Upgrading the GRAVITY fringe tracker for GRAVITY+ - Tracking the white-light fringe in the non-observable optical path length state-space

M Nowak, S Lacour, R Abuter, J Woillez, R Dembet, M.S Bordoni, G Bourdarot, B Courtney-Barrer, D Defrère, A Drescher, et al.

A&A, 684, A184 (2024)
https://doi.org/10.1051/0004-6361/202348771

Astronomy & Astrophysics

Upgrading the GRAVITY fringe tracker for GRAVITY+

Tracking the white-light fringe in the non-observable optical path length state-space

M. Nowak[1,2,⋆], S. Lacour[3,4,⋆], R. Abuter[4], J. Woillez[4], R. Dembet[3], M. S. Bordoni[5], G. Bourdarot[5], B. Courtney-Barrer[4,6], D. Defrère[10], A. Drescher[5], F. Eisenhauer[5], M. Fabricius[5], H. Feuchtgruber[5], R. Frahm[4], P. Garcia[13,14], S. Gillessen[5], V. Gopinath[5], J. Graf[5], S. Hoenig[7], L. Kreidberg[8], R. Laugier[10], J. B. Le Bouquin[9], D. Lutz[5], F. Mang[5], F. Millour[11], N. More[5], N. Morujão[13,14], T. Ott[5], T. Paumard[3], G. Perrin[3], C. Rau[5], D. C. Ribeiro[5], J. Shangguan[5], T. Shimizu[5], F. Soulez[12], C. Straubmeier[15], F. Widmann[5], and B. Wolff[4]

(Affiliations can be found after the references)

Received 28 November 2023 / Accepted 5 February 2024

ABSTRACT

Context. The GRAVITY beam-combiner at the Very Large Telescope Interferometer has recently made important contributions to many different fields of astronomy, from observations of the Galactic centre to the study of massive stars, young stellar objects, exoplanet atmospheres, and active galactic nuclei. These achievements were only made possible by the development of several key technologies, including the development of reliable and high-performance fringe trackers. These systems compensate for disturbances ranging from atmospheric turbulence to vibrations in the optical system, enabling long exposures and ensuring the stability of interferometric measurements.

Aims. As part of the ongoing GRAVITY+ upgrade of the Very Large Telescope Interferometer infrastructure, we aim to improve the performance of the GRAVITY fringe tracker, and to enable its use by other instruments.

Methods. We modified the group-delay controller to consistently maintain tracking in the white-light fringe, which is characterised by a minimum group delay. Additionally, we introduced a novel approach in which fringe-tracking is performed in the non-observable optical path length state-space using a covariance-weighted Kalman filter and an auto-regressive model of the disturbance. We outline this new state-space representation and the formalism we used to propagate the state vector and generate the control signal. While our approach is presented specifically in the context of GRAVITY/GRAVITY+, it can easily be adapted to other instruments or interferometric facilities.

Results. We successfully demonstrate phase-delay tracking within a single fringe, with any spurious phase jumps detected and corrected in less than 100 ms. We also report a significant performance improvement, as shown by a reduction of ∼30 to 40% in phase residuals, and a much better behaviour under sub-optimal atmospheric conditions. Compared to what was observed in 2019, the median residuals have decreased from 150 nm to 100 nm on the Auxiliary Telescopes and from 250 nm to 150 nm on the Unit Telescopes.

Conclusions. The improved phase-delay tracking combined with white-light fringe tracking means that from now on, the GRAVITY fringe tracker can be used by other instruments operating in different wavebands. The only limitation remains the need for an adjustment of the optical path dispersion.

Key words. instrumentation: high angular resolution – instrumentation: interferometers – techniques: interferometric

1. Introduction

By combining the light from multiple telescopes, long-baseline optical interferometry can achieve a much higher angular resolution than is possible with a single-dish telescope. Even in the upcoming era of telescopes exceeding 30 m, the Very Large Telescope Interferometer (VLTI) will continue to outperform single-dish observations in terms of angular resolution (for an up-to-date overview, see Eisenhauer et al. 2023). This is particularly useful for many applications, as shown by the broad range of subjects to which it has been applied with groundbreaking results (GRAVITY Collaboration 2018a,b,c, 2019a,b,c, 2020a,b,c,d,e,f, 2021, 2022).

One of the challenges that arise at this high angular resolution is the need to manage disturbances, which can include atmospheric turbulence, telescope vibrations, and other noise sources, and which remain the primary limitations to the precision of interferometric measurements. These disturbances lead to variations in the optical path length (OPL) of light travelling through each telescope to the beam combiner, resulting in an unstable interference pattern (unstable fringes).

To obtain stable interferometric measurements, these fringes need to be controlled in real time, and any OPL fluctuations need to be compensated for. In modern instruments, this is the role of a specific subsystem: the fringe tracker. The crucial role of fringe-tracking in optical interferometry has justified the extensive development efforts made over the past decades (Shao et al. 1988; Sorrente et al. 2001; Delplancke et al. 2006; Le Bouquin et al. 2008; Cassaing et al. 2008; Houairi et al. 2008; Colavita et al. 2010; Lozi et al. 2011; Menu et al. 2012; Choquet et al. 2014; Lacour et al. 2019). Various techniques have been proposed,

⋆ These authors contributed equally.

based either on hardware or software solutions, but they all share a common objective: to measure and stabilise the fringes, thereby maximising the signal-to-noise ratio (S/N) of interferometric measurements.

For the GRAVITY instrument, the fringe-tracker was initially designed to perform control in the six-dimensional optical path delay (OPD) state-space, where measurements are available. The algorithm, described in Lacour et al. (2019, subsequently Paper I), was based on the use of an auto-regressive (AR) model to model the disturbance in OPD space and on a Kalman filter to propagate an OPD state-vector using an asymptotic approximation of the optimal gain.

Deviating from this initial concept, we hereafter introduce a novel approach in which fringe-tracking is performed in the four-dimensional non-observable OPL state-space. The reduction in dimensionality alleviates the need for an asymptotic approximation of the Kalman gain, whose optimal value is now dynamically calculated at each iteration. This work is part of a general upgrade of the GRAVITY Fringe-Tracker that was made in late 2022 and 2023, and it can be considered as a follow-up to Paper I.

The paper is divided as follows. Section 2 gives a general description of how measurements are made within the fringe tracker, and which quantities are observable. Section 3 describes the details of the implementation of the upgraded fringe tracker, in particular, the state-space model, and the two control loops (phase and group delay). Section 4 reports the first on-sky results obtained with this upgraded fringe tracker, with an emphasis on the behaviour of the group-delay control loop. Finally, Sect. 5 gives some prospects and avenues for future work, and Sect. 6 gives our final conclusions.

2. Observables

2.1. From pixels to fluxes

In GRAVITY, the fringe-tracking beam combiner, like the science combiner, is a silica-on-silicium integrated optics component that uses an ABCD beam-combining scheme. This means that for any given baseline j,k linking telescopes T_j to T_k, the beam combiner registers a set of four outputs, corresponding to four values of additional phase shifts introduced between the beams (0, $\pi/2$, π, and $3\pi/2$),

$$a_{j,k} = \left\langle |E_j|^2 + |E_k|^2 + 2\mathrm{Re}\left(E_j E_k^*\right)\right\rangle, \tag{1}$$

$$b_{j,k} = \left\langle |E_j|^2 + |E_k|^2 + 2\mathrm{Re}\left(E_j E_k^* e^{i\pi/2}\right)\right\rangle, \tag{2}$$

$$c_{j,k} = \left\langle |E_j|^2 + |E_k|^2 + 2\mathrm{Re}\left(E_j E_k^* e^{i\pi}\right)\right\rangle, \tag{3}$$

$$d_{j,k} = \left\langle |E_j|^2 + |E_k|^2 + 2\mathrm{Re}\left(E_j E_k^* e^{3\pi/2}\right)\right\rangle, \tag{4}$$

where E is the complex amplitude of the electric field at the input of the combiner, $\langle . \rangle$ denotes the average over the integration time, and the asterisk denotes the complex conjugate.

In principle, both the coherent flux $\Gamma_{j,k} = \left\langle E_j E_k^* \right\rangle$ and the incoherent flux $F_j + F_k = \left\langle |E_j|^2 \right\rangle + \left\langle |E_k|^2 \right\rangle$ can be recovered by linear combinations of these four outputs,

$$F_j + F_k = \left(a_{j,k} + b_{j,k} + c_{j,k} + d_{j,k}\right)/4, \tag{5}$$

$$\Gamma_{j,k} = (a_{j,k} - c_{j,k})/2 + i\,(d_{j,k} - b_{j,k})/2. \tag{6}$$

Each individual F_j for $j \in [1 \ldots 4]$ can subsequently be recovered by a second linear combination of the different baselines,

$$F_j = \sum_{k \neq j} (F_j + F_k)/3 - \sum_{\substack{k,l \neq j \\ k<l}} (F_k + F_l)/6. \tag{7}$$

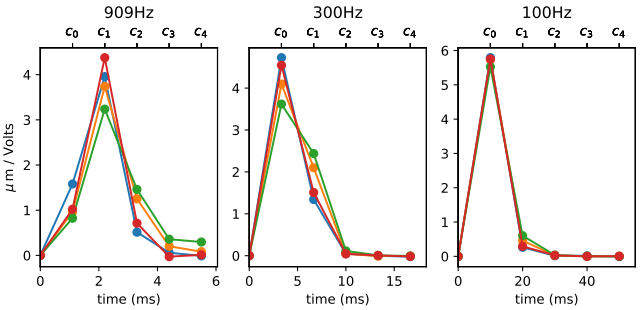

Fig. 1. Open-loop response of the piezo control chain. This corresponds to the response of the piezo in micrometers per volt as a function of time (bottom axis) when submitted to an impulse command. Equivalently, this also gives the values of the coefficients of the fifth-order model used in Eq. (20) (top axis). Each colour represents a distinct actuator.

In the GRAVITY fringe tracker, the 24 outputs (4 outputs for each of the six ABCD combiners) are dispersed over six wavelength channels in the K band and are recorded on a SAPHIRA detector (Finger et al. 2014). We refer to Fig. 1 in Paper I for an overview of the pixel arrangement on the detector real time display (RTD). As shown by Eqs. (1)–(7), for a given wavelength channel, the column vector $(q_1, \ldots, q_{24})^T$ containing the 24 different intensity outputs and the column vector $(F_j, \Gamma_{j,k})^T$ containing the ten incoherent and coherent flux values are related by a multiplication by a single matrix: the P2VM (pixel to visibility matrix),

$$\mathrm{P2VM} \cdot (q_1, \ldots, q_{24})^T = (F_1, \ldots, F_4, \Gamma_{1,2}, \ldots, \Gamma_{3,4})^T. \tag{8}$$

The incoherent and coherent fluxes are computed in real time for each of the wavelength channels, which are treated independently. Similarly, when the polarisation is split, the two polarisations are treated independently.

In practice, the ABCD combiners are not perfectly balanced, and the phase offsets can be different from their fiducial values, which means that the coefficients linking the ABCD outputs to the complex amplitudes of the electric field in Eqs. (1)–(4) can vary. Therefore, the P2VM must be calibrated, which in GRAVITY is done during daytime using an internal source. This P2VM formalism was initially introduced by Tatulli et al. (2007) for the AMBER instrument (Petrov et al. 2007). It was then adapted to ABCD combiners by Lacour et al. (2008), and more details about its implementation can be found in these papers, along with Paper I.

2.2. From fluxes to phase/group delays and closures

After the P2VM calculation, the resulting wavelength-dependent coherent fluxes $\Gamma_{j,k,\lambda}$ are still affected by dispersion (atmospheric dispersion, and dispersion in the fibre delay lines, FDDL), which at first order introduces a phase curvature of the form $e^{iD(1-\lambda_0/\lambda)^2}$, where $\lambda_0 = 2.2\,\mu$m. This is corrected for by using an empirical value for D that depends on the position of the star in the sky and on the position of the FDDLs. This correction was calibrated during the first year of on-sky observations.

