OPEN ACCESS
IOP Publishing
Classical and Quantum Gravity
Class. Quantum Grav. **39** (2022) 195016 (20pp) https://doi.org/10.1088/1361-6382/ac8a88

Wavefront error based tilt-to-length noise analysis for the LISA transmitted beam

Alexander Joseph Weaver*, Guido Mueller and Paul J Fulda

Department of Physics, University of Florida, 2001 Museum Road, PO Box 118440, Gainesville, FL 32611-8440, United States of America

E-mail: aweaver7@ufl.edu

Received 23 March 2022, revised 19 July 2022
Accepted for publication 17 August 2022
Published 8 September 2022

Abstract

The laser interferometer space antenna (LISA) will open the signal-rich 100 μHz to 1 Hz gravitational wave window. LISA is expected to be limited by acceleration noise in the low frequency range and noise associated with the optical measurement system above a few mHz. Of the latter, apparent length changes due to spacecraft (SC) angular jitter are among the most critical contributors. One of the coupling mechanisms is via wavefront error in the transmitted beam. Utilizing a Zernike polynomial decomposition of such wavefront error, we introduce and explore the validity of extremely fast best fit polynomial expansion based noise recreation tools that provide a clear picture for which transmit beam perturbations couple most strongly with SC jitter into LISA noise.

Keywords: LISA, Zernike, WFE, interferometry

(Some figures may appear in colour only in the online journal)

1. Introduction

The laser interferometer space antenna (LISA) will be the first space-based laser interferometric gravitational wave observatory [1]. It complements current and future ground-based observatories such as LIGO [2, 3, 6, 7], VIRGO [4–7], KAGRA [6–8], Einstein telescope [9, 10], and Cosmic Explorer [11, 12]. LISA will listen in the 100 μHz to 1 Hz frequency range to signals from a very diverse range of sources, from massive black hole binaries with

*Author to whom any correspondence should be addressed.

constituent masses from $\mathcal{O}(10^7)$–$\mathcal{O}(10^8)$ solar masses at the high mass end, to thousands of potentially resolvable galactic white dwarf binaries at the low mass end [1, 14].

LISA consists of three spacecraft (SC) in a near equilateral triangular configuration with a 2.5 Gm baseline. Laser beams will be exchanged between the SC to measure (in-band) length changes at nearly the pm $\sqrt{\text{Hz}}^{-1}$ level [1, 21]. As in ground-based observatories [13], tilt-to-length (TTL) coupling might be one of the dominant noise contributors. The *tilt* in LISA is caused by angular SC motion a.k.a. SC jitter which become apparent length changes in readout due to optical or instrumental imperfections. Unwanted aberrations in the phase of transmitted beams, denoted wavefront errors (WFE), generate wavefront gradients at the receiving SC. As the sending SC jitters, this wavefront sweeps over the receiving telescope and changes the phase of the received field. This paper introduces fast, transparent modeling techniques for the purpose of estimating TTL-noise associated with various WFE, and discusses the accuracy of such techniques.

Our analysis of TTL noise induced by WFE begins in section 2 as a first order expansion in SC jitter components coupling to the received field's phase gradient, yielding two TTL coupling components to two jitter degrees of freedom. This first order expansion in jitter allows for actual recovery of expected TTL noise as a function of receiving SC location, in contrast to methodologies of prior studies [18, 19] where the TTL is defined more as a statistic over the entire map. We discuss how gradients are obtained via a Hermite–Gaussian (HG) modal decomposition based propagation of an approximation in our initial beam, and the limitations of this approximation. Restricting to a cone of the far field (FF) out to 50 nrad, section 3 details how TTL expressions in terms of the field's phase gradient can be written accurately as first or second order polynomials in amplitudes of the Zernike polynomials contributing to WFE. Having verified the accuracy of these expansions, the maps associated with polynomial terms show only some Zernikes within WFE, specified in section 4, couple significantly to TTL noise. We can then neglect a significant number of unimportant terms in the expansions, further speeding up this fast modeling technique without a loss of accuracy, allowing for more extensive simulations than those of prior studies [18, 19].

2. Mechanism of wavefront error based tilt-to-length noise within LISA

The three SC of the LISA constellation form the arms of a huge interferometer of side roughly 2.5 Gm in length. The laser beams exchanged within these arms are expected to be nearly Gaussian, with spherical wavefronts centered on the test masses (TM) about which any SC jitter will occur. As jitter of the transmit SC sweeps the spherical wavefront across the receiving SC the phase of the received light, and thus measured arm length, should remain constant.

WFE produced by imperfections of real optics within the telescope generate deviations from a spherical wavefront at the receiving SC. Now a varying phase sweeps across the receiving SC whenever the transmit SC jitters. The change in received phase induces TTL noise in length measurements (figure 1).

2.1. Tilt to length calculations

As the transmit SC jitters by angle J, the transmitted field remains the same in the transmit frame while the receiving SC appears to moves over an arc of the FF of length $J * L$ where L is the 2.5 Gm inter-SC distance. We can recover the length change to first order by taking the

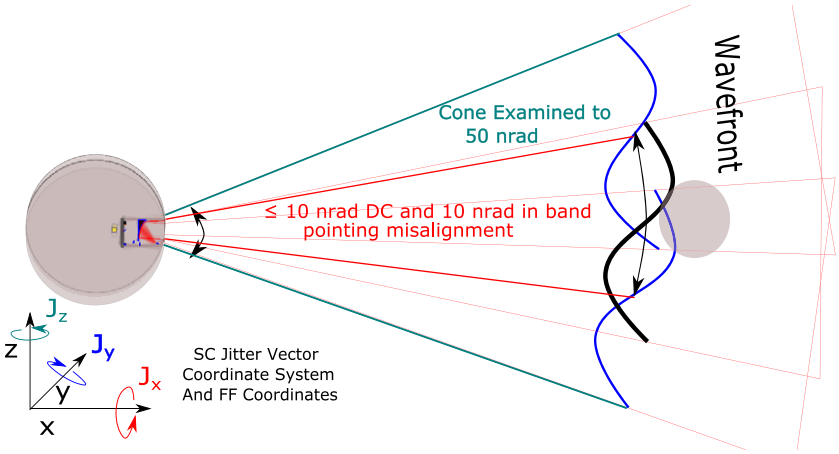

Figure 1. Transmitting SC rotation shifting the wavefront off the sphere of at the receiving SC, thus inducing TTL noise. We show here the cone examined our simulations.

