - MM_\eta^\dagger)s_{ol} \qquad (92)\end{aligned}$$

where

$$M = \begin{bmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} \end{bmatrix} \qquad (93)$$

a matrix with size of slope measurements by 2 (modes) and

$$M_\eta^\dagger = (M^\intercal \mathbf{\Sigma}_\eta^{-1} M)^{-1} M^\intercal \mathbf{\Sigma}_\eta^{-1} \qquad (94)$$

the noise-weighted pseudo-inverse of $M$ with $\mathbf{\Sigma}_\eta^{-1}$ the inverse measurement noise covariance matrix.

# B LQG controller transfer functions

**LQG transfer functions.**

Let the general transfer

$$\mathbf{y}(z) = \mathbf{C}(z)\mathbf{u}(z). \tag{95}$$

For a LQG using the predictor version $\mathbf{u}_k = -\mathcal{K}_\infty \widehat{\mathbf{x}}_{k+1|k}$ the transfer function $\mathbf{C}_p(z)$ is

$$\mathbf{C}_p(z) = -\left(\mathbf{I} + \mathcal{K}_\infty \Lambda_p z^{-d} \mathcal{L}_\infty \mathcal{D}\right)^{-1} \mathcal{K}_\infty \Lambda_p \mathcal{A}_{\mathsf{tur}} \mathcal{H}_\infty, \tag{96}$$

with

$$\Lambda_p = \left(\mathbf{I} - z^{-1}\left(\mathcal{A}_{\mathsf{tur}} - \mathcal{L}_\infty \mathcal{C}_{\mathsf{tur}}\right)\right)^{-1}. \tag{97}$$

If instead we were to compute $\mathbf{u}_k = -\mathcal{K}_\infty \widehat{\mathbf{x}}_{k|k-1}$ then one would get

$$\mathbf{C}_p(z) = -\left[\mathbf{I} + z^{-1}\mathcal{K}_\infty \Lambda_p \left(z^{-d}\mathcal{L}_\infty \mathcal{D} + \mathcal{B}_{\mathsf{tur}}\right)\right]^{-1} z^{-1}\mathcal{K}_\infty \Lambda_p \mathcal{A}_{\mathsf{tur}} \mathcal{H}_\infty. \tag{98}$$

For a command $\mathbf{u}_k = -\mathcal{K}_\infty \widehat{\mathbf{x}}_{k|k}$,

$$\mathbf{C}_e(z) = -\left\{\mathbf{I} + \mathcal{K}_\infty \Lambda_e \left[z^{-1}(\mathbf{I} + \mathcal{H}_\infty \mathcal{C}_{\mathsf{tur}})\mathcal{B}_{\mathsf{tur}} + z^{-d}\mathcal{H}_\infty \mathcal{D}\right]\right\}^{-1} \mathcal{K}_\infty \Lambda_e \mathcal{H}_\infty, \tag{99}$$

where

$$\Lambda_e = \left(\mathbf{I} - z^{-1}\mathcal{A}_{\mathsf{tur}} + z^{-1}\mathcal{H}_\infty \mathcal{C}_{\mathsf{tur}} \mathcal{A}_{\mathsf{tur}}\right)^{-1}. \tag{100}$$

△

**Demonstration.**

The state-update and state-estimate equations of the LQG are

$$\begin{cases} \widehat{\mathbf{x}}_{k|k} & = \widehat{\mathbf{x}}_{k|k-1} + \mathcal{H}_\infty(\mathbf{s}_k - \mathcal{C}_{\mathsf{tur}}\widehat{\mathbf{x}}_{k|k-1} + \mathcal{D}\mathbf{u}_{k-2}) \\ \widehat{\mathbf{x}}_{k+1|k} & = \mathcal{A}_{\mathsf{tur}}\widehat{\mathbf{x}}_{k|k} + \mathcal{B}_{\mathsf{tur}}\mathbf{U}_k \end{cases}, \tag{101}$$

The second line of Eq. (101) can be likewise written, using $\widehat{\mathbf{X}}_p$ and $\widehat{\mathbf{X}}_e$ the $\mathcal{Z}$-transfers of $\widehat{\mathbf{x}}_{k|k}$ and $\widehat{\mathbf{x}}_{k|k-1}$,

$$z\widehat{\mathbf{X}}_p = \mathcal{A}_{\mathsf{tur}}\widehat{\mathbf{X}}_e + \mathcal{B}_{\mathsf{tur}}\mathbf{U}. \tag{102}$$

I assume next that the state cannot be attained by the commands, i.e. $\mathcal{B}_{\mathsf{tur}} = 0$.

