NFIRAOS RTC ALGORITHM DESCRIPTION

TMT.AO.DRD.08.002.REL05

November 2, 2011

Contents

List of Figures

List of Tables

1 Introduction

1.1 Introduction

This document is the Thirty Meter Telescope (TMT) NFIRAOS Real Time Controller (RTC) Algorithm Description Document. This document is a supporting document of the NFIRAOS RTC Requirement document [AD2].

1.2 Purpose

This document contains a description of the processing algorithms to be used by the Thirty Meter Telescope NFIRAOS RTC.

The intended audience for this document is primarily the designers of the NFIRAOS RTC and also:

- the NFIRAOS team,
- the NFIRAOS instrument teams,
- the reviewers of the NFIRAOS RTC.

1.3 Scope

The NFIRAOS RTC is a sub-system of NFIRAOS, which is the early light TMT AO facility. The RTC comprises the computer(s) and software responsible for the wavefront correction. In particular, the RTC processes the outputs from a suite of WFS (located within NFIRAOS and within the NFIRAOS client instruments) and computes the commands for the wavefront correctors.

The NFIRAOS RTC Algorithm Description document does not describe the algorithms used within the NFIRAOS Component Controller or the NFIRAOS Truth Wavefront Sensor Detector Controller.

Section 1 contains information about this document. Section 2 contains information about the NFIRAOS system and the RTC sub-system, and section 3 contains the algorithm descriptions.

1.4 Applicable Documents

AD1 NFIRAOS Design Requirement Document - TMT.AOS.DRD.07.002

AD2 NFIRAOS RTC Requirement Document - TMT.AOS.DRD.08.001

1.5 Reference Documents

RD1 NFIRAOS to NFIRAOS RTC ICD - TMT.AOS.ICD.08.001

RD2 NFIRAOS RTC to NFIRAOS DME ICD - TMT.AOS.ICD.08.003

RD3 NFIRAOS RTC to NFIRAOS TTS ICD - TMT.AOS.ICD.08.004

RD4 NFIRAOS VCAM to NFIRAOS RTC ICD - TMT.AOS.ICD.08.002

RD5 NFIRAOS to IRIS ICD - TMT.SEN.ICD.07.035

RD6 NFIRAOS to IRMS ICD - TMT.SEN.ICD.07.036

RD7 TCS to NFIRAOS ICD - TMT.SEN.ICD.07.037

RD8 NFIRAOS to LGSF ICD - TMT.SEN.ICD.07.027

1.6 Change Record

Revision	Date	Section	Modifications
DRF01	January 2008	All	Initial draft
REL01	February 2008	3.2	Remove Linear Algebra Concept Section
		3.2	Update block diagrams
		3.4.2.3	Update LGS WFS Gradient Gain Estimation section
		3.5.2.3	Update NGS WFS Gradient Gain Estimation section
		3.6.2.2, 3.6.2.4	Update TT(F) WFS Pixel Processing and TT(F) WFS Gradient Gain Estimation sections
		3.6.2.4	Add TT(F) WFS Gradient Noise Covariance Matrix Computation Section
		3.7	Rewrite LGS Wavefront Reconstruction section
		3.9.2.5	Add Centering TT(F) WFS section
		3.11	Update Turbulence Parameter Estimation Section
		3.12	Remove DM uncontrolled Mode Removal section
		3.14	Add PSF Reconstruction section
		3.*	Added computation and memory requirements for all subsections
REL02	May 2008	All	Replace TT(F) WFS with OIFWS
		2.1	Update Figure 2
		3.2	Update block diagrams
		3.4.2.2	Update C matrix
		3.4	Add a new background process to the LGS WFS processing section (LGS WFS drift modes computation)
		3.9.2.1	Remove reference to focus mode sent to NFIRAOS CC
		3.11.2	Add computation of θ_2
		3.12.2.3	Update telescope mode computation

Revision	Date	Section	Modifications
REL03			Skipped
REL04	December 2008	3.2	Update RTC figures 3 to 9
		3.4.2.5	Clarify how the mean and mean square of the total intensity for each LGS WFS sub-aperture are computed
		3.4.2.6	Fix typo in formula 3.4.26
		3.6.2.2	Modify the OIWFS matched filter algorithm
		3.6.2.5	Modify the OIWFS spot detection algorithm
		3.7.1	Differential mode removal modifications
		3.7.2.1	Update LGS pseudo-open loop gradient computation for bicubic spline influence function and dead/detached actuators
		3.7.2.2.1	Differential mode removal modifications
		3.7.2.2.3	Differential mode removal modifications
		3.7.2.2.3.1.1	Differential mode removal modifications
		3.7.2.2.3.1.2	Differential mode removal modifications
		3.7.2.3.2	Updated DM fitting computation for bicubic spline influence functions and dead/detached actuators
		3.7.2.4	Fix DM command subtraction
		3.7.3.1	Update computation and memory requirements for BGS-CBS and BGS-CG20
		3.7.3.2	Update computation and memory requirements for CG30 and FD3
		3.9.2.4	Rotate OIWFS modes
		3.9.2.5	Remove the section "Centering the OIWFS"
		3.11.2	Fix typo in formula 3.11.8
		3.12.2.1.1	Fix DM command subtraction
		3.12.2.1.3	Introduced DM reference "flat" command and physical DM actuator command $c_{DM,phys}$
		3.12.2.1.4	Replaced notation "clipped" and "extra" by "physical" and "edge"
		3.12.2.1.5	Removed temperature dependence of DM gain and offset coefficients
		3.12.2.1.6	Expressed "clipped" DM actuator command $c_{DM,clipped}$ in terms of the physical command vector $c_{DM,phys}$
		3.12.3	Update computation and memory requirement for wavefront control

REL05	October 2011	3.4.2.3	Updated LGS WFS matched filter dithering
		3.7.3	Included FD3 with only 2 layers oversampled
		3.9.2.2	Low-order mode temporal filtering (Type II control)
		3.11.2	Turbulence parameter estimation (SLODAR)
		3.12.2.1	Updated DM control
		3.12.2.2	Updated TTS control

1.7 Abbreviations

aO	Active Optics
AO	Adaptive Optics
AOSQ	Adaptive Optics Sequencer
AOESW	Adaptive Optics Executive Software
BGS	Block Gauss Seidel
CBS	Cholesky Back Substitution
CG	Conjugate Gradient
DM	Deformable Mirror
DME	Deformable Mirror Electronics
DMS	Data Management System
FDPCG	Fourier Domain Preconditioned Conjugate Gradient
FFT	Fast Fourier Transform
FoV	Field of View
FSM	Fast Steering Mirror
IR	InfraRed
IRIS	InfraRed Imaging Spectrograph
IRMS	InfraRed Multi Object Spectrograph
LGS	Laser Guide Star
LGSF	Laser Guide Star Facility
LIS	Laser Interlock System
LUT	Look-Up-Table
MCAO	Multi Conjugate Adaptive Optics
MVM	Matrix-Vector Multiplication
NCPA	Non common path aberration
NFIRAOS	Narrow Field Infrared Adaptive Optics System
NGS	Natural Guide Star
NIRES	Near InfraRed Echelle Spectrograph
OIWFS	On Instrument Wavefront Sensor
PCG	Preconditioned Conjugate Gradient
PSD	Power Spectral Density
PSF	Point Spread Function
RPG	Reconstructor Parameter Generator
RTC	Real Time Controller
SH	Shack Hartmann
SNR	Signal to Noise Ratio

TCS	Telescope Control System
TMT	Thirty Meter Telescope
TTF	Tip Tilt Focus
TTS	Tip Tilt Stage
TWFS	Truth Wavefront Sensor
WFS	Wavefront Sensor
WIRC	Wide Field InfraRed Camera

2 Overall Description

2.1 Perspective

The TMT will implement a first light adaptive optics (AO) System called NFIRAOS, which will feed 3 science instruments on the telescope Nasmyth platform. This *Narrow Field Infrared Adaptive Optics System* is a Multi Conjugate Adaptive Optics (MCAO) system, which provides turbulence compensation over a moderately large field of view (1-2 arcmin) in order to sharpen the images of natural guide stars and improve the sky coverage. NFIRAOS includes:

- two Deformable Mirrors (DM) conjugated at 0km (DM0) and 11.2km (DM11.2),
- one Tip/Tilt Stage (TTS) serving as a mount for DM0,
- six Laser Guide Star (LGS) wavefront sensors observing an asterism as illustrated in Figure 1 below,
- up to three low order Natural Guide Star wavefront sensors[1] within each NFIRAOS instrument,
- one high order visible Natural Guide Star (NGS) wavefront sensor[2], which is used for operation without LGS,
- one Truth Wavefront Sensor (TWFS), which is used to calibrate for slow-varying biases due to flexure effects and variations in the sodium layer profile in LGS AO mode.
- and the RTC, which processes the inputs from the LGS or NGS and on instrument wavefront sensors to compute the commands for the deformable mirrors and the tip/tilt stage.

The RTC will interface with additional telescope and AO sub-systems, including the AO Sequencer, the NFIRAOS Component Controller, the Laser Guide Star Facility System (LGSF), the NFIRAOS instruments, the NFIRAOS TWFS Detector Controller and the Data Management System (DMS).

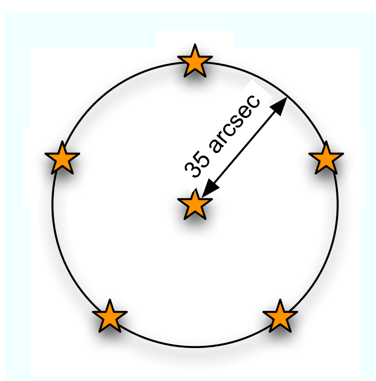

Figure 1: NFIRAOS asterism. NFIRAOS will utilize 6 laser guide stars, 5 equally spaced on a radius of 35 arcsec, and one on-axis.

[1] Referenced as OIWFS in the rest of the document
[2] Referenced as NGS WFS in the rest of the document

The following figure is a block diagram of the NFIRAOS facility.

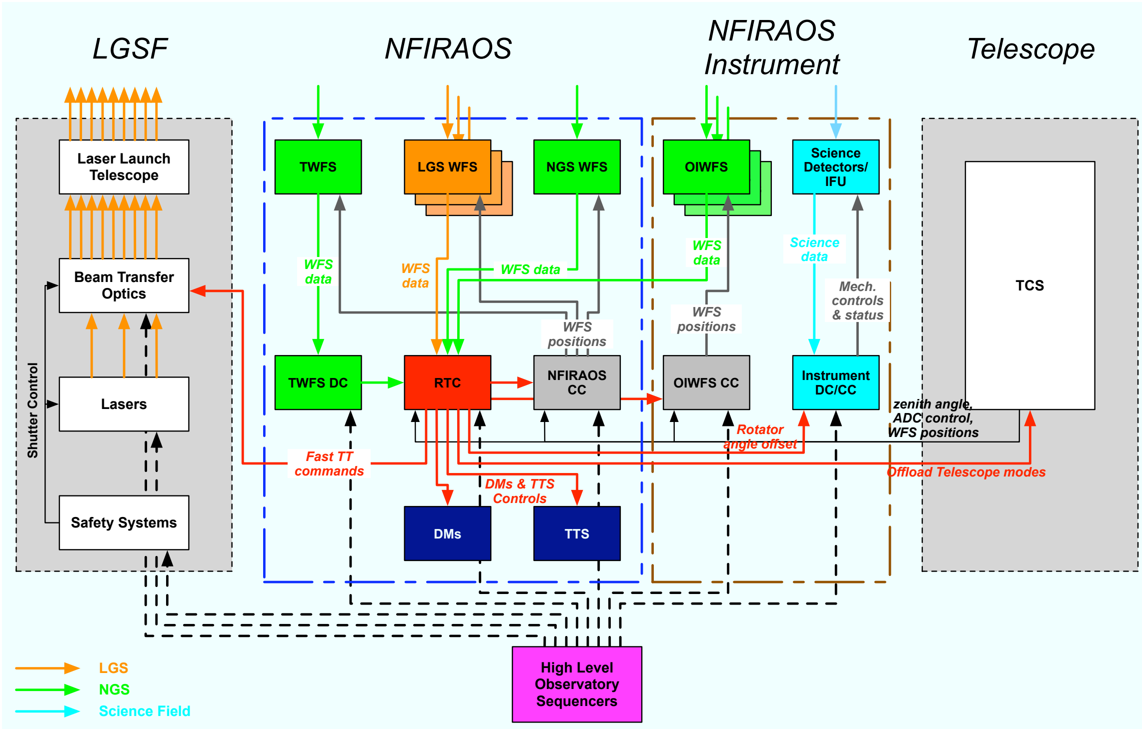

Figure 2: NFIRAOS AO LGS block diagram

2.2 System Functions

The NFIRAOS RTC system functions are described in the NFIRAOS RTC Requirement Document [AD2].

2.3 User and Operator Characteristics

The NFIRAOS RTC User and Operator Characteristics are described in the NFIRAOS RTC Requirement Document [AD2].

2.4 External Interfaces

The NFIRAOS RTC external interfaces are described in the NFIRAOS RTC Requirement Document [AD2]

2.5 Constraints

The following areas may be updated in the next release:

- Computation of the LGS WFS reference vector. In general the architecture for the computation of the WFS reference vector should be revisited (within TWFS versus RPG).
- Estimation of the residual uncorrected error.
- Real time processing of any data required by PSF reconstruction.
- DM and TTS control: treatment of clipping and inluding more sophisticated predictive algorithms for OIWFS modes.
- OIWFS pixel processing: alternative algorithms to improve linear dynamic range.
- OIWFS reconstruction: non sidereal objects tracking case should be reviewed (derotation).
- Computation/memory requirement for pixel processing, DM/TTS control, OIWFS reconstruction and SLODAR.
- Atmospheric time constant (Greenwood frequency) and outer scale parameter estimation.

2.6 Assumptions and Dependencies

The RTC algorithm definition depends heavily on the NFIRAOS AO System design. Any changes in the NFIRAOS design may affect the RTC algorithm definitions and the computation/memory requirements.

The estimated computation and memory requirements are included for guidance only. They are based upon particular implementations for each algorithm, and may be subject to further optimization.

3 RTC Algorithm Description

3.1 Definitions and Notations

N_{OIWFS}	Number of low order OIWFS
N_{LGS}	Number of LGS WFS
N_{DM}	Number of DMs
n_{pix}^{lgs}	Total number of pixels per LGS WFS
n_{sa}^{lgs}	Number of illuminated sub-apertures per LGS WFS
n_{grad}^{lgs}	Number of gradients per LGS WFS
N_{grad}^{LGS}	Total number of LGS WFS gradients
n_{pix}^{ngs}	Total number of pixels of the NGS WFS
n_{grad}^{ngs}	Total number of gradients of the NGS WFS
N_{grad}^{OIWFS}	Total number of OIWFS gradients
N_{vact}^{DM}	Total number of DM active actuators
N_{act}^{DM}	Total number of DM actuators (including active and edge actuators)

The following rules are used in the document:

– Real time RTC operations are highlighted with a solid box around the corresponding equations or pseudo-code.

– Background and optimization tasks are highlighted with a dashed box around the corresponding equations or pseudo-code.

3.2 RTC Block Diagrams

The RTC will implement 3 operating modes:

- LGS AO mode. In this mode, the RTC uses the six LGS WFS and up to three OIWFS to compute the commands for the two DMs and the TTS. The commands for the DMs and for the TTS are offloaded to the telescope active optics modes.
- NGS AO mode. In this mode, the RTC uses the high-order NGS WFS and (if selected) the 2x2 OIWFS of the instrument to compute the commands for the DM0 and the TTS (DM11.2 being flattened). The commands for the DM0 and the TTS are offloaded to the telescope active optics modes.
- Seeing limited mode. In this mode, the RTC uses the TWFS and the 2x2 OIWFS to compute the TTS commands and the telescope active optics modes (DM0 and DM11.2 being flattened).

The RTC block diagrams for the LGS AO mode are given in the following figures:

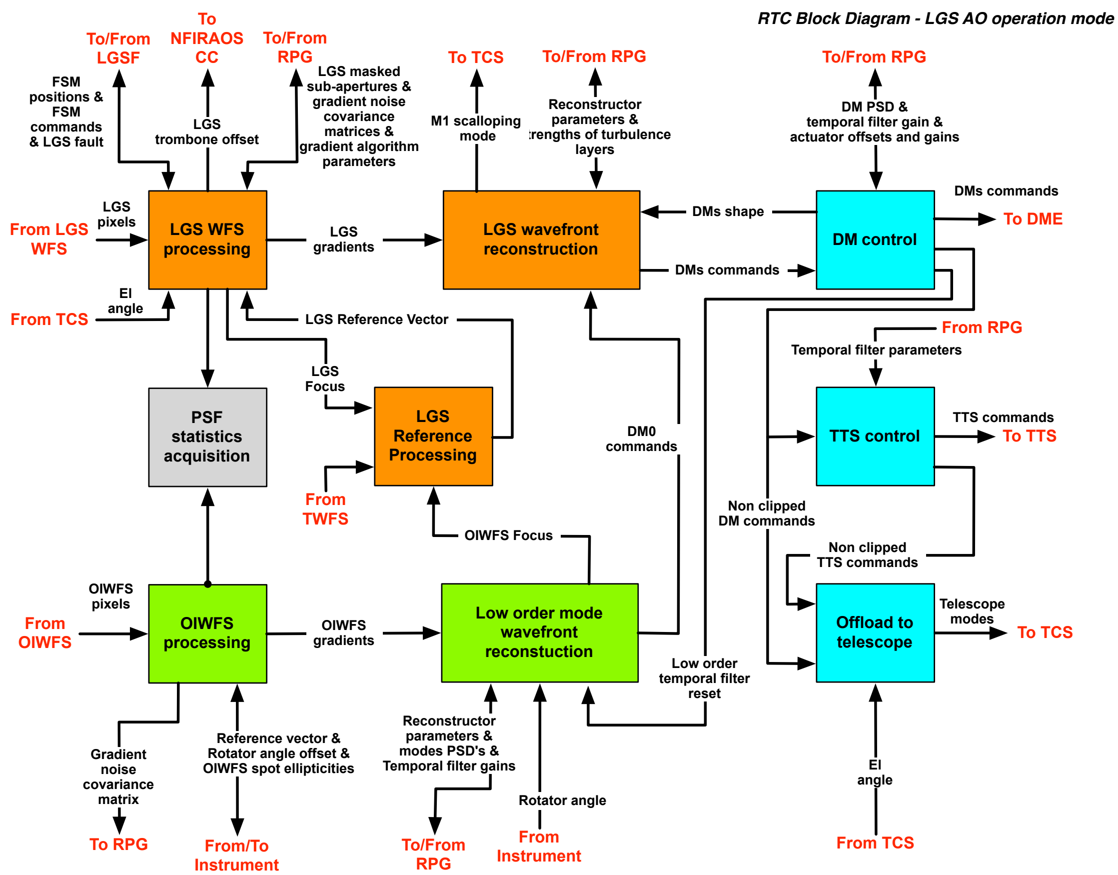

Figure 3: RTC block diagram for the LGS AO mode

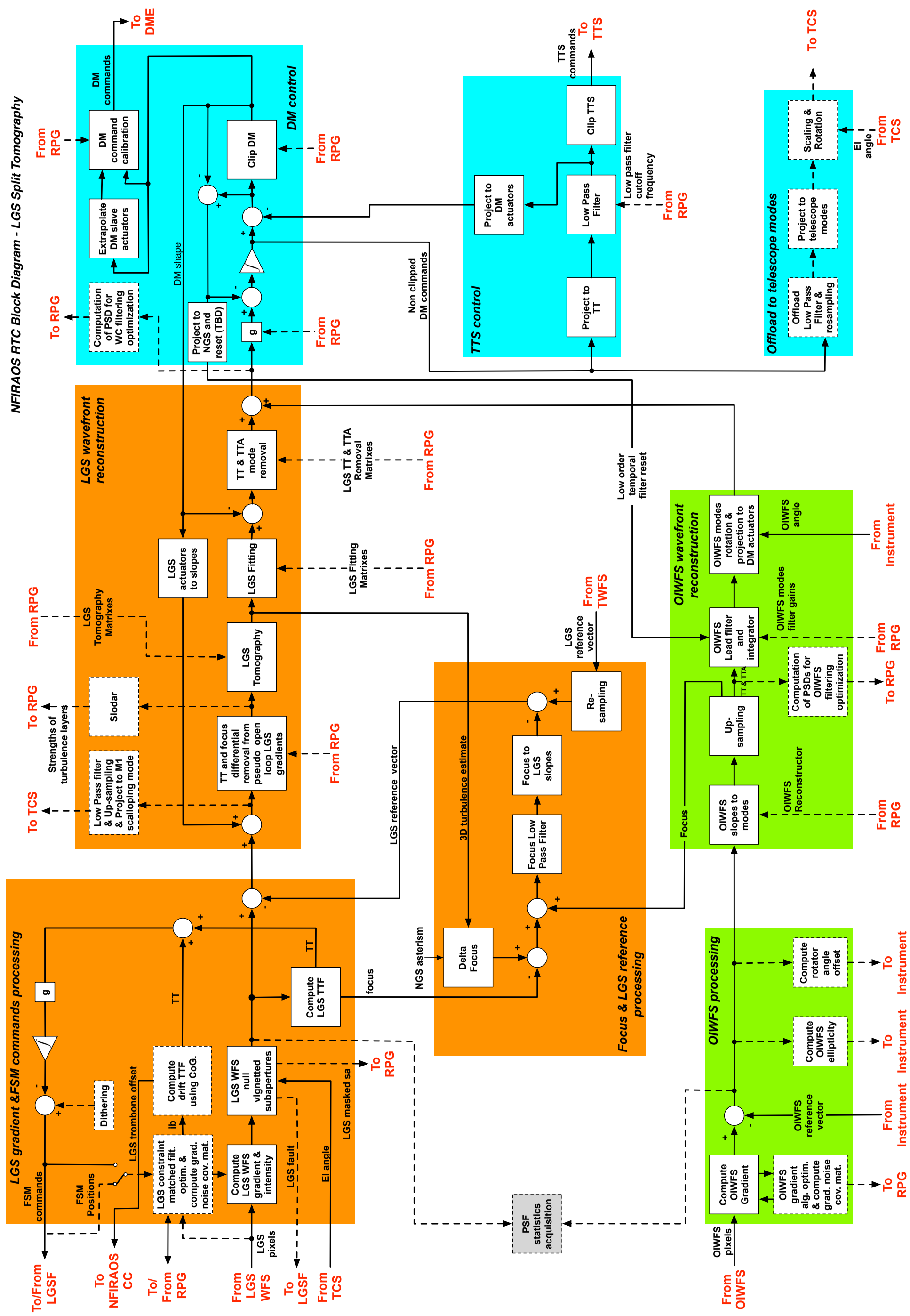

Figure 4: RTC detailed block diagram for the LGS AO mode

The RTC block diagrams for the NGS AO mode are given in the following figures:

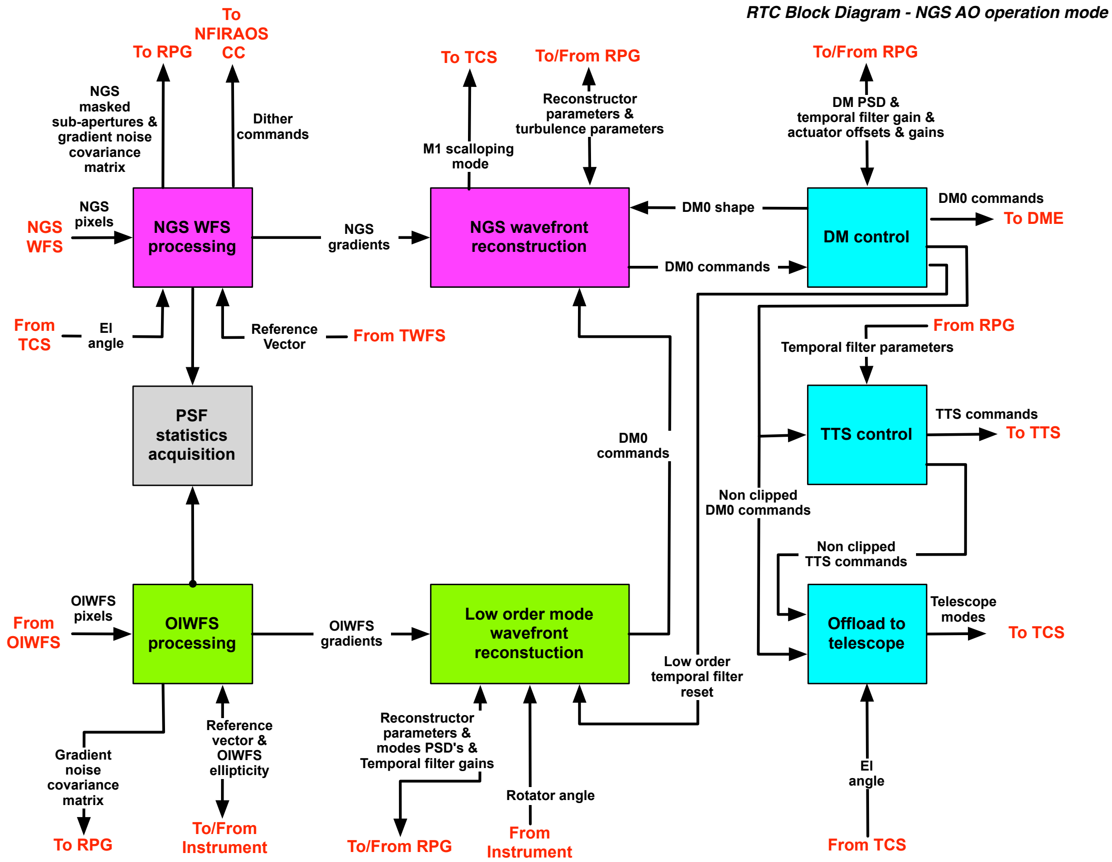

Figure 5: RTC block diagram for the NGS AO mode

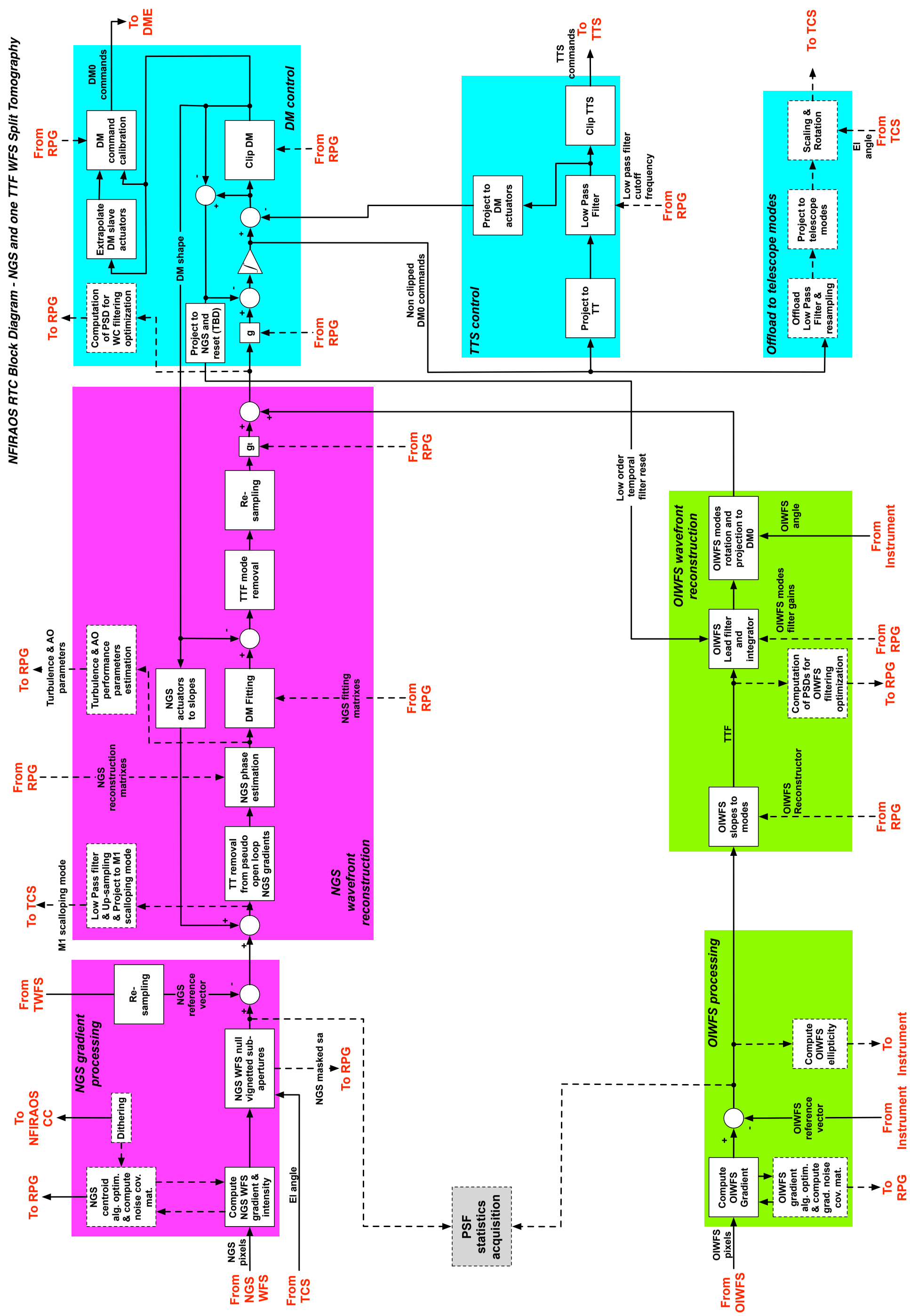

Figure 6: RTC detailed block diagram for the NGS AO mode with an additional 2x2 OIWFS

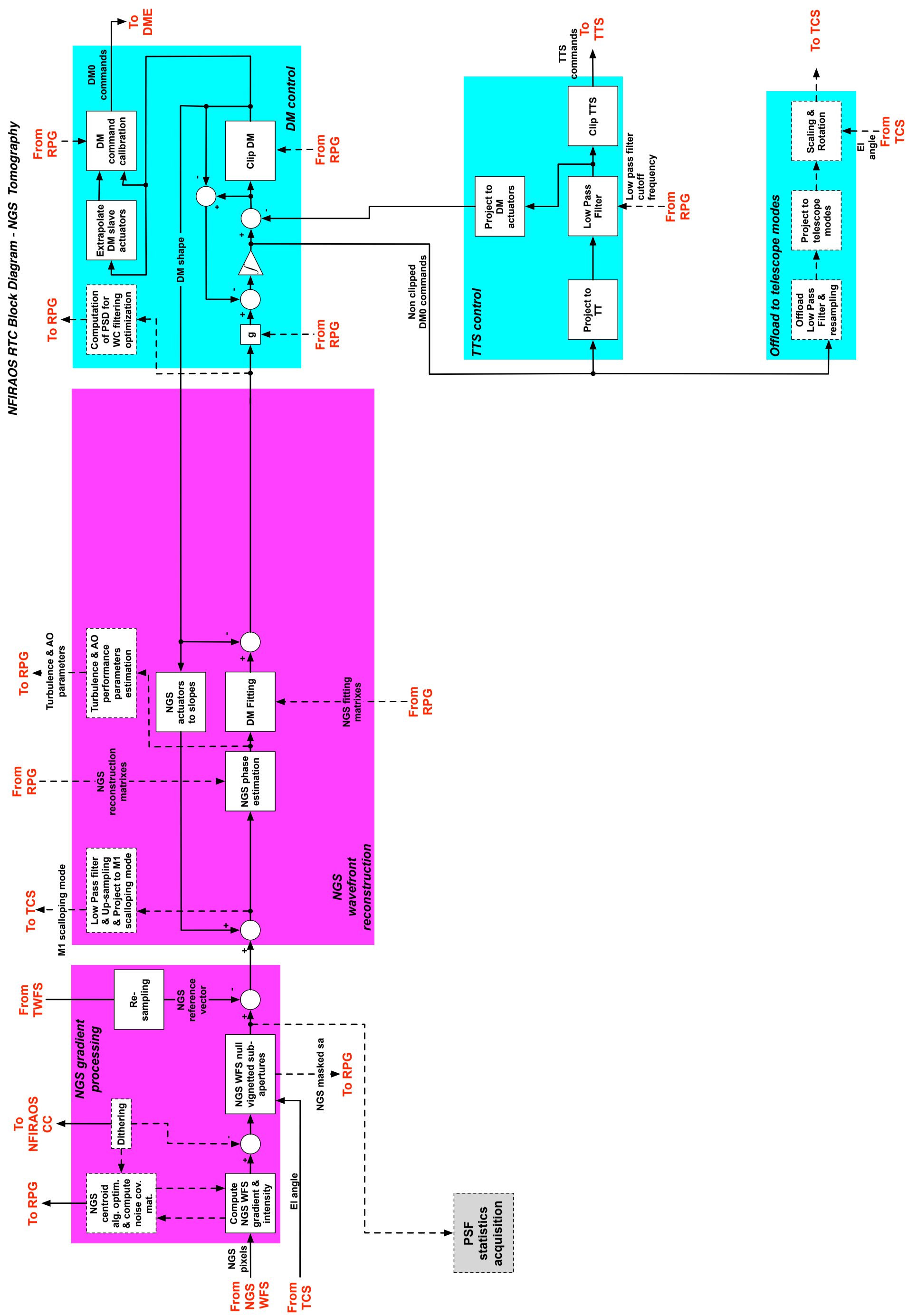

Figure 7: RTC detailed block diagram for the NGS AO mode without an additional 2x2 OIWFS

The RTC block diagrams for the seeing limited mode are given in the following figures:

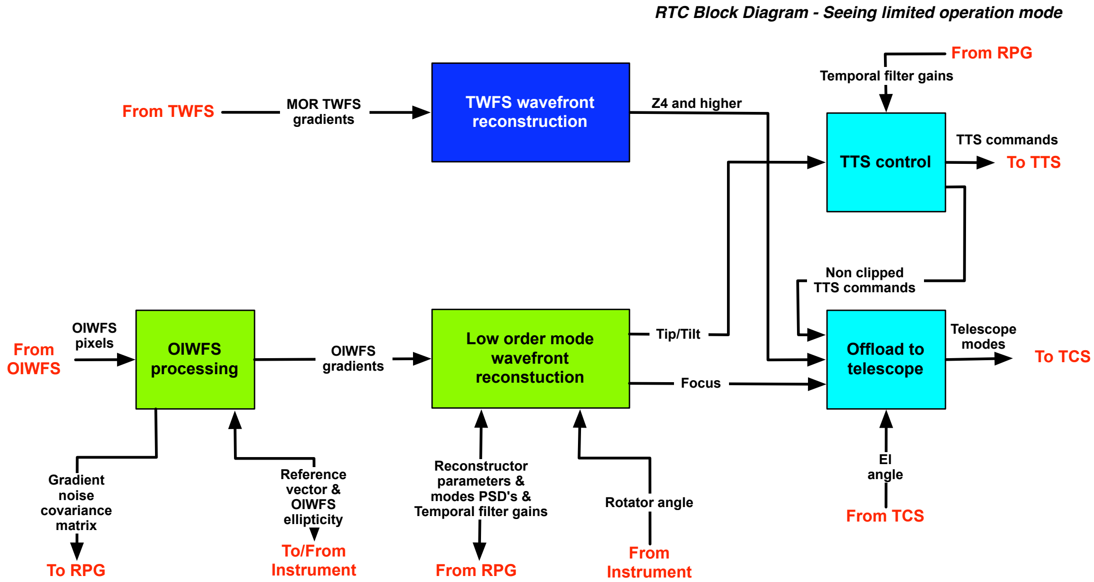

Figure 8: RTC block diagram for the seeing limited mode

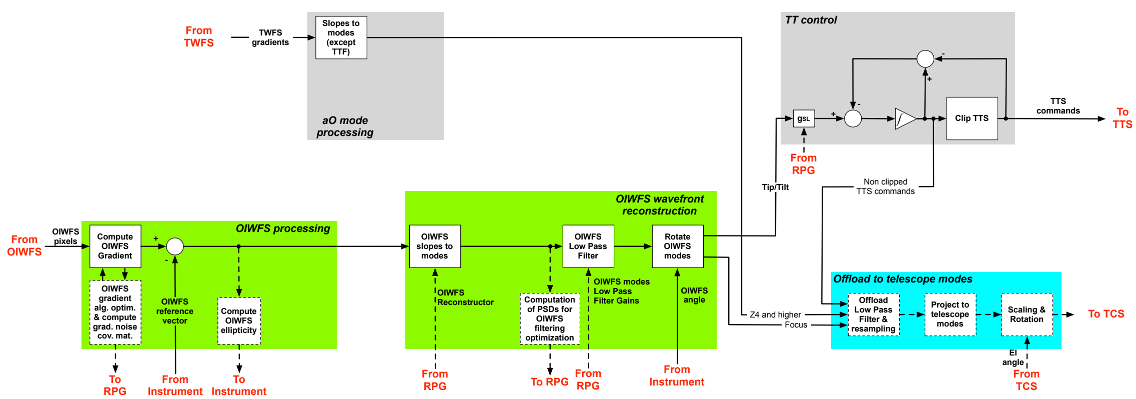

Figure 9: RTC detailed block diagram for the seeing limited mode

3.3 WFS Synchronization

NFIRAOS and its client instruments will include the following WFS:

- Six 60x60 LGS WFS, located within NFIRAOS, used in LGS AO mode only and working at 800Hz. These WFS will be controlled and read by the RTC.
- One high order 60x60 NGS WFS, located within NFIRAOS, used in NGS AO mode only and working at a frame rate varying from 10Hz to 800Hz. The NGS WFS will be controlled and read by the RTC.

– Up to three OIWFS located in each of the NFIRAOS instruments and working at frame rates varying from 10Hz to 800Hz. These WFS will be controlled and read by the RTC. The NFIRAOS instruments IRIS, NIRES, and WIRC will implement 3 OIWFS, each of which may be configured as either a 2x2 SH WFS or a Tip/Tilt WFS. LGS AO observations with these instruments will employ one 2x2 SH WFS, and up to two Tip/Tilt WFS.

  one 2x2 SH WFS and two tip/tilt SH WFS (sometimes described as TT WFS). The NFIRAOS instrument IRMS will implement only one 2x2 SH WFS.

– One TWFS, located within NFIRAOS. This WFS is controlled and read by the TWFS Detector Controller, which passes WFS reference vectors to the RTC.

In LGS AO mode, the start of the integration time of the OIWFS will be synchronized with the LGS WFS. The OIWFS may have different frame rates to take into account differences in the brightness of the guide stars. The centroid measurements of each OIWFS will be computed independently at their individual frame rates. The low order modal reconstruction will be performed at the rate of the fastest OIWFS, using the most recent centroids from the other two WFS. Finally, the low order mode-to-actuator projection will be performed at the LGS WFS frame rate, using the most recently computed modes. Note that in the IRIS case, if only one IR guide star is available, it will be used as the 2x2 OIWFS guide star.

In NGS AO mode, if the 2x2 OIWFS is used in addition to the NGS WFS, the start of the integration time of the NGS WFS will be synchronized with the 2x2 OIWFS, which is assumed to operate at a higher frame rate on account of the much smaller number of sub-apertures. In this case, the centroid computation, the modal reconstruction and the mode-to-actuator projection of the 2x2 OIWFS will be computed at the rate of the 2x2 OIWFS, the outputs of the NGS wavefront reconstruction will be resampled to the higher 2x2 OIWFS frame rate.

In seeing limited mode, no WFS synchronization will be required.

3.4 LGS WFS Processing

3.4.1 Overview

This section describes:

(i) the LGS WFS pixel processing algorithm, which uses a constrained matched filter algorithm to estimate the sub-aperture wavefront gradients,

(ii) the optimization of the parameters of the LGS WFS pixel processing algorithm,

(iii) and the computation of the commands for the fast tip-tilt mirrors located in the Laser Guide Star Facility (LGSF); these mirrors are used to stabilize the LGS pointing on the sky.