derivative of the FF phase at the receiving SC with respect to the introduced jitter:

$$\Delta l = J\frac{\partial \Phi_{\mathrm{FF}}}{\partial J} \tag{1}$$

where J dependence comes through the coordinates of the receiving SC in the transmitting SC frame. Writing rotational jitter of the transmit SC by angle J about vector $\hat{j}$ as $\vec{J} = J\hat{j}$, we show in appendix A that by decomposing this jitter vector into components (J_x, J_y, J_z) the change in length signal can be written:

$$\Delta l \longrightarrow \overbrace{\frac{-\lambda}{2\pi}\frac{\partial \Phi_{\mathrm{FF}}}{\partial \theta}}^{\mathrm{TTL}_\theta} J_y + \underbrace{\frac{-\lambda}{2\pi}\frac{\partial \Phi_{\mathrm{FF}}}{\partial \phi}}_{\mathrm{TTL}_\phi} J_z. \tag{2}$$

This gives the TTL noise induced by WFE associated with transmit SC jitter[4]. This decomposition, with coordinate system definitions, can be found in appendix A. The effective roll term J_x enters only to second order in the combination of DC pointing offset and jitter components. The mission DFACS system constrains DC pointing offset to 10 nrad and in band misalignment jitter J to 10 nrad $\sqrt{\mathrm{Hz}}^{-1}$ [1, 21], so that this roll term would be suppressed by an order of 10^{-8} relative to other terms above.

The calculation of these components in the FF given some initial WFE written as a sum of Zernike polynomials (orthogonal polynomials over the unit disk ideal for representing phase aberrations over an aperture [15]), is performed via Hermite Gauss (HG) modal decomposition [16]. Appendix B more thoroughly explains how this, and the Fresnel diffraction integrals of 2.2, are utilized in TTL calculations.

[4] Technically this is only a part of the TTL noise associated with the transmit SC jitter, the portion due to a net piston effect on the received beam phase. Tilts introduced to the received beam through this transmit SC WFE and jitter may also contribute to TTL noise.

2.2. A note on validity of TTL calculations

As mentioned, TTL estimates of this paper are all initially extracted from fields found with an HG modal decomposition based propagation scheme [16]. These utilize an expansion for the initial outgoing beam:

$$U_0\, \mathrm{e}^{\mathrm{i}\Phi_{\mathrm{WFE}}} \longrightarrow U_0(1 + \mathrm{i}\Phi_{\mathrm{WFE}}) \tag{3}$$

with initial field U_0 the clipped Gaussian transmit beam and Φ_{WFE} the phase of the additional WFE. We originally expected this to be accurate for LISA because of the 35 nm RMS WFE requirements [20–22]. It turns out that peak to valley phase variations, even in these WFE, can be large enough to significantly reduce the accuracy of FF recreations based on a first order expansion in Φ_{WFE}.

To asses for what WFE such simulations remained accurate, we examined the error in TTL recreation as a function of the order to which the initial $\mathrm{e}^{\mathrm{i}\Phi_{\mathrm{WFE}}}$ was Taylor expanded. Numerical integration was used to evaluate TTL expressions based on Fresnel diffraction of the Taylor expanded beams[5] (detailed in appendix B). We looked at WFE composed of the first 36 Zernike terms in Noll indexing. TTL recreation error was then examined over a FF grid of potential receiving SC locations with points every 12.5 m in the y–z directions of figure 1 and extending out to angular offsets of 50 nrad (or transversely 125 m). For LISA our concern is TTL noise at or above the picometer level. With up to 10 nrad $\sqrt{\mathrm{Hz}}^{-1}$ of in band SC jitter allowed by mission constraints, the error in recovered TTL coefficients must be below 0.1 pm nrad^{-1} for sub-pm level error in noise simulations.

The error introduced in TTL estimates by expanding the initial phase to various order was looked at for 650 different WFE, using a uniform random number generator for the Z_4 through Z_{36} amplitudes in each map[6]. Figure 2 illustrates that first order expansion based propagation can be inaccurate for relatively small WFE. For general WFE with RMS heights as low as 14 nm the error in TTL coefficients recreated via our HG modal based propagation could be larger than 0.1 pm nrad^{-1}. It is important to note that by expanding the initial $\mathrm{e}^{\mathrm{i}\Phi_{\mathrm{WFE}}}$ to only second order, such simulations appeared accurate for WFE up to the LISA requirement.

Thus techniques of this paper should only be applied to analysis of sufficiently small WFE. More specifically, this can be used for LISA WFE based FF TTL studies involving WFE with RMS heights up to 10 nrad, but is not sufficient for the full range of potential LISA WFE.

[5] Fresnel diffraction and HG modal propagation of an initial beam are equivalent when utilizing an infinite number of HG modes [23]. Although often more accurate than the finite mode representation in an actual simulations, numerical Fresnel diffraction is often significantly slower for all practical purposes, especially when recreating the FF at a significant number of points.

[6] The 35 nrad requirement corresponds to Z_4 and higher Zernikes. Z_1 or piston is a constant phase offset which is taken out by the differential phase measurements employed in LISA, while Z_2–Z_3 correspond to beam tilts i.e. part of LISA's 10 nrad pointing accuracy requirement. This is accounted for in our simulation by allowing for an angular offset of the receiving SC from beam center, although we've allowed for offsets up to 50 nrad which are significantly larger than the LISA requirement.

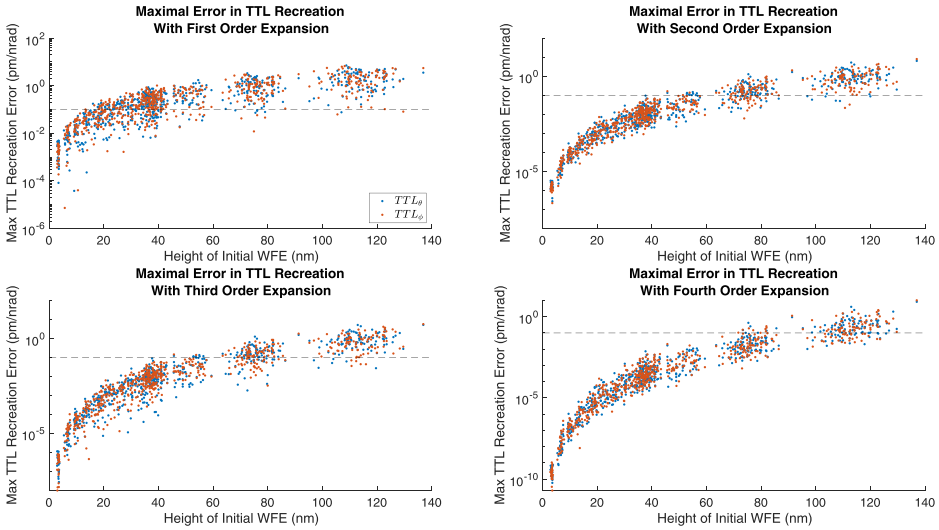

Figure 2. Agreement of various order expansions of the WFE term in recreating FF TTL couplings for general WFE as a function of the WFE RMS height. The first order outgoing beam expansion utilized in this paper appears accurate for WFE RMS heights as large as 14 nm.