Multipying out by $z^{-1}$ on both sides, one gets

$$\widehat{\mathbf{X}}_p = z^{-1}\mathcal{A}_{\mathsf{tur}}\widehat{\mathbf{X}}_e. \tag{103}$$

With this result and the second line in Eq. (101)

$$\begin{aligned}\widehat{\mathbf{X}}_p &= z^{-1}\mathcal{A}_{\mathsf{tur}}\left(\widehat{\mathbf{X}}_p + \mathcal{H}_\infty\left(\mathbf{S} - \mathcal{C}_{\mathsf{tur}}\widehat{\mathbf{X}}_p + z^{-2}\mathcal{D}\mathbf{U}\right)\right)\\ &= \left(\mathbf{I} - z^{-1}\mathcal{A}_{\mathsf{tur}}\left(\mathbf{I} - \mathcal{H}_\infty\mathcal{C}_{\mathsf{tur}}\right)\right)^{-1} z^{-1}\mathcal{A}_{\mathsf{tur}}\mathcal{H}_\infty\left(z^{-2}\mathcal{D}\mathbf{U} + \mathbf{S}\right).\end{aligned} \tag{104}$$

Using a command $\mathbf{u}_k = -\mathcal{K}_\infty\widehat{\mathbf{x}}_{k+1|k}$

$$\mathbf{U} = -\mathcal{K}_\infty\left(\mathbf{I} - \mathcal{A}_{\mathsf{tur}}\left(\mathbf{I} - \mathcal{H}_\infty\mathcal{C}_{\mathsf{tur}}\right)\right)^{-1}\left(z^{-2}\mathcal{D}\mathbf{U} + \mathcal{A}_{\mathsf{tur}}\mathcal{H}_\infty\mathbf{S}\right). \tag{105}$$

Grouping terms, one gets

$$\mathbf{U} = -\left(\mathbf{I} + \mathcal{K}_\infty\Lambda_p z^{-2}\mathcal{L}_\infty\mathcal{D}\right)^{-1}\mathcal{K}_\infty\Lambda_p\mathcal{A}_{\mathsf{tur}}\mathcal{H}_\infty\mathbf{S}, \tag{106}$$

where

$$\Lambda_p = \left(\mathbf{I} - z^{-1}\mathcal{A}_{\mathsf{tur}}\left(\mathbf{I} - \mathcal{H}_\infty\mathcal{C}_{\mathsf{tur}}\right)\right)^{-1}. \tag{107}$$

The controller transfer function is finally written as

$$\mathbf{C}_p(z) = -\left(\mathbf{I} + \mathcal{K}_\infty\Lambda_p z^{-2}\mathcal{L}_\infty\mathcal{D}\right)^{-1}\mathcal{K}_\infty\Lambda_p\mathcal{A}_{\mathsf{tur}}\mathcal{H}_\infty. \tag{108}$$

The general case with an arbitrary delay and $\mathcal{B}_{\mathsf{tur}} \neq 0$ one gets

$$\mathbf{C}_p(z) = -\left[\mathbf{I} + \mathcal{K}_\infty\Lambda_p\left(z^{-d}\mathcal{L}_\infty\mathcal{D} + \mathcal{B}_{\mathsf{tur}}\right)\right]^{-1}\mathcal{K}_\infty\Lambda_p\mathcal{A}_{\mathsf{tur}}\mathcal{H}_\infty. \tag{109}$$