From this dispersion-corrected coherent flux measurement, the OPD $\phi_{j,k}$ is calculated for baseline (j,k) using

$$\phi_{j,k} = \frac{\lambda_0}{2\pi} \times \arg\left(\sum_{\lambda=1}^{N_\lambda} \Gamma_{j,k,\lambda}\right). \tag{9}$$

These values are concatenated to form the OPD vector Φ, using the convention $\Phi = (\phi_{21}, \phi_{31}, \phi_{41}, \phi_{32}, \phi_{42}, \phi_{43})^T$. The closure phase $\theta^{\mathrm{PD}}_{j,k,l}$ over the triangle linking telescopes T_j, T_k, and T_l is calculated in units of length using

$$\theta^{\mathrm{PD}}_{j,k,l} = \frac{\lambda_0}{2\pi} \times \arg\left(\left\langle \sum_{\lambda=1}^{N_\lambda} \Gamma_{j,k,\lambda} \sum_{\lambda=1}^{N_\lambda} \Gamma_{k,l,\lambda} \sum_{\lambda=1}^{N_\lambda} \Gamma^*_{j,l,\lambda} \right\rangle_{350\mathrm{DITs}}\right), \tag{10}$$

where the incoherent flux is first averaged over 350 detector integration times (DITs, i.e., bout 385 ms at 909 Hz) to boost the S/N, the phase closure being a notoriously noisy quantity.

The group delay $\psi_{j,k}$ and group closure $\theta^{\mathrm{GD}}_{j,k,l}$ are obtained from the coherent flux, which is first corrected for any phase offset constant in wavelength by subtracting the phase delay, and then averaged over 150 DITs (about 165 ms at 909 Hz), again to boost S/N,

$$\Gamma'_{j,k,\lambda} = \left\langle \Gamma_{j,k,\lambda} \exp\left[-i\frac{2\pi}{\lambda_0}\phi_{j,k}\right]\right\rangle_{150\ \mathrm{DITs}}. \tag{11}$$

The group delay is then given by

$$\psi_{j,k} = \frac{\lambda - 0^2/\Delta\lambda}{2\pi} \times \arg\left(\sum_{\lambda=1}^{N_\lambda - 1} \Gamma'_{j,k,\lambda+1} \Gamma'^*_{j,k,\lambda}\right), \tag{12}$$

where $\Delta\lambda$ is the difference between the effective wavelength of two pixel bins. From this, we also constructed another vector in OPD space: $\Psi = (\psi_{21}, \psi_{31}, \psi_{41}, \psi_{32}, \psi_{42}, \psi_{43})^T$. The closure of the group delay can be written as

$$\theta^{\mathrm{GD}}_{j,k,l} = \frac{\lambda - 0^2/\Delta\lambda}{2\pi} \times \arg\left(\left\langle \sum_{\lambda=1}^{N_\lambda - 1} \Gamma'_{\substack{j,k,\\ \lambda+1}} \Gamma'^*_{\substack{j,k,\\ \lambda}} \sum_{\lambda=1}^{N_\lambda - 1} \Gamma'_{\substack{k,l,\\ \lambda+1}} \Gamma'^*_{\substack{k,l,\\ \lambda}} \sum_{\lambda=1}^{N_\lambda - 1} \Gamma'^*_{\substack{j,l,\\ \lambda+1}} \Gamma'_{\substack{j,l,\\ \lambda}} \right\rangle_{350\ \mathrm{DITs}}\right). \tag{13}$$

The last quantity of interest for the control algorithm is the uncertainty on the OPD estimate, $\sigma_{j,k}$. This uncertainty can be computed by first introducing the interferometric S/N, defined as the ratio of the modulus of the coherent flux and the square root of its variance. This variance is approximated as the half sum of the variance of the real and imaginary parts of the coherent flux, and the S/N is averaged over three DITs[1],

$$S/N_{j,k} = \left\langle \frac{\left|\sum_\lambda \Gamma_{j,k,\lambda}\right|}{\sqrt{\frac{1}{2}\sum_\lambda \mathrm{Var}(\Re\Gamma_{j,k,\lambda}) + \frac{1}{2}\sum_\lambda \mathrm{Var}(\Im\Gamma_{j,k,\lambda})}} \right\rangle_{3\ \mathrm{DITs}}. \tag{14}$$

The uncertainty on the phase can then be computed using a similar approach as in the appendix of Shao et al. (1988),

$$\frac{1}{\sigma_{j,k}} = \frac{2\pi}{\lambda_0} \times S/N_{j,k}. \tag{15}$$

[1] We thank the referee for pointing out that calculating this S/N as the ratio between the average of the numerator and the average of the denominator (as opposed to the average of the ratio) would likely result in a better estimate. We plan on testing this and implementing it in a future version of the fringe tracker.

In Eq. (14), the variance of the real and imaginary parts of the coherent flux are estimated for each frame from the background, detector, and photon noises. Using these OPD noise estimators, and under the assumption that each baseline noise is uncorrelated, we also write the covariance matrix W on Φ,

$$W = \mathrm{diag}\left(\sigma^2_{21}, \sigma^2_{31}, \sigma^2_{41}, \sigma^2_{32}, \sigma^2_{42}, \sigma^2_{43}\right). \tag{16}$$

The variance of Γ is estimated from the variance on the individual pixels q_j, multiplied by the P2VM and P2VM^T matrix, as described in Sect. 3.4 of Paper I. Compared to Paper I, where the $\sigma_{j,k}$ were computed coherently on five DITs, the new uncertainties are now computed on individual DITs and then averaged over only three DITs. This change was motivated by the use of an adaptive gain in the new control algorithm, which strongly benefits from a better estimate of the instantaneous S/N.

3. Control loop

3.1. Three challenges of building a fringe tracker

A first challenge that needs to be solved in the implementation of a fringe tracker is the dichotomy between two distinct spaces. The fringe tracker uses a set of four piezoelectric delay lines (Pfuhl et al. 2014) to change the OPL of the four beams (one per telescope). The feedback is provided by measurements of the coherent fluxes (see Sect. 2), however, which are transformed into a set of six measurements of OPD ϕ (one per baseline) and group delay ψ. The OPL space itself is not observable, and this mismatch between the control space and the measurement space creates a paradox that needs to be solved by the system.

A second challenge arises from the need to regularly update the model that is used to represent and predict the disturbance. To do this, we chose to use a set of AR models, whose parameters are updated every few seconds via a model-fitting routine performed on a separate computer (see the hardware description in Sect. 4.1). In order to perform the model fitting, this computer needs to be fed with measurements that only represent the disturbance, excluding the additional piezo command. The control must therefore work with a state model that explicitly separates the OPL into two components: a disturbance component L that represents both atmospheric and mechanical turbulence, and the OPL introduced by the piezo actuators X. Only the total OPD, which derives from the total OPL $L + X$, can be observed by the fringe tracker. In order to be able to update the disturbance model, however, the two components must be kept separated.

A third challenge comes from the fact that the main phase control loop is fed by phase measurements, which are only known to a modulo 2π (or λ_0, in terms of OPD). Thus, this loop is blind to any potential fringe shift, and we therefore need the additional group delay loop. In Paper I, the controller consisted of two completely independent loops running in parallel, with the potential of issuing conflicting commands. The algorithm developed at the time ensured that the integrator of the two controllers would not diverge too widely, but it did not strictly constrain the absolute group delay. The consequence was a tendency of the fringe tracker to jump between fringes during a single observation. The updated architecture detailed in this work provides a much more unified approach to controlling both the group and phase delays, thereby avoiding conflicting commands, mitigating phase jumps, and maintaining tracking of the white-light fringe.

3.2. State model

To solve these challenges, we adopted a state model in OPL space that explicitly separates the disturbance and the OPL introduced by the actuators. We used a set of two state vectors that both have the unit of length and that we denote X_n and L_n. For simplicity, these state vectors can be broken down into four independent vectors, corresponding to each telescope. For a given telescope T_k, the vector of the last N values of OPL introduced by the piezo actuators ($X_{n,k}$) and by the disturbance ($L_{n,k}$) can be expressed by

$$X_{n,k} = [x_k(t_n), x_k(t_{n-1}), \ldots, x_k(t_{n-N+1})]^T, \tag{17}$$

$$L_{n,k} = [l_k(t_n), l_k(t_{n-1}), \ldots, l_k(t_{n-N+1})]^T, \tag{18}$$

where $x_k(t_j)$ and $l_k(t_j)$ denote the OPL introduced by the piezo actuator and the disturbance on telescope T_k at time t_j, respectively. These vectors have a total length of 150, which matches the smoothing length of the group delay. The four telescopes can be concatenated to give our two final state vectors X_n and L_n of dimension $4 \times 150 = 600$,

$$X_n = \begin{bmatrix} X_{n,1} \\ X_{n,2} \\ X_{n,3} \\ X_{n,4} \end{bmatrix} \quad L_n = \begin{bmatrix} L_{n,1} \\ L_{n,2} \\ L_{n,3} \\ L_{n,4} \end{bmatrix}. \tag{19}$$

Each of the piezo control chains (which include the actuator response as well as the pure delay introduced by the processing) has a response function that we modelled as a fifth order, where the position of the piezo at time t_n was derived from the command sent at the five last time steps,

$$x_k(t_{n+1}) = c_{k,0}u_k(t_n) + c_{k,1}u_k(t_{n-1}) + \ldots + c_{k,4}u_k(t_{n-4}). \tag{20}$$

The c coefficients can be empirically measured by sending an impulse command (in Volts) through the control chain and measuring how the response evolves with time (Fig. 1).

The command vectors $U_{n,k} = [u_k(t_n), \ldots, u_k(t_{n-4})]$ are then defined such that

$$x_k(t_{n+1}) = \begin{bmatrix} c_{k,0} & c_{k,1} & c_{k,2} & c_{k,3} & c_{k,4} \end{bmatrix} \cdot U_{n,k}. \tag{21}$$

Again, the four telescopes can be concatenated to give U_n of dimension $4 \times 5 = 20$,

$$U_n = \begin{bmatrix} U_{n,1} \\ U_{n,2} \\ U_{n,3} \\ U_{n,4} \end{bmatrix}. \tag{22}$$

With these notations, the propagation of the piezo state vector X_n takes the form

$$X_{n+1} = A_X \cdot X_n + C \cdot U_n, \tag{23}$$

where A_X is a 600×600 block-diagonal matrix that just shifts the xs for each telescope,

$$A_X = \mathrm{diag}(A, A, A, A), \tag{24}$$

with

$$A = \begin{bmatrix} 0 & 0 & \ldots & 0 & 0 \\ 1 & 0 & \ldots & 0 & 0 \\ 0 & 1 & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0 & 0 \\ 0 & \ldots & 0 & 1 & 0 \end{bmatrix}, \tag{25}$$

and C is also a block-diagonal matrix, but of size 600×20, which contains the coefficients of the piezo responses,

$$C = \mathrm{diag}(C_1, C_2, C_3, C_4), \tag{26}$$

with

$$C_k = \begin{bmatrix} c_{k,0} & c_{k,1} & c_{k,2} & c_{k,3} & c_{k,4} \\ 0 & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}. \tag{27}$$

For the disturbance state vector L_n, we used a model similar to what was already in use on the GRAVITY fringe tracker (see Paper I), which gave satisfactory results, and which consisted of a set of six AR models in OPD space (see Appendix A). Since the upgraded version of the fringe tracker now uses a state model in OPL rather than in OPD space, this AR model needed to be converted into OPL space. To do this, we first note that for a four-telescope (six-baseline) interferometer such as GRAVITY, an OPD vector $\Phi = (\phi_{1,2}, \phi_{1,3}, \phi_{1,4}, \phi_{2,3}, \phi_{2,4}, \phi_{3,4})^T$ and an OPL vector $(l_1, l_2, l_3, l_4)^T$ are linked by the matrix equation

$$\Phi = M \cdot l, \tag{28}$$

where M is a 6×4 matrix, defined as

$$M = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 1 & 0 & -1 & 0 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & -1 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -1 \end{bmatrix}. \tag{29}$$

Therefore, computing an OPL vector from the knowledge of the OPD vector is under-constrained and yields multiple solutions. A solution entails using the Moore-Penrose pseudo-inverse of M,

$$l = M^+ \cdot \Phi. \tag{30}$$

For reference, the value of M^+ is

$$M^+ = \frac{1}{4} \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 1 & 1 & 0 \\ 0 & -1 & 0 & -1 & 0 & 1 \\ 0 & 0 & -1 & 0 & -1 & -1 \end{bmatrix}. \tag{31}$$

This particular choice ensures that the mean value of the OPLs on each telescope is 0, which is beneficial for a fringe tracker. It helps to avoid runaway situations in which all the actuators move toward a large offset and saturate the system.