3. Far field TTL components: linear and quadratic Zernike contribution expansions

To speed up simulations for large scale WFE analysis, we examined whether it was possible at relevant FF point to accurately represent TTL components of equation (2) as a linear:

$$\mathrm{TTL}_{\theta/\phi,\,\mathrm{FULL}} \approx \mathrm{TTL}_{\theta/\phi,\,1\mathrm{st.}} = \mathrm{TTL}_{\theta/\phi,0} + \sum_m c_m A_{\theta/\phi,m}, \tag{4}$$

or second order[7]:

$$\mathrm{TTL}_{\theta/\phi,\,2\mathrm{nd.}} = \mathrm{TTL}_{\theta/\phi,0} + \sum_m c_m \left[A_{\theta/\phi,m} + \sum_{n \geqslant m} c_n B_{\theta/\phi,m,n} \right], \tag{5}$$

expansions in Zernike contributions to initial transmit WFE:

$$\Phi_{\mathrm{WFE}} = \sum_m c_m Z_m. \tag{6}$$

Here Z_m refer to normalized Zernike polynomials written with Noll indexing [15]. Although beam propagation is done with a linearized WFE contribution to the initial field, discussed in section 2.2, this does not guarantee obtained TTL can be accurately written as a linear expansion in WFE Zernike amplitudes c_m. That depends on linearizability of the complex argument

[7] Note the existence of the $B_{\theta/\phi,m,n}$ terms mean the optimal $A_{\theta/\phi,m}$ obtained via least squares for equation (5) can differ from the $A_{\theta/\phi,m}$ minimizing the squared TTL recreation error using equation (4).

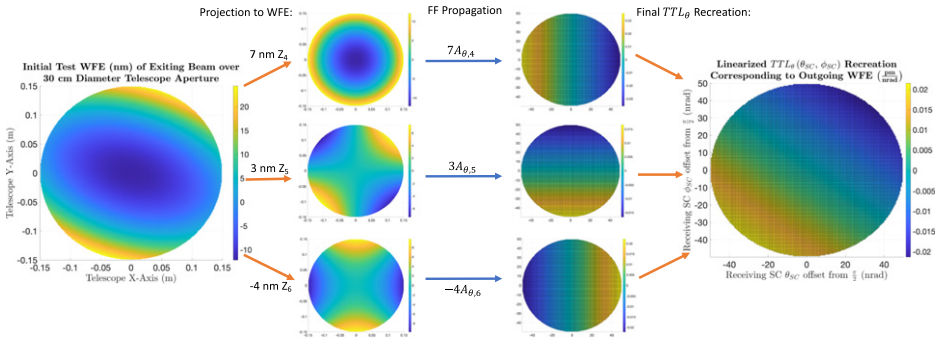

Figure 3. Example of how the linearized recovery of TTL_θ works given an initial WFE composed of only Z_4–Z_6. As mentioned maps are generated at a grid of points occurring every 1.25 m along transverse dimensions over the FF extending out to 50 nrad. We've explicitly included grid point boundaries in the FF maps (images to the right).

or phase function about the phase at each point in the FF. To avoid confusion with the terminology of section 2.2, for the remainder of this paper first and second order refers to the expansion order of TTL in terms of Zernike contributions to WFE, i.e. whether results come from using equation (4) or equation (5). It does not refer to the order $e^{i\Phi_{\mathrm{WFE}}}$ has been expanded to in our initial beams, which is first order.

To obtain the coefficient maps $\mathrm{TTL}_{\theta/\phi,0}$, $A_{\theta/\phi,m}$, and $B_{\theta/\phi,m,n}$, we first generate TTL components as described in appendix B for some generating set of initial WFE. This is done for each FF point on a grid with 1.25 m transverse resolution extending out to 50 nrad (125 m) to account for DC pointing misalignment (10 times higher resolution than the simulations of section 2.2, this grid is visualized in figure 3). The coefficient maps at each FF point are then found by minimizing the sum of all generated WFE of squared errors of equation (4) or equation (5) at that point. Appendix C more fully describes this least-squares (LS) minimization process, as well as our choice for the generating set of WFE. A visual of the linearized TTL_θ recovery associated with a single WFE realization is provided by figure 3.

We tested for 1123 randomly generated 10 nm RMS WFE[8] the linear and second order TTL approximations against actual TTL obtained from FF propagation based on HG modal decomposition [16, 17]. These simulations are binned in histograms of figure 4 by greatest error in the magnitude of TTL_θ components recovered (ϕ component errors are nearly identical) over the FF grid of receiving SC positions. As an example of how this would be read, in the top histogram the bin at -1 fm nrad^{-1} extending to nearly 0.04 tells us that for about 44 test maps the greatest error produced by linear TTL_θ recreation occurred when it underestimated the magnitude of the actual TTL_θ by between 0.95 and 1.05 fm nrad^{-1} at some possible receiving SC location. The top two histograms detail linear recreation performance while the bottom two detail the second order recreation performance. Both first and second order recreations yield order of magnitude improvement in recreation time over the original HG modal based propagation methods. More so, errors in even the linear recreation were below 0.1 pm nrad^{-1} for any WFE that we could utilize our HG modal propagation to investigate.

[8] Each Zernike coefficient in a map for Z_4–Z_{36} was generated from a uniform distributions, with the total maps scaled to an RMS height of 10 nm.