For a command $\mathbf{u}_k = -\mathcal{K}_\infty\widehat{\mathbf{x}}_{k|k}$, following an analogous reasoning

$$\widehat{\mathbf{X}}_e = \left(\mathbf{I} - z^{-1}\mathcal{A}_{\mathsf{tur}} + z^{-1}\mathcal{H}_\infty\mathcal{C}_{\mathsf{tur}}\mathcal{A}_{\mathsf{tur}}\right)^{-1}\mathcal{A}_{\mathsf{tur}}\mathcal{H}_\infty\left(z^{-2}\mathcal{D}\mathbf{U} + \mathbf{S}\right), \tag{110}$$

yielding

$$\mathbf{U} = -\mathcal{K}_\infty\left(\mathbf{I} - z^{-1}\mathcal{A}_{\mathsf{tur}} + z^{-1}\mathcal{H}_\infty\mathcal{C}_{\mathsf{tur}}\mathcal{A}_{\mathsf{tur}}\right)^{-1}\mathcal{A}_{\mathsf{tur}}\mathcal{H}_\infty\left(z^{-2}\mathcal{D}\mathbf{U} + \mathbf{S}\right) \tag{111}$$

Grouping terms

$$\mathbf{C}_e(z) = -\left(\mathbf{I} + \mathcal{K}_\infty\Lambda_e z^{-2}\mathcal{H}_\infty\mathcal{D}\right)^{-1}\mathcal{K}_\infty\Lambda_e\mathcal{H}_\infty, \tag{112}$$

where

$$\Lambda_e = \left(\mathbf{I} - z^{-1}\mathcal{A}_{\mathsf{tur}} + z^{-1}\mathcal{H}_\infty\mathcal{C}_{\mathsf{tur}}\mathcal{A}_{\mathsf{tur}}\right)^{-1}. \tag{113}$$

The general case when $\mathcal{B}_{\mathsf{tur}} \neq 0$ and arbitrary delay equates to

$$\mathbf{C}_e(z) = -\left\{\mathbf{I} + \mathcal{K}_\infty\Lambda_e\left[z^{-1}(\mathbf{I} + \mathcal{H}_\infty\mathcal{C}_{\mathsf{tur}})\mathcal{B}_{\mathsf{tur}} + z^{-d}\mathcal{H}_\infty\mathcal{D}\right]\right\}^{-1}\mathcal{K}_\infty\Lambda_e\mathcal{H}_\infty, \tag{114}$$

which reduces to the previous case if $\mathcal{B}_{\mathsf{tur}} = 0$ and $d = 2$. ∎

## B.1 Script to use

Use function `LQG_controller_TF.m` to compute above functions. An exemple is included in the function headed to compare the TFs to matlab built-in Z-transform TFs. All cases work perfectly well.

# References


[Anderson and Moore, 1995a] Anderson, B. D. O. and Moore, J. B. (1995a). *Optimal Control, Linear Quadratic Methods*. Dover Publications Inc. 3, 5

[Anderson and Moore, 1995b] Anderson, B. D. O. and Moore, J. B. (1995b). *Optimal Filtering*. Dover Publications Inc. 17

[Bar-Shalom and Tse, 1974] Bar-Shalom, Y. and Tse, E. (1974). Dual effect, certainty equivalence, and separation in stochastic control. *IEEE Transactions on Automatic Control*, 19:494–500. 3, 5

[Chassat, 1989] Chassat, F. (1989). Calcul du domaine d'isoplanetisme d'un systeme d'optique adaptative fonctionnant a travers la turbulence atmospherique. *J. Optics (Paris)*, 20(1):13–23. 17

[Clare et al., 2006] Clare, R. M., Ellerbroek, B. L., Herriot, G., and Véran, J.-P. (2006). Adaptive optics sky coverage modeling for extremely large telescopes. *Appl. Opt.*, 45(35):8964–8978. 12

[Conan and Correia, 2014] Conan, R. and Correia, C. (2014). Object-oriented matlab adaptive optics toolbox. In *Proc. of the SPIE*, volume 9148, pages 91486C–91486C–17. 17

[Correia et al., 2014] Correia, C., Jackson, K., Véran, J.-P., Andersen, D., Lardière, O., and Bradley, C. (2014). Static and predictive tomographic reconstruction for wide-field multi-object adaptive optics systems. *J. Opt. Soc. Am. A*, 31(1):101–113. 19