These equations can be extended to our OPL state-space vector L_n as defined in Eq. (19), and a similarly defined OPD vector $\Phi_n = (\Phi_{n,1,2}, \ldots, \Phi_{n,3,4})^T$ using similar matrices in which the 1s are replaced by identity matrices of size 150,

$$\mathcal{M} = \begin{bmatrix} I_{150} & -I_{150} & 0 & 0 \\ I_{150} & 0 & -I_{150} & 0 \\ I_{150} & 0 & 0 & -I_{150} \\ 0 & I_{150} & -I_{150} & 0 \\ 0 & I_{150} & 0 & -I_{150} \\ 0 & 0 & I_{150} & -I_{150} \end{bmatrix}, \tag{32}$$

$$\mathcal{M}^+ = \frac{1}{4} \begin{bmatrix} I_{150} & I_{150} & I_{150} & 0 & 0 & 0 \\ -I_{150} & 0 & 0 & I_{150} & I_{150} & 0 \\ 0 & -I_{150} & 0 & -I_{150} & 0 & I_{150} \\ 0 & 0 & -I_{150} & 0 & -I_{150} & -I_{150} \end{bmatrix}. \tag{33}$$

With these matrices, the relations remain

$$\Phi_n = \mathcal{M} \cdot L_n, \tag{34}$$
$$L_n = \mathcal{M}^+ \cdot \Phi_n. \tag{35}$$

Our AR model fitting provides a set of six matrices $A_{j,k}$ that can be used to propagate the individual $\Phi_{n,j,k}$. These can be combined into a block-diagonal matrix to propagate Φ_n itself,

$$\Phi_{n+1} = \begin{bmatrix} A_{1,2} & 0 & \dots & 0 \\ 0 & A_{1,3} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & A_{3,4} \end{bmatrix} \cdot \Phi_n. \tag{36}$$

Switching to OPL space, we have

$$\mathcal{M} \cdot L_{n+1} = \begin{bmatrix} A_{1,2} & 0 & \dots & 0 \\ 0 & A_{1,3} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & A_{3,4} \end{bmatrix} \cdot \mathcal{M} \cdot L_n. \tag{37}$$

At this point, we can note that

$$\mathcal{M}^+ \cdot \mathcal{M} = I_{600} - \frac{1}{4} \begin{bmatrix} I_{150} & I_{150} & I_{150} & I_{150} \\ I_{150} & I_{150} & I_{150} & I_{150} \\ I_{150} & I_{150} & I_{150} & I_{150} \\ I_{150} & I_{150} & I_{150} & I_{150} \end{bmatrix}, \tag{38}$$

which means that as long as our OPL vectors L_{n+1} contain values where the average on all telescopes is 0, we can consider that $\mathcal{M} \cdot \mathcal{M}$ is the identity matrix. In other words, as long as

$$\begin{bmatrix} I_{150} & I_{150} & I_{150} & I_{150} \\ I_{150} & I_{150} & I_{150} & I_{150} \\ I_{150} & I_{150} & I_{150} & I_{150} \\ I_{150} & I_{150} & I_{150} & I_{150} \end{bmatrix} \cdot L_{n+1} = 0, \tag{39}$$

we have

$$\mathcal{M}^+ \cdot \mathcal{M} \cdot L_{n+1} = L_{n+1}. \tag{40}$$

Assuming temporarily that L_{n+1} fulfils the condition given by Eq. (39), we can left-multiply Eq. (37) by $\mathcal{M}^+$ to obtain

$$L_{n+1} = \left(\mathcal{M}^+ \cdot \begin{bmatrix} A_{1,2} & 0 & \dots & 0 \\ 0 & A_{1,3} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & A_{3,4} \end{bmatrix} \cdot \mathcal{M} \right) \cdot L_n. \tag{41}$$

A block matrix multiplication inside the brackets gives the final form of our model in OPL state-space,

$$\hat{L}_{n+1} = A_L \cdot L_n, \tag{42}$$

with

$$A_L = \begin{bmatrix} \frac{A_{1,2}+A_{1,3}+A_{1,4}}{4} & -A_{1,2}/4 & -A_{1,3}/4 & -A_{1,4}/4 \\ -A_{1,2}/4 & \frac{A_{1,2}+A_{2,3}+A_{2,4}}{4} & -A_{2,3}/4 & -A_{2,4}/4 \\ -A_{1,3}/4 & -A_{2,3}/4 & \frac{A_{1,3}+A_{2,3}+A_{3,4}}{4} & -A_{3,4}/4 \\ -A_{1,4}/4 & -A_{2,4}/4 & -A_{3,4}/4 & \frac{A_{1,4}+A_{2,4}+A_{3,4}}{4} \end{bmatrix}. \tag{43}$$

We can formulate two important remarks on the propagation model given by Eq. (42). Firstly, since

$$\begin{bmatrix} I_{150} & I_{150} & I_{150} & I_{150} \\ I_{150} & I_{150} & I_{150} & I_{150} \\ I_{150} & I_{150} & I_{150} & I_{150} \\ I_{150} & I_{150} & I_{150} & I_{150} \end{bmatrix} \cdot A_L = 0, \tag{44}$$

our propagation model always guarantees that the condition given by Eq. (39) is fulfilled. Physically, this stems from the fact that our propagation model really is the conversion of an OPD model into OPL space, where we used the freedom given by the under-constrained nature of the conversion to explicitly guarantee that the average piston is always 0,through the definition of M^+ in Eq. (31). Secondly, the presence of $\hat{L}$ with a hat in Eq. (42) emphasises the fact that this is only a model-propagated estimate of L. There is no hat for X in Eq. (23), since we consider our piezo model to be perfect (or at least much better than our atmospheric model).

Equation (42) is linear, and the covariance matrix on L_n is propagated accordingly, using an additional 600×600 matrix Q, which represents the process noise,

$$\hat{P}_{n+1} = A_L \cdot P_n \cdot {A_L}^T + Q. \tag{45}$$

The process noise Q is derived from a set of 150×150 matrices $Q_{j,k}$ using a similar block construction,

$$Q = \begin{bmatrix} \frac{Q_{1,2}+Q_{1,3}+Q_{1,4}}{4} & -Q_{1,2}/4 & -Q_{1,3}/4 & -Q_{1,4}/4 \\ -Q_{1,2}/4 & \frac{Q_{1,2}+Q_{2,3}+Q_{2,4}}{4} & -Q_{2,3}/4 & -Q_{2,4}/4 \\ -Q_{1,3}/4 & -Q_{2,3}/4 & \frac{Q_{1,3}+Q_{2,3}+Q_{3,4}}{4} & -Q_{3,4}/4 \\ -Q_{1,4}/4 & -Q_{2,4}/4 & -Q_{3,4}/4 & \frac{Q_{1,4}+Q_{2,4}+Q_{3,4}}{4} \end{bmatrix}. \tag{46}$$

The matrices $A_{j,k}$ and $Q_{j,k}$ are recalculated every 5 s by an asynchronous machine (`wgvkalm` described in Sect. 4.1) based on OPD measurements. This computation makes use of the Python toolbox time-series analysis (Seabold & Perktold 2010). A detailed explanation of its implementation is provided in Appendix A.

3.3. Phase control

Without any feedback, that is, without any additional information about the state of the system, the internal model of the fringe tracker would simply continue to indefinitely propagate the state vector L using Eq. (42) and to dispatch commands to the piezo actuators to respond to its own predictions. This behaviour would probably be acceptable over a few iterations, but given the uncertainties on the predictions of the disturbance, the internal state vector L_n would quickly diverge from the real values. The internal model is aware of this, as Eq. (45) shows that in this scenario, the covariance matrix would slowly inflate due to the accumulating process noise Q.

To properly track the fringes over long periods, the fringe tracker therefore needs to integrate measurements into the loop. These measurements should not be taken at face value either, however, as they are also affected by varying levels of uncertainties. Integrating noisy measurements into the loop without careful consideration will create additional noise in the system, thereby reducing its overall performance. The main task of the Kalman controller in phase-delay loop is to determine how to best balance the information from the state model propagation and from the additional measurements performed on the system. In Paper I, this was done using a fixed asymptotic Kalman gain,

which could not react in real time to the quality of the model and measurements. This new version of the Kalman provides an optimal gain, calculated at each iteration to combine the two sources of information in the best possible way.

At each time step t_n, a set of six measurements of the phase (or OPD) are obtained on each baseline, with an estimate of their error bars. The measurement process is detailed in Sect. 2, and for the context of this section, we simply denote $\Phi_n = (\phi_{j,k}(t_n))$ the column vector of size 6 containing the individual phase measurements (see Eq. (9)), and W_n the associated 6×6 diagonal covariance matrix, as defined from Eq. (15) in Eq. (16).

These measurements are related to our state vectors L_n and X_n through the measurement equations,

$$\phi_{j,k}(t_n) = \left[l_k(t_n) - l_j(t_n) + x_k(t_n) - x_j(t_n) \right]_{\mathrm{mod}\ \lambda_0}. \tag{47}$$

This means that although our state vectors have a length of 600 to accommodate for the group-delay measurements, which are averaged over 150 DITs, the OPD measurement itself only relies on the first values (at time t_n) for each telescope. In matrix form, this can be written as

$$\Phi_n = \left[H_\phi \cdot (L_n + X_n) \right]_{\mathrm{mod}\ \lambda_0}, \tag{48}$$

with H_ϕ defined by

$$H_\phi = \begin{bmatrix} \boldsymbol{v}_1 & -\boldsymbol{v}_1 & 0 & 0 \\ \boldsymbol{v}_1 & 0 & -\boldsymbol{v}_1 & 0 \\ \boldsymbol{v}_1 & 0 & 0 & -\boldsymbol{v}_1 \\ 0 & \boldsymbol{v}_1 & -\boldsymbol{v}_1 & 0 \\ 0 & \boldsymbol{v}_1 & 0 & -\boldsymbol{v}_1 \\ 0 & 0 & \boldsymbol{v}_1 & -\boldsymbol{v}_1 \end{bmatrix}, \tag{49}$$

where $\boldsymbol{v}_1 = (1, 0, \ldots, 0)^T$ is the first basis vector with a length of 150.

Given our assumption that the piezo state-vector X_n is perfectly known, the difference between the actual measurement and the expected measurement $\hat{\Phi}_n = H_\phi \cdot (\hat{L}_n + X_n)$ is directly related to the difference between the actual state of the disturbance L_n and its estimate $\hat{L}_n$,

$$\Phi_n - \hat{\Phi}_n = \left[H_\phi \cdot \left(L_n - \hat{L}_n \right) \right]_{\mathrm{mod}\ \lambda_0}, \tag{50}$$

and ignoring the modulo, it is therefore tempting to simply update the state vector according to $L_n = \hat{L}_n + {H_\phi}^+ \cdot (\hat{\Phi}_n - \Phi)$. This would reinject all the noise from the measurement into the control loop, however, without any consideration for the respective levels of confidence of the measurement and the prediction from the model propagation.

These levels of confidence are represented by the matrices $\hat{P}_n$ and W_n, which can therefore be used to calculate a Kalman gain that properly weights the model prediction and the measurement (Nowak 2019, Sect. 4.3, for a proper derivation of this gain),

$$K_n = \left(\hat{P}_n \cdot {H_\phi}^T \right) \cdot \left(H_\phi \cdot \hat{P}_n \cdot {H_\phi}^T + W_n \right)^{-1}. \tag{51}$$

The update state vector and its covariance matrix are then given by

$$L_n = \hat{L}_n + K_n \cdot \left[\hat{\Phi}_n - \Phi_n + \Phi_{\mathrm{CP},n} \right]_{\mathrm{mod}\ \lambda_0}, \tag{52}$$

$$P_n = \hat{P}_n + K_n \cdot (H_\phi \cdot \hat{P}_n \cdot {H_\phi}^T + W_n) \cdot {K_n}^T, \tag{53}$$

which indeed reduces to $L_n = \hat{L}_n + {H_\phi}^+ \cdot (\hat{\Phi}_n - \Phi_n + \Phi_{\mathrm{CP},n})$ in the $W_n \ll \hat{P}_n$ regime and also to $L_n = \hat{L}_n$ in the $\hat{P}_n \ll W_n$ regime.