3.4.2 Algorithm Description

3.4.2.1 LGS WFS Pixel Calibration

For each LGS WFS, the LGS WFS pixel intensities are calibrated (flat fielded and bias subtracted) before computing the gradients. Flat fielding is an element-by-element vector multiplication, and bias subtraction is a simple vector subtraction. The vector of calibrated pixels is obtained as follows:

$$\boxed{i = (p - d) \cdot * f} \quad (3.4.1)$$

where:

- i is the pixel intensity vector for a LGS WFS, the dimension of this vector is n_{pix}^{lgs},
- p is the raw pixel intensity vector for the corresponding LGS WFS before calibration,
- f is the flat field vector for the corresponding LGS WFS,
- d is the dark vector for the LGS WFS.

3.4.2.2 LGS WFS Pixel Processing

To derive the matched filter gradient estimation algorithm, the vector of measured pixel intensities i_{model} in a sub-aperture is modeled by the following vector form expression:

$$i_{model} = (1 + \nu) * i_0 + G \,.\, \theta + n \quad (3.4.2)$$

where:

- ν is the relative change in signal level for the LGS WFS (identical for each sub-aperture)
- i_0 is the vector of reference pixel intensities in the sub-aperture with no spot displacement. The dimension of the vectors i_{model} and i_0 is n_{pix}, where n_{pix} is the number of pixels per sub-aperture.
- G is the $(n_{pix}, 2)$ Jacobian matrix of partial derivatives of pixel intensities with respect to spot displacement in the sub-aperture,
- θ is the image displacement in the sub-aperture, expressed in a common X-Y coordinate system for all sub-apertures
- and n corresponds to the vector of measurement noise in the sub-aperture (detector read noise and photon noise).

The estimate of the image displacement θ in the sub-aperture is obtained by applying a filter R_c:

$$\boxed{\widehat{\theta} = R_c \,.\, i} \quad (3.4.3)$$

The coefficients of the filter R_c are selected to minimize the RMS error due to noise, subject to a set of constraints which improve linearity. Specifically, the R_c filter is updated at a slow rate to minimize the trace of the matrix $R < nn^T > R^T$, where R satisfies the following constraints:

- $R \cdot G = Id$ where Id is the 2x2 identity matrix,
- $R \cdot i_0 = 0$
- $R \cdot i_j = \theta_j$, where j varies from 1 to 4, θ_j are the angular displacements corresponding to image shifts of +/- 1 pixel in either the radial or azimuthal direction within the sub-aperture, and i_j is a copy of i_0 shifted in the same direction.

The expression to compute R_c can be formulated as follows:

$$R_c = C\,(\tilde{G}^T \Sigma^{-1} \tilde{G})^{-1} \tilde{G}^T \Sigma^{-1} \tag{3.4.4}$$

where:

- $\tilde{G}$ is the augmented sensitivity matrix, $\tilde{G} = \left[\begin{array}{cccccc} G & i_0 & i_1 & i_2 & i_3 & i_4 \end{array}\right]$
- C is the constraint matrix defined as
  $$C = \begin{pmatrix} 1 & 0 & 0 & \cos\theta_{sa,j} & -\cos\theta_{sa,j} & \sin\theta_{sa,j} & -\sin\theta_{sa,j} \\ 0 & 1 & 0 & -\sin\theta_{sa,j} & \sin\theta_{sa,j} & \cos\theta_{sa,j} & -\cos\theta_{sa,j} \end{pmatrix}$$
  where $\theta_{sa,j}$ is the azimuthal angle of the sub-aperture j.
- $\Sigma =< nn^T >$ is the covariance matrix of the pixel intensity measurement noise vector n in the sub-aperture.

3.4.2.3 LGS WFS Gradient Gain Estimation

In practice, the quantities $\tilde{G}$, Σ, and i_0 will depend upon seeing, the sodium layer profile, and laser beam quality. The following paragraphs describe how they will be initialized to nominal values, and then updated in real time to recompute the matched filter R_c.

1. The reference vector i_0 is initialized for each sub-aperture as a convolution of the following terms:

   - the two dimensional transverse cross section of the laser profile in the sodium layer, as estimated from nominal values for the laser beam profile and atmospheric seeing,
   - the point spread function for a LGS WFS sub-aperture, as estimated from the seeing and the nominal LGS WFS design parameters,
   - the laser beam elongation for each sub-aperture, as estimated from the launch-telescope-to-subaperture separation and a nominal estimate of the sodium layer profiles

   The initialization of the vector i_0 is performed by the Reconstructor Parameter Generator (RPG) and the initial vector is transferred to the RTC during the configuration process.

   The reference intensity vector is then updated in real time by the RTC with a time constant of 1 to 10 seconds based on the intensity vectors i obtained while in closed loop. The boxcar-averaged intensity vector i_b is first computed as follows:

   $$(i_b)_q = \frac{1}{n_d n_c} \sum_{k=(q-1)(n_d n_c)+1}^{q n_d n_c} i_k \tag{3.4.5}$$

   where:

   - q is the update cycle number;
   - i_k is the subaperture intensity vector at RTC frame number k;

- n_d is the number of RTC frames per dither period ($n_d = 4$ i.e. 1 dither period every 4 frames) ; and
- n_c is the number of dither periods to average in computing i_b ($n_c = 0.3s * 800Hz/n_d$).
- $n'_c = 10$ is the number of values of i_b used to generate i_0 (used for focus tracking computation).

i_0 is then computed by averaging i_b over multiple frames:

$$i_0 = \sum_{q=(p-1)n'_c+1}^{pn'_c} (i_b)_q \tag{3.4.6}$$

2. The matrix Σ is the covariance matrix of the pixel intensity measurement noise, and is a diagonal matrix of dimension (n_{pix}, n_{pix}). If the vectors n and i are expressed in units of photo-detection events, it is described by the expression:

$$\Sigma = diag(\sigma_E^2 + i_0) \tag{3.4.7}$$

where σ_E is the detector read noise.

3. The G matrix is first computed analytically by the RPG based upon the initial value of i_0 derived from the nominal estimates of the seeing, the laser beam profile, and the sodium layer profile. The initial value of G is transferred to the RTC during the configuration process. The G matrix is then updated in real time by dithering the Fast Steering Mirror (FSM) of the Laser Guide Star Facility system as described in detail below.

Sinusoidal signals are injected into the commands of the FSM computed in section 3.4.2.7 below:

$$c_{dit,x} = a_1 \sin(2\pi k/n_d) \tag{3.4.8}$$

$$c_{dit,y} = a_1 \cos(2\pi k/n_d) \tag{3.4.9}$$

The frequency of the dither must be chosen carefully so that the (discrete time) signals are truly orthogonal to each other. For example, $n_d = 4$, giving a 200 Hz nutation for k incrementing at 800 Hz. In this case, the two equations may be implemented by a four-row table. The amplitudes of the signals shall be chosen to be significantly smaller than the errors being compensated for, but large enough to be detectable in the presence of noise.

The correlations between the injected signals and WFS pixel intensities are then averaged over n_c dither cycles, i.e., $n_c n_d$ RTC frames, to obtain the current (q^{th}) boxcar average estimates of each element of the matrix G:

$$\left(\frac{\partial i_b(j)}{\partial \theta_x}\right)_q = \frac{2}{a_2 n_d n_c} \sum_{k=(q-1)n_d n_c+1}^{q n_d n_c} \sin(2\pi k/n_d + \Delta_{m-1}) i_k(j) \tag{3.4.10}$$

$$\left(\frac{\partial i_b(j)}{\partial \theta_y}\right)_q = \frac{2}{a_2 n_d n_c} \sum_{k=(q-1)n_d n_c+1}^{q n_d n_c} \cos(2\pi k/n_d + \Delta_{m-1}) i_k(j) \tag{3.4.11}$$

Here $\frac{\partial i_b(j)}{\partial \theta}$ represents the first-order variation of the intensity in the pixel number j of the boxcar-averaged image i_b based upon the LGS image displacement θ in the sub-aperture.

The actual amplitude of the dither, a_2, is not necessarily identical to the input dither signal a_1, and is computed in step 5 below. The quantity Δ_{m-1} represents the most recent estimate of the phase delay in the dither due to the round-trip time-of-flight to the sodium layer, and is determined for each WFS by a delay-locked loop as described in section 4 below.

Finally, an average is applied to each element of the G matrix to update the R_c matrix:

$$\frac{\partial i_0}{\partial \theta} = \sum_{q=(p-1)n'_c+1}^{pn'_c} \left(\frac{\partial i_b}{\partial \theta}\right)_q \quad (3.4.12)$$

4. Delay Locked Loop

Using the tip/tilt values $e_{FSM}(k)$ calculated at cycle k using Eq. (3.4.26) for each LGS WFS, phase errors in the dither signal are accumulated as follows:

$$\begin{aligned}\delta(k) &= \delta(k-1) + \\ &\quad g_{dll}\left[\cos(2\pi k/n_d + \Delta_{m-1})e_{FSM,x}(k) - \sin(2\pi k/n_d + \Delta_{m-1})e_{FSM,y}(k)\right] \quad (3.4.13)\end{aligned}$$

where

- g_{dll} the gain for the delay-locked loop,
- and Δ_{m-1} is the previous estimate of the phase shift between the FSM dither signal and the tip/tilt dither sensed by the LGS WFS.

Note that the sines and cosines generated for use in Eq. (3.4.13) may be reused in Eq.'s (3.4.10) and (3.4.11).

The phase shift delay is updated after an integral number of dither periods $n_{dll} \geq 1$:

$$\Delta_m = \delta(k) \qquad \text{if } k/n_d n_{dll} = m. \quad (3.4.14)$$

5. Determine the actual dither amplitude a_2 from the LGSF FSM Position sensors (if selected), or the FSM commands, for use in Eq.'s (3.4.10) and (3.4.11):

$$d_{FSM} = \begin{cases} p_{FSM} & \text{if FSM position sensors are used;} \\ c_{FSM} & \text{otherwise.} \end{cases} \quad (3.4.15)$$

Here d_{FSM} is the selected data source, c_{FSM} are the fast steering mirror commands from Eq. (3.4.28), and p_{FSM} is the reading from the position sensors (if implemented) within each FSM. If the dynamics induced by the FSM are unacceptable (i.e., attenuate the commands too much at the dither frequency), the RTC will be configured (as indicated by the switch shown in the upper left of Fig. 4) to employ the position sensors within the FSMs to determine the dither amplitudes a_2 for use in Eq. (3.4.10) and (3.4.11).

The correlations between the injected signals and the FSM positions or commands are averaged over n_{dll} dither periods as follows:

$$(ip)_v = \sum_{k=(v-1)n_d n_{dll}+1}^{v n_d n_{dll}} d_{FSM,x}(k)c_{dit,x}(k) + d_{FSM,y}(k)c_{dit,y}(k) \quad (3.4.16)$$

$$(qp)_v = \sum_{k=(v-1)n_d n_{dll}+1}^{v n_d n_{dll}} d_{FSM,x}(k) c_{dit,y}(k) - d_{FSM,y}(k) c_{dit,x}(k) \tag{3.4.17}$$

$$(a_2)_v = \frac{1}{n_d n_{dll} a_1} \left(ip_v^2 + qd_v^2 \right)^{1/2}, \tag{3.4.18}$$

where v denotes update cycles of the dither amplitude estimate, and p_{FSM} are the FSM positions or commands, ip is the in-phase component of the FSM motions, and qd is the quadrature component

3.4.2.4 LGS WFS Gradient Noise Covariance Matrix Computation

The LGS WFS gradient noise covariance matrices C_k are computed by the RTC for each LGS WFS, concurrently with estimating the gains. These matrices are used later by the LGS wavefront reconstruction process. The C_k matrices are block diagonal matrices composed of 2x2 blocks (each block corresponds to a sub-aperture). Each block is obtained as follows:

$$[C_k]_{j,j} = R_c \, \Sigma \, R_c^T \tag{3.4.19}$$

where R_c and Σ are the sub-aperture matched filter matrix and the covariance matrix of the pixel intensity measurement noise, corresponding to the sub-aperture j of the LGS WFS k.

The inverses of the C_k matrices are computed by individually inverting each of these 2x2 blocks. The inverted blocks corresponding to masked sub-apertures (see below) are then set to zero.

3.4.2.5 LGS WFS Subaperture Masking and Reference Subtraction

Once an unmasked gradient vector θ has been estimated for each LGS WFS, the LGS WFS gradient vector is obtained as follows:

– First, the unmasked vector is multiplied by a mask vector (element-by-element vector multiplication) to filter the gradients of the sub-apertures which are unusable because of the telescope obscuration.

$$s_{masked} = v_{mask} \cdot * \theta \tag{3.4.20}$$

The mask vector v_{mask} elements are either 1 or 0 depending up whether the sub-aperture is usable or not. It is pre-computed by the Reconstructor Parameter Generator (RPG) for all zenith angles varying from 0 to 65 degrees with steps of 0.1 degrees. The resulting mask vectors are stored into the RTC upon initialization. At a rate of 1 Hz, the RTC gets the actual zenith angle from the TCS and selects the corresponding mask vector.

– The total intensity for each sub-aperture I_{sa} is estimated at 800Hz as the sum of the pixel intensities within the sub-aperture. The total intensity is then compared to a defined threshold T_1 to detect any additional un-illuminated or partially illuminated sub-apertures, and to set their corresponding gradients to their values on the previous frame.

– The LGS WFS gradient vector is obtained at 800Hz as follows:

$$\boxed{s = s_{masked} - v_{mask} \cdot *r} \quad (3.4.21)$$

where r is the LGS WFS reference vector.

– Finally, variations in the intensity level of each sub-aperture due to M3 misalignment or Rayleigh backscatter are monitored by a background process at a rate of 1 Hz. This process computes the mean and standard deviation of the intensity level for each sub-aperture ($< I_{sa} >$ and σ_{sa}) and compares the mean to the threshold T_1 and the standard deviation to another threshold T_2. The list of additional masked sub-apertures is output to the RPG for use in the next update of the LGS wavefront reconstructor.

$$\sigma_{sa} = \sqrt{< I_{sa} >^2 - < I_{sa}^2 >} \quad (3.4.22)$$

Note that the mean and mean square of the total intensity for each sub-aperture are accumulated at 800Hz.

3.4.2.6 LGS WFS drift modes computation

The purpose of this task is to compute at a rate of 3.33 Hz, the LGS tip/tilt and focus drift terms, which will be used (i) to stabilize the commands of the fast tip/tilt mirrors located in the Laser Guide Star Facility and (ii) to control the LGS trombone mechanism.

For each LGS WFS, the processing steps consists of:

– computing the LGS reference gradients s_0 from the vector of reference pixel intensities i_b. For each sub-aperture, the centroid algorithm consits of two steps: (i) computing the reference gradients using a standard center of gravity algorithm and then (ii) rotating the centroids into a common (x,y) coordinates system.

– filtering the unusable sub-apertures due to telescope obscuration as described in the section 3.4.2.5. The vector of filtered LGS WFS gradients is $s_{0,masked}$.

– computing the LGS tip/tilt/focus drift modes as described below:

$$m_0 = \begin{pmatrix} t_0 \\ z_0 \end{pmatrix} = S^{\dagger}_{TTF} \cdot s_{0,masked} \quad (3.4.23)$$

where $S^{\dagger}_{TTF}$ is the pseudo inverse of the tip/tilt/focus LGS sensitivity matrix. This matrix is updated as described in the following equation when new LGS WFS gradient noise inverse covariance matrices are available:

$$S^{\dagger}_{TTF} = (S^T_{TTF}\, W_k\, S_{TTF})^{-1}\, S^T_{TTF}\, W_k \quad (3.4.24)$$

The matrix S_{TTF} has 3 columns. The first two columns contain 0,1-valued tip/tilt modes measured by the LGS WFS, and the third column contains the focus mode measured by the LGS WFS. The weighting matrix W_k is equal to C_k^{-1} with the blocks corresponding to the masked sub-apertures zeroed out. The matrices $S^T_{TTF}\, W_k$ and $(S^T_{TTF}\, W_k\, S_{TTF})^{-1}$ are updated as a background task at the same rate as the gradient noise covariance matrices C_k.

The LGS focus drift modes (one per LGS WFS), z_0, should be averaged and then sent to the NFIRAOS Component Controller to adjust the position of the LGS trombone mechanism.

The LGS tip/tilt drift modes t_0 are used to stabilize the commands of the FSM as described in the next section.

3.4.2.7 FSM Command Processing

The purpose of this task is to compute the commands for the fast tip/tilt mirrors (FSM) located in the Laser Guide Star Facility. These mirrors are used to stabilize the pointing of the laser guide stars on the sky. At each LGS WFS frame, the processing steps consist of:

- computing the overall noise-weighted tip, tilt and focus values (corresponding to t and z respectively) of each LGS WFS for the valid and non masked sub-apertures as described below:

$$\boxed{m = \begin{pmatrix} t \\ z \end{pmatrix} = S^{\dagger}_{TTF} \cdot s_{masked}} \tag{3.4.25}$$

  where $S^{\dagger}_{TTF}$ is the pseudo inverse of the tip/tilt/focus LGS sensitivity matrix defined in 3.4.2.6.

- The FSM tip/tilt vector is computed as follows:

$$\boxed{e_{FSM} = g_{LF} t_0 + t} \tag{3.4.26}$$

  where t and t_0 represent the tip/tilt modes and tip/tilt drift modes of the m and m_0 vectors respectively and where g_{LF} is the gain. [3]

- The tip/tilt vector e_{FSM} is integrated (scaled by a gain g_{FSM} and then temporally filtered) as follows:

$$\boxed{f_{FSM}(k) = g_{FSM} e_{FSM}(k) + f_{FSM}(k-1).} \tag{3.4.27}$$

- The filtered tip/tilt vector is then subtracted from the dither command vector c_{dit} computed in Eq.(3.4.8) and (3.4.9):

$$\boxed{c_{FSM} = c_{dit} - f_{FSM}} \tag{3.4.28}$$

- The commands c_{FSM} are then sent to the LGSF.

[3] g_{LF} could be 1.

3.4.3 Computation and Memory Requirements

Tasks	Number of MMAC per LGS WFS			Memory per LGS WFS (MB)
	Per Frame	Frequency (Hz)	Allowed computation time (s)	
LGS WFS pixel calibration	0.2	800	500E-6	1.2
LGS WFS pixel processing	0.41	800	500E-6	0.8
LGS WFS gradient gain estimation computations at 800Hz	0.6	800		2.34
LGS WFS gradient gain estimation computations at 1Hz	27	1		4.84
LGS WFS gradient noise covariance matrix computation	1.25	1		0.03
LGS WFS sub-aperture masking and reference subtraction	0.22	800	500E-6	0.25
LGS WFS drift modes computation	0.6	1		0.8
FSM command computation	0.006	800	50E-6	0
FSM control matrix computation	0.05	1		0.04
	Average MMAC/s			
Estimated total per WFS	1194			10.25

Table 1: Operation and memory requirements for the LGS WFS processing - Please review FSM computation and add new background process

Please note that the average computation rate is not representative of the peak computation rate, which is higher on account of the real time processes which must be completed in the first $550\,\mu$s of each frame. The peak computation rate will be at least 1660 MMAC/s per WFS during this interval.

3.5 NGS WFS Processing

3.5.1 Overview

This section describes the NGS WFS quadrant detector processing algorithm, and the optimization of the parameters of this algorithm in real time.

3.5.2 Algorithm Description

3.5.2.1 NGS WFS Pixel Calibration

The NGS WFS pixel intensities are calibrated (flat fielded and bias subtracted) before computing the gradients. This operation has already been described for the LGS WFS in section 3.4.2.1.