[Correia et al., 2010] Correia, C., Raynaud, H.-F., Kulcsár, C., and Conan, J.-M. (2010). On the optimal reconstruction and control of adaptive optical systems with mirror dynamics. *J. Opt. Soc. Am. A*, 27(2):333–349. 20

[Correia et al., 2012] Correia, C., Véran, J.-P., and Herriot, G. (2012). Advanced vibration suppression algorithms in adaptive optics systems. *J. Opt. Soc. Am. A*, 29(3):185–194. 21

[Correia et al., 2013] Correia, C., Véran, J.-P., Herriot, G., Ellerbroek, B., Wang, L., and Gilles, L. (2013). Increased sky coverage with optimal correction of tilt and tilt-anisoplanatism modes in laser-guide-star multiconjugate adaptive optics. *J. Opt. Soc. Am. A*, 30(4):604–615. 14

[Correia et al., 2015] Correia, C. M., Jackson, K., Véran, J.-P., Andersen, D., Lardière, O., and Bradley, C. (2015). Spatio-angular minimum-variance tomographic controller for multi-object adaptive-optics systems. *Appl. Opt.*, 54(17):5281–5290. 1, 17, 19

[Ellerbroek and Rigaut, 2001] Ellerbroek, B. L. and Rigaut, F. (2001). Methods for correcting tilt anisoplanatism in laser-guide-star-based multiconjugate adaptive optics. *J. Opt. Soc. Am. A*, 18(10):2539–2547. 13

[Flicker and Rigaut, 2002] Flicker, R. and Rigaut, F. (2002). Tilt anisoplanatism and PSF retrieval in LGS MCAO using a predictive controller. In E. Vernet, R. Ragazzoni, S. Esposito & N. Hubin, editor, *European Southern Observatory Conference and Workshop Proceedings*, volume 58 of *European Southern Observatory Conference and Workshop Proceedings*, page 377. 13

[Flicker et al., 2003] Flicker, R. C., Rigaut, F. J., and Ellerbroek, B. L. (2003). Tilt anisoplanatism in laser-guide-star-based multiconjugate adaptive optics. Reconstruction of the long exposure point spread function from control loop data. *Astronomy and Astrophysics*, 400:1199–1207. 15

[Gavel and Wiberg, 2003] Gavel, D. T. and Wiberg, D. (2003). Toward strehl-optimizing adaptive optics controllers. In Wizinowich, P. L. and Bonaccini, D., editors, *Proc. of the SPIE*, volume 4839, pages 890–901. SPIE. 19

[Gilles and Ellerbroek, 2008] Gilles, L. and Ellerbroek, B. L. (2008). Split atmospheric tomography using laser and natural guide stars. *J. Opt. Soc. Am. A*, 25(10):2427–2435. 14

[Gilles et al., 2011] Gilles, L., Wang, L., and Ellerbroek, B. L. (2011). Minimum variance split tomography for laser guide star adaptive optics. *European Journal of Control*, 17(3). 14

[Herrmann, 1992] Herrmann, J. (1992). Phase variance and Strehl ratio in adaptive optics. *J. Opt. Soc. Am. A*, 9:2257–2258. 4

[Jackson et al., 2015] Jackson, K., Correia, C., Lardière, O., Andersen, D., and Bradley, C. (2015). Linear prediction of atmospheric wavefronts for tomographic adaptive optics systems: modelling and robustness assessment. *Opt. Lett.*, 40(2):143–146. 19

[Kulcsár et al., 2006] Kulcsár, C., Raynaud, H.-F., Petit, C., Conan, J.-M., and de Lesegno, P. V. (2006). Optimal control, observers and integrators in adaptive optics. *Opt. Express*, 14(17):7464–7476. 6

[Le Roux et al., 2004] Le Roux, B., Conan, J.-M., Kulcsár, C., Raynaud, H.-F., Mugnier, L. M., and Fusco, T. (2004). Optimal control law for classical and multiconjugate adaptive optics. *J. Opt. Soc. Am. A*, 21:1261–1276. 6

[Meimon et al., 2010] Meimon, S., Petit, C., Fusco, T., and Kulcsar, C. (2010). Tip–tilt disturbance model identification for Kalman-based control scheme: application to XAO and ELT systems. *J. Opt. Soc. Am. A*, 27(11):A122–A132. 9