The introduction of Φ_{CP} in this equation ensures a zero closure-phase on the measurement error. It is crucial to understand that because the model is in OPL space, it cannot account for any non-zero closure phase. This closure-phase issue was already present in the first version of the fringe tracker, and we refer to Sect. 4.2 of Paper I for a detailed explanation. We simply recall here that Φ_{CP} is constructed by setting three of the four closure phases $\theta^{\mathrm{PD}}_{j,k,l}$ to the appropriate baselines, ignoring the closure phase on the triangle of the lowest S/N,

$$\Phi_{\mathrm{CP}} = \begin{pmatrix} \theta^{\mathrm{PD}}_{1,2,4} \\ \theta^{\mathrm{PD}}_{1,3,4} \\ 0 \\ \theta^{\mathrm{PD}}_{2,3,4} \\ 0 \\ 0 \end{pmatrix} \text{or} \begin{pmatrix} \theta^{\mathrm{PD}}_{1,2,3} \\ 0 \\ -\theta^{\mathrm{PD}}_{1,3,4} \\ 0 \\ -\theta^{\mathrm{PD}}_{2,3,4} \\ 0 \end{pmatrix} \text{or} \begin{pmatrix} 0 \\ -\theta^{\mathrm{PD}}_{1,2,3} \\ -\theta^{\mathrm{PD}}_{1,2,4} \\ 0 \\ 0 \\ \theta^{\mathrm{PD}}_{2,3,4} \end{pmatrix} \text{or} \begin{pmatrix} 0 \\ 0 \\ 0 \\ \theta^{\mathrm{PD}}_{1,2,3} \\ \theta^{\mathrm{PD}}_{1,2,4} \\ \theta^{\mathrm{PD}}_{1,3,4} \end{pmatrix}. \tag{54}$$

It is also important to note with this definition, two fundamentally different types of variations of Φ_{CP} can occur: (1) Φ_{CP} can change because of slowly varying closures phases in the measurements, in which case it simply tracks actual variations. (2) Φ_{CP} can switch between options given in Eq. (54), which occurs when the triangle of the lowest S/N changes. The second case corresponds to a change in the configuration, leading to an abrupt variation of Φ_{CP}, and it is therefore banned while the science detector is integrating. This behaviour is controlled by an internal Boolean parameter, called the mobility flag, which is "False" during science integrations.

3.4. Group-delay control

The initial version of the fringe tracker described in Paper I was designed for an instrument performing prolonged exposures at a consistent wavelength of 2.2 µm. As long as it stayed in the coherence envelope of the fringes, the fringe tracker had the freedom to transition from one fringe to the next. This is unacceptable if the GRAVITY fringe tracker is to be used to feed a different science instrument operating at a different wavelength, as this would blur the fringes on the science detector. In an effort to make the GRAVITY fringe tracker compatible with the L-band combiner MATISSE (Lopez et al. 2022), the group-delay controller was modified to consistently track the same fringe during a scientific exposure. This new MATISSE mode is called GRA4MAT (Woillez et al., in prep.), and is now routinely offered to the community.

For this to work, the group-delay control loop must determine in real time which fringe is being tracked by the phase-control loop, and it updates the state vector L_n by integer multiple of λ_0 if required, depending on which fringe it wishes to track. These changes are completely transparent to the phase-control loop due to the modulo λ_0 in Eq. (50).

To implement this behaviour, we used a formalism similar to the phase-delay loop. We first introduced the 6×600 measurement matrix H_ψ to relate the group delay Ψ to the state vectors L_n and X_n,

$$\Psi_n = H_\psi \cdot (L_n + X_n), \tag{55}$$

with

$$H_\psi = \frac{1}{150}\begin{bmatrix} \mathbf{1} & -\mathbf{1} & 0 & 0 \\ \mathbf{1} & 0 & -\mathbf{1} & 0 \\ \mathbf{1} & 0 & 0 & -\mathbf{1} \\ 0 & \mathbf{1} & -\mathbf{1} & 0 \\ 0 & \mathbf{1} & 0 & -\mathbf{1} \\ 0 & 0 & \mathbf{1} & -\mathbf{1} \end{bmatrix}, \tag{56}$$

and where $\mathbf{1} = (1, 1, \ldots, 1)$ is a row vector with a length of 150 full of 1s. This is similar to the definition of H_ϕ, but includes an averaging of the past 150 DITs.

We also introduced $\Psi_{\rm CP}$, an analogue of $\Phi_{\rm CP}$, which is calculated using $\theta^{\rm GD}_{j,k,l}$ instead of $\theta^{\rm PD}_{j,k,l}$ in Eq. (54), and an additional vector $\Psi_{\rm ZeroGD}$. This additional setpoint is used to select which specific fringe is set at a group delay of 0, to be tracked by the fringe tracker. Again, to avoid fringe jumps during a science integration, a mobility flag is used to freeze $\Psi_{\rm ZeroGD}$ during integrations on the science detector. When the mobility flag allows changes, the $\Psi_{\rm Zero\,GD}$ setpoint is estimated so that

$$\left[\Psi_n - \hat{\Psi}_n - \Psi_{\rm CP} - \Psi_{\rm ZeroGD}\right]_{\rm mod\ lambda_0} = 0. \tag{57}$$

This condition is valid modulo λ_0, which means that Eq. (57) does not exactly determine which fringe is tracked. To lift this uncertainty, two requirements must be added:
R1: Vector $\Psi_{\rm ZeroGD}$ must be as small as possible.
R2: Vector $\Psi_{\rm ZeroGD}$ must not have a closure phase.
The first requirement is to ensure that the system tracks the white-light fringe. The second requirement is needed because the closure phase has already been nulled by the $Psi_{\rm CP}$ term.

Therefore, the zero group-delay setpoint is estimated by the real-time computer as

$$\Psi_{\rm ZeroGD} = M \cdot \left[M^\dagger \cdot (\Psi_n - \hat{\Psi}_n - \Psi_{\rm CP})\right]_{\rm mod\ \lambda_0}, \tag{58}$$

where M was introduced in Eq. (29), and $M^\dagger$ is a pseudo-inverse weighted by the covariance matrix W_n, that is, the matrix that solves the linear system $Y = MX$ in the least-square sense, considering that W_n from by Eq. (16) is the covariance matrix of X. This matrix $M^\dagger$ can be calculated using

$$M^\dagger = (M^T \cdot W_n \cdot M)^{-1} \cdot M^T \cdot W_n. \tag{59}$$

The controller ensures that the group delay never deviates by more than $\lambda_0/2$ from this setpoint. To do this, the controller calculates at each cycle an error in the form of $\Psi - \hat{\Psi} - \Psi_{\rm CP} - \Psi_{\rm Zero\,GD}$. It converts this group-delay error into OPL space using $M^\dagger$ from Eq. (59). When this error exceeds $\lambda_0/2$ on any telescope, the controller subtracts λ_0 to all the 150 values corresponding to this telescope in the state vector L. Similarly, when an OPL value falls below $-\lambda/2$, the elements of L corresponding to this telescope are increased by λ_0. These adjustments by integer multiples of λ_0 never disturb the phase controller due to the presence of a modulo λ_0 in Eq. 52, and because the propagation model is stationary (more details in Appendix A).

3.5. Commands to the actuators

The command for the actuators is obtained from

$$U_{n+1} = F \cdot L_n + U_{{\rm setpoint},\,n} + U_{{\rm search},\,n}\,. \tag{60}$$

In this equation, $U_{\rm setpoint}$ is used to modulate the position of the fringes in a process detailed in Paper I[2]. This setpoint can now also be controlled externally by another machine, for example, to correct non-common OPD, as described in Sect. 4.1.

The term $U_{\rm search}$ is used to search for the fringes when they are lost by one or more telescopes. This is made to ensure that the fringe tracker continues to track the fringes on the baselines where they remain visible, while searching on the others.

The term $F \cdot L_n$ cancels out the disturbance calculated by the fringe tracker. More precisely, this term is used to negate the influence of the disturbance within a given number of future DITs. This number of DITs, d, is made to be adjusted to the mean response speed of the fringe tracker. Practically, it is $d = 2$ at 909 Hz, and $d = 1$ at lower frequencies. Hence, F is calculated such that

$$C \cdot F \cdot L_n = H_\phi \cdot A^d \cdot L_n. \tag{61}$$

This is obtained to the best of our capabilities by taking

$$F = \frac{1}{4}\begin{bmatrix} \frac{1}{C_{T_1}} & \frac{1}{C_{T_1}} & \frac{1}{C_{T_1}} & 0 & 0 & 0 \\ \frac{-1}{C_{T_2}} & 0 & 0 & \frac{1}{C_{T_2}} & \frac{1}{C_{T_2}} & 0 \\ 0 & \frac{-1}{C_{T_3}} & 0 & \frac{-1}{C_{T_3}} & 0 & \frac{1}{C_{T_3}} \\ 0 & 0 & \frac{-1}{C_{T_4}} & 0 & \frac{-1}{C_{T_4}} & \frac{-1}{C_{T_4}} \end{bmatrix} \cdot H_\phi \cdot A^d \tag{62}$$

with C_{T_k} representing the integrated impulse response of the piezo actuator at telescope k,

$$C_{T_k} = \sum_{n=0}^{4} c_{k,n}, \tag{63}$$

where the c coefficients are the same as in Eq. (20).

3.6. Summary of the new control loop

The architecture of the new control loop is illustrated in Fig. 2. In this updated architecture, the two control loops no longer run in parallel, as was the case in Paper I. To some extent, the phase loop can now be seen as the main control loop and the group-delay control as an auxiliary loop that provides an additional update of the state parameters to ensure proper tracking of the group delay. The phase tracking is performed using a Kalman filter that no longer relies on asymptotic gain, but uses a real-time covariance-weighted combination of a model prediction (in OPL space) and a measurement update (in OPD space). The control algorithm consists of the following steps:

1. At iteration n, we start with an estimate of the state vector $\hat{L}_n$, propagated from the previous iteration. Our state vector L contains the path length at each telescope over a given number $N = 150$ previous iterations. This estimate, which also comes with a covariance matrix $\hat{P}_n$, represents our best knowledge of the state parameters. We also have a state vector X_n that stores the position of the piezo actuators over the last 150 iterations.
2. From these state vectors, we construct the total OPL at each telescope, which is simply $\hat{L}_n + X_n$, and then calculate the expected values for the phase and group delay ($\hat{\Phi}$ and $\hat{\Psi}_n$) using a set of measurement matrices H_ϕ and H_ψ, given in Eqs. (49) and (56), respectively, and which convert our OPL-based vectors into our OPD-based measurements,

$$\hat{\Phi}_n = H_\phi \cdot (\hat{L}_n + X_n) \tag{64}$$

$$\hat{\Psi}_n = H_\psi \cdot (\hat{L}_n + X_n). \tag{65}$$

[2] $U_{\rm setpoint}$ is called $U_{\rm modulation}$ in Paper I.

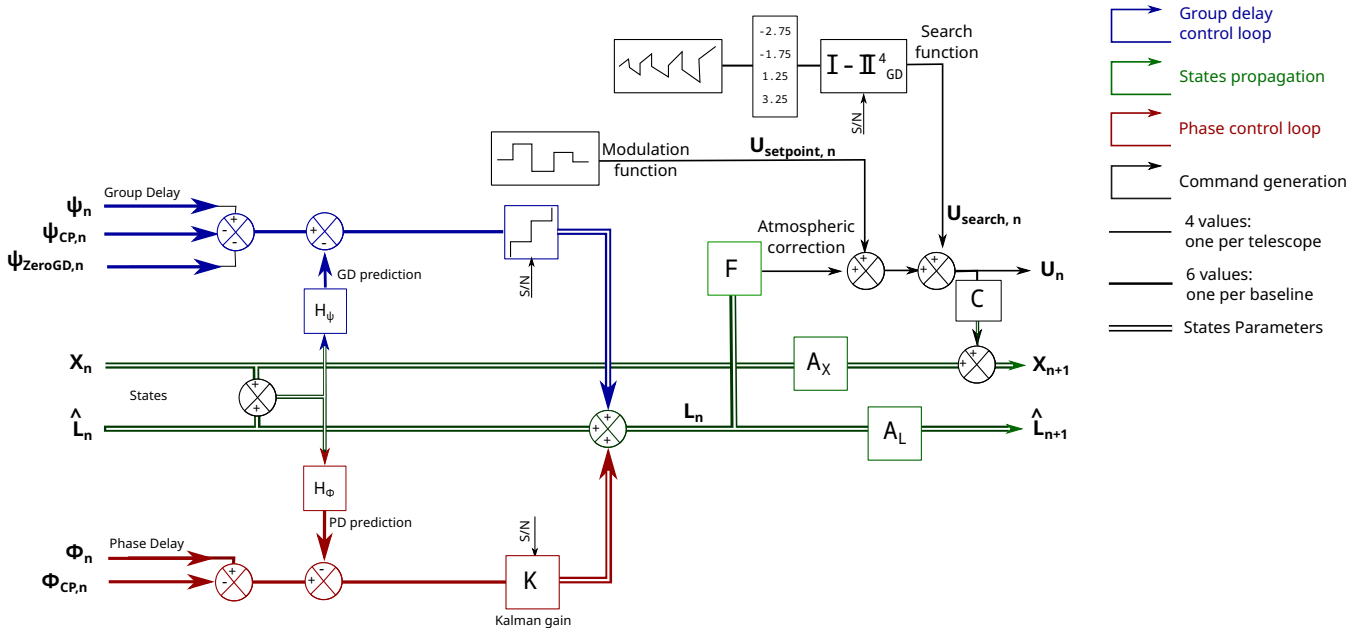

Fig. 2. Block diagram of the GRAVITY fringe-tracking controller within a single integration time, denoted n. The parameter states, illustrated in green, are central to the fringe tracker. They represent the disturbance conditions (L_n) and actuator positions (X_n). The blue segment indicates the group-delay feedback applied to the disturbance state, the red segment denotes the phase-delay feedback on L_n, and the black segments represent the command send to the actuators.