3.5.2.2 NGS WFS Pixel Processing

The NGS WFS gradients $s_{x,j}$ and $s_{y,j}$ in subaperture j are computed using a standard quadrant detector algorithm. The algorithm used to compute the slopes is given by:

$$\boxed{c_{x,j} = \frac{i_1 + i_4 - i_2 - i_3}{\sum_{k=1}^{4} i_k}} \quad \text{and} \quad \boxed{s_{x,j} = c_{x,j}/g_{x,j}} \tag{3.5.1}$$

$$\boxed{c_{y,j} = \frac{i_1 + i_2 - i_3 - i_4}{\sum_{k=1}^{4} i_k}} \quad \text{and} \quad \boxed{s_{y,j} = c_{y,j}/g_{y,j}} \tag{3.5.2}$$

where i_k is the pixel intensity for the pixel k in the quadrant detector, k varies from 1 to 4, and $g_{x,j}$ and $g_{y,j}$ are the centroids gains for the quadrant detector algorithm.

3.5.2.3 NGS WFS Gradient Gain Estimation

The centroids gains must be updated at a rate of 0.1Hz to account for change in the atmospheric turbulence conditions. The method used consists of injecting a known dither signal into the AO control loop and tracking the response to this pattern in the output of the NGS WFS. Sinusoidal signals are injected into the tip tilt commands of a dedicated tip/tilt mirror located in the NGS WFS path to minimize additional distortions of the science wavefront:

$$\boxed{d_x = a_1 \sin\left(2\pi k/n_d\right)} \tag{3.5.3}$$

$$\boxed{d_y = a_1 \cos\left(2\pi k/n_d\right)} \tag{3.5.4}$$

The frequency of the dither must be chosen carefully so that the signals are truly orthogonal to each other. The amplitudes of the signals shall be chosen to be significantly smaller than the errors being compensated for, but large enough to be detectable in the presence of noise.

The correlations between the injected sinusoid and the WFS measurements are then averaged over n_c dither cycles, to obtain the current boxcar-averaged signals h_x and h_y :

$$\boxed{\left(h_{x,j}\right)_q = \sum_{k=(q-1)n_d n_c+1}^{q n_d n_c} \left(d_x\right)_k \left(c_{x,j}\right)_k} \tag{3.5.5}$$

$$\boxed{\left(h_{y,j}\right)_q = \sum_{k=(q-1)n_d n_c+1}^{q n_d n_c} \left(d_y\right)_k \left(c_{y,j}\right)_k} \tag{3.5.6}$$

where:

- q is the update cycle number,
- n_d is the number of RTC frames per dither period,
- and n_c is the number of dither periods to average.

Finally, a temporal filter with a time constant of from 1 to 10 seconds is applied to provide a smooth estimate of the centroids gains:

$$\left(g_{x,j}\right)_q = \beta * \left(g_{x,j}\right)_{q-1} + (1-\beta) * \frac{2}{a_1 n_d n_c}\left(h_{x,j}\right)_q \tag{3.5.7}$$

$$\left(g_{y,j}\right)_q = \beta * \left(g_{y,j}\right)_{q-1} + (1-\beta) * \frac{2}{a_1 n_d n_c}\left(h_{y,j}\right)_q \tag{3.5.8}$$

where β is temporal filter gain.

The dither signals are removed from the instantaneous WFS measurements when the tip tilt modes of the wavefront are obtained from the NGS WFS directly instead of an additional 2x2 OIWFS located in the instrument. In such case, the slopes for each sub-aperture are computed as follows:

$$s_{x,j} = c_{x,j}/g_{x,j} - d_x \tag{3.5.9}$$

$$s_{y,j} = c_{y,j}/g_{y,j} - d_y \tag{3.5.10}$$

3.5.2.4 NGS WFS Gradient Noise Covariance Matrix Computation

The NGS WFS gradient noise covariance matrix C is computed by the RTC, concurrently with estimating the NGS WFS gains. This matrix is used later by the NGS wavefront reconstruction process. The C matrix is a block diagonal matrix composed of 2x2 blocks (each block corresponds to a sub-aperture). Each block j is obtained as follows:

$$[C]_j = \begin{pmatrix} \gamma_j/g_{x,j}^2 & 0 \\ 0 & \gamma_j/g_{y,j}^2 \end{pmatrix} \tag{3.5.11}$$

where $\gamma_j = \frac{I_{T,j}+4\sigma_E^2}{I_{T,j}^2}$ and $I_{T,j} =< \sum_{k=1}^{4} i(k) >$ corresponds to the time averaged total intensity level for the sub-aperture j, and σ_E is the detector read noise.

3.5.2.5 NGS WFS Subaperture Masking and Reference Subtraction

The process to identify and remove the unusable sub-apertures due to the telescope structure obscuration and/or due to M3 alignment problem is identical to the LGS process described in section 3.4.2.5 with s_{masked} being the vector of NGS WFS gradients obtained after applying the mask of unusable sub-apertures.

Finally, the NGS WFS gradient vector is obtained as follows:

$$s = s_{masked} - v_{mask} \cdot * r \tag{3.5.12}$$

where v_{mask} is the sub-aperture mask vector and r is the NGS WFS reference vector. The NGS WFS reference vector is updated from the TWFS at a rate of 0.1Hz and resampled to the NGS WFS frame rate (the NGS WFS reference vector is computed from the science non common path aberration vector which is derotated into the NFIRAOS coordinate system based on the instrument rotator angle, and the NGS path NCPA vector).

3.5.2.6 NGS WFS Spot Detection

During NGS AO acquisition, the RTC checks whether the NGS guide star falls within the FOV of the NGS WFS:

- First the AOSQ sets the NGS WFS frame rate to the minimum value of 10Hz. Then, the RTC computes the total time averaged intensities for each sub-aperture and compares it to the threshold T_1. The list of adequately illuminated sub-apertures is then compared against the value of the illumination mask v_{mask} for the current elevation angle. The NGS guide star falls within the FOV of the NGS WFS if nearly [4] all of the expected the sub-apertures are illuminated.
- The list of additional unusable sub-apertures is output to the RPG for initialization of the NGS reconstructor.
- The RTC then computes the average sub-aperture intensity, and passes it to the RPG which uses it as an input together with the most recent Fried parameter r_0 and the most recent atmospheric time constant τ_0 (computed by the RTC during the previous observation) to determine the recommended initial NGS WFS frame rate from a LUT populated during the commissioning of the NFIRAOS system.

[4] To allow for a few sub-apertures to be obscured due to a possible M3 alignment problem

3.5.3 Computation and Memory Requirements

Tasks	Number of MMAC			Memory
	Per Frame	Frequency (Hz)	Allowed computation time (s)	(MB)
NGS WFS pixel calibration	0.01	up to 800	500E-6	0.07
NGS WFS pixel processing	0.03	up to 800	500E-6	0.03
NGS WFS gradient gain estimation computations at 800Hz	0.01	up to 800		0.02
NGS WFS gradient gain estimation computations at 0.1Hz	0.01	0.1		0
NGS WFS gradient noise covariance matrix computation	0.03	0.1		0.03
NGS WFS sub-aperture masking and reference subtraction	0.02	up to 800	500E-6	0.25
	Average MMAC/s			
Estimated total	58			0.40

Table 2: Operation and memory requirements for the NGS WFS processing

Please note that the average computation rate is not representative of the peak computation rate, which is higher on account of the real time processes which must be completed in the first $500\,\mu$s of each frame. The peak computation rate will be at least 120 MMAC/s during this interval.

3.6 OIWFS Processing

3.6.1 Overview

This section describes (i) the OIWFS gradient processing, which uses a constrained matched filter algorithm, (ii) the real time optimization of the parameters used by this algorithm and (iii) the computation of the residual field rotation error when three OIWFS are used.

3.6.2 Algorithm Description

3.6.2.1 OIWFS Pixel Calibration

The OIWFS pixel intensities of each OIWFS are calibrated (flat fielded and bias subtracted) before computing the gradients. This operation has already been described for the LGS WFS in section 3.4.2.1. Note however, that for the OIWFS, the dark subtraction may be frame rate dependent as described below:

$$\boxed{i = (p - \alpha(\tau)\, *\, d) \cdot *f} \tag{3.6.1}$$

where $\alpha(\tau)$ is dependent upon the integration time τ.

3.6.2.2 OIWFS Pixel Processing

The computation of the OIWFS gradients uses a constrained, matched filter algorithm as described above in section 3.4.2.2 for the case of the LGS WFS, except that the vector i_0 of reference pixel intensities and the Jacobian matrix G of pixel intensities gains are computed assuming a Gaussian model for the Shack-Hartmann spots. The reference image i_0 is defined by the equation

$$i_0(k) = \alpha \exp\left\{-\left[x^2(k) + y^2(k)\right]/2\sigma^2\right\}, \tag{3.6.2}$$

where

- k is a pixel index within the subaperture;
- α is an intensity scale factor;
- $x(k)$ and $y(k)$ are pixel coordianates within the subaperture, expressed in units of pixel widths; and
- σ is the width of the Gaussian profile, again expressed in units of pixels.

The background estimation of the parameters α and σ is described in section 3.6.2.3 below. The elements of the Jacobian gain matrix G are given by

$$\frac{\partial i(k)}{\partial\{x,y\}} = -\left(\frac{\alpha\{x,y\}(k)}{\sigma^2}\right) \exp\left\{-\left[x^2(k) + y^2(k)\right]/2\sigma^2\right\}. \tag{3.6.3}$$

The quantities $\widetilde{G}$, W, and R_c are all computed from G and i_0 as described in section 3.4.2.2 above for the case of the LGS WFS, and all of these quantities are updated at the same rate as the parameters i_0 and α. Note that the constrained matched filter is applied on a subarray of pixels of size $n_f \times n_f$ which may be smaller than the OIWFS subaperture window size $n_w \times n_w$.

Once the constrained matched filter R_c has been computed, the estimate of the wavefront gradient s in the subaperture is obtained according to

$$s = \begin{cases} R_c i_s(p) + p - r & \text{if sum}(i_s(p)) > t \\ 0 & \text{otherwise} \end{cases} \tag{3.6.4}$$

Here p are the coordinates of the peak pixel in the image (TBC), i_s is a subarray centered around the peak pixel, and r are the reference gradients for the subaperture and t is a threshold on the minimum total signal level.

3.6.2.3 OIWFS Gradient Gain Estimation

The parameter σ is initialized from a LUT, which is based upon the WFS frame rate, the WFS signal level, the most recent estimate of the Fried parameter r_0, and the atmospheric time constant

τ_0. Then σ is updated in real time at a rate of 0.1Hz to account for change in the atmospheric turbulence parameters in order to minimize the function $e(\sigma)$ described as follows:

$$e(\sigma) = \sum_{k=1}^{n_{pix}} (< i(k) > -\alpha i_0(k,\sigma))^2, \quad (3.6.5)$$

where $< i(k) >$ represents the time-averaged intensity for the pixel k computed as a rate of 0.1Hz, and where scaling factor α is given by

$$\alpha = \frac{\sum_{k=1}^{n_{pix}} (< i(k) > i_0(k,\sigma))}{\sum_{k=1}^{n_{pix}} i_0^2(k,\sigma)}. \quad (3.6.6)$$

3.6.2.4 OIWFS Gradient Noise Covariance Matrix Computation

The OIWFS gradient noise covariance matrix C is computed by the RTC, concurrently with estimating the OIWFS gradient gains. This matrix is then passed to the RPG to update the low order modal reconstruction matrix. The C matrix is a block diagonal matrix composed of 2x2 blocks, with each block corresponding to a sub-aperture. Each block j is obtained as follows:

$$\boxed{[C]_{jj} = R_c W R_c^T} \quad (3.6.7)$$

where R_c is the matched filter for subaperture j,and W is the noise covariance matrix for the pixel intensities in subaperture j. Note that each of these blocks will actually be a multiple of the 2x2 identity matrix on account of the symmetric Gaussian Shack-Hartmann spot model used to derive the matched filter, so there is no need to transform the matrix to account for the coordinate system rotation between NFIRAOS and the OIWFS.

3.6.2.5 OIWFS Spot Detection

The algorithm used to detect the OIWFS spots during an acquisition sequence is similar to the one used for the high order NGS WFS. The RTC works in concert with the AOSQ and RPG to check if the OIWFS guide stars fall within the FOV of the OIWFS's by setting the frame rate for each active OIWFS to the minimum, which is 10Hz. Then, the RTC computes the unweighted total intensities for each sub-aperture and compares it to a threshold t_3. The OIWFS guide stars fall within the FOV of the WFS if all the sub-apertures are illuminated.

3.6.2.6 Instrument Rotator Angle Offset

This algorithm is used only when the three OIWFS are used in the LGS AO mode of IRIS (and eventually for WIRC and NIRES).

Let $\overline{s_x^l}$ and $\overline{s_y^l}$ be the x and y time and sub-aperture averages of the OIWFS gradient vectors s, with l varying from 1 to 3. The time averages are computed at a rate of 20 Hz. The rotation component of the OIWFS centroids is obtained as follows:

$$\boxed{\delta\,\theta_R = \sum_{l=1}^{3} \frac{1}{r_l^2} (-r_l \sin\theta_l, r_l \cos\theta_l).(\overline{s_x^l}, \overline{s_y^l})} \quad (3.6.8)$$

where $(r_l \cos\theta_l, r_l \sin\theta_l)$ are the nominal locations of the OIWFS guide stars in the NFIRAOS FOV.

The instrument component controller inputs the residual field rotation $\delta\,\theta_R$ provided by the RTC, applies a temporal filter and outputs the filtered angle to the instrument rotator mechanism.

3.6.3 Computation and memory requirements

Tasks	**Number of MMAC**			**Memory**
	Per Frame	**Frequency (Hz)**	**Allowed computation time (s)**	**(MB)**
OIWFS pixel calibration in LGS mode	0.098	up to 800	500E-6	0.28
OIWFS pixel calibration in NGS mode	0.033	up to 800	500E-6	0.09
OIWFS pixel processing in LGS mode	0.098	up to 800	500E-6	0.19
OIWFS pixel processing in NGS mode	0.033	up to 800	500E-6	0.03
OIWFS gradient gain estimation in LGS mode	7.7	0.1		1.13
OIWFS gradient gain estimation in NGS mode	2.55	0.1		0.09
OIWFS gradient noise covariance matrix computation in LGS mode	0.3	0.1		0.0
OIWFS gradient noise covariance matrix computation in NGS mode	0.098	0.1		0.0
OIWFS instrument rotator angle offset computation in LGS mode	0	20		0.0
	Average MMAC/s			
Estimated total in LGS mode	158			1.6
Estimated total in NGS mode	52			0.22

Table 3: Operation and memory requirements for the OIWFS processing

Please note that the average computation rate is not representative of the peak computation rate, which is higher on account of the real time processes which must be completed in the first $500\,\mu$s of each frame. The peak computation rate will be at least 392 MMAC/s during this interval.

Additionally, these computation and memory requirements have been computed for the full OIWFS image size (128x128 pixels/image). Both computation and memory reuirements would be propotionally smaller if windowing was applied.

3.7 LGS Wavefront Reconstruction

3.7.1 Overview

LGS wavefront reconstruction refers to the computation of wavefront values at DM actuator locations from closed loop LGS gradients in the LGS AO mode. The approach adopted is based on computationally efficient algorithms that provide iterative solutions to open loop minimum variance wavefront reconstruction with a sparse approximation of the regularization term. In this framework, wavefront values are first estimated on a set of phase screens (tomography step), and this estimate is then projected onto DM actuator locations (fitting step) to optimize performance (wavefront variance and Strehl ratio) averaged over a specified field of view (FoV). However, the proposed reconstruction architecture involves the following important modifications from conventional open loop minimum variance reconstruction:

(i) 5 modes, namely tip/tilt (TT) and 3 plate scale or tilt anisoplanatism (TA) modes, are controlled at a lower bandwidth by a second servo loop using a noise-weighted least-squares reconstruction matrix operating on closed loop OIWFS gradients. The computation of these modes is described in section 3.9.

(ii) All modes orthogonal to these 5 TT/TA modes are controlled at high bandwidth based upon pseudo open loop LGS gradient measurements using computationally efficient algorithms which approximately solve a sparse-plus-low-rank tomography matrix system and a sparse fitting matrix system.

(iii) The pseudo open loop LGS gradients are tip/tilt (TT) and differential focus (DF) removed on account of LGS position uncertainty, and these modes are also removed inside the tomography matrix.

(iv) The 5 TT/TA modes are projected from the LGS DM actuator vector at the output of the fitting step to reduce the cross-coupling between the two loops (although it is possible to derive a theoretical model of fully decoupled LGS/NGS loops, simulation results indicate that these loops still cross-couple due to system nonlinearities).

(vi) The tomography matrix incorporates a sparse, scaled, Laplacian squared regularization matrix (high-pass spatial filter) which approximates the atmospheric turbulence inverse covariance matrix, summed with a scaled identity matrix to prevent the system from being singular.

(vii) Piston removal and piston regularization constraints are not imposed on the fitting system.

The main benefits of the split tomography approach are:

- a simpler formulation of the LGS tomography step,
- a simpler, more flexible control of the TT and TA modes,
- a reduced coupling betweeen the LGS and TT/TA modes.

LGS wavefront reconstruction performs operations on 3 types of wavefront grids: a set of atmospheric phase screens, an aperture-plane grid, and two DM grids. In what follows, four algorithms will be discussed to perform atmospheric tomograhy. Three of them are conveniently implemented on square grids and atmospheric grids in the cone coordinate system where mesh size is squeezed with range according to the cone compression factor:

$$\Delta_k = \alpha_k \, \xi_k \, \Delta_0 \quad (3.7.1)$$
$$\xi_k = 1 - h_k/h_{\mathsf{lgs}} \leq 1 \quad (3.7.2)$$
$$\alpha_k = 1 \text{ or } 2 \quad (3.7.3)$$
$$\Delta_0 = d_{\mathsf{sa}}/2 = 1/4\,\text{m} \quad (3.7.4)$$

where Δ_0 denotes the aperture-plane grid mesh size and h_{lgs} the range to the sodium layer. Note that regardless of the choice of coordinate system, any LGS tomographic wavefront reconstruction algorithm needs to be updated slowly with time due to sodium layer range temporal variability. In the cone coordinate system, this update is most conveniently accomplished by updating the atmospheric grids themselves, i.e. updating the above ξ_k parameters.

Four core operations are involved in the overall LGS wavefront reconstruction process:

- Bilinear and bicubic interpolation. This operation computes wavefront values at the intercepts of rays traced through phase screens (atmospheric and DM grids) to/from the aperture plane, for sources either at finite or infinite range. Rays traced to/from the aperture plane correspond to forward/backward ray tracing operations, which are transposes of one another. Projected aperture-plane grid point coordinates on a plane at range h for a source at range h_{src} in direction θ are given by the following expression:

  $$x' = (1 - h/h_{\mathsf{src}})\, x + h\, \theta\,, \quad (3.7.5)$$

  where x denotes aperture-plane grid points coordinates.

- Aperture-plane gradient and gradient transpose. The gradient operator computes each LGS WFS gradient measurement from wavefront values at up to 9 points of the aperture-plane grid on a 3×3 stencil covering the associated subaperture. Stencil weights are fully determined by the amplitude profile defining a subaperture. At least 80% of the subapertures in each WFS are expected to be fully illuminated, and the gradient operator associated with these subapertures has a regular stencil coupling only the 6 points that border opposite subaperture edges (3 points on each side with opposite weights). Similarly, the gradient transpose operator associated with these subapertures has a regular stencil coupling either 0, 2 or 4 subapertures into each grid point.

- Atmospheric wavefront curvature squared. This operation regularizes the tomography step. The curvature squared operator couples 13 points lying on a 5×5 stencil, with 3 points from each of the 4 corners removed. The 13 weights are identical for all atmospheric phase screens, and are identical throughout each screen, except near the boundary.

- Aperture weighting. This operation is performed in the DM fitting step to (i) capture aperture edge effects and (ii) yield a lower fitting error by cross-coupling aperture-plane grid points with up to 8 nearest neighbors on a 3×3 stencil. Stencil weights are fully determined by the amplitude profile defining the aperture, and are identical throughout the aperture, except near the bounday. Fully interior stencils have Simpson weights.

All four of these operations shall be described in greater detail below. Grid-based as opposed to sparse matrix-based operations shall be implemented since they have the highest potential to (i) fully exploit the underlying structure of these operators, (ii) minimize storage requirements, and (iii) most importantly, exploit hardware parallelism. Both grid-based and sparse matrix implementations shall be investigated for the gradient/gradient transpose and aperture weighting operators with somewhat irregular, but highly sparse, stencils.