[Noll, 1976] Noll, R. J. (1976). Zernike polynomials and atmospheric turbulence. *J. Opt. Soc. Am. A*, 66:207–211. 14, 17

[Oppenheim and Willsky, 1997] Oppenheim, A. V. and Willsky, A. S. (1997). *Signals & Systems*. Prentice-Hall,Inc., 2nd edition. 9

[Petit, 2006] Petit, C. (2006). *Etude de la commande optimale en OA et OAMC, validation numérique et expérimentale*. PhD thesis, E.D. Galilée, Univ. Paris XIII. 6

[Petit et al., 2009] Petit, C., Conan, J.-M., Kulcsár, C., and Raynaud, H.-F. (2009). Linear quadratic gaussian control for adaptive optics and multiconjugate adaptive optics: experimental and numerical analysis. *J. Opt. Soc. Am. A*, 26(6):1307–1325. 19

[Petit et al., 2008] Petit, C., Conan, J.-M., Kulcsár, C., Raynaud, H.-F., and Fusco, T. (2008). First laboratory validation of vibration filtering with LQG control law for adaptive optics. *Opt. Express*, 16(1):87–97. 9, 20

[Petit et al., 2004] Petit, C., Quirós-Pacheco, F., Conan, J.-M., Kulcsár, C., Raynaud, H.-F., Fusco, T., and Rousset, G. (2004). Kalman-filter-based control for adaptive optics. In *Proc. of the SPIE*, volume 5490, pages 1414–1425. 6

[Piatrou and Roggemann, 2007] Piatrou, P. and Roggemann, M. C. (2007). Performance study of Kalman filter controller for multiconjugate adaptive optics. *Appl. Opt.*, 46(9):1446–1455. 19

[Poyneer and Véran, 2008] Poyneer, L. and Véran, J.-P. (2008). Predictive wavefront control for adaptive optics with arbitrary control loop delays. *J. Opt. Soc. Am. A*, 25(7):1486–1496. 19

[Sivo et al., 2014] Sivo, G., Kulcsár, C., Conan, J.-M., Raynaud, H.-F., Éric Gendron, Basden, A., Vidal, F., Morris, T., Meimon, S., Petit, C., Gratadour, D., Martin, O., Hubert, Z., Sevin, A., Perret, D., Chemla, F., Rousset, G., Dipper, N., Talbot, G., Younger, E., Myers, R., Henry, D., Todd, S., Atkinson, D., Dickson, C., and Longmore, A. (2014). First on-sky scao validation of full LQG control with vibration mitigation on the canary pathfinder. *Opt. Express*, 22(19):23565–23591. 19

[Takato and Yamaguchi, 1995] Takato, N. and Yamaguchi, I. (1995). Spatial correlation of zernike phase-expansion coefficients for atmospheric turbulence with finite outer scale. *J. Opt. Soc. Am. A*, 12(5):958–963. 17

[Valley and Wandzura, 1979] Valley, G. C. and Wandzura, S. M. (1979). Spatial correlation of phase-expansion coefficients for propagation through atmospheric turbulence. *J. Opt. Soc. Am.*, 69(5):712–717. 17

[Wallner, 1983] Wallner, E. P. (1983). Optimal wave-front correction using slope measurements. *J. Opt. Soc. Am.*, 73(12):1771–1776. 20

[Whiteley et al., 1998a] Whiteley, M. R., Roggemann, M. C., and Welsh, B. M. (1998a). Temporal properties of the zernike expansion coefficients of turbulence-induced phase aberrations for aperture and source motion. *J. Opt. Soc. Am. A*, 15:993–1005. 17

[Whiteley et al., 1998b] Whiteley, M. R., Welsh, B. M., and Roggemann, M. C. (1998b). Optimal modal wave-front compensation for anisoplanatism in adaptive optics. *J. Opt. Soc. Am. A*, 15:2097–2106. 17

[Winker, 1991] Winker, D. M. (1991). Effect of a finite outer scale on the zernike decomposition of atmospheric optical turbulence. *J. Opt. Soc. Am. A*, 8(10):1568–1573. 17