3. A measurement Φ of the OPDs is performed, which comes with an estimate of its uncertainties, in the process described in Sect. 2. The difference between this measurement and the expected value $\hat{\Phi}_n$ encodes some information on the state of the disturbance. This information is combined with the a priori knowledge of the state vector $\hat{L}_n$ to give an updated a posterori estimate L_n. This is a form of optimal data fusion that uses the Kalman gain K_n, defined in Eq. (51)

$$L_n = \hat{L}_n + K_n \cdot \left[\hat{\Phi}_n - (\Phi_n - \Phi_{\mathrm{CP},n})\right]_{\mathrm{mod}\,\lambda_0}. \tag{66}$$

   The use of Φ_{CP} arises from fact that the controller, which works in OPL, cannot account for a non-zero phase closure. At this step, the covariance P_n on L_n is propagated from $\hat{P}_n$, using Eq. (53).
4. The measured and expected group delays Ψ and $\hat{\Psi}$ are used to determine whether L_n needs to be further modified by an increment of integer multiples of λ_0. This is the group-delay tracking explained in Sect. 3.4.
5. The vector L_n contains our best estimate of the disturbance and is used to generate a command U_{n+1} in the process described by Eq. (60).
6. The state vectors are propagated according to their respective model,

$$X_{n+1} = A_X \cdot X_n + C \cdot U_{n+1} \tag{67}$$
$$\hat{L}_{n+1} = A_L \cdot L_n. \tag{68}$$

   The covariance matrix P_n on L_n is also propagated into $\hat{P}_{n+1}$ using

$$\hat{P}_{n+1} = A_L \cdot P_n \cdot {A_L}^T + Q. \tag{69}$$

   After this propagation, the next iteration can start.

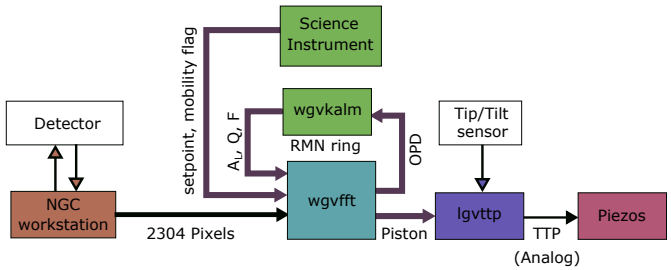

Fig. 3. Hardware architecture of the GRAVITY fringe tracker. The real-time controller operates within the `wgvfft` Linux workstation. Another Linux workstation, `wgvkalm`, is dedicated to calculating the Kalman filter parameters based on the observed OPDs. A Motorola CPU in the `lgvttp` workstation integrates the piston information with tip-tilt sensor data to accurately adjust the mirrors mounted on piezoelectric actuators in the tip-tilt piston (TTP). Additionally, an external workstation can provide the fringe tracker with piston setpoints and a mobility flag, which influences the controller behaviour.

4. On-sky performances

4.1. Hardware implementation

In November 2022, a substantial upgrade of the GRAVITY fringe-tracker hardware was carried out to increase the computing power, as outlined by Abuter et al. (2016). As a result, the computing time was markedly reduced, opening up the possibility of implementing the updated and more complicated control algorithm. This new hardware setup is depicted in Fig. 3.

The upgraded real-time hardware of the controller, termed `wgvfft`, is now based on a Linux workstation. The application cycle is initiated by the arrival of detector data frames from the ESO New General detector Control (NGC) workstation, facilitated through an sFPDP communication link. This application is executed on dedicated cores, with a portion of the RAM isolated

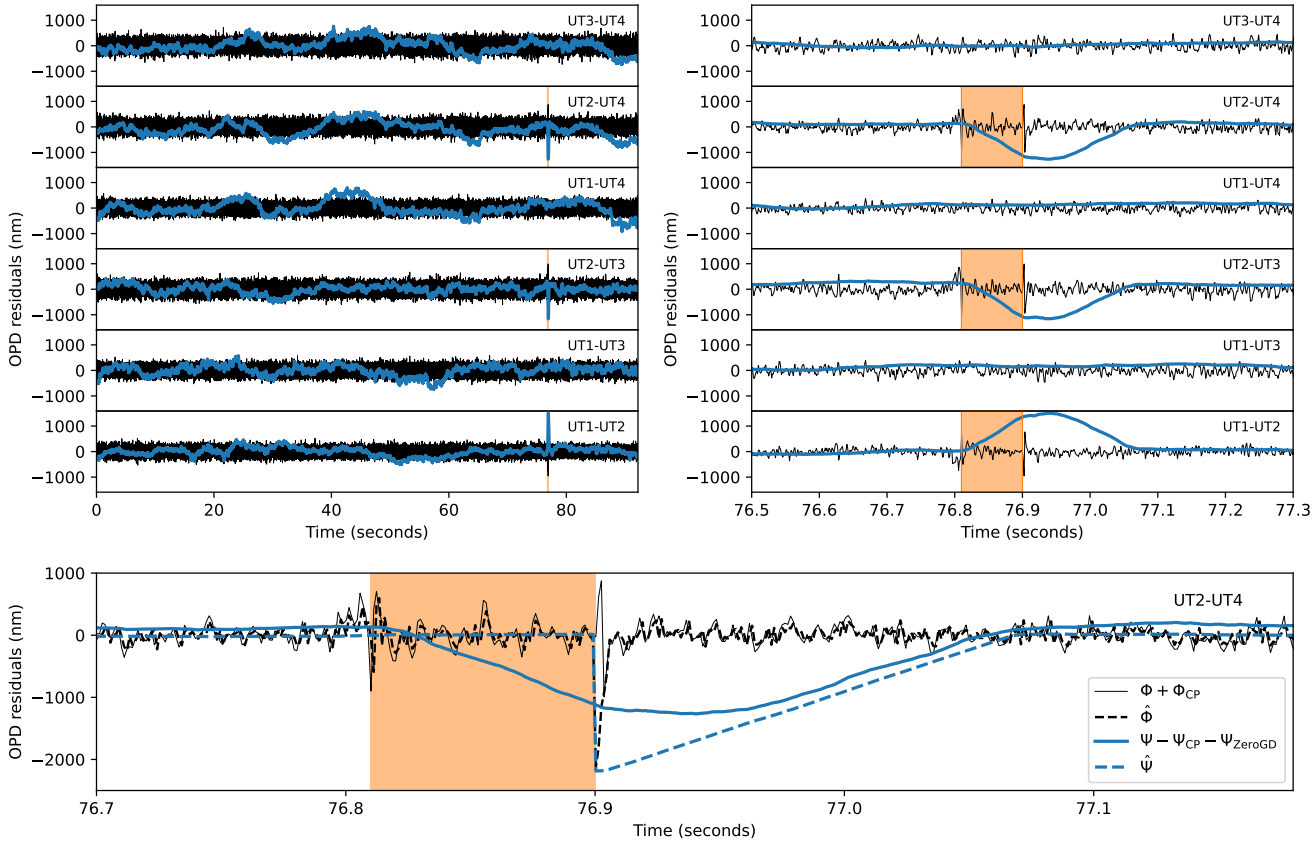

Fig. 4. Observation of star HR 8799 using GRAVITY. The data are taken from a single DIT exposure on the spectrometer, spanning 100 s. The black curves denote the OPD, derived from the phase delay ($\Phi - \Phi_{CP}$). The blue curves illustrate observations from the group delay ($\Psi - \Psi_{CP} - \Psi_{ZeroGD}$). Each panel represents a different baseline. At $t_1 = 76.81$ s, a noticeable fringe jump occurred on UT2. The effect of this jump across all baselines is depicted in the right panels. Impressively, the group delay identified and corrected the jump within 100 ms. The lower panel offers a magnified perspective of the baseline between UT2 and UT4, and the predictions from the controller state $\hat{\Phi}$ and $\hat{\Psi}$ are overlaid. Notably, at $t_2 = 76.9$ s, all predictions shifted by 2200 nm, marking the detection of the jump. To shift the OPD by one λ, the phase delay took 3 DITs (3.3 ms), while the group delay took 150 DITs (165 ms), which is attributed to the smoothing length of the observable.

from the Linux kernel, ensuring efficient access by the reflective memory (PCIE-5565PIORC) via direct memory access (DMA).

The Kalman filter parameters (A_L, Q, and F) are recalculated every 5 s by another Linux workstation (`wgvkalm`). This workstation is also fitted with a PCIE reflective memory card, which allows it to record the OPDs calculated by `wgvfft` through the VLTI reflective memory network (RMN) ring. This RMN ring is also used to transfer the Kalman filter parameters back to `wgvfft`.

An optional external machine, also connected to the RMN ring, can be used by another instrument to control the fringe tracker in real time. A workstation like this would have the possibility to send to `wgvfft` fringe-tracking setpoints (the U_{setpoint}). It can also be used to send the mobility flag introduced after Eq. (54).

The piston-correction values calculated by the `wgvfft` workstation are transmitted to another workstation named `lgvttp` using the RMN ring network. The `lgvttp` workstation operates with a Motorola CPU and is built around an mv6100 single-board computer, incorporating a VME bus for system communication. There, they are combined with external measurements of tip-tilt to form a comprehensive correction signal. This signal in analog form is then sent out to control four active mirrors mounted on piezoelectric tip-tilt platforms provided by Physik Instrumente.

The commissioning of the updated hardware was conducted in late 2022 as part of the GRAVITY+ upgrade (Eisenhauer 2019; GRAVITY+ Collaboration 2022) and was promptly made available to the scientific community. We could enable the white-light fringe tracking a few months later, allowing the GRA4MAT mode of MATISSE (Woillez et al., in prep.). Similarly, the new Kalman filter state model, introduced in June 2023, was immediately offered to the community upon commissioning.

4.2. Performance of the group-delay control loop: Analysis of a fringe jump

On 2 July 2023, we observed the exoplanet HR 8799 e during a scientific run of the ExoGRAVITY large programme (ESO LP 1104.C-0651). The atmospheric conditions were average, with a coherence time of 4 ms and a seeing of 0.8″. The planet was at a separation of 440 mas, and we observed using the off-axis dual-field mode of the fringe tracker. We used the roof mirror to split the field, so that all the flux from the star was injected in the fringe tracker fibre and all the flux of the planet in the science fibre.

Because the exoplanets are faint, these observations require extended integration times on the science channel, specifically, DITs of 100 s in this example. To optimise contrast, the fringe tracker was instructed to remain on the same fringe for the entire 100 s duration. Figure 4 shows that the fringe tracker was able to properly track the fringes on all six baselines. The residual

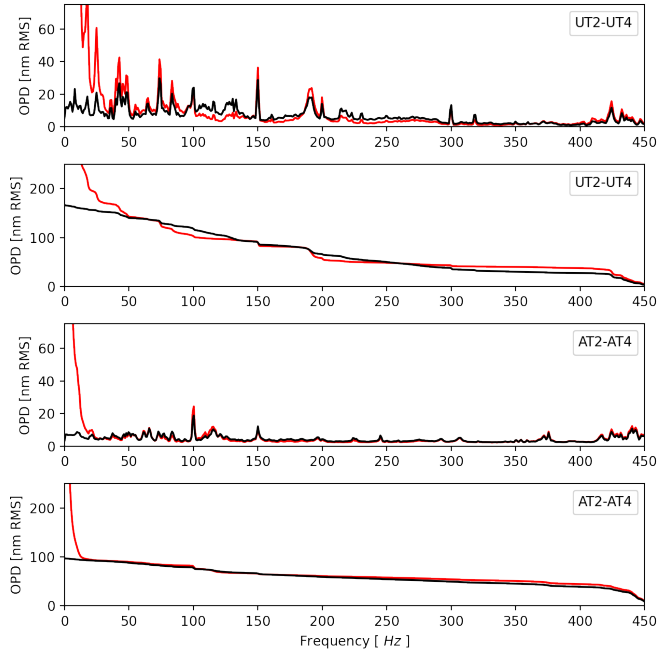

Fig. 5. Spectra of the residual OPD (shown in black) and the reconstructed perturbation ($H_\phi \cdot L$, depicted in red) for two observations. The upper panels present data on HR 8799 for one baseline from the UTs, corresponding to the data shown in Fig. 4. The standard deviation of the OPD is 151 nm. The panel below shows a dataset from the ATs for HR 7672 with an OPD standard deviation of 86 nm. The cumulative spectrum for these two dataset are also shown. The data for all baselines are presented in Appendix B.