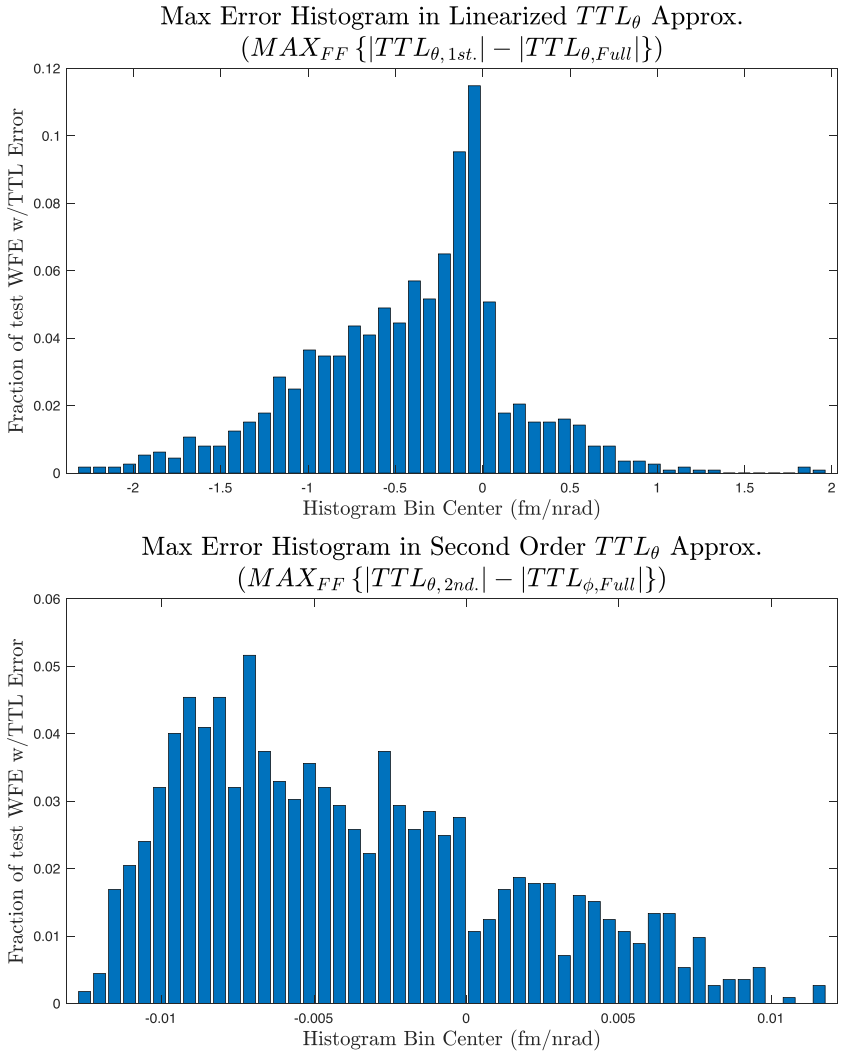

Figure 4. Histogram organized by greatest errors (from full HG mode based propagation simulations) in TTL_θ component recreation based on linear and second order TTL approximations (TTL_ϕ matches this well). We see both linear and second order expansions produce errors below 0.1 pm nrad^{-1} error for 10 nm RMS WFE and thus should be accurate enough for simulations with any such WFE.

4. TTL contributions of individual or coupled Zernikes

With the polynomial in Zernike amplitude for TTL component recreation, we can easily identify which Zernike terms generate problematic TTL, potentially guiding the polishing process for LISA telescope mirrors. We can also see which contributions can be completely neglected in order to speed up computation. We consider only simulations involving WFE up to an RMS height of 10 nm, guaranteeing accuracy of results.

For a WFE with RMS height h, the maximal individual Zernike amplitude $|c_m|$ of equation (6) is also h (all WFE comes from Zernike Z_m). This yields a value of h^2 for the

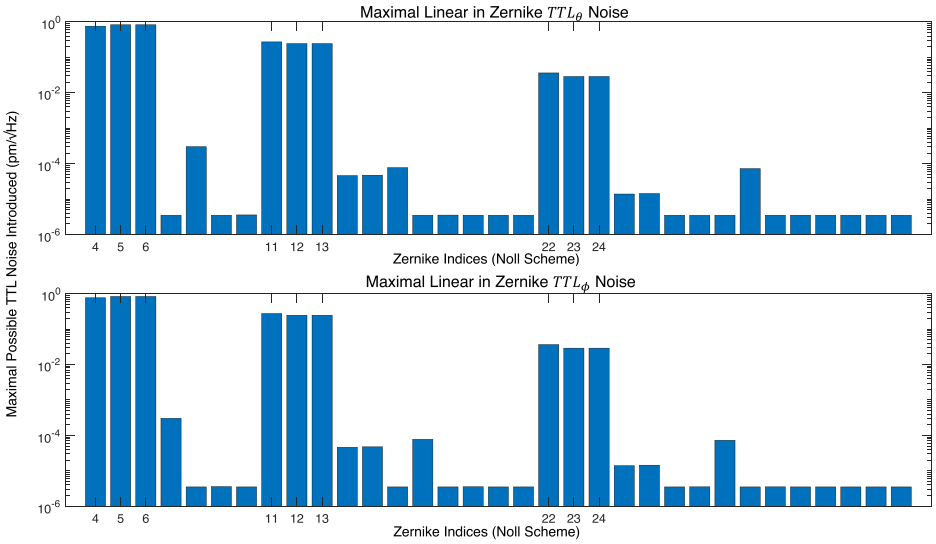

Figure 5. Maximal contribution of linear couplings within TTL expressed as a polynomial to second order in Zernike WFE contributions over a 50 nrad FF cone. Given expected WFE and jitter tolerances, only Zernike polynomials Z_4–Z_6, Z_{11}–Z_{13}, and Z_{22}–Z_{24} have the potential to contribute to TTL through these.

maximal c_m^2 amplitude, while the maximal value of $|c_m c_n|$ can be shown to be $\frac{h^2}{2}$ for $m \neq n$ (i.e. when all WFE comes from two equally contributing Zernike polynomials). Thus we can write the maximal contribution to TTL possible from individual maps $A_{\theta/\phi,m}$ and $B_{\theta/\phi,m,n}$ for any WFE having RMS heights up to 10 nm as:

$$\mathrm{TTL}_{\max}(A_{\theta/\phi,m}) = (10\ \mathrm{nmWFE})\mathrm{MAX}\left\{A_{\theta/\phi,m}\left(\frac{\mathrm{pm\ nrad}^{-1}}{\mathrm{nm\ WFE}}\right)\middle|50\ \mathrm{nrad\,FF}\right\} \tag{7}$$

$$\mathrm{TTL}_{\max}\ (B_{\theta/\phi,m,n}) = \left(\frac{10^2}{2-\delta_{m,n}}\ \mathrm{nm^2WFE}\right)\mathrm{MAX}\left\{B_{\theta/\phi,m,n}\left(\frac{\mathrm{pm\ nrad}^{-1}}{\mathrm{nm^2\ WFE}}\right)\middle|50\ \mathrm{nrad\,FF}\right\} \tag{8}$$

Figure 5 shows noise contributions are only significant from those $A_{\theta/\phi,m}$ maps associated with three tiers of Zernike polynomials, Z_4–Z_6 (first tier), Z_{11}–Z_{13} (second tier), and Z_{22}–Z_{24} (third tier). Figure 6 illustrates maximal contributions of the 595 $B_{\theta/\phi,m,n}$ maps to TTL based length noise generated by 10 nrad of SC jitter about the corresponding axis (i.e. $10 * \mathrm{TTL}_{\max}\left(B_{\theta/\phi,m,n}\right)$). Thus the error in TTL noise estimates by neglecting any such term is below 115 fm for any WFE we consider here. Additionally, the most significant second order contributions are those cross-couplings to azimuthally symmetric polynomials within these tiers (Z_4, Z_{11}, and Z_{22}). These couple into TTL_ϕ noise significantly through combination with Z_7, Z_{17}, and Z_{29}, while rotated versions of these (Z_8, Z_{16}, and Z_{30}) couple into TTL_θ, as would be expected. However both show any contribution in the regime considered depends most heavily on the nine Zernike polynomials contributing to the linear recreation terms.