3.7.2 Algorithm Description

3.7.2.1 LGS Pseudo Open Loop Gradient Computation

For stability reasons, the LGS loop must operate on pseudo open loop gradients. In the absence of measurement noise and systematic errors (WFS pupil distortion, misregistration/calibration errors), the poles of the LGS closed loop transfer matrix are then equal to those of a conventional minimum variance reconstructor with closed loop constraints. Pseudo open-loop gradients represent an estimate of the open loop gradients that would have been obtained with ideal flat DMs and ideal fully linear wavefront sensors. As a result, the pseudo open loop architecture eliminates the need for an invisible mode removal step as in a conventional closed loop system. The pseudo open loop gradients associated to LGS WFS k are computed as follows:

- The influence of each DM upon the guidestar wavefronts is determined using bicubic spline interpolation on both DM grids at intercepts of rays traced from a finite range source in the direction of LGS number k to the aperture-plane grid. The bicubic spline interpolation operator for DM grid l will be denoted $\left[\widetilde{H}_{\mathsf{dm}}\right]_{kl}$, where tilde indicates ray tracing from a finite range source. The width of actuator influence function is $4d_{sa}$, where d_{sa} is the inter-actuator spacing. Since the aperture-plane grid is $2\times$ oversampled with respect to both DM grids, bicubic interpolation on the ground-level DM grid couples the nodes of the aperture-plane grid to either 9, 12, or 16 actuators with fixed weights depending upon the inter-actuator coupling coefficient, c. Typical values take the form:

$$w = \begin{cases} u\,u^{\mathsf{T}} & \text{or} \\ v\,v^{\mathsf{T}} & \text{or} \\ u\,v^{\mathsf{T}} & \text{or} \\ v\,u^{\mathsf{T}} & \end{cases}$$

$$\text{where} \quad u = \begin{bmatrix} c \\ 1 \\ c \end{bmatrix} \quad \text{and} \quad v = \frac{1}{16}\begin{bmatrix} 6c-1 \\ 9+10c \\ 9+10c \\ 6c-1 \end{bmatrix} \qquad (3.7.6)$$

  These weights will be downloaded to the RTC during system initialization.

  On the other hand, on the upper DM grid, LGS ray intercepts are surrounded by 4×4 nearest actuators with bicubic spline weights which take the form:

$$
\begin{aligned}
w &= u(\beta_y)\, v^{\mathsf{T}}(\beta_x) \\
\text{where} \quad u(\beta_y) &= \begin{bmatrix} (2c-\frac{1}{2})(1-\beta_y)^2 + (\frac{1}{2}-c)(1-\beta_y)^3 \\ 1 + (4c-\frac{5}{2})\beta_y^2 + (\frac{3}{2}-3c)\beta_y^3 \\ 1 + (4c-\frac{5}{2})(1-\beta_y)^2 + (\frac{3}{2}-3c)(1-\beta_y)^3 \\ (2c-\frac{1}{2})\beta_y^2 + (\frac{1}{2}-c)\beta_y^3 \end{bmatrix} \\
v(\beta_x) &= P\, u(\beta_x) \qquad (3.7.7)
\end{aligned}
$$

where (β_x, β_y) is the ray intercept offset (expressed in terms of the DM grid mesh size d_{sa}) from the lower left nearest neighbor actuator in the 4×4 stencil, and P permutes u in reverse order and arises from the fact that the reference point is the lower left nearest neighbour actuator. Note that this offset is not constant throughout the upper DM grid, since that grid is not in the cone coordinate system and the intercepts are from finite range sources. The x- and y-weights for each LGS can be stored separately since the x-weights are only a function of the x-intercept of a ray in the DM conjugate plane (and similarly for the y-weights). Finally, note that the values of these weights need to be downloaded to the RTC at a slow rate of 0.1 Hz as the range to the sodium layer varies.

- Any dead or detached DM actuators do not contribute to the pseudo-open-loop gradient computation. The list of such actuators is fixed and will be downloaded to the RTC during system initialization.
- Additionally, the influence functions of the 8 (TBC) nearest neighbors of a detached actuator will be perturbed by the absence of the constraint normally imposed by that actuator, and hence their interpolation weight will be changed. Formally, the DM influence operator can then be expressed as $\widetilde{H}_{dm} = \widetilde{H}^0_{dm} + \widetilde{H}^\delta_{dm}$, where the nominal operator $\widetilde{H}^0_{dm}$ is computed with the stencil weights given in (3.7.6) and (3.7.7), and $\widetilde{H}^\delta_{dm}$ contains the pertubed stencil weights of the nearest neighbors of up to 10 (TBC) detached actuators.
- Summation of the interpolated wavefront values on both DM grids.
- Aperture-plane gradient computations for the summed wavefronts. This operator will be denoted Γ. The gradient operator associated with fully illuminated subapertures has the following regular stencil:

$$
w_x = \frac{1}{d_{\mathsf{sa}}} \begin{bmatrix} -1/4 & 0 & 1/4 \\ -1/2 & 0 & 1/2 \\ -1/4 & 0 & 1/4 \end{bmatrix}, \quad w_y = -w_x^{\mathsf{T}} \qquad (3.7.8)
$$

where $d_{\mathsf{sa}} = 2\,\Delta_0 = 1/2$ m denotes the subaperture width. Note that the central point of the stencil has a zero weight for both measurement directions. For the gradient transpose operator (which will appear in future subsections), grid points that are the common vertex to 4 fully illuminated subapertures, receive a non-zero contribution from both x- and y-gradients with the followiing weights:

$$
w_x = \frac{1}{d_{\mathsf{sa}}} \begin{bmatrix} 1/2 & -1/2 \\ 1/2 & -1/2 \end{bmatrix}, \quad w_y = -w_x^{\mathsf{T}} \qquad (3.7.9)
$$

whereas grid points that are the common mid-point node bordering 2 fully illuminated sub-apertures, receive a non-zero contribution from either the x- or y-gradients with the followiing weights:

$$w_x = \frac{1}{d_{\mathsf{sa}}} \begin{bmatrix} 1 & -1 \end{bmatrix}, \quad w_y = -w_x^{\mathsf{T}} \tag{3.7.10}$$

In condensed notation, the above operations may be written as follows:

$$[s_{\mathsf{dm}}]_k = \Gamma \begin{bmatrix} \left[\widetilde{H}_{\mathsf{dm}}\right]_{k1} & \left[\widetilde{H}_{\mathsf{dm}}\right]_{k2} \end{bmatrix} \begin{bmatrix} [\widehat{x}]_1 \\ [\widehat{x}]_2 \end{bmatrix}, \quad 1 \leq k \leq N_{\mathsf{lgs}} \tag{3.7.11}$$

$$[s_{ol}]_k = [s]_k + [s_{\mathsf{dm}}]_k \tag{3.7.12}$$

where s_{ol} and s denote respectively the pseudo open loop and closed loop gradient vectors, and $\widehat{x} = c_{DM,clipped}$ is the DM actuator command vector relative to the nominal “flat” command (see section 3.12). The aperture-plane gradient operator Γ and the upper DM interpolation operator $\left[\widetilde{H}_{\mathsf{dm}}\right]_{k2}$ (interpolation weights or sparse matrix) are updated by the RPG as the LGS range and the pupil rotation angle vary, at a rate of 0.1 Hz.

The RTC computes the telescope M1 scalloping mode coeffcient as the inner product of the averaged pseudo open loop LGS WFS gradients, $\langle s_{ol} \rangle$, with the gradient scalloping mode vector:

$$m_{scall} = \langle s_{ol}^T \rangle s_{scall} \tag{3.7.13}$$

The pseudo open loop gradients s_{ol} are averaged during a period of 5 sec, and m_{scall} is then sent to the TCS.

3.7.2.2 LGS Tomography

LGS tomography is decomposed into 3 main steps:

- LGS pseudo open loop gradient tip/tilt removal.
- LGS tomography matrix system right hand side vector computation.
- LGS tomography matrix system approximate solution computation.

These steps are detailed below.

3.7.2.2.1 LGS Pseudo Open Loop Gradient Tip/Tilt and Differential Focus Removal

The first step of LGS tomography consists in removing global tip/tilt and differential focus from the LGS pseudo open loop gradients. Denoting $M = \begin{bmatrix} M_{\mathsf{TT}} & M_{\mathsf{DF}} \end{bmatrix}$ the matrix whose first 12 columns, M_{TT}, are the gradient tip/tilt modes and last 5 columns, M_{DF}, are the gradient differential focus modes, we have that:

$$M_{\mathsf{TT}} = \mathsf{Diag}(T, \cdots, T) \quad \text{with} \quad T = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{3.7.14}$$

where 0 and 1 indicate respectively vectors of zeros and ones, and M_{DF} is non block-diagonal and of the form:

$$M_{\text{DF}} = \begin{bmatrix} f & f & f & f & f \\ -f & 0 & 0 & 0 & 0 \\ 0 & -f & 0 & 0 & 0 \\ 0 & 0 & -f & 0 & 0 \\ 0 & 0 & 0 & -f & 0 \\ 0 & 0 & 0 & 0 & -f \end{bmatrix} \tag{3.7.15}$$

where 0 indicates again a vector of zeros and f denotes a vector whose components are the x- and y- coordinates of the subaperture centers of a LGS WFS.

The gradient tip/tilt and differential focus mode removal operator is defined as $I - P_{\text{M}}$, where $P_{\text{M}} = M\, S_{\text{M}}^{-1}\, M^{\mathsf{T}}\, W$. $S_{\text{M}} = M^{\mathsf{T}}\, W\, M$ is the 17×17 mode cross-coupling matrix (which is symmetric and will in general be fully populated), and $W = \text{Diag}(W_1, \cdots, W_{N_{\text{lgs}}})$ denotes the block-diagonal thresholded LGS measurement noise inverse covariance matrix (sparse symmetric weighting matrix with 2 non-zero elements per row). Application of the mode removal operator to the pseudo open loop gradient vector s_{ol} (or other vectors, as described in the following paragraphs) is trivially implemented by first computing mode coefficients and then removing modes:

$$\widehat{m} = S_{\text{M}}^{-1}\, M^{\mathsf{T}}\, W\, s_{ol} \tag{3.7.16}$$

$$\overline{s_{ol}} = s_{ol} - M\, \widehat{m} \tag{3.7.17}$$

3.7.2.2.2 LGS Tomography Matrix System Right Hand Side Vector Computation

The second step of LGS tomography consists in assembling the right hand side vector of the tomography matrix system from the tip/tilt removed LGS pseudo open loop gradients, which is accomplished by the following operations:

$$[y]_k = \Gamma^{\mathsf{T}}\, W_k\, [\overline{s_{ol}}]_k\,, \qquad 1 \le k \le N_{\text{lgs}} \tag{3.7.18}$$

$$[b]_l = \begin{bmatrix} \left[\widetilde{H}_{\text{atm}}\right]_{1l}^{\mathsf{T}} & \cdots & \left[\widetilde{H}_{\text{atm}}\right]_{N_{\text{lgs}}l}^{\mathsf{T}} \end{bmatrix} \begin{bmatrix} [y]_1 \\ \vdots \\ [y]_{N_{\text{lgs}}} \end{bmatrix}, \qquad 1 \le l \le N_{\text{ps}} \tag{3.7.19}$$

where:

- $[b]_l$ denotes the right hand side tomography sub-vector corresponding to phase screen l,
- $\left[\widetilde{H}_{\text{atm}}\right]_{kl}$ is the bilinear interpolation operator between the nodes of atmospheric layer l and intercepts of rays traced to the aperture-plane grid from the direction of LGS WFS k. Tilde indicates again ray tracing from a finite range source. Note that the interpolation offset (β_x, β_y) is constant throughout the grid when grid nodes are in the cone coordinate system, which is slowly updated to reflect sodium layer range variability. This property is particularly well suited for a parallel grid-based implementation.

The bilinear interpolation operators $\left[\widetilde{H}_{\mathsf{atm}}\right]_{kl}$ are updated by the RPG as the range to the screens (i.e. zenith angle) and range to the LGS vary, at a rate of 0.1 Hz. The nodes of the aperture-plane grid can always be chosen to coincide with nodes of the ground-layer phase screen, such that no interpolation is needed for that screen.

3.7.2.2.3 LGS Tomography Matrix System Approximate Solution Computation

The tomography matrix system takes the following form:

$$A\,x \quad = \quad b\,, \tag{3.7.20}$$

where:

- x denotes the concatenated tomography vector of unknowns (atmospheric phase screens)
- A is the block structured tomography operator
- b is the tomography right hand side vector.

All tomography solvers discussed below perform $[A]_{kl}\ [v]_l$ operations ($1 \leq k, l \leq N_{\mathsf{ps}}$), which can be expressed as follows:

$$\psi = \mathsf{Diag}(\Gamma^{\mathsf{T}}, \cdots, \Gamma^{\mathsf{T}})\, W\, (I - P_{\mathsf{M}})\, \mathsf{Diag}(\Gamma, \cdots, \Gamma) \begin{bmatrix} \left[\widetilde{H}_{\mathsf{atm}}\right]_{1l} \\ \vdots \\ \left[\widetilde{H}_{\mathsf{atm}}\right]_{N_{\mathsf{lgs}}l} \end{bmatrix} [v]_l \tag{3.7.21}$$

$$[p]_k = \begin{bmatrix} \left[\widetilde{H}_{\mathsf{atm}}\right]^{\mathsf{T}}_{1k} & \cdots & \left[\widetilde{H}_{\mathsf{atm}}\right]^{\mathsf{T}}_{N_{\mathsf{lgs}}k} \end{bmatrix} \psi \tag{3.7.22}$$

$$[q]_k = \delta_{kl}\,(\gamma_k\, L^2 + \epsilon)\ [v]_l \tag{3.7.23}$$

$$[A]_{kl}\ [v]_l = [p]_k + [q]_k \tag{3.7.24}$$

where:

- $I - P_{\mathsf{M}}$ is implemented as described for the pseudo open loop gradient vector
- δ_{kl} denotes the Kronecker delta
- L^2 is the curvature squared regularization operator, which couples 13 grid points on a 5×5 stencil on each screen with the following weights:

$$w = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 2 & -8 & 2 & 0 \\ 1 & -8 & 20 & -8 & 1 \\ 0 & 2 & -8 & 2 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix} \tag{3.7.25}$$

– γ_k is a scaling constant proportional to Δ_k^{-4} and to the 5/6 th power of the Fried parameter of layer k, which is computed by the RTC in real time and sent to the RPG (refer to 3.11),

– and ϵ is a small regularization parameter to prevent A from being singular.

Four algorithms falling in two categories are proposed for solving the LGS tomography system: two block oriented algorithms operating on the blocks (i.e. layers) of the system one by one and two system-oriented algorithms operating on the whole system (i.e. all layers) at once. These algorithms are respectively:

- BGS-CBS: Block Gauss-Seidel with Cholesky Back Substitutions. The core of this algorithm performs forward/back substitutions on N_{ps} smaller systems $[A]_{kk}\ [x]_k = [v]_k$, where the right hand side vector $[v]_k$ depends on the solution obtained for all diagonal blocks above or below k (the later direction depends on whether forward or backward BGS is implemented). Memory requirements and operation counts are minimal when the algorithm is implemented on uncompressed non-square atmospheric grids which are matched in size to the “metapupil” associated with the NFIRAOS LGS/NGS asterisms and the science FoV. Note that interpolation is still well suited for a parallel grid-based implementation on such grids with some more book-keeping on account of edge effects.
- BGS-CG20: Block Gauss-Seidel with twenty iterations of Conjugate Gradient instead of the above Cholesky backsubstitutions. The algorithm is well suited to operate on square grids in the cone coordinate system.
- CG30: Thirty iterations of Conjugate Gradient operating on the whole tomography system. The algorithm is well suited to operate on square grids in the cone coordinate system.
- FD3: Three iterations of Conjugate Gradient with a Fourier Domain preconditioning matrix. The algorithm must operate on square grids in the cone coordinate system.

The following points are worth mentioning:

(i) Warm restart is used in all these iterative algorithms, and is crucial since it reduces the number of iterations significantly. The solution from frame k is used as the initial guess for frame $k+1$.

(ii) With the exception of FD3 and BGS-CBS, all the above algorithms have a virtually memory-free implementation and are order N, where N denotes system size.

Each algorithm is detailed below.

3.7.2.2.3.1 Block Gauss-Seidel

In what follows, forward BGS is assumed without loss of generality. In condensed notation, starting with a warm initial guess $x^{(0)}$ the algorithm approximately solves the following system:

$$(A_{\text{D}} + A_{\text{L}})\,x = y\,, \tag{3.7.26}$$

$$y = b - A_{\text{U}}\,x^{(0)}\,, \tag{3.7.27}$$

where A_{D}, A_{L} and A_{U} denote respectively the block diagonal, block lower and block upper parts of the tomography operator, and b is the tomography right hand side vector.

The algorithm is organized in two main steps: (i) assembling the transformed right hand side vector y, (ii) block lower triangular system solution layer by layer. Each step is detailed below.

3.7.2.2.3.1.1 Transformed Right Hand Side Vector Computation

The transformed right hand side vector (3.7.27) has the following block decomposition:

$$[y]_k = [b]_k - \left[A_{\mathsf{U}}\, x^{(0)}\right]_k , \quad (3.7.28)$$

$$\left[A_{\mathsf{U}}\, x^{(0)}\right]_k = \left[\begin{array}{ccc} \left[\widetilde{H}_{\mathsf{atm}}\right]^{\mathsf{T}}_{1k} & \cdots & \left[\widetilde{H}_{\mathsf{atm}}\right]^{\mathsf{T}}_{N_{\mathsf{lgs}}k} \end{array}\right] \mathsf{Diag}(\Gamma^{\mathsf{T}},\cdots,\Gamma^{\mathsf{T}})\, W\,(I - P_{\mathsf{M}})\,\mathsf{Diag}(\Gamma,\cdots,\Gamma) \times \left[\begin{array}{ccc} \left[\widetilde{H}_{\mathsf{atm}}\right]_{1k+1} & \cdots & \left[\widetilde{H}_{\mathsf{atm}}\right]_{1N_{\mathsf{ps}}} \\ \vdots & \vdots & \vdots \\ \left[\widetilde{H}_{\mathsf{atm}}\right]_{N_{\mathsf{lgs}}k+1} & \cdots & \left[\widetilde{H}_{\mathsf{atm}}\right]_{N_{\mathsf{lgs}}N_{\mathsf{ps}}} \end{array}\right] \left[\begin{array}{c} \left[x^{(0)}\right]_{k+1} \\ \vdots \\ \left[x^{(0)}\right]_{N_{\mathsf{ps}}} \end{array}\right] . \quad (3.7.29)$$

Eq.(3.7.29) illustrates that the number of block components processed increases from 0 for $k = N_{\mathsf{ps}}$ to $N_{\mathsf{ps}} - 1$ for $k = 1$. This suggests the following sequential pseudo code:

$[y]_{N_{\mathsf{ps}}} = [b]_{N_{\mathsf{ps}}}$; $\psi = 0$
For $k = N_{\mathsf{ps}} - 1, \cdots, 1$

$$\psi := \psi + \mathsf{Diag}(\Gamma^{\mathsf{T}},\cdots,\Gamma^{\mathsf{T}})\, W\,(I - P_{\mathsf{M}})\,\mathsf{Diag}(\Gamma,\cdots,\Gamma) \left[\begin{array}{c} \left[\widetilde{H}_{\mathsf{atm}}\right]_{1k+1} \\ \vdots \\ \left[\widetilde{H}_{\mathsf{atm}}\right]_{N_{\mathsf{lgs}}k+1} \end{array}\right] \left[x^{(0)}\right]_{k+1}$$

$$[y]_k = [b]_k - \left[\begin{array}{ccc} \left[\widetilde{H}_{\mathsf{atm}}\right]^{\mathsf{T}}_{1k} & \cdots & \left[\widetilde{H}_{\mathsf{atm}}\right]^{\mathsf{T}}_{N_{\mathsf{lgs}}k} \end{array}\right] \psi$$

End

3.7.2.2.3.1.2 Block Gauss-Seidel Matrix System Approximate Solution Computation

Processing of the block lower triangular system is organized as follows:

$[\widehat{x}]_1 = \mathsf{Solve}([A]_{11}, [y]_1)$
$\psi = 0$
For $k = 2, \cdots, N_{\mathsf{ps}}$

$$\psi := \psi + \mathsf{Diag}(\Gamma^{\mathsf{T}},\cdots,\Gamma^{\mathsf{T}})\, W\,(I - P_{\mathsf{M}})\,\mathsf{Diag}(\Gamma,\cdots,\Gamma) \left[\begin{array}{c} \left[\widetilde{H}_{\mathsf{atm}}\right]_{1k-1} \\ \vdots \\ \left[\widetilde{H}_{\mathsf{atm}}\right]_{N_{\mathsf{lgs}}k-1} \end{array}\right] [\widehat{x}]_{k-1}$$

$$[v]_k = [y]_k - \left[\begin{array}{ccc} \left[\widetilde{H}_{\mathsf{atm}}\right]^{\mathsf{T}}_{1k} & \cdots & \left[\widetilde{H}_{\mathsf{atm}}\right]^{\mathsf{T}}_{N_{\mathsf{lgs}}k} \end{array}\right] \psi$$

$$[\widehat{x}]_k = \mathsf{Solve}([A]_{kk}, [v]_k)$$

End

where the function $x = \mathsf{Solve}(A, b)$ provides either an exact or approximate solution of the generic system $A\,x = b$, with one of the following two methods:

1. Layer Solution via Cholesky Forward/Back Substitutions:

   Layer solution via Cholesky forward/back substitutions (CBS) does not operate on A but rather on the Cholesky factor L of a re-ordered version of the sparse component of A. It is by far the most economical algorithm in terms of number of operations on account of the high sparsity of L, however it requires (i) to store in cache N_{ps} Cholesky factors and N_{ps} low-rank matrices U' (further described below), and (ii) partially sequential forward/back substitutions for each layer. The Cholesky factors and low-rank matrices are updated by the RPG as turbulence conditions and range to the LGS change, at a rate of 0.1 Hz.