OPD of this constant tracking is plotted in the upper left panel of Fig. 4. HR 8799 is a bright star with a magnitude of 5.24 in K band and is no challenging target for the fringe tracker, which easily tracks these fringes at an interferometric S/N of ~45. The power spectrum of the OPD residuals is shown in Fig. 5 and in Appendix B.

Over a 100-second integration period, the phase-delay controller consistently tracked constant OPD values, maintaining a standard deviation of 150 nm in the residual OPDs. Intriguingly, the group-delay residuals are notably higher, but they manifest at lower frequencies and vary over timescales of several seconds. The peak-to-peak amplitude of these variations remains below 1000 nm. As a result, the group-delay controller perceived no need for fringe adjustment and avoids any false triggers for a phase-jump detection. Consequently, the phase-delay fringe tracker operated smoothly and successfully maintained its tracking in the same phase.

Nonetheless, sporadic events can sometimes prompt the fringe tracker to transition abruptly from tracking one fringe to the next a full wavelength, λ_0, apart. This phenomenon, unnoticed by the phase delay, is referred to as a fringe jump. A quintessential instance of this can be observed in Fig. 4. This fringe jump manifested itself at $t_1 = 76.81$ s across all baselines related to UT2. The lower panel of the figure gives the phase delay ($\Phi - \Phi_{CP}$), represented by solid black curves. This phase delay has already accounted for the closure phase vector, as outlined in Eq. (10). The dashed black curve displays the estimation $\hat{\Phi}$ derived from the state parameters. Surrounding t_1, the OPD fluctuates, but the Kalman filter overlooks the phase jump, and the λ_0 deviation in the OPD is not registered by $\hat{\Phi}$. This jump manifests itself in the group delay, represented by the solid blue curve in the figure. Owing to its smoothing over 150 DITs, the group delay gradually moves to a value close to $-\lambda/2$ at $t_2 = 76.9$. At this point, all the UT2-related OPL state parameters of the phase control loop undergo a shift by λ_0, induced by the group-delay controller. This is evidenced in the figure by the abrupt λ_0 shift of the phase- and group-delay state predictions ($\hat{\Phi}$ and $\hat{\Psi}$).

This example perfectly illustrates the behaviour of the group-delay controller and shows that it required a span of 90 ms to detect the fringe jump. This interval represents the response time of the group-delay control loop and is dictated by the group-delay smoothing length (the 150 DITs). This setting represents a balance between responsiveness and the potential for false positives due to noise[3]. As illustrated in the top right panels of Fig. 4, it is challenging to predict this noise because it does not appear to be purely random. The value of 150 retained in the algorithm has a mostly empirical basis.

4.3. Performance of the phase-delay control loop: OPD residuals

In Paper I, we discussed the OPD residuals observed in 2019, noting a significant dependence on the coherence time, τ_0. During unfavourable conditions ($\tau_0 < 3$ mas), the values typically exceeded 380 nm. Under average conditions (3 mas $< \tau_0 <$ 7 mas), these values ranged around 150 nm (ATs) or 250 nm (UTs), and under optimal conditions ($\tau_0 > 7$ mas), the UTs observed values down to 220 nm. We hypothesised that the inability of the UTs to reach residuals below 220 nm even in the best atmospheric conditions was due to vibrations.

Over the past 4 ye, significant efforts have been made to mitigate these vibrations. Alongside the upgraded fringe tracker, these efforts have resulted in notable improvements. Figure 5 displays the spectrum of the OPD residuals for two targets, HR 8799 and HR 7672, over a single representative baseline. The description of the dataset and spectrum for all baselines is presented in Appendix B.

For the UTs, vibrations are predominant, exhibiting significant components at high frequencies, such as 20 nm vibrations at 300 Hz and 40 nm at 150 Hz. The control system strives to cancel these out, achieving partial success. However, it inadvertently reintroduces noise at other frequencies, resulting in no substantial improvement in the OPD residuals down to 50 Hz. This is evident in the cumulative spectrum, where the red and black curves exhibit similar amplitudes from 50 Hz upwards. Below 50 Hz, the fringe tracker is highly effective, cancelling even broad-frequency vibrations around 40 Hz. For the UTs, the most detrimental range is between 50 and 100 Hz, where a forest of spectral lines challenges the controller's ability to correct them, contributing to the observed ≈150 nm of residual. Potentially, a faster response time in the control loop could further reduce residuals within this range.

On the ATs, vibrations are not the dominant factor in the residuals, which are characterised by a white-noise pattern that aligns with the theoretically expected S/N. For HR 7672, the dataset presented in Fig. 5, the interferometric S/N

[3] In principle, this could be a tunable parameter that could be adjusted on a per target basis. In the case of HR 8799, given the brightness of the star and the high interferometric S/N obtained, decreasing this value would likely result in better performances. However, in the framework in which the GRAVITY fringe tracker is implemented, this change is hard, and this value remains currently fixed at 150 DITs.

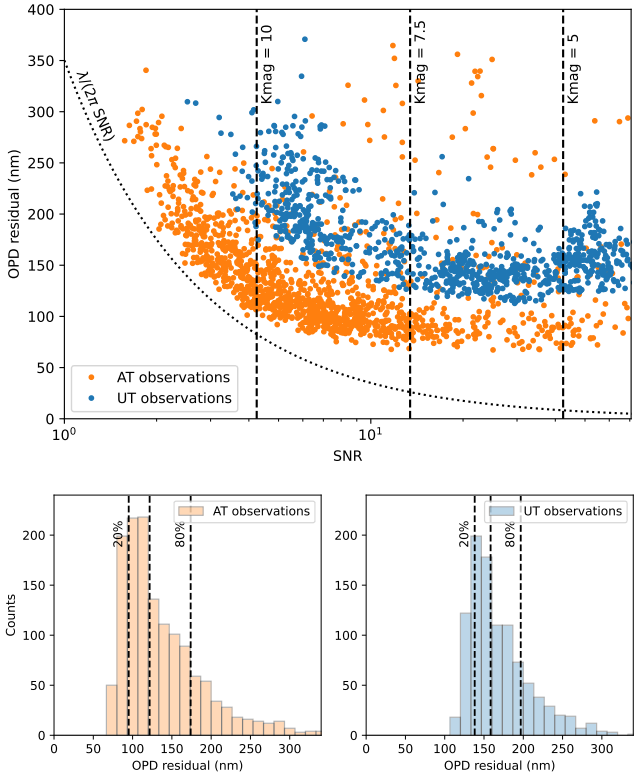

Fig. 6. OPD residuals. Top panel: OPD residuals plotted against the S/N per baseline for all data collected from June to August 2023. The orange dots represent AT observations, and the blue dots denote UT observations. The dotted curve represents the theoretical limitation imposed by the measurement noise ($\lambda/2\pi\sigma_{\mathrm{j,k}}$). The vertical dashed lines indicate the S/N observed with the UTs of a star at a given star magnitude. It should be shifted by about 2.5 magnitude for the ATs. Bottom panels: histogram of the same OPD residuals as in the upper panel for the ATs (left) and UTs (right).

is around 8, equating to a white-noise level of 60 nm. The residuals for different baselines range between 71 and 93 nm, which means that the system OPD residuals are close to the photon- and background-noise limitation. In this scenario, we observed minimal vibrations that did not degrade the interferometric S/N.

This behaviour of the ATs is also substantiated statistically, as demonstrated by the dataset present in the ESO archive on datasets with a low S/N (S/N below 3). Figure 6 shows the OPD residuals for all calibrators observed from June to August 2023. To expand the dataset, additional observations within the Exo-GRAVITY large programme during this period were included. The upper panel of Fig. 6 indicates that at lower fluxes for the ATs, the GRAVITY fringe tracker closely aligns with the theoretical phase error that arises from noise alone, indicating a reach towards the theoretical limit of the S/N. However, when the S/N surpasses 3 for both UTs and ATs, the performance begins to diverge from this limit, potentially due to piston noise from external factors such as the atmosphere.

In conclusion, the AT residuals at the 20%, 50%, and 80% percentiles are 95 nm, 120 nm, and 170 nm (1σ), respectively, and for the UTs, the corresponding values are 135 nm, 150 nm, and 185 nm. These values are derived from the histograms in the lower panels of Fig. 6. The dependence on the coherence time for the performance is reduced. An optimal performance can now be achieved even at a low τ_0 of 2 ms.

5. Future prospects

5.1. Areas of improvement

At least three key areas require further attention to enhance the capabilities of fringe tracking:

- Reducing the OPD residuals at high S/N;
- Enhancing the S/N for a given photon count;
- Boosting the sensitivity by operating at slower speeds.

OPD residuals. The critical question is whether OPD residuals can be further reduced at high S/N. There appears to be potential for improvement on the UTs, as their performance despite notable advances still lags that of the ATs. If we ignore injection-related issues, the atmospheric piston impact on an 8 m telescope should in theory be lower than on a 1.8 meter telescope (Conan et al. 1995). Because the baselines in the astrometric configuration used in our AT observations of HR 7672 are very similar to the UT baselines, the fringe tracker should have achieved better results on the UTs and it should be possible to reduce the OPD residuals below 80 nm on the UTs, with an ultimate atmospheric limitation of 5 nm according to Courtney-Barrer et al. (2022). The discrepancy in performance might be attributed to instrumental vibrations, which are particularly high above 50 Hz (Fig. 5). The control loop struggle to correct for this although they seem to be properly captured by the model. Amplitude fluctuations induced by the performance of the adaptive optics (AO), with lower Strehl ratios observed on the UTs, are another major factor to take into account. A better control of the vibrations, as well as a better AO correction, will certainly help us to further decrease the OPD residuals on the UTs.

The signal-to-noise ratio. Another question worthy of interest is whether the measured interferometric S/N matches the expectations given the magnitude of our targets. The observations of the bright target HR 8799 in July 2023 yielded an event rate of $n_{\mathrm{e^-/s/UT}} = 10^7$ photons per telescope per second, suggesting a combined transmission for GRAVITY and VLTI of about 1%. The photometric S/N can be calculated as

$$S/N_{\mathrm{photometric}} = \sqrt{\frac{4 n_{\mathrm{e^-/s/UT}} \delta t}{\gamma}}. \tag{70}$$

Here, the factor of 4 arises from the use of four telescopes, and γ represents the efficiency of the recombination architecture, with lower values being more desirable. At minimum, γ is inherently limited by the number of degrees of freedom in the system. Given the 16 degrees of freedom of our architecture (four real fluxes and six complex coherent fluxes), an optimised beam combiner could theoretically achieve a γ as low as 16. Assuming $\gamma = 16$, $\delta t = 0.00085$[4] and given the fact that the visibility is close to one for this unresolved target, the interferometric S/N for HR 8799 should be around 46, which indeed matches our measured value of ~45. This implies that the γ value for GRAVITY is indeed close to 16, which is indicative of a well-designed combiner. Further improvements will require innovative approaches such as the hierarchical fringe tracker, which might also decrease the γ value by reducing the system degrees of freedom (Petrov et al. 2022).

The ultimate sensitivity. During the June commissioning run, the exact sensitivity limit of the fringe tracker on the UTs

[4] The discrepancy between this integration time and the running frequency of 909 Hz is explained by the detector readout and reset time.

remained undetermined because the AO performance started to degrade at magnitudes below 10 (see the UT data in Fig. 6). The introduction of laser guide stars in GRAVITY+ is anticipated to extend the AO limiting magnitude, potentially allowing us to ascertain the true sensitivity of the enhanced GRAVITY fringe tracker. When the S/N limit at 1 kHz is achieved, exploring a 100 Hz mode becomes a viable option for further boosting sensitivity by mitigating the detector readout noise. In theory, as previously discussed, the theoretical atmospheric piston on an 8-meter telescope should be minimal, which could even enable an operation as slow as 10 Hz. However, achieving this would necessitate a reduction in the instrumental piston caused by vibrations. Mechanical solutions, such as directly removing vibrations or compensating for them using accelerometers (as in the ongoing MANHATTAN-II project, Lieto et al. 2008; Woillez et al. 2018), offer one approach. Alternatively, as these vibrations are intrinsic to the VLTI infrastructure, the optical path could be monitored and adjusted with a dedicated metrology system, as proposed in the now-abandoned VibMET project (Woillez et al. 2018).