The magnitude of both first and second order couplings decreases as we move from tier one to three, with the magnitude of noise decreasing by a factor of three to five with an increase in tier. As only second order cross couplings are significant, this corresponds directly

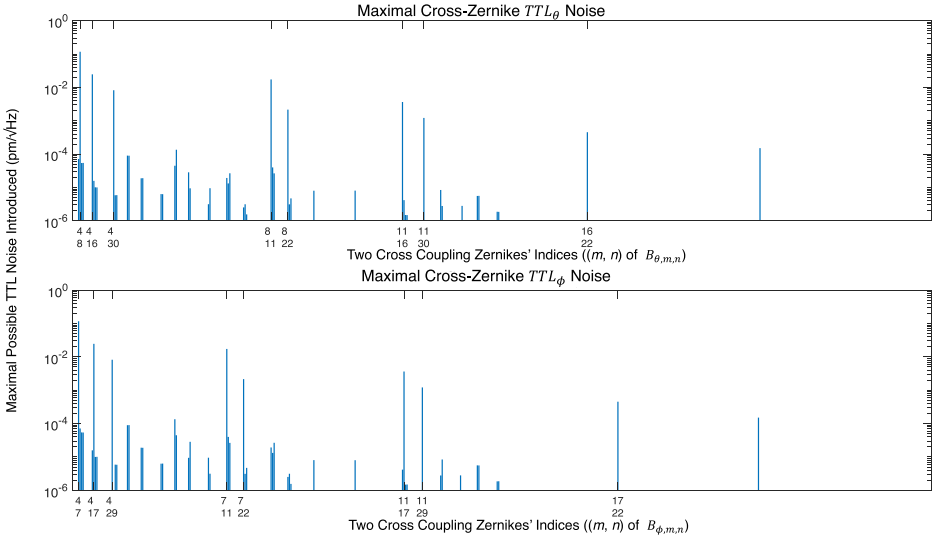

Figure 6. Maximal contribution of second order couplings within TTL expressed as a polynomial to second order in Zernike WFE contributions over a 50 nrad FF cone. For significant cross couplings we list both Zernikes in each pair that contribute. We see that for 10 nm WFE it appears utilizing only the first order expansion in Zernike coefficients introduce errors on the order of 100 fm $\sqrt{\text{Hz}}^{-1}$ in TTL noise estimates.

to relative allowable RMS contributions between different tiers. Thus in the regime of WFE considered the LISA mission can tolerate three to five times more second tier contribution to WFE than first, and similarly three to five times more third tier than second. It's important to note that although we are considering TTL generated by smaller WFE, and over a smaller FF region, than [18], their prediction of 0.02 pm nrad^{-1} of TTL given a $\lambda/40$ initial WFE RMS height seems reasonable when compared to our prediction of a maximal coupling of 0.005 pm nrad^{-1} of TTL generated per nanoradian of defocus and astigmatism Zernike polynomials.

We thus significantly speed up simulations by summing only the nine significant linear terms of figure 5:

$$\{Z_4, Z_5, Z_6, Z_{11}, Z_{12}, Z_{13}, Z_{22}, Z_{23}, Z_{24}\}, \tag{9}$$

for both ϕ and θ TTL components. Simple Cauchy–Schwartz based inequalities using the maximal values of coefficient maps can be used to show that TTL component estimates using only these terms recreation will differ by less than 40 fm nrad^{-1} from component estimates using the full 595 term second order summation, thus having a maximal error in predicted TTL noise below 0.6 pm. This is accurate enough for LISA simulations involving any WFE up to 10 nm in RMS height.

5. Conclusions

LISA wavefront error (WFE) based TTL noise may present a significant obstacle in meeting mission sensitivity requirements. We've shown our ability to write TTL coefficients accurately as a first order polynomial in Zernike contributions to the WFE. This is accurate for WFE up

to 10 nm in RMS height, providing an extremely fast TTL recreation scheme for which it is clear which Zernikes couple most heavily to TTL noise.

Initial simulations with WFE having large contributions from those Zernikes coupling heavily into TTL noise showed certain combinations of these Zernikes were capable of producing noise far below LISA requirements, indicating a degree of TTL cancellation between terms. Ongoing work aims to improve representations of our initial field to guarantee suitable recreation of TTL noise for even the largest WFE LISA telescopes may introduce. Currently we've been able to show second order expansions for TTL components in Zernike amplitude, now based on Fresnel diffraction of initial beams without simplification of the initial $e^{\Phi_{\text{WFE}}}$ term, are accurate for nearly any WFE LISA may have. Using these results we hope to examine which Zernike combinations maximally and minimally couple to TTL noise, exploiting such TTL cancellation. Such work may produce a more ideal basis for representing defects within LISA WFE. Simple projections to terms in this basis may then be sufficient to identify whether or not LISA TTL requirements can be satisfied without post-processing noise removal. The aim is to provide definitive metrics for mirror manufacturers to use in the development of telescope components to meet LISA requirements, and workarounds in case they do not.

Acknowledgment

This research was funded by NASA Grants 80NSSC20K0126 and 80NSSC19K0324.

Data availability statement

All data that support the findings of this study are included within the article (and any supplementary files).

Appendix A. Jitter-to-TTL coordinate system

We first begin with our transmitting SC outgoing beam reference system, defining the optical axis as oriented along the x axis of this frame which we initially align with a global frame for the purpose of working with rotations. These are defined via

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix}_T = x\,\hat{x}_T + y\,\hat{y}_T + z\,\hat{z}_T \tag{10}$$

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix}_G = x\,\hat{x}_G + y\,\hat{y}_G + z\,\hat{z}_G \tag{11}$$

where T as a subscript denotes coordinates in the transmit telescope frame, which can change with respect to our unchanging global coordinate frame specified by the G subscript. The beam center falls along the x axis in the transmit SC coordinate system, and falls on the receiving SC in the case of zero pointing alignment at $(L,0,0)_T = (L,0,0)_G$ with L the 2.5 Gm inter-SC distance. The receiving SC will be offset from the beam center by an angle θ_{DC} but still a distance L away, so that we can write this in general as:

$$\vec{r}_R = L \begin{pmatrix} \cos\,\theta_{\text{DC}} \\ \sin\,\theta_{\text{DC}}\,\cos\,\phi_{\text{DC}} \\ \sin\,\theta_{\text{DC}}\,\sin\,\phi_{\text{DC}} \end{pmatrix}_{T/G} \tag{12}$$

there $\theta_{\rm DC}$ is the pointing offset from beam center, limited to 10 nrad, while $\phi_{\rm DC}$ further specifies the orientation but can take on any value. Note these receiving SC coordinates will only change in the T basis.