   The following paragraph describes the precomputation of the low-rank matrices U' in the RPG. First, $[A]_{kk}$ is rewritten in sparse-plus-low-rank form: $[A]_{kk} = [M]_{kk} - U\,U^{\mathsf{T}}$, and U is assembled as follows:

$$U = \left[\; \left[\widetilde{H}_{\mathsf{atm}}\right]^{\mathsf{T}}_{1k} \quad \cdots \quad \left[\widetilde{H}_{\mathsf{atm}}\right]^{\mathsf{T}}_{N_{\mathsf{lgs}}k} \; \right] \mathsf{Diag}(\Gamma^{\mathsf{T}} \; \cdots \; , \Gamma^{\mathsf{T}})\, W\, M\, S_{\mathsf{M}}^{-1/2} \tag{3.7.30}$$

   Next, U' is precomputed by calling the CBS routine sketched below for each column of U, i.e.

$$U' = \mathsf{CBS}(L, p, U) \tag{3.7.31}$$

   where the CBS function is given by the following pseudo code:

```
function x = CBS(L, p, b)

b := b(p) % permuted right hand side vector
Forward substitutions on L z = b
Back substitutions on L^T x = z
x(p) := x % inverse permuted solution vector
```

   U' is updated by the RPG at 0.1Hz as the turbulence conditions and range to the LGS change.

   The following 17×17 symmetric cross-coupling matrix needs also to be precomputed in the RPG:

$$S_{\mathsf{U}} = U'^{\mathsf{T}}\, U - I \tag{3.7.32}$$

   where I denotes the 17×17 identity matrix.

   Finally, the BGS-CBS solution is assembled as follows:

$$\boxed{x = \mathsf{CBS}(L, p, b) + x' \qquad (3.7.33)}$$

   where the low-rank contribution is given by the following expression:

$$\boxed{x' = U'\, S_{\mathsf{U}}^{-1}\, U'^{\mathsf{T}}\, b \qquad (3.7.34)}$$

2. Layer Solution via Conjugate Gradient:

   $K = 20$ iterations are required (weak layers have worse convergence on account of the weaker Laplacian square regularization term) and it is not clear wether this approach could be faster than the above partially sequential substitutions. A generic CG algorithm is provided below.

```
function x = CG(A, b, x)

r = b − A x % initial residual
p = r % initial search direction
For k = 1, ··· , K
        q = A p
        γ = r^T r
        α = γ/(p^T q)
        x := x + α p % solution update
        If k < K
                r := r − α q % residual update
                β = r^T r/γ
                p := β p + r % search direction update
        End
End
```

3.7.2.2.3.2 Conjugate Gradient

The CG algorithm presented in Section 3.7.2.2.3.1.2 as bullet number 2 can be applied to the whole tomography operator. Simulation results indicate that $K = 30$ iterations are required.

3.7.2.2.3.3 Fourier Domain Preconditioned Conjugate Gradient

The FDPCG algorithm operates in the spatial domain with a sparse pre-computed Fourier Domain Hermitian preconditioning matrix $\widehat{M}$ updated by the RPG as the range to the layers, range to the LGS, SNR, and turbulence conditions change, at a rate of 0.1 Hz. A generic PCG algorithm is provided below.

```
function x = PCG(A, b, x)

r = b − A x % initial residual
r′ = Precond(A, r) % initial residual of preconditioned system
p′ = r % initial search direction
For k = 1, ··· , K
        q = A p′
        γ = r′^T r
        α = γ/(p′^T q)
        x := x + α p′; % solution update
        If k < K
                r := r − α q % residual update
                r′ = Precond(A, r)
                β = r′^T r/γ
                p′ := β p′ + r′;
        End
End
```

Simulation results indicate that $K = 3$ iterations are required. The preconditioning step is implemented in parallel via N_{ps} ffts, a Hermitian matrix vector multiplication and N_{ps} inverse ffts, which in condensed notation can be written as follows:

function $x = \text{Precond}(A, b)$

$$x = \mathcal{F}^{-1}\,\widehat{M}\,\mathcal{F}\,b$$

where $\mathcal{F}$ denotes the block-diagonal 2D fft operator.

3.7.2.3 LGS fitting step

LGS fitting is decomposed into the following 3 main steps:

- LGS tomography estimate propagation along evaluation directions.
- LGS fitting matrix system right hand side vector computation.
- LGS fitting matrix system approximate solution computation.

Each item is discussed below.

3.7.2.3.1 LGS Tomography Estimate Propagation

The first step of LGS fitting consists in the geometrical propagation of the LGS tomography estimate (ray trace interpolation/accumulation) along N_{evl} evaluation directions from sources at infinite range, which can be represented as follows:

$$[y]_k = \begin{bmatrix} [H_{\text{atm}}]_{k1} & \cdots & [H_{\text{atm}}]_{kN_{\text{ps}}} \end{bmatrix} \begin{bmatrix} [\widehat{x}]_1 \\ \vdots \\ [\widehat{x}]_{N_{\text{ps}}} \end{bmatrix}, \qquad 1 \le k \le N_{\text{evl}} \quad (3.7.35)$$

where:

- $[\widehat{x}]_k$: tomography estimate for layer k
- $[H_{\text{atm}}]_{kl}$: sparse bilinear interpolation operator between the nodes of atmospheric grid l and intercepts of rays traced to the aperture-plane grid from the evaluation direction k. The absence of tilde on H indicates ray tracing from infinite range sources. Note that the interpolation offsets (β_x, β_y) are not constant on a given phase screen, since atmospheric phase screens are in the cone coordinate system and ray tracing is from infinite range sources. However, the offsets need to be stored for only 1 row/column of grid points for each infinite range source, and are updated by the RPG as the range to the LGS and the choice of evaluation directions vary, at a rate of 0.1 Hz. Typically, ray tracing is done in $N_{\text{evl}} = 9$ evaluation directions.

3.7.2.3.2 LGS Fitting Matrix System Right Hand Side Vector Computation

The fitting right hand side vector is computed as follows:

$$[q]_k \quad = \quad \omega_k \, W \, [y]_k \; , \qquad 1 \leq k \leq N_{\mathsf{evl}} \tag{3.7.36}$$

$$[b]_l \quad = \quad \left[\; [H_{\mathsf{dm}}]^{\mathsf{T}}_{1l} \quad \cdots \quad [H_{\mathsf{dm}}]^{\mathsf{T}}_{N_{\mathsf{evl}}l} \; \right] \begin{bmatrix} [q]_1 \\ \vdots \\ [q]_{N_{\mathsf{evl}}} \end{bmatrix} , \qquad 1 \leq l \leq N_{\mathsf{dm}} = 2 \tag{3.7.37}$$

where:

- $[b]_l$ is the right hand side sub-vector for DM grid l
- $[H_{\mathsf{dm}}]_{kl} = \left[H^0_{\mathsf{dm}}\right]_{kl} + \left[H^\delta_{\mathsf{dm}}\right]_{kl}$ is a sparse interpolation operator between the actuators of DM grid l and intercepts of the rays traced to the aperture-plane grid from the evaluation direction k. The absence of tilde on H indicates projection from infinite range sources. The term H^0_{dm} is the nominal bicubic spline interpolation operator, and H^δ_{dm} contains the perturbed stencil weights of the nearest neighbor actuators of up to 10 (TBC) detached actuators. The width of each influence function is $4d_{sa}$.

  The interpolation operator for DM number 1 at $h = 0$ km is fixed and independent of the evaluation direction; the interpolation operator for DM number 2 is a function of the evaluation direction, but the bicubic spline weights $\left[H^0_{\mathsf{dm}}\right]_{k2}$ are constant for all points in the aperture plane grid. The stencil weights defining both $\left[H^0_{\mathsf{dm}}\right]_{k2}$ and $\left[H^\delta_{\mathsf{dm}}\right]_{k2}$ are updated by the RPG at a slow rate of 0.1 Hz as the evaluation directions vary (e.g., as the science FoV rotates in the NFIRAOS coordinate system).
- ω_k is the scalar weight for evaluation direction k, which is updated by the RPG as the importance of evaluation directions vary.
- W is a fixed sparse aperture-plane weighting operator coupling each grid point with up to 8 nearest neighbors on a 3×3 stencil. Fully interior stencils have Simpson weights:

$$w = \begin{bmatrix} 1/6 \\ 2/3 \\ 1/6 \end{bmatrix} \begin{bmatrix} 1/6 & 2/3 & 1/6 \end{bmatrix} = \begin{bmatrix} 1/36 & 1/9 & 1/36 \\ 1/9 & 4/9 & 1/9 \\ 1/36 & 1/9 & 1/36 \end{bmatrix} . \tag{3.7.38}$$

3.7.2.3.3 LGS Fitting Matrix System Approximate Solution Computation

The LGS fitting matrix system to solve is the following:

$$A\,x \quad = \quad b\,, \tag{3.7.39}$$

where:

- x is the concatenated DM actuator vector of unknowns
- b is the concatenated fitting right hand side vector

- A is the block-structured fitting operator.

The proposed fitting solver is $K = 4$ CG iterations operating on the whole fitting system. The solver performs $[A]_{kl}\ [v]_l$ operations ($1 \leq k, l \leq N_{\text{dm}} = 2$), which can be expressed as follows:

$$\psi = \text{Diag}(\omega_1 W, \cdots, \omega_{N_{\text{evl}}} W) \begin{bmatrix} [H_{\text{dm}}]_{1l} \\ \vdots \\ [H_{\text{dm}}]_{N_{\text{evl}} l} \end{bmatrix} [v]_l \quad (3.7.40)$$

$$[A]_{kl}\ [v]_l = \begin{bmatrix} [H_{\text{dm}}]^{\mathsf{T}}_{1k} & \cdots & [H_{\text{dm}}]^{\mathsf{T}}_{N_{\text{evl}} k} \end{bmatrix} \psi \quad (3.7.41)$$

Note that if the solution $\hat{x}$ of the fitting step has been computed on square grids, an additional step is needed to remove the virtual actuators outside the metapupil.

3.7.2.4 LGS Error Computation and Tip/Tilt and Tilt Anisoplanatism Mode Removal

The final step of LGS wavefront reconstruction consists in computing the LGS error vector:

$$e = \widehat{x} - c_{DM,clipped} \quad (3.7.42)$$

where $\widehat{x}$ denotes the fitting estimate and $c_{DM,clipped}$ the clipped DM commands. The 5 TT/TA modes are then removed from this error vector, by first computing mode coefficients as follows:

$$c = \Sigma\, M^{\mathsf{T}}\, e \quad (3.7.43)$$

where Σ is the inverse of the 5×5 modal cross-coupling matrix and M is the modal matrix. The Mode removal is then performed as follows:

$$\overline{x} \;=\; e - M\, c \quad (3.7.44)$$

The matrices M and Σ are a function of a single parameter, which is updated from the RPG at a rate of 0.1Hz as the range to the sodium layer varies.

3.7.3 Computation and Memory Requirements

This section describes computation and memory requirements of the LGS wavefront reconstruction processes for the proposed 4 different tomography algorithms: BGS-CBS, BGS-CG20, CG30 and FD3. DM Fitting is done using CG4. All algorithms use a $n_0 = 123$-points wide aperture-plane grid with mesh size $\Delta_0 = d_{\text{sa}}/2 = 1/4$ m, and 2 staggered DM grids, respectively $63-$ and 75-actuators wide with $1/2$ m mesh size. Requirements are given for the following most stressing atmospheric grid sizes that allow for a telescope zenith angle of up to 60 deg, a circular science field of view of diameter of up to 120 arcsec and fixed oversampling:

Note that the BGS-CBS grids are non-square, uncompressed, and fully contained within the sizes reported in table 4, which provides a 20% reduction in the number of grid points compared to the requirement for square grids.

Algorithm	Grid Size	α_k
BGS-CBS	$\begin{bmatrix} 123 & 135 & 75 & 83 & 95 & 101 \end{bmatrix}$	$\begin{bmatrix} 1 & 1 & 2 & 2 & 2 & 2 \end{bmatrix}$
BGS-CG20 and CG30	$\begin{bmatrix} 123 & 143 & 161 & 179 & 110 & 122 \end{bmatrix}$	$\begin{bmatrix} 1 & 1 & 1 & 1 & 2 & 2 \end{bmatrix}$
FD3	$\begin{bmatrix} 256 & 256 & 256 & 256 & 256 & 256 \end{bmatrix}$	$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}$
FD3-OS2	$\begin{bmatrix} 256 & 256 & 128 & 128 & 128 & 128 \end{bmatrix}$	$\begin{bmatrix} 1 & 1 & 2 & 2 & 2 & 2 \end{bmatrix}$

Table 4: Tomography atmospheric algorithms and associated grids. $N_{\text{ps}} = 6$ atmospheric layers are reconstructed by all algorithms at the following altitudes: 0, 2.58, 5.15, 7.73, 12.89 and 15.46 km. FD3-OS2 refers to FD3 with only $N_{\text{os}} = 2$ atmospheric layers oversampled.

All of the following computation and memory requirements have been obtained from a simple MATLAB code which counts the number of nonzero elements of a sparse matrix associated to each operator. This code as well as all sparse matrices are available upon request. Regarding the BGS-CBS algorithm, a simple MATLAB code is available upon request that computes for each layer the number of clock cycles required to complete the forward and back-substitutions.

The storage requirement for the gradient TT and DF mode projection operator P_{M} is minimal and is for the noise weighting matrix W, the 17×17 S_{M}^{-1} matrix, and the f vector. The only non-neglegible extra storage requirement arises from the rank 17 precomputed matrices U' and the 17×17 S_{U}^{-1} matrix that are used in the BGS-CBS algorithm. Application of the $I - P_{\text{M}}$ mode removal operator costs $100\, n_{sa}^{lgs} + 17^2$ multiplications/accumulations (macs), where $n_{sa}^{lgs} = 2576$ denotes the number of active subapertures in a LGS WFS for the sample case used to derive computation and memory requirements. This number has been estimated as follows: application of the gradient TT/DF modal matrix M costs $(12 + 20)\, n_{sa}^{lgs}$ macs, and application of the noise weighting matrix W costs $24\, n_{sa}^{lgs}$ macs. Hence, the total cost of applying P_{M} is $(24 + 32 + 32)\, n_{sa}^{lgs} + 17^2 = 88\, n_{sa}^{lgs} + 17^2$. Finally, the cost of applying $I - P_{\text{M}}$ is equal to the cost of applying P_{M} plus $12\, n_{sa}^{lgs}$.

3.7.3.1 BGS-CBS and BGS-CG20 Algorithms

Operator	Memory (MB)	Nb. of Operations (MMAC)
Pseudo open loop gradients	$\widetilde{H}_{\text{dm}} : < 0.03$ $\Gamma : < 0.4$	BGS-CBS: 2.66 BGS-CG20: 3.2
Gradient TT and DF Removal	$W_k, S_{\text{M}}^{-1}, f :\ 0.18$	0.26
Tomography right hand side vector	$\widetilde{H}_{\text{atm}} : < 0.09$	BGS-CBS: 1.8 BGS-CG20: 1.85
Transformed right hand side vector	0	BGS-CBS: 7.37 BGS-CG20: 7.56
Layer solution	BGS-CBS $(L, p, U', S_{\text{U}}^{-1})$: 48.6 BGS-CG20: 0	BGS-CBS: 20.5 BGS-CG20: 212
Total Tomography	BGS-CBS: 49.3 BGS-CG20: 0.8	BGS-CBS: 32.6 BGS-CG20: 224.9
H_{atm}	< 0.09	BGS-CBS: 1.79 BGS-CG20: 2.71
W	0.5	0.94
H_{dm}^{T}	< 0.04	BGS-CBS: 3.52 BGS-CG20: 4.3
A		BGS-CBS: 7.98 BGS-CG20: 9.58
Total DM Fitting	0.6	BGS-CBS: 46.4 BGS-CG20: 56.1
TT/TA Modes Removal	0.19	0.1
Estimated Grand Total	BGS-CBS: 50.1 BGS-CG20: < 1.6	BGS-CBS: 79.1 BGS-CG20: 281.1

Table 5: LGS wavefront reconstruction computation and memory requirements for the BGS-CBS and BGS-CG20 algorithms for a 60 deg zenith angle pointing and the grid sizes reported in table 4.

The only operators that require sparse matrix format storage are the Cholesky factors used to run the backsolves of the BGS-CBS algorithm. The compressed sparse row (CSR) format has been assumed, which requires storage of (i) a real-valued vector containing the non-zero elements of the sparse matrix, (ii) an integer-valued vector containing the column indices of these elements, and (iii) an integer-valued vector containing pointers to the beginning of each row. If nnz denotes the number of non-zero elements of the sparse matrix, and n the number of rows, a total of nnz real numbers and $nnz + n$ integer numbers have to be stored. 4 bytes per real and integer have been used. Memory requirements for the aperture-plane operators Γ and W are listed for a sparse matrix implementation with CSR storage format. For the BGS-CBS algorithm, the cost to run the back-substitutions on a sparse triangular Cholesky factor with nnz non-zero elements has been calculated as nnz. Since there is a pair of triangular systems to solve for each layer, the total cost of the back-substitutions per layer is $2\,nnz$. Note that the memory requirement for the noise weighting operator W_k has been listed but is accounted for in table 1. Finally, memory requirements for each interpolation operator (bilinear and bicubic) has been calculated for a grid based implementation. If $n_0 = 123$ denotes the aperture-plane grid size, up to n_0 integers and n_0 $(4n_0)$ real numbers have to be stored for the $x-$ direction, and similarly for the $y-$ direction for the bilinear (bicubic) interpolation operations respectively. The $4n_0$ real numbers to be stored per direction for the bicubic interpolation arise from the fact that stencil weights rather than interpolation offsets should be stored. There is no storage for ground level grids (DM and atmospheric screen).