5.2. Unexplored possibilities

The use of the GRAVITY fringe tracker alongside other instruments introduces a number of challenges in maintaining the fringe-tracking accuracy across different wavelengths. At a basic level, this requires the fringe tracker to avoid fringe jumps, a critical improvement implemented by the new algorithm described in this paper. A notable observation also pertains to the issue of varying atmospheric dispersion, however. The blue curves in Fig. 4 display slowly changing dispersion patterns. While the GD smoothing is approximately 150 ms, the blue curves exhibit correlations over significantly longer durations, up to 10 s. This can be attributed to fluctuations in water vapour, as investigated by Colavita et al. (2013) and Müller et al. (2014). These signals are relevant for infrared instruments, such as ELT/METIS (Absil et al. 2022), but might also be addressed at the fringe-tracker level to adjust the setpoint, ensuring stable fringes at another wavelength despite the changing dispersion.

Machine-learning algorithms represent another unexplored avenue in the realm of fringe tracking. These algorithms have the potential to integrate a broader spectrum of information compared to our current Kalman filter-based loop. Data from AO wavefront sensors and/or telescope vibration sensors, for instance, could substantially enhance the fringe-tracking capabilities (Perera et al. 2022). A model trained on a comprehensive dataset could leverage this additional information for a more accurate performance. The success observed in AO, particularly with reinforcement learning (Pou et al. 2022), indicates the potential advancements that machine learning could bring to fringe tracking.

6. Conclusion

We have introduced a new approach to fringe tracking that is based on a Kalman filter working in OPL state-space. This state-space is non-observable, which makes the integration of measurements harder than in OPD space. In particular, this leads to a somewhat cumbersome transformation of the AR model that is used to predict the evolution of the OPD disturbances to OPL space. The reward is a reduced dimensionality of the state-space from six (the number of baselines, or OPDs) to four (the number of telescopes, or OPLs), however. In the case of the new GRAVITY fringe tracker, this allows us to propagate covariance matrices on the state vector and hence to use an optimal Kalman gain that is dynamically calculated at each iteration.

We described this new algorithm specifically in the context of GRAVITY, but the conceptual framework can seamlessly be extended to various other instruments. As opposed to the number of OPDs, which grows quadratically with the number of telescopes in the array, the number of OPLs only grows linearly. Therefore, the gain offered by this novel approach increases with the size of the array, thereby offering a compelling solution for future concepts such as the Planet Formation Imager (Monnier et al. 2018).

In the meantime, with this new version of the GRAVITY fringe tracker, ESO is now able to offer the interferometric community a facility fringe tracker that not only delivers a competitive performance, but also ensures seamless integration within the VLTI environment. This new fringe tracker is already in use with MATISSE (the GRA4MAT mode, described in Woillez et al., in prep.). Beyond this, it could also accommodate visitor instruments, becoming a useful tool for a much larger community.

Acknowledgements. The authors wish to express their gratitude to the referee, F. Cassaing for an excellent report, which contained both useful suggestions to improve the paper, corrections of important mistakes, and ideas for future work. GRAVITY was developed in a collaboration of the Max Planck Institute for Extraterrestrial Physics, LESIA of Paris Observatory, IPAG of Université Grenoble Alpes / CNRS, the Max Planck Institute for Astronomy, the University of Cologne, the Centro Multidisciplinar de Astrofisica from Lisbon and Porto, and the European Southern Observatory. This work used observations collected at the European Southern Observatory under ESO programme 1104.C-0651 and 109.238N.003. S.L., J.B.L., and F.M. acknowledge the support of the French Agence Nationale de la Recherche (ANR), under grant ANR-21-CE31-0017 (project ExoVLTI). D.D. and R.L. have received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement CoG – 866070). P.G. acknowledges the financial support provided by FCT/Portugal through grants PTDC/FIS-AST/7002/2020 and UIDB/00099/2020. S.L. would like to warmly thank Sylvain Rousseau for spotting many typographical errors in the equations of Paper I.

References

Absil, O., Delacroix, C., Orban de Xivry, G., et al. 2022, SPIE Conf. Ser., 12185, 1218511
Abuter, R., Dembet, R., Lacour, S., et al. 2016, SPIE Conf. Ser., 9907, 990721
Cassaing, F., Le Duigou, J.-M., Houairi, K., et al. 2008, in Optical and Infrared Interferometry, eds. W. C. Danchi, F. Delplancke, & M. Schöller, Proc. Soc. Photo-Opt. Instrum. Eng., 7013
Choquet, É., Menu, J., Perrin, G., et al. 2014, A&A, 569, A2
Colavita, M. M., Booth, A. J., Garcia-Gathright, J. I., et al. 2010, PASP, 122, 795
Colavita, M. M., Wizinowich, P. L., Akeson, R. L., et al. 2013, PASP, 125, 1226
Conan, J. M., Rousset, G., & Madec, P.-Y. 1995, J. Opt. Soc. Am. A, 12, 1559
Courtney-Barrer, B., Woillez, J., Laugier, R., et al. 2022, SPIE Conf. Ser., 12183, 1218328
Delplancke, F., Derie, F., Lévêque, S., et al. 2006, Proc. SPIE, 6268, 62680U
Eisenhauer, F. 2019, in The Very Large Telescope in 2030 (Astronomy Australia Ltd)
Eisenhauer, F., Monnier, J. D., & Pfuhl, O. 2023, ARA&A, 61, 237
Finger, G., Baker, I., Alvarez, D., et al. 2014, SPIE Conf. Ser., 9148, 914817
GRAVITY Collaboration (Abuter, R., et al.) 2018a, A&A, 615, A15
GRAVITY Collaboration (Abuter, R., et al.) 2018b, A&A, 618, A10
GRAVITY Collaboration (Sturm, E., et al.) 2018c, Nature, 563, 657
GRAVITY Collaboration (Abuter, R., et al.) 2019a, A&A, 625, L10
GRAVITY Collaboration (Lacour, S., et al.) 2019b, A&A, 623, L11
GRAVITY Collaboration (Perraut, K., et al.) 2019c, A&A, 632, A53
GRAVITY Collaboration (Abuter, R., et al.) 2020a, A&A, 636, L5
GRAVITY Collaboration (Amorim, A., et al.) 2020b, A&A, 643, A154
GRAVITY Collaboration (Bauböck, M., et al.) 2020c, A&A, 635, A143
GRAVITY Collaboration (Dexter, J., et al.) 2020d, A&A, 635, A92
GRAVITY Collaboration (Nowak, M., et al.) 2020e, A&A, 633, A110

GRAVITY Collaboration (Pfuhl, O., et al.) 2020f, A&A, 634, A1
GRAVITY Collaboration (Abuter, R., et al.) 2021, A&A, 647, A59
GRAVITY Collaboration (Abuter, R., et al.) 2022, A&A, 657, L12
GRAVITY+ Collaboration (Abuter, R., et al.) 2022, The Messenger, 189, 17
Houairi, K., Cassaing, F., Perrin, G., et al. 2008, in Optical and Infrared Interferometry, eds. W. C. Danchi, F. Delplancke, & M. Schöller, Proc. Soc. Photo-Opt. Instrum. Eng., 7013
Lacour, S., Jocou, L., Moulin, T., et al. 2008, Proc. SPIE, 7013, 701316
Lacour, S., Dembet, R., Abuter, R., et al. 2019, A&A, 624, A99
Le Bouquin, J.-B., Bauvir, B., Haguenauer, P., et al. 2008, A&A, 481, 553
Lieto, N., Haguenauer, P., Sahlmann, J., & Vasisht, G. 2008, Proc. SPIE, 7013, 193
Lopez, B., Lagarde, S., Petrov, R. G., et al. 2022, A&A, 659, A192
Lozi, J., Cassaing, F., Le Duigou, J.-M., et al. 2011, in Techniques and Instrumentation for Detection of Exoplanets V, ed. S. Shaklan, Proc. Soc. Photo-Opt. Instrum. Eng., 8151
Menu, J., Perrin, G., Choquet, E., & Lacour, S. 2012, A&A, 541, A81
Monnier, J. D., Kraus, S., Ireland, M. J., et al. 2018, Exp. Astron., 46, 517
Müller, A., Pott, J. U., Mérand, A., et al. 2014, A&A, 567, A98
Nowak, M. 2019, Ph.D. Thesis, Université Paris sciences et lettres, France
Perera, S., Pott, J.-U., Woillez, J., et al. 2022, MNRAS, 511, 5709
Petrov, R. G., Malbet, F., Weigelt, G., et al. 2007, A&A, 464, 1
Petrov, R. G., Allouche, F., Boskri, A., et al. 2022, SPIE Conf. Ser., 12183, 121830I
Pfuhl, O., Haug, M., Eisenhauer, F., et al. 2014, Proc. SPIE, 9146, 914623
Pou, B., Ferreira, F., Quinones, E., Gratadour, D., & Martin, M. 2022, Opt. Express, 30, 2991
Quirk, M., & Liu, B. 1983, in ICASSP '83. IEEE International Conference on Acoustics, Speech, and Signal Processing, 8, 1095
Seabold, S., & Perktold, J. 2010, in 9th Python in Science Conference
Shao, M., Colavita, M. M., Hines, B. E., et al. 1988, A&A, 193, 357
Sorrente, B., Cassaing, F., Rousset, G., RobbeDubois, S., & Rabbia, Y. 2001, A&A, 365, 301
Tary, J. B., Herrera, R. H., & van der Baan, M. 2014, Geophys. J. Int., 196, 600
Tatulli, E., Millour, F., Chelli, A., et al. 2007, A&A, 464, 29
Woillez, J., Darré, P., Egner, S., et al. 2018, SPIE Conf. Ser., 10701, 1070103

[1] Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK
e-mail: mcn35@cam.ac.uk
[2] Kavli Institute for Cosmology, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK
[3] LESIA, Observatoire de Paris, Université PSL, CNRS, Sorbonne Université, Univ. Paris Diderot, Sorbonne Paris Cité, 5 place Jules Janssen, 92195 Meudon, France
[4] European Southern Observatory, Karl-Schwarzschild-Straße 2, 85748 Garching, Germany
[5] Max Planck Institute for extraterrestrial Physics, Giessenbachstraße 1, 85748 Garching, Germany
[6] Research School of Astronomy & Astrophysics, Australian National University, ACT 2611, Australia
[7] Department of Physics & Astronomy, University of Southampton, Hampshire SO17 1BJ, Southampton, UK
[8] Max Planck Institute for Astronomy, Konigstuhl 17, 69117 Heidelberg, Germany
[9] Univ. Grenoble Alpes, CNRS, IPAG, 38000 Grenoble, France
[10] Institute of Astronomy, KU Leuven, Celestijnenlaan 200D, 3001 Leuven, Belgium
[11] Laboratoire Lagrange, Observatoire de la Cote d'Azur, CNRS, Boulevard de l'Observatoire, CS 34229, 06304 Nice, France
[12] Universite de Lyon, CNRS, Centre de Recherche Astrophysique de Lyon UMR 5574, 69230 Saint-Genis-Laval, France
[13] Universidade do Porto, Faculdade de Engenharia, Rua Dr. Roberto Frias, 4200-465 Porto, Portugal
[14] CENTRA – Centro de Astrofísica e Gravitação, IST, Universidade de Lisboa, 1049-001 Lisboa, Portugal
[15] 1st Institute of Physics, University of Cologne, Zülpicher Straße 77, 50937 Cologne, Germany

Appendix A: Determination of the propagation matrix $A_{j,k}$ and process noise $Q_{j,k}$

The goal of this appendix is to derive the $A_{j,k}$ and $Q_{j,k}$ representing the AR models in OPD space that are used in Equation 43 and 46. The $A_{j,k}$ and $Q_{j,k}$ matrices were both obtained using the `statsmodels` Python library (Seabold & Perktold 2010) from a sequence of 10,000 Φ_n vectors (≈ 10 s at 909 Hz). The challenge in estimating the propagation matrix lies in its need to be resilient to unwrapping errors. It should also not produce aberrant values in situations where the fringes are undetected or lost.

To address this, the matrices are determined based on phase differences rather than the phases themselves. Explicitly, for each baseline, a quantity $\Delta\phi_{j,k,n} = \left[\phi_{j,k,n} - \phi_{j,k,n-1}\right]_{\mathrm{mod}\ \lambda_0}$ is computed. To ensure reliable default operation during low signal-to-noise ratios on this baseline (e.g. when fringes are absent or lost), we set $\Delta\phi_{j,k,n} = 0$.