Now we apply a jitter to our transmit SC as shown in figure 1:

$$\vec{J} = \begin{pmatrix} J_x \\ J_y \\ J_z \end{pmatrix}_G = J\,\hat{j} \tag{13}$$

corresponding again to a rotation of $\sqrt{J_x^2 + J_y^2 + J_z^2} \equiv J$ radians about the corresponding unit vector $\hat{j}$. For computational convenience we write out without loss of generality:

$$\begin{pmatrix} J_x \\ J_y \\ J_z \end{pmatrix}_G = J \begin{pmatrix} \sin\,\theta_j \cos\,\phi_j \\ \sin\,\theta_j \sin\,\phi_j \\ \cos\,\theta_j \end{pmatrix}_G = J\,\hat{j} \tag{14}$$

noting this is just temporary for moving forwards. Rotating a vector about this will keep the inner product between it and $\hat{j}$ constant, while the projection to the orthogonal plane will have polar coordinates (where we assume right-handed systems) advance by angle J.

We do this naturally by first choosing an orthogonal vector to $\hat{j}$, lets call it:

$$\hat{\gamma}^{\,\rm T} = (\sin\,\phi_j, -\cos\,\phi_j, 0)_G, \tag{15}$$

where we have just used the T superscript to denote transpose or row vector for writing in line. This $\hat{\gamma}$ will be equivalent to the x axis of our polar coordinate plane (with the y axis of this plane naturally defined by the cross product $\hat{j} \times \hat{\gamma}$, naturally guaranteeing the polar angle correspond to what would be expected in a right-handed). Denoting

$$\hat{\beta}^{\,\rm T} \equiv (\,\hat{j} \times \hat{\gamma})^{\rm T} = (\cos\,\theta_j \cos\,\phi_j, \cos\,\theta_j \sin\,\phi_j, -\sin\,\theta_j), \tag{16}$$

the polar coordinates of some vector $\vec{r}$ are then defined by

$$r_P \cos\,\phi_P = \vec{r} \cdot \hat{\gamma} \tag{17}$$

$$r_P \sin\,\phi_P = \vec{r} \cdot \hat{\beta} \tag{18}$$

where P denotes the polar coordinate component. Rotation about $\hat{j}$ corresponds to a shift in this polar angle, so for general vector $\vec{r}$ rotated about $\vec{J}$ to $\vec{r}\,'$ we must have

$$\begin{aligned} \vec{r}\,' \cdot \hat{\gamma} &= r_P \cos\,\phi'_P = r_P \cos(\phi_P + J) = r_P(\cos\,\phi_P \cos\,J - \sin\,\phi_P \sin\,J) \\ &= \vec{r} \cdot \left(\cos\,J\,\hat{\gamma} - \sin\,J\,\hat{\beta}\right) \end{aligned} \tag{19}$$

$$\begin{aligned} \vec{r}\,' \cdot \hat{\beta} &= r_P \sin\,\phi'_P = r_P \sin(\phi_P + J) = r_P(\sin\,\phi_P \cos\,J + \cos\,\phi_P \sin\,J) \\ &= \vec{r} \cdot \left(\cos\,J\,\hat{\beta} + \sin\,J\,\hat{\gamma}\right) \end{aligned} \tag{20}$$

so that

$$\vec{r}\,' = \hat{j}\left(\vec{r} \cdot \hat{j}\right) + \hat{\gamma}\left(\vec{r} \cdot \left(\cos\,J\,\hat{\gamma} - \sin\,J\,\hat{\beta}\right)\right) + \hat{\beta}\left(\vec{r} \cdot \left(\cos\,J\,\hat{\beta} + \sin\,J\,\hat{\gamma}\right)\right) \tag{21}$$

$$= \hat{j}(\vec{r}\cdot\hat{j}) + (\vec{r}\cdot\hat{\gamma})\Big(\cos J\,\hat{\gamma} + \sin J\,\hat{\beta}\Big) + \Big(\vec{r}\cdot\hat{\beta}\Big)\Big(\cos J\,\hat{\beta} - \sin J\,\hat{\gamma}\Big) = \hat{\mathbf{R}}(\cdot,\vec{r}) \tag{22}$$

where the matrix $\hat{\mathbf{R}}$ is defined by

$$\hat{j}\otimes\hat{j} + \hat{\gamma}\otimes\Big(\cos J\,\hat{\gamma} - \sin J\,\hat{\beta}\Big) + \hat{\beta}\otimes\Big(\sin J\,\hat{\gamma} + \cos J\,\hat{\beta}\Big) \tag{23}$$

This is exactly the coordinate transformation jitter induces to each basis vector of the transmit frame, with the components for the location of the receiving SC in the transmitting frame as a function of this rotation given by the dot product of $\vec{r}_R$ with each of these rotated basis vectors.

$$\begin{pmatrix} r_{R,T,x'} \\ r_{R,T,y'} \\ r_{R,T,z'} \end{pmatrix} = \begin{pmatrix} \hat{x}'_T\cdot\vec{r}_R \\ \hat{y}'_T\cdot\vec{r}_R \\ \hat{z}'_T\cdot\vec{r}_R \end{pmatrix} = \begin{pmatrix} \hat{\mathbf{R}}(\vec{r}_R,\hat{x}_T) \\ \hat{\mathbf{R}}(\vec{r}_R,\hat{y}_T) \\ \hat{\mathbf{R}}(\vec{r}_R,\hat{z}_T) \end{pmatrix} \tag{24}$$