3.7.3.2 CG30 and FD3 Algorithms

Operator	Memory (MB)	Nb. of Operations (MMAC)
Pseudo open loop gradients	$\widetilde{H}_{\text{dm}} : < 0.03$ $\Gamma : < 0.4$	3.2
Gradient TT and DF Removal	$W_k, S_{\text{M}}^{-1}, f :\ 0.18$	0.26
Tomography right hand side vector	$\widetilde{H}_{\text{atm}} : < 0.09$	CG30: 1.85 FD3: 1.78 FD3-OS2: 1.9
A		CG30: 5.2 FD3: 8.1 FD3-OS2: 6.1
Precond	FD3: 37.5 FD3-OS2: 9.8	FD3: 50.3 FD3-OS2: 15.5
Total Tomography	CG30: 0.8 FD3: 38.3 FD3-OS2: 10.5	CG30: 188.9 FD3: 194.8 FD3-OS2: 79.2
H_{atm}	< 0.09	2.71
W	0.5	0.9
H_{dm}^{T}	< 0.04	4.32
A		9.58
Total DM Fitting	0.6	56.1
TT/TA Modes Removal	0.19	0.1
Estimated Grand Total	CG30: < 1.6 FD3: 39.3 FD3-OS2: 11.5	CG30: 245.1 FD3: 251 FD3-OS2: 135.4

Table 6: LGS wavefront reconstruction computation and memory requirements for the CG30 and FD3 algorithms for a 60 deg zenith angle pointing and the grid sizes reported in table 4.

The memory requirements for the aperture-plane operators Γ and W are for a sparse matrix implementation with CSR storage format and 4 bytes per real and integer number. The Fourier Domain preconditioning matrix used by the FD3 algorithm does not require sparse matrix storage. This matrix is block-diagonal with N/b full square blocks, where N denotes the total system size and $b = 4N_{\text{os}} + (N_{\text{ps}} - N_{\text{os}})$ the number of rows of the blocks. Note that each block is Hermitian, therefore only $N\,(b+1)/2$ complex numbers need to be stored. The cost to apply the FD preconditioning matrix has been calculated as $4\,N\,b + 2\,\sum_{k=1}^{N_{\text{ps}}} n_k\,\log_2(n_k)$, where $n_k = 256$ or 128 denotes the 1D size of the atmospheric grids used by the FD3 algorithm, and $N = \sum_k n_k^2$. Note that the memory requirement for the noise weighting operator W_k has been listed but is accounted for in table 1.

3.8 NGS Wavefront Reconstruction

3.8.1 Overview

NGS wavefront reconstruction refers to the computation of wavefront values at DM actuator locations from closed loop NGS gradients. The proposed architecture follows the model developed for the LGS AO mode of operation. By abuse of terminology, NGS tomography will refer to the estimation of wavefront values on a single ground level phase screen oversampling twice the DM actuator grid, and NGS fitting will refer to the subsequent least-squares fit of this estimate to DM actuator locations to optimize performance on-axis. All block-structured operators/vectors used in LGS AO mode reduce to standard operators/vectors and the bilinear interpolation operators $\widetilde{H}_{\text{atm}}$ and H_{atm} reduce to the identity.

3.8.2 Algorithm Description

3.8.2.1 NGS Pseudo Open Loop Gradient Computation

NGS wavefront reconstruction shall operate on pseudo open loop gradients. These gradients are obtained as in LGS AO mode but with on 1 DM grid and 1 WFS:

$$s_{\text{dm}} = \Gamma\, H_{\text{dm}}\, \widehat{x}\,, \quad (3.8.1)$$

$$s_{ol} = s + s_{\text{dm}} \quad (3.8.2)$$

where s_{ol} and s denote respectively the pseudo open loop and closed loop gradient vectors, $\widehat{x} = c_{DM,clipped}$ is the DM actuator vector (see section 3.12), and H_{dm} is the (fixed) bicubic interpolation operator for the ground-level DM grid.

As in the LGS mode, the RTC computes the M1 scalloping ,pde by averaging the open loop gradients as described in eq.(3.7.13).

3.8.2.2 NGS Tomography

Note that the term tomography refers here to only a single layer at the ground. NGS tomography is decomposed into 3 main steps:

- NGS pseudo open loop gradients tip/tilt removal. This step is performed only if the tip/tilt/focus measurements are obtained separately from an instrument 2x2 OIWFS.
- NGS tomography matrix system right hand side vector computation.

- NGS tomography matrix system approximate solution computation.

These steps are detailed below.

3.8.2.2.1 NGS Pseudo Open Loop Gradient Tip/Tilt Removal

NGS pseudo open loop gradient tip/tilt removal follows the exact same procedure as that detailed for the LGS AO mode.

3.8.2.2.2 NGS Tomography Matrix System Right Hand Side Vector Computation

The NGS tomography matrix system right hand side vector is expressed as follows:

$$\boxed{b \quad = \quad \Gamma^{\mathsf{T}} W \, \overline{s_{ol}} \,, \quad (3.8.3)}$$

where $\overline{s_{ol}}$ denotes the tip/tilt removed NGS pseudo open loop gradient vector, W is the pre-computed thresholded NGS measurement noise inverse covariance matrix and Γ is the aperture-plane gradient matrix. W is computed by the RTC as SNR conditions change (refer to 3.5.2.4), and Γ is updated by the RPG as the pupil rotation angle varies, at a rate of 0.1 Hz.

3.8.2.2.3 NGS Tomography Matrix System Approximate Solution Computation

The NGS tomography matrix system takes the following form:

$$A\,x \quad = \quad b\,, \tag{3.8.4}$$
$$A \quad = \quad \Gamma^{\mathsf{T}} W\,\Gamma + \gamma\,L^2 + \epsilon\,I\,, \tag{3.8.5}$$

where L is a sparse curvature squared regularization operator. All of the algorithms proposed for LGS tomography can also be used for NGS tomography.

3.8.2.3 NGS Fitting

NGS fitting is decomposed into the following 2 steps:

- NGS fitting matrix system right hand side vector computation.
- NGS fitting matrix system approximate solution computation.

Each item is discussed below.

3.8.2.3.1 NGS Fitting Matrix System Right Hand Side Vector Computation

The NGS fitting right hand side vector is computed by sequential sparse matrix vector multiplications/accumulations that can be represented in condensed form as follows:

$$b = H_{\mathsf{dm}}^{\mathsf{T}}\, W\, \widehat{x}\,, \tag{3.8.6}$$

where $\widehat{x}$ denotes the NGS tomography solution.

3.8.2.3.2 NGS Fitting Matrix System Approximate Solution Computation

The NGS fitting matrix system takes the following form:

$$A\,x = b\,, \tag{3.8.7}$$
$$A = H_{\mathsf{dm}}^{\mathsf{T}}\,W\,H_{\mathsf{dm}}\,, \tag{3.8.8}$$

As in LGS AO mode, $K = 4$ iterations of the CG algorithm shall be used to compute an approximate solution of the NGS fitting step. Note that if the solution $\hat{x}$ of the fitting step has been computed on square grids, an additional step is needed to remove the actuators outside the metapupil.

3.8.2.4 Tip/Tilt/Focus Removal

The final step of NGS wavefront reconstruction consists in removing tip/tilt/focus from the fitting estimate if these modes are obtained separately using a 2x2 OIWFS. This task is performed as in the LGS case with only 3 instead of 5 modes.

3.8.3 Computation and Memory Requirements

The NGS wavefront reconstruction computation and memory requirements have not been computed because they are be bounded by the LGS wavefront reconstruction computation and memory requirements, which will determine the hardware architecture requirements for the RTC.

3.9 Low order wavefront reconstruction

3.9.1 Overview

This section describes the OIWFS modal wavefront reconstruction, the temporal filtering and the projection of the low order modes onto the DM actuators.

3.9.2 Algorithm Description

3.9.2.1 Low Order Modal Reconstruction

The OIWFS gradients (represented by the vector s) are used as input signals to compute the tip/tilt/focus modes in either NGS AO or seeing limited mode. They are also used to compute the tilt anisoplanatism modes in LGS AO mode, if more than one OIWFS are utilized.

In all cases, this corresponds to a simple matrix multiplication using a preloaded control matrix M_{OIWFS} of size $(N_{mod}^{OIWFS}, N_{grad}^{OIWFS})$ where N_{mod}^{OIWFS} is the number of modes (Three in NGS AO mode or seeing limited mode and up to six in LGS AO mode):

$$\boxed{m = M_{OIWFS}\,.\,s} \tag{3.9.1}$$

The modal reconstruction is performed at the fastest OIWFS frame rate, using the most recent centroid measurements from the other two slower OIWFS in the LGS AO mode. It is performed at the 2x2 OIWFS frame rate in the NGS AO mode or seeing limited mode.

In LGS AO, NGS AO or seeing limited modes, the modal control matrix is initialized by the RPG during the AO observation configuration process based upon the OIWFS asterism. The control matrix is then updated at a rate of 0.1Hz to account for changes (i) in the OIWFS measurement noise and (ii) the telescope line-of-sight due to science observations with dithering.

The low order modes are resampled at the LGS WFS frame rate and the resampled focus mode is sent to the LGS Reference Processing task described in 3.13. After that, the remaining tip/tilt and tilt anisoplanatism modes are filtered and projected to the DM actuators.

3.9.2.2 Low Order Modal Temporal Filtering

A lead filter is applied to each mode m_i of the low order mode vector m:

$$\boxed{m_{lead,i}(n) = \frac{1}{2ag+1}\left[(2ag-1)m_{lead,i}(n-1) + (2g+1)m_i(n) - (2g-1)m_i(n-1)\right]} \quad (3.9.2)$$

where a is the lead filter phase parameter, $g = \nu_s/(2\pi\sqrt{a}\nu_0)$, ν_s is the sampling frequency, and ν_0 is the cross-over frequency parameter (the frequency at which the open loop transfer function has unit gain). An integrator with gain α_i is then applied to the output of the lead filter:

$$\boxed{m_{filt,i}(n) = \alpha_i m_{lead,i}(n) + m_{filt,i}(n-1)} \quad (3.9.3)$$

3.9.2.3 Low Order Modal Temporal Filter Optimization

The temporal filters are optimized at a rate of 0.1Hz by a RPG to minimize the closed loop error in all the low order modes. The supporting computations performed in the RTC, by a background process running at the same rate, are described further below:

To perform this optimization, it is necessary to save a time history of at least 512 values of the OIWFS modes before filtering. For each OIWFS mode, the squared modulus of the FFT is computed using its time history. The resulting PSDs for each mode are averaged, and then passed to the RPG, which performs the optimization of the servo coefficients.

We will use the following notation to describe this process:

- m_i refers to the component i of the mode vector m
- The optimization cycle is divided in N_{seg} FFT segments (typically 8-16)
- Each segment corresponds to N_{samp} sampling times of the loop (typically 512)

The PSD computations consist of the following:

- As each segment (of N_{samp} elements) of m_i becomes available in the buffer, compute the closed loop PSD, $PSD_{cl,i}$, as the squared modulus of the Fourier transform of m_i:

$$PSD_{cl,i} = |\mathcal{FFT}(m_i)|^2 \quad (3.9.4)$$

- average the PSD's over N_{seg} consecutive segments,
- transfer the PSD's to the RPG.

3.9.2.4 Low Order Mode Projection

In LGS and NGS AO mode, the filtered mode vector is projected to the DM actuator commands (DM0 and DM11.2 in the LGS AO mode, and only DM0 in NGS AO mode) as described below:

$$\boxed{e_{OIWFS} = P_{OIWFS} \,.\, R(\theta_R) \,.\, m_{filt}} \tag{3.9.5}$$

where $R(\theta_R) = \begin{pmatrix} \cos\theta_R & \sin\theta_R & 0 & 0 & 0 \\ -\sin\theta_R & \cos\theta_R & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & \cos 2\theta_R & \sin 2\theta_R \\ 0 & 0 & 0 & -\sin 2\theta_R & \cos 2\theta_R \end{pmatrix}$ is the modal rotation matrix and θ_R is the instrument rotator angle. The instrument rotator angle is read from the OIWFS component controller at a rate of 20Hz. Note that the OIWFS probe arm angle (one per OIWFS) is fixed and taken into account in the modal reconstructor provided by the RPG. P_{OIWFS} is the projection matrix. The number of rows of the matrix corresponds to the number of active actuators of DM0 and DM11.2 in LGS AO mode N_{vact}^{DM} or to the number of active actuators of DM0 only in NGS AO mode N_{vact}^{DM0}. The number of columns corresponds to the number of low order modes (Tip, tilt and focus in the NGS AO case, and tip, tilt and tip tilt anisoplanatism modes in the LGS AO case).

3.9.3 Computation and Memory Requirements

Tasks	Number of MMAC			Memory
	Per Frame	Frequency (Hz)	Allowed computation time (s)	(MB)
Low order modal reconstruction in LGS mode	0.0	up to 800	0.001	0.0
Low order modal reconstruction in NGS mode	0.0	up to 800	0.001	0.0
Low order modal temporal filtering in LGS mode	0.0	up to 800	0.001	0.0
Low order modal temporal filtering in NGS mode	0.0	up to 800	0.001	0.0
Low order modal temporal filtering optimization in LGS mode	0.45	0.1		0.02
Low order modal temporal filtering optimization in NGS mode	0.27	0.1		0.012
Low order mode projection in LGS mode	0.033	up to 800	0.001	0.077
Low order mode projection in NGS mode	0.008	up to 800	0.001	0.02
	Average MMAC/s			
Estimated total in LGS mode	27			0.096
Estimated total in NGS mode	6.8			0.033

Table 7: Operation and memory requirements for the Low order wavefront reconstruction

Please note that the average computation rate is not representative of the peak computation rate, which is higher on account of the real time processes which must be completed in the first $1000\,\mu$s with a goal of $400\,\mu$s of each frame. The peak computation rate will be at least 34 MMAC/s in LGS mode and 8.5 MMAC/s in NGS mode during this interval.

3.10 Active Optics Mode Computation in Seeing Limited Mode

3.10.1 Overview

This section describes the NFIRAOS active optics (aO) computations for the seeing limited mode. In this mode, the RTC uses the 2x2 OIWFS and the Truth WFS (TWFS) to compute the aO telescope modes.

3.10.2 Algorithm Description

The tip, tilt and focus modes are computed from the 2x2 OIWFS and temporally filtered at the 2x2 OIWFS frame rate as described in section 3.9. The tip tilt modes are sent to the TTS control process (refer to 3.12.2.2), while the focus is directly sent to the telescope mode offload process (refer to 3.12.2.3).

The higher order modes (from astigmatism up to radial fifth order modes) are computed from the TWFS gradients which are received at a rate of up to 100 Hz from the TWFS DC. The modes are computed with a simple matrix vector multiplication as described by the following equation:

$$\boxed{m_{TWFS} = M_{TWFS} \,.\, s_{TWFS}} \tag{3.10.1}$$

where:

- s_{TWFS} is the TWFS gradient vector of dimension $n_{grad}^{TWFS} = 48$ (TBC),
- m_{TWFS} is the vector of telescope modes excluding piston, tip, tilt and focus; the dimension of this vector is $n_{mod}^{aO} = 20$,
- M_{TWFS} is the fixed modal control matrix of size (n_{mod}^{aO}, n_{grad}^{TWFS}).

The higher order modes are then sent to the telescope offload process (refer to 3.12.2.3).

3.10.3 Computation and Memory Requirements

The computation and memory requirements for the active optics mode computation are on the order of 0.1MMAC/s and 2 KBytes.

3.11 Turbulence Parameter Estimation

3.11.1 Overview

This section describes the turbulence profile estimation algorithm.

3.11.2 Algorithm Description

1. For the LGS AO mode, a SLODAR-like algorithm is used to compute an estimate of the turbulence profile. This algorithm estimates the turbulence strength at n_l layers whose ranges are given by:

$$h_k = k\Delta/(\theta + k\Delta/H)\,, k = 0, 1, \cdots, n_l - 1 \tag{3.11.1}$$

where $\Delta = 0.5$ is the LGS WFS subaperture size, H is the range to the sodium layer and $\theta = 67''$ is the angular separation of the LGS pair oriented along the $y-$axis in the NFIRAOS coordinate system with greatest separation. A correlation vector b with $2n_l$ components is accumulated in real-time by cross-correlating properly selected pseudo open loop LGS WFS gradients from the LGS WFS pair. Vector b is the concatenation of two subvectors, b_x and

b_y, each with n_l components. Valid subapertures for component $j = 0, \cdots, n_l - 1$ of b_α (α representing either x or y) are all subaperture pairs separated by a distance $j\Delta$ along the $y-$axis, and whose illumination is at least 97%. In order to eliminate sensitivity to global tip/tilt and focus, it is necessary to correlate the local curvature of the pseudo open loop measurements, i.e.

$$b_j^\alpha = \frac{1}{N_{\text{valid subap}}} \sum_{x_p \in \text{valid subap}} m_\alpha(x_p) m'_\alpha(x_p + (0, j\Delta)) \tag{3.11.2}$$

where

$$m_\alpha(x_p) = s_{ol,\alpha}(x_p - \delta_\alpha) - 2s_{ol,\alpha}(x_p) + s_{ol,\alpha}(x_p + \delta_\alpha) \tag{3.11.3}$$

and $\delta_x = (\Delta, 0)$, $\delta_y = (0, \Delta)$.

The turbulence profile vector p with components $p_k = r_{0,k}^{-5/3}$ is estimated by applying a reconstruction matrix of size equal to $n_l \times 2n_l$ to b:

$$p = (A^T A)^{-1} A^T b \tag{3.11.4}$$

Matrix A depends only on the range to the sodium layer, and is computed/updated in the RPG using the following 2D Fourier transform:

$$A_{jk}^\alpha = \xi_k^{-2} 16\pi^2 (0.0229) \int df_x df_y K(f_x, f_y) \text{PSD}(f_x, f_y) e^{2i\pi(j-k) f_\alpha \Delta} \tag{3.11.5}$$

where

$$\xi_k = 1 - h_k/H = 1/(1 + k\Delta/(\theta H)), \tag{3.11.6}$$

and

$$K(f_x, f_y) = (\cos(2\pi f_\alpha \Delta) - 1)^2 f_\alpha^2 \text{sinc}^2(f_x \Delta) \text{sinc}^2(f_y \Delta) \tag{3.11.7}$$

and

$$\text{PSD}(f_x, f_y) = (\xi_k^{-2}(f_x^2 + f_y^2) + L_0^{-2})^{-11/6} \tag{3.11.8}$$

This 2D Fourier transform is computed using a $n \times n$ 2D FFT, where $n = 4n_l = 48$, with resolution $df = 1/(ndx) = 1/12$ m^{-1}, where $dx = \Delta/2 = 1/4$. When negative entries occur in the estimate p, the corresponding columns are removed in A and a new reconstructor is computed from the reduced rank matrix A. Finally, the n_l-layer turbulence estimate is adaptively binned down to 6 layers by application of a binning matrix (typically defined with triangular MASS-like influence functions):

$$p' = Bp. \tag{3.11.9}$$

The global Fried parameter r_0 is then computed as follows:

$$r_0 = \left(\sum_l r_{0,l}^{-5/3} \right)^{-3/5} \tag{3.11.10}$$

The strength of the layers γ_l is computed from the Fried parameter for layer l as described in the following equation:

$$\gamma_l = \left(\frac{r_{0,l}}{r_0}\right)^{-5/3} \tag{3.11.11}$$

The atmospheric time constant parameter $\tau_{0,l}$ for layer l is estimated by computing the temporal structure function as described in the following equation:

$$< \frac{1}{k}\sum_{k=1}^{K}[\phi_l(r_k,t) - \phi_l(r_k,t+\tau_j)]^2 >\sim \left(\frac{\tau_j}{\tau_{0,l}}\right)^{5/3} + 2\sigma^2 \tag{3.11.12}$$

where ϕ_l is the LGS tomographic phase estimate for layer l, τ_j is one of 8 time delays between 25 and 200 ms, the points r_k are represent one of 30×30 sample points on the phase screen grid, and σ^2 is the variance due to the WFS measurement noise. The temporal structure function is accumulated at a rate of between 10 and 100 Hz, and then averaged spatially (over k) and temporally over one minute. $\tau_{0,l}$ is then estimated from Eq. (3.11.12) using a TBD fit to the measured structure functions.

The global atmoshpheric time constant τ_0 is then computed using the same formula as r_0 (3.11.10) by replacing r_0 by τ_0.