The statsmodel time-series analysis is then employed to fit an AR model to the 9,999 $\Delta\phi_{j,k,n}$ values, one baseline at a time. The `statsmodels.tsa.ar_model.AutoReg` class is used to fit an AR model of order 22 using conditional maximum likelihood. The AR(22) model was chosen as it strikes a balance between computational efficiency and its ability to fit low frequencies. With 22 samples at 909 Hz, the minimum frequency and resolution of the predictor can be expected to be around 40 Hz. In practice, however, we find that the resolution is much better, and we are able to fit for frequencies as low as ~ 10 Hz, which is in line with the fact that AR models are notoriously better in terms of resolution than Fourier-transform based algorithm (Quirk & Liu 1983; Tary et al. 2014).

The `statsmodels.tsa.ar_model.AutoReg` class fit gives six vectors of 22 values for the AR model, $g_{l,j,k}$, along with six scale parameters $q_{l,j,k}$. The AR(22) model of the phase difference is then integrated to give an AR(23) model g' of the phase itself, using

$$g'_{0,j,k} = g_{0,j,k}, \tag{A.1}$$

$$g'_{23,j,k} = -g_{22,j,k}, \tag{A.2}$$

$$g'_{l,j,k} = g_{l,j,k} - g_{l-1,j,k} \quad \text{for } l \text{ between 1 and 22.} \tag{A.3}$$

This ensures the stationnarity of our model since $\sum_{l=0}^{23} g'_{l,j,k} = 0$. The propagation matrices representing the AR model are then

$$A_{j,k} = \begin{bmatrix} g'_{0,j,k} & g'_{1,j,k} & \cdots & g'_{21,j,k} & g'_{22,j,k} & 0 & \cdots & 0 \\ 1 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 & 1 & 0 \end{bmatrix}, \tag{A.4}$$

These matrices have a dimension of 150×150, but only the first 23 values and the sub-diagonal are non-zero. The AR(23) could be represented by a 23×23 matrix, but the additional dimensions are required by our state-space model, which also needs to account for the group delay.

The matrix $Q_{j,k}$ also has the dimension 150×150. It is more straightforwardly generated from the scale parameter $q_{0,jk}$, which corresponds to the variance of the residuals in the `statsmodels` fit,

$$Q_{j,k} = \begin{bmatrix} q_{0,j,k} & 0 & \ldots & 0 \\ 0 & 0 & \ldots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & \ldots & \ldots & 0 \end{bmatrix} \tag{A.5}$$

Appendix B: Power spectrum density

In Figs. B.1 and B.2, we present the Fourier transforms of 100 seconds of fringe-tracking OPD, Φ, acquired with the UTs and ATs, respectively. The first dataset was obtained using the UTs to observe HR 8799 (Kmag= 5.44) on July 2, with an effective K magnitude of 4.4. The second dataset was captured using the ATs in the astrometric configuration (A0-G1-J2-K0) to observe HR 7672 (Kmag= 4.4) on 21 July. The exact file IDs for these observations are GRAVI.2023-07-02T09:52:22.630 and GRAVI.2023-07-21T06:16:19.030. The atmospheric conditions were good during the HR 8799 observations with a seeing of 0.8″ and a wind speed of 3.9–4.6 m s^{-1}, and excellent during the HR 7672 observations with a seeing of 0.4″ and wind speed of ~ 2.8 m s^{-1}. We used Welch's method from the scipy library to compute the spectrum and cumulative spectrum of the OPD. The overall standard deviations of the OPDs are provided in the figure captions.

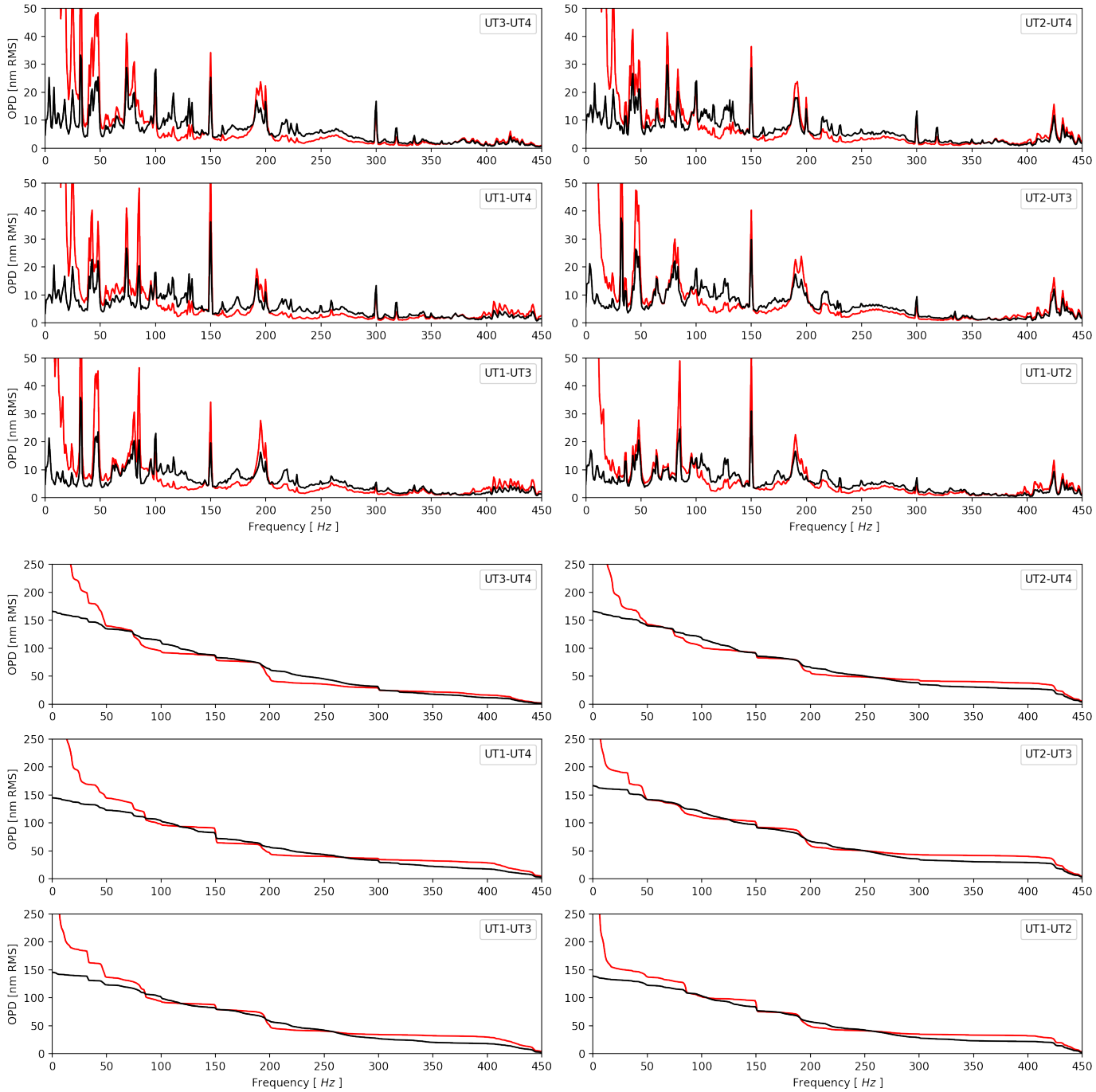

Fig. B.1. Power spectrum density (PSD) and cumulative sum of the PSD of phase residuals ($\mathbf{\Phi}_n$) for target HR 8799. The back curves represent the close-loop values. The red curves represent the pseudo-open-loop values, signifying the scenario without a fringe tracker ($H_\phi \cdot L$). The six lower plots show the reverse cumulative sum of the power spectrum. Over 100 s of a single scientific DIT, the residual standard deviation is 158 nm (UT3-UT4), 150 nm (UT2-UT4), 133 nm (UT1-UT4), 152 nm (UT2-UT3), 138 nm (UT1-UT3), and 124 nm (UT1-UT2).

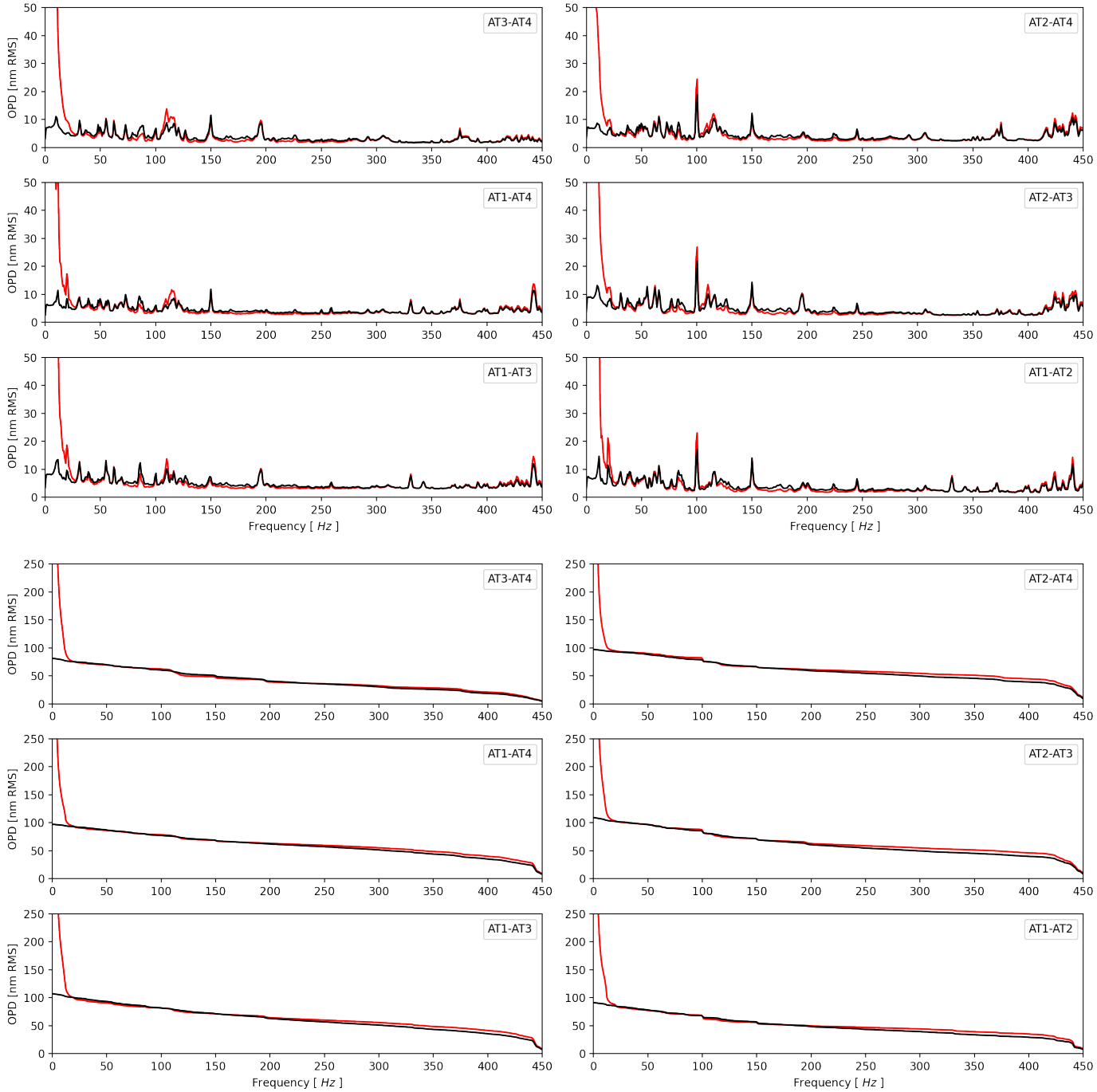

Fig. B.2. Power spectrum density and cumulative sum of the PSD of phase residuals ($\mathbf{\Phi}_n$) for target HR 7672. The back curves represent the close-loop values. The red curves represent the pseudo-open-loop values, signifying the scenario without a fringe tracker ($H_\phi \cdot L$). The six lower plots show the reverse cumulative sum of the power spectrum. Over 100 s of a single scientific DIT, the residual standard deviation is 71 nm (UT3-UT4), 86 nm (AT2-AT4), 86 nm (AT1-AT4), 93 nm (AT2-AT3), 93 nm (AT1-AT3), and 80 nm (AT1-AT2).