$$= \begin{pmatrix} \hat{x}_T^T \\ \hat{y}_T^T \\ \hat{z}_T^T \end{pmatrix}\hat{\mathbf{R}}^T\vec{r}_R = \begin{pmatrix} \hat{x}_T\cdot\hat{j} & \hat{x}_T\cdot\hat{\gamma} & \hat{x}_T\cdot\hat{\beta} \\ \hat{y}_T\cdot\hat{j} & \hat{y}_T\cdot\hat{\gamma} & \hat{y}_T\cdot\hat{\beta} \\ \hat{z}_T\cdot\hat{j} & \hat{z}_T\cdot\hat{\gamma} & \hat{z}_T\cdot\hat{\beta} \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos J & \sin J \\ 0 & -\sin J & \cos J \end{pmatrix}\begin{pmatrix} \vec{r}_R\cdot\hat{j} \\ \vec{r}_R\cdot\hat{\gamma} \\ \vec{r}_R\cdot\hat{\beta} \end{pmatrix} \tag{25}$$

$$= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} \sin\theta_j\cos\phi_j & \sin\phi_j & \cos\theta_j\cos\phi_j \\ \sin\theta_j\sin\phi_j & -\cos\phi_j & \cos\theta_j\sin\phi_j \\ \cos\theta_j & 0 & -\sin\theta_j \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos J & \sin J \\ 0 & -\sin J & \cos J \end{pmatrix} \tag{26}$$

$$\times\begin{pmatrix} \sin\theta_j\cos\phi_j & \sin\theta_j\sin\phi_j & \cos\theta_j \\ \sin\phi_j & -\cos\phi_j & 0 \\ \cos\theta_j\cos\phi_j & \cos\theta_j\sin\phi_j & -\sin\theta_j \end{pmatrix}\begin{pmatrix} L\cos\theta_{\mathrm{DC}} \\ L\sin\theta_{\mathrm{DC}}\cos\phi_{\mathrm{DC}} \\ L\sin\theta_{\mathrm{DC}}\sin\phi_{\mathrm{DC}} \end{pmatrix} \tag{27}$$

This is still completely general. We get rid of the identity matrix at the start which arises because our transmit SC basis originally aligned with our global basis. We now write out this in terms of typical spherical coordinates, i.e. defining angular θ and ϕ by

$$\begin{pmatrix} \sin\theta(J)\cos\phi(J) \\ \sin\theta(J)\sin\phi(J) \\ \cos\theta(J) \end{pmatrix} \equiv \begin{pmatrix} \sin\theta_j\cos\phi_j & \sin\phi_j & \cos\theta_j\cos\phi_j \\ \sin\theta_j\sin\phi_j & -\cos\phi_j & \cos\theta_j\sin\phi_j \\ \cos\theta_j & 0 & -\sin\theta_j \end{pmatrix} \tag{28}$$

$$\times\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos J & \sin J \\ 0 & -\sin J & \cos J \end{pmatrix}\begin{pmatrix} \sin\theta_j\cos\phi_j & \sin\theta_j\sin\phi_j & \cos\theta_j \\ \sin\phi_j & -\cos\phi_j & 0 \\ \cos\theta_j\cos\phi_j & \cos\theta_j\sin\phi_j & -\sin\theta_j \end{pmatrix}$$

$$\times\begin{pmatrix} \cos\theta_{\mathrm{DC}} \\ \sin\theta_{\mathrm{DC}}\cos\phi_{\mathrm{DC}} \\ \sin\theta_{\mathrm{DC}}\sin\phi_{\mathrm{DC}} \end{pmatrix} \tag{29}$$

So that we can now write out the effective length change due to jitter via

$$\Delta l = \frac{J}{k}\frac{\partial \Phi_{\rm FF}}{\partial J}_{J=0} = \frac{1}{k}\left(J\frac{\partial\theta(J)}{\partial J}\frac{\partial\Phi_{\rm FF}}{\partial\theta} + J\frac{\partial\phi(J)}{\partial J}\frac{\partial\Phi_{\rm FF}}{\partial\phi}\right)_{J=0} \tag{30}$$

To calculate these derivatives we do so implicitly, taking the derivatives of the left and right hand sides above at $J = 0$ giving

$$\begin{pmatrix} J\cos\theta_0\cos\phi_0\dfrac{\partial\theta}{\partial J} - J\sin\theta_0\sin\phi_0\dfrac{\partial\phi}{\partial J} \\ J\cos\theta_0\sin\phi_0\dfrac{\partial\theta}{\partial J} + J\sin\theta_0\cos\phi_0\dfrac{\partial\phi}{\partial J} \\ -J\sin\theta_0\dfrac{\partial\theta}{\partial J} \end{pmatrix}$$

$$= J\begin{pmatrix} \sin\theta_j\cos\phi_j & \sin\phi_j & \cos\theta_j\cos\phi_j \\ \sin\theta_j\sin\phi_j & -\cos\phi_j & \cos\theta_j\sin\phi_j \\ \cos\theta_j & 0 & -\sin\theta_j \end{pmatrix}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix} \tag{31}$$

$$\times\begin{pmatrix} \sin\theta_j\cos\phi_j & \sin\theta_j\sin\phi_j & \cos\theta_j \\ \sin\phi_j & -\cos\phi_j & 0 \\ \cos\theta_j\cos\phi_j & \cos\theta_j\sin\phi_j & -\sin\theta_j \end{pmatrix}\begin{pmatrix} \cos\theta_{\rm DC} \\ \sin\theta_{\rm DC}\cos\phi_{\rm DC} \\ \sin\theta_{\rm DC}\sin\phi_{\rm DC} \end{pmatrix} \tag{32}$$

$$= \begin{pmatrix} 0 & J_z & -J_y \\ -J_z & 0 & J_x \\ J_y & -J_x & 0 \end{pmatrix}\begin{pmatrix} \cos\theta_{\rm DC} \\ \sin\theta_{\rm DC}\cos\phi_{\rm DC} \\ \sin\theta_{\rm DC}\sin\phi_{\rm DC} \end{pmatrix} \tag{33}$$

where we have rewritten in terms of our original definitions for the jitter vector. Multiplying both sides by a common matrix to isolate terms we get

$$\begin{pmatrix} J\dfrac{\partial\theta}{\partial J} \\ J\dfrac{\partial\phi}{\partial J} \end{pmatrix} = \begin{pmatrix} \cos\theta_0\cos\phi_0 & \cos\theta_0\sin\phi_0 & -\sin\theta_0 \\ \dfrac{-\sin\phi_0}{\sin\theta_0} & \dfrac{\cos\phi_0}{\sin\theta_0} & 0 \end{pmatrix}\begin{pmatrix} 0 & J_z & -J_y \\ -J_z & 0 & J_x \\ J_y & -J_x & 0 \end{pmatrix}$$

$$\times\begin{pmatrix} \cos\theta_{\rm DC} \\ \sin\theta_{\rm DC}\cos\phi_{\rm DC} \\ \sin\theta_{\rm DC}\sin\phi_{\rm DC} \end{pmatrix} \tag{34}$$