Note that τ_0 is only computed in the direction of the observation and is not zenith corrected because the direction of the wind is not known.

The isoplanatic angle θ_0 is computed as follows:

$$\theta_0 = \left(6.88\sum_l h_l^{5/3}\, r_{0,l}^{-5/3}\right)^{-3/5} \tag{3.11.13}$$

The generalized isoplanatic angle θ_2 is computed as follows

$$\theta_2 = \left(6.88\sum_l F_{2,l}\, r_{0,l}^{-5/3}\right)^{-3/5} \tag{3.11.14}$$

where $F_{2,l}$ is computed as follows:

$$\begin{aligned} F_{2,l} &= 0.5|h_l - H_1|^{5/3} \\ &+0.5|h_l - H_2|^{5/3} \\ &-0.25|H_2 - H_1|^{5/3} \\ &-0.25|H_2 - H_1|^{-5/3}\left(|h_l - H_1|^{5/3} - |h_l - H_2|^{5/3}\right)^2 \end{aligned} \tag{3.11.15}$$

and where H_1 and H_2 correspond to the zenith corrected altitudes at which the DMs are conjugated: $H_1 = 0$ and $H_2 = 11200\cos\psi$.

Finally, the value of the outer scale parameter L_0, l for phase screen layer number l is computed as the following fit:

$$L_{0,l} = \arg\min_L \sum_j w'_j \left[f_l(j) - f_{mod}(j; r_{0,l}, L)\right]^2 . \tag{3.11.16}$$

where $f_l(j)$ is the structure function at separation number j computed as

$$f_l(j) = \left\langle \frac{1}{k_j} \sum_{k=1}^{K} W_l(r_k) W_l(r_k + \Delta_j) \left[\phi_l(r_k, t) - \phi_l(r_k + \Delta_j, t)\right]^2 \right\rangle_t , \tag{3.11.17}$$

where:

- Δ_j represents one of up to 100 structure function separations,
- $W_l(r)$ is a coarsely sampled, $\{0, 1\}$-valued "windowing" function, indicating which points on the phase screen grid are within the metapupil sampled by the LGS wavefront sensors,
- $k_j = \sum_k W_l(r_k) W_l(r_k + \Delta_j)$ is the number of valid points used in the structure function at separation number j.

Here the fixed weights w'_j are choosen to emphasize the contribution of the structure function values at moderate to large separations. The function f_{mod} is a model structure function that is computed by interpolating into a 2-dimensional table precomputed for 21 values of r_0 and 11 values of L_0.

2. For the NGS AO mode, r_0 is estimated using the formula

$$r_0 = (\sec\psi)^{3/5} \arg\min_r \sum_j w_j \left[f_0(j) - 6.88 \left|\frac{\Delta_j}{r}\right|^{5/3}\right]^2 , \tag{3.11.18}$$

where the fixed weights w_j are selected to emphasize the importance of the value of the structure function at relatively small separations and where ψ is the zenith angle.

3.11.3 Computation and Memory Requirements

The computation and memory requirements for the turbulence profile estimation process are approximately 211 M Ops/sec and 1.43 MBytes, respectively. Nearly all of the computations are associated with the accumulation of the spatial structure functions at a rate of 100 Hz. Nearly all of the memory is associated with maintaining the phase profile time histories which are required to compute the temporal structure functions.

3.12 Wavefront Corrector Control

3.12.1 Overview

This section describes the control of the DMs and the tip/tilt stage (TTS), as well as the computation of the active optics modes offloaded to the telescope.

A type II controller is implemented for the control of the low order modes (TTF in NGS mode, TT and TT anisoplanatism in LGS mode). The TTS is used as the woofer (low-bandwidth-high-stroke device) and the DM0 is used as the tweeter (high-bandwidth-low-stroke device) for the Tip/Tilt modes, using complementary low-pass and high-pass filters.

3.12.2 Algorithm Description

3.12.2.1 Control of the DMs

3.12.2.1.1 Computation of the combined DM actuator command vector provided by the wavefront reconstruction processes

Following the wavefront reconstruction processes, the current error e in the DM actuator command vector is computed as a combination of 2 terms: (i) the vector of DM actuator commands obtained by the LGS or NGS wavefront reconstruction process e_{tom}, and (ii) the vector of DM actuator commands obtained by the low order wavefront reconstruction process e_{OIWFS}. This operation is described by the following equation:

$$\boxed{e = e_{tom} + e_{OIWFS}} \quad (3.12.1)$$

The dimension of the e vectors corresponds to the number of active actuators of both DM0 and DM11.2 in the LGS AO case, and of DM0 only in the NGS AO case.

In LGS AO case, the error vector e is computed at the frame rate of the LGS WFS. In NGS AO case, when an additional 2x2 OIWFS is used, the error vector e is computed at the frame rate of either the higher order WFS or the 2x2 OIWFS, which ever is fastest.

In the NGS AO case, however, when no additional 2x2 OIWFS is used, the current error vector e in the DM0 actuator commands is simply equals to the vector of DM0 actuator commands obtained from the NGS wavefront reconstruction process e_{tom}. Again the dimension of the e vector corresponds to the number of active actuators of DM0 and in this case, the vector e is computed at the NGS WFS frame rate.

In all cases (LGS AO, NGS AO, NGS AO with additional 2x2 OIWFS), the resulting vector e of DM actuator errrors is then temporally filtered, and the filtered vector is clipped before the commands are extrapolated to the DM edge actuators and then applied to the DMs. All these operations are performed at the same frame rate used to compute the e vector (LGS WFS, NGS WFS, or 2x2 OIWFS frame rate). This frame rate will be refered to as the wavefront corrector frame rate in the rest of the document.

3.12.2.1.2 DM Temporal Filtering

The temporal filter consists of a simple integrator with gain. Integrator windup is prevented by using the clipped commands actually applied to the DMs instead of the outputs of the integrator. The temporal filter is described by the following equation:

$$\boxed{c_{DM}(n) = g_I * e(n) + c_{DM,clipped}(n-1)} \quad (3.12.2)$$

where $c_{DM,clipped}$ is the vector of DM actuator commands after clipping.

The gain g_I is optimized in the RPG at a slow rate of 0.1Hz. This task is similar to the low order modal temporal filtering optimization task described in 3.9.2.3, except that only one global filter coefficient is optimized for all the DM actuator commands. This global coefficient is calculated from a mean PSD of the actuator errors e, which is computed in the RTC by averaging over a subset of TBD actuators.

To avoid integrator windup in the low order WFS reconstruction block in case of low order mode saturation, some TBD process may need to be implemented.

3.12.2.1.3 DM Actuator Clipping

The clipping process consists of checking that the amplitudes, velocities and nearest neighbor differences of the commands do not exceed pre-specified thresholds. The process is performed at the wavefront corrector frame rate according to the function:

$$\boxed{c_{DM,clipped} = \text{clip}(c_{DM} - M_{TT}c_{TT}),} \tag{3.12.3}$$

where C_{TT} are the contributions from the low pass filtered tip/tilt modes, M_{TT} is the tip/tilt mode-to-actuator matrix, and the function "clip" performs the following checks:

- If the command of an actuator reaches its limit amplitude, the command is clipped to the corresponding limit,
- If the frame-to-frame change in the command of a DM actuator is above a specified threshold, the command is similarly clipped at the maximum allowable adjustment,
- If the difference between the commands for a pair of nearest-neighbor actuators exceeds a specified threshold, both commands are adjusted toward their average value to respect the threshold (TBC)
- The number of clipped actuators (for whatever reason) should be monitored and the control loop should be opened if this number remains above a specified threshold for a specified number of control loop cycles.
- The actuator limit amplitudes are computed by the RPG based on the DM physical actuator limit and the non-common path aberration flat vector for a given observation

3.12.2.1.4 DM Extrapolation to Edge Actuators

Finally, the commands to the DM edge actuators are computed at each frame from the actively controlled actuators. The extrapolation is perfomed independently for each deformable mirror (DM0 only in NGS AO mode) and consists of a matrix-vector multiplication with a matrix E_{DM} as described below:

$$\boxed{c_{DM,edge} = E_{DM} \cdot c_{DM,clipped}} \tag{3.12.4}$$

The E_{DM} matrix is a sparse matrix, as the command to each edge actuator is determined from a small number (8) of nearby active actuators.

3.12.2.1.5 DM Gain and Offset Compensation

This is the final step that is performed at the wavefront corrector frame rate before the DM actuator commands are sent to the DM Electronics. It consists of converting the DM actuator commands into the units acceptable to the DM electronics.

For example, the DM actuator commands are converted as follows:

$$\boxed{v_{DM} = k_{DM} \cdot * c_{DM} + o_{DM}} \tag{3.12.5}$$

where $c_{DM} = \begin{pmatrix} c_{DM,clipped} \\ c_{DM,edge} \end{pmatrix}$ is the combined vector of both actively controlled and extrapolated actuators, k_{DM} is the conversion gain vector and o_{DM} is the offset vector. Note that the calibration coefficients are updated by the RPG at a very slow rate of 0.001 Hz to 0.1 Hz based on the DM temperatures and the science non-common path aberration, which are derotated to account for instrument rotation.

The offset and gain vectors for the current temperature are obtained from the RPG, by interpolating into a pair of lookup tables and by filtering according to the formula:

$$o_{DM}(n) = o_{DM}(n-1) + \alpha * (o_{DM}(LUT) - o_{DM}(n-1)) \tag{3.12.6}$$

$$k_{DM}(n) = k_{DM}(n-1) + \beta * (k_{DM}(LUT) - k_{DM}(n-1)) \tag{3.12.7}$$

with $o_{DM}(n = 0)$ equal to the flat.

3.12.2.2 Control of the TTS

3.12.2.2.1 DM to TTS Projection

The output of the DM integrator before clipping is used to compute the commands to the TTS actuators at the wavefront corrector frame rate. This corresponds to a simple matrix-vector multiplication as described by the following equation:

$$\boxed{c_{DM0toTT} = P_{TT} \,.\, c_{DM0}} \tag{3.12.8}$$

where P_{TT} is the pre-loaded DM0 to TTS projection matrix. The number of rows corresponds to the number of TTS actuators (which is 2) and the number of columns corresponds to the number of DM0 active actuators.

3.12.2.2.2 TTS Temporal Filtering

The temporal filter consists of a low pass filter proportional integral filter described by the following equation:

$$\boxed{c_{TT}(n) = \alpha_{TT} * c_{DM0toTT}(n) + (1 - \alpha_{TT}) * c_{DM0toTT}(n-1)} \tag{3.12.9}$$

where:

- α_{TT} is the low pass filter parameter.

In the seeing limited case, however, the TTS is used to provide medium bandwidth tip/tilt correction. The TTS temporal filter consists of a simple integrator with gain. Integrator windup is prevented by using the clipped commands actually applied to the TTS:

$$\boxed{c_{TT}(n) = g_{SL} * m_{filt}(n) + c_{TT,clipped}(n-1)} \tag{3.12.10}$$

where:

- g_{SL} is the integrator gain,
- m_{filt} is the vector of tip/tilt modes computed in section 3.9,
- $c_{TT,clipped}$ is the vector of clipped TTS actuator commands

3.12.2.2.3 TTS actuator clipping

The TTS clipping process consists of checking that the amplitudes and the velocities of the TTS commands do not exceed pre-specified thresholds. The process is performed at the wavefront corrector frame rate:

- If the command of an actuator reaches its limit amplitude, the command is clipped to the corresponding limit,
- If the frame-to-frame change in the command of a TTS actuator is above a specified threshold, the command is similarly clipped at the maximum allowable adjustment,
- The control loop should be opened if one or both TTS actuators remain clipped for more than a specified number of control loop cycles.

3.12.2.3 Computation of the Telescope Modes

The commands to the DM0 and TTS actuators before clipping are used to compute the low-order modes to offload to the telescope.

The non-clipped commands of the TTS are (i) low-pass filtered, (ii) rotated into the telescope coordinates system, (iii) multiplied by a scale factor and then output to the AO Sequencer at a rate of 5Hz as described by the set of equations below:

$$\boxed{a_{TT}(n) = \alpha \ * \ c_{TT}(n) + (1-\alpha) \ * \ a_{TT}(n-1)} \tag{3.12.11}$$

$$o_{TT} = g \cdot *R_{TT} \, . \, a_{TT} \tag{3.12.12}$$

where:

- a_{TT} is the vector of low-pass-filtered TTS commands
- α is low-pass filter gain
- g is the vector of scale factors

- R_{TT} is the 2x2 rotation matrix based upon the telescope elevation angle
- o_{TT} is the vector of tip/tilt modes to offload to the telescope

Higher order quasi-static abberations other than piston and tip/tilt are computed from the unclipped DM0 commands. The unclipped DM0 commands are low-pass filtered over a period of 1s, and the filtered commands are used to compute the focus, coma, and M1 modes as described by the following equations:

$$\boxed{a_{DM0}(n) = \alpha\, *\, c_{DM0}(n) + (1-\alpha)\, *\, a_{DM0}(n-1)} \tag{3.12.13}$$

$$m_{tel} = M_{tel}\, .\, a_{DM0} \tag{3.12.14}$$

where:

- a_{DM0} is the vector of low-pass-filtered DM0 commands
- α is low-pass-filter gain
- m_{tel} is the vector of telescope modes to be offloaded before rotation and scaling
- M_{tel} is the DM0-actuators-to-telescope-modes matrix

The telescope modes are then rotated and scaled on a mode-by-mode basis using an equation similar to 3.12.12. The rotated and scaled modes are sent to the AO Sequencer.

In the seeing limited case, the non-clipped TTS commands are used to compute the tip tilt modes to offload to the telescope. However, the focus mode is provided directly by the low order wavefront reconstruction process and the higher order modes are computed directly from the TWFS gradient measurements as described in 3.10. The low pass filter is applied directly to the focus and higher order mode measurements over a period of 1s as described below:

$$\boxed{a_{focus}(n) = \alpha * m_{filt,focus}(n) + (1-\alpha) * a_{focus}(n-1)} \tag{3.12.15}$$

$$\boxed{a_{homod}(n) = \alpha * m_{TWFS}(n) + (1-\alpha) * a_{homod}(n-1)} \tag{3.12.16}$$

As for the LGS and NGS cases, the telescope modes are then rotated and scaled on a mode-by-mode basis. The rotated and scaled modes are sent to the AO Sequencer.

For LGS AO observations in which all 3 OIWFS are used, the unclipped commands of both DM0 and DM11.2 are used to compute the curvature plate scale modes. The computation of the curvature plate scale modes follows the same process: (i) apply a low pass filter to the DM11.2 commands, (ii) and compute the curvature plate scale mode via a simple matrix multiplication using the low pass filtered commands of both DMs. The curvature plate scale mode is a combination of primary mirror curvature and secondary mirror focus. It is TBD whether the computation of the M2 focus mode also depends upon the DM11.2 commands in this case.

3.12.3 Computation and Memory Requirements

The computation and memory requirements for the wavefront corrector control are on the order of 109MMAC/s in average and 1.25MBytes for the LGS mode and 47 MMAC/s average and 1.17MBytes for the NGS case based on an analysis done for each steps of the wavefront corrector control (TBC).

3.13 LGS Reference Processing

3.13.1 Overview

This section describes the LGS reference vector processing, which is computed based upon the OIWFS defocus, the NGS-to-LGS delta focus derived from the 3D turbulence profile estimate, the LGS WFS defocus measurements, and the TWFS measurements.

3.13.2 Algorithm Description

For each LGS WFS, the WFS defocus measurement, the NGS-to-LGS delta defocus term and the OIWFS defocus measurement are combined together at the LGS WFS frame rate, as described below:

$$\boxed{z_k = z_{OIWFS} + z_{tom,k} - z_{LGS,k}} \tag{3.13.1}$$

where:

- z_{OIWFS} is the focus component of the OIWFS mode vector m, as described in section 3.9,
- $z_{tom,k}$ is the NGS-to-LGS delta focus term for the LGS WFS k:
  $z_{tom} = \Delta \, . \, \widehat{x}$,
  where $\widehat{x}$ is the 3D turbulence profile estimated in section 3.7.2.2.3, Δ is the delta-focus matrix of dimension (6,N), and N is the dimension of the vector $\widehat{x}$.
- $z_{LGS,k}$ is the defocus term computed from the LGS WFS gradients of the LGS WFS k in 3.4.25.
- z_k is the combined defocus term.

The combined defocus term is low pass filtered as described in the following equation:

$$\boxed{z_{filt,k}(n) = \alpha_k \, * \, z_k(n) + (1 - \alpha_k) \, * \, z_{filt,k}(n-1)} \tag{3.13.2}$$

Separate coefficients α_k may be used for each defocus component.

Finally, the LGS reference vector r_k of the LGS WFS k is obtained as follows:

$$\boxed{r_k = r_{TWFS,k} - F \, . \, z_{filt,k}} \tag{3.13.3}$$

where $r_{TWFS,k}$ is the LGS WFS reference vector from the TWFS. The TWFS combines the reference vector contribution from the TWFS measurements, the science NCPA vector, which is derotated at 20Hz to account for instrument rotation, and the LGS path NCPA vector.

3.13.3 Computation and Memory Requirements

The computation and memory requirements for thr LGS reference processing are on the order of 68MMAC/s and 0.16MB per LGS WFS.

3.14 PSF Reconstruction

3.14.1 Overview

As a baseline, the RTC telemetry system will acquire and store the time histories of the WFS measurements and DM actuator commands which are needed to reconstruct, or estimate, the PSF for the NFIRAOS science object in postprocessing. It is an RTC *goal* to compute (i) LGS WFS and NGS WFS structure functions and (ii) OIWFS gradient covariance matrices in real time, so that "quick-look" estimates of the science object PSF can be provided to the observer shortly following each science observation. These optional RTC computations are described in the following paragraphs. (Note that the actual estimation of the PSF is still performed separately by the PSF reconstruction software, which is not part of the RTC. The further details of the PSF post processing algorithm are consequently not described in this document.)

3.14.2 Algorithm Description

The computation of the WFS structure functions consists of the following steps, which are identical for both the NGS and LGS case.

1. The start of the statistical accumulation of the WFS structure functions is synchronized with the science observation.
2. At each frame, compute the high order (LGS or NGS) and low order (On Instrument) WFS gradient covariance matrices $COV(s_k)$ independently for each WFS k. The dimension of each gradient covariance matrix is equal to the dimension of the gradient vector s_k for the corresponding WFS.
3. Average the WFS gradient covariance matrices over the N consecutive frames of the science observation to obtain $\langle COV(s_k)\rangle$.
4. In parallel, average the high order (LGS or NGS) and low order (On Instrument) WFS noise gradient covariance matrices computed in 3.4.2.4, 3.5.2.4 and 3.6.2.4 over the same interval to obtain $\langle C_k\rangle$.
5. At the end of the science exposure, condition the time-averaged gradient covariance matrices to remove the noise contribution by subtracting off the time-averaged noise covariance matrices:

$$C_{s,k} = \langle COV(s_k)\rangle - \langle C_k\rangle \,. \tag{3.14.1}$$

6. Next, compute the high order (LGS or NGS) WFS structure functions as follows:

$$D_{\phi,k}(x,y) = \sum_i \sum_j C_{s,k}(i,j) * U_{i,j}(x,y) + (D/r_0)^{5/3} * D_{\phi,k,0}(x,y), \quad (3.14.2)$$

   where $U_{i,j}(x,y)$ are precomputed slope-to-structure-function "influence matrices" with 128 by 128 sample points, and where $D_{\phi,k,0}(x,y)$ is a precomputed estimate of the "fitting error" in the WFS structure function for the normalized value $D/r_0 = 1$.

7. Pass the high order WFS structure functions $D_{\phi,k}(x,y)$ and the low order WFS gradient covariance matrices $C_{s,k}$ to the Data Management System to be stored (and processed) together with the observation science data.

3.14.3 Computation and Memory Requirements

The computation and memory requirements of the $C_{s,k}$ matrices are on the order of 27GOp/s and 32MBytes per LGS WFS. The computation and memory requirements of the LGS WFS structure functions are on the order of 137GOp/s and 0.03MBytes per LGS WFS.