Now these expressions are first order in the jitter variable, on the order of 10 nanoradians, while $\theta_{\rm DC}$ is of the same order of magnitude, so that any first order term in $\theta_{\rm DC}$ is suppressed by O(10^{-8}) relative to zeroth order terms. Thus only including the zeroth order $\theta_{\rm DC}$ expansion, evaluating the θ_0 and ϕ_0 here at $J = 0$ now gives

$$\sin\theta_0\cos\phi_0 = \cos\theta_{\rm DC} \longrightarrow 1 \tag{35}$$

$$\sin\theta_0\sin\phi_0 = \sin\theta_{\rm DC}\cos\phi_{\rm DC} \longrightarrow 0 \tag{36}$$

$$\cos\theta_0 = \sin\theta_{\rm DC}\sin\phi_{\rm DC} \longrightarrow 0 \tag{37}$$

and our expression becomes

$$J\begin{pmatrix}\frac{\partial\theta}{\partial J}\\ \frac{\partial\phi}{\partial J}\end{pmatrix} = \begin{pmatrix}0 & 0 & -1\\ 0 & 1 & 0\end{pmatrix}\begin{pmatrix}0 & J_z & -J_y\\ -J_z & 0 & J_x\\ J_y & -J_x & 0\end{pmatrix}\begin{pmatrix}1\\0\\0\end{pmatrix} \tag{38}$$

$$= \begin{pmatrix}-J_y\\ -J_z\end{pmatrix} \tag{39}$$

yielding the final expression:

$$\Delta l = \frac{-\lambda}{2\pi}\left(J_y\frac{\partial\Phi_{\mathrm{FF}}}{\partial\theta} + J_z\frac{\partial\Phi_{\mathrm{FF}}}{\partial\phi}\right) \tag{40}$$

Appendix B. Far field phase and phase derivative derivations

B.1. HG modal decomposition based propagation and TTL extraction using it

The majority of beam recreations in this paper are based on the HG modal decomposition based propagation described in [16]. This exploits that for sufficiently paraxial light, such as the LISA transmit beams we are concerned with, the projection to each HG mode is unchanging (for more on HG modes and the parameters they depend on such as waist size and waist size location, see [16, 17, 23]). Thus if the amplitude of the initial outgoing LISA beam in the $(0,0)$ or fundamental Gaussian HG mode in some basis is $a_{0,0}$, the amplitude of the FF in the propagated form of this mode is still $a_{0,0}$. This allows us to write the transmitted field

$$E(x,y,z) = \mathrm{e}^{-\mathrm{i}k(z-z_0)}\sum_{m,n} a_{m,n}\mathrm{HG}_{m,n}(x,y,z;w_0,z_b) \tag{41}$$

where

$$a_{m,n} = \iint_S \mathrm{d}x\mathrm{d}y E(x,y,z_0)\mathrm{HG}_{m,n}(x,y,z_0;w_0,z_b)^*. \tag{42}$$

Here S is taken to be the exit aperture of the transmit telescope at plane $z = z_0$, with outgoing field there $E(x,y,z_0)$. $\mathrm{HG}_{m,n}(x,y,z;w_0,z_b)$ is HG mode (m,n) at position (x,y,z) taken from HG mode basis of proper wavelength having waist size w_0 at location z_b along the optical axis. [16] gives an analytic formula for such $a_{m,n}$ obtained for fields of the form $\mathrm{HG}_{m,n}(x,y,z;w_0,z_b)\tilde{Z}_n^m\left(\frac{x}{R},\frac{y}{R}\right)$ passing through a circular aperture of radius R. $\tilde{Z}_n^m\left(\frac{x}{R},\frac{y}{R}\right)$ here are the Zernike polynomials[9] of angular index m and radial index n scaled for orthogonality over the disk of radius R, exactly what will be used to decompose the WFE[10]. The initial un-aberrated LISA beam is an $\mathrm{HG}_{0,0}$ mode, and thus for Φ_{WFE} decomposed as in equation (6) we have the outgoing beam over the telescope exit aperture of

[9] Use of $\tilde{Z}$ as opposed to simply Z distinguishes we have written the Zernike polynomial in terms of Cartesian coordinates and not the traditional polar coordinates these are typically functions of.

[10] Zernike polynomials labeled by both upper and lower indices are traditionally defined Zernikes with upper angular index and lower radial index. As mentioned in earlier sections, Zernike polynomials labeled with only a single subscript refer to the corresponding Zernike polynomial in the Noll indexing scheme. For example, Z_0^0 and Z_1 refer to the same polynomial.

radius R:

$$\mathrm{HG}_{0,0}(x,y,0;w_{\mathrm{LISA}},0)\mathrm{e}^{\mathrm{i}\sum_m c_m \bar{Z}_m\left(\frac{x}{R},\frac{y}{R}\right)} \longrightarrow \mathrm{HG}_{0,0}(x,y,0;w_{\mathrm{LISA}},0) \times \left(Z_1\left(\frac{x}{R},\frac{y}{R}\right)+\sum_m c_m Z_m\left(\frac{x}{R},\frac{y}{R}\right)\right) \tag{43}$$

when expanded to first order in the WFE phase (w_{LISA} is taken as the expected LISA Gaussian transmit beam waist size). Thus each $a_{m,n}$ used in writing the total field when expanding to first order in Φ_{WFE} is simply a sum of terms obtained from [16].

Utilizing an HG modal basis (choice of w_0) optimized to represent the transmit beam with no WFE present in terms of as few modes as possible with waist at the center or rotation i.e. the TM center or origin of our coordinate system, we obtain $a_{m,n}$ for the transmit beam with WFE as described above yielding beam at the receiving SC having angular coordinates θ/ϕ from appendix A:

$$E(\theta,\phi)=\mathrm{e}^{-\mathrm{i}kL\ \sin\ \theta\ \sin\ \phi}\sum_{m,n} a_{m,n}\mathrm{HG}_{m,n}(L\ \cos\ \theta, L\ \sin\ \theta\ \cos\ \phi,\ L\ \sin\ \theta\ \sin\ \phi; w_0, 0) \tag{44}$$

The phase can then be written using the complex argument function with some choice of branch Arg as:

$$\Phi_{\mathrm{FF}}(\theta,\phi)=\arctan\left(\frac{2L\ \sin\ \theta\ \sin\ \phi}{kw_0^2}\right) - \frac{kL\ \sin\ \theta\ \sin\ \phi\left(1+\sin^2\theta\sin^2\phi+\frac{k^2w_0^2}{2L^2}\right)}{2\sin^2\theta\sin^2\phi+\frac{k^2w_0^2}{2L^2}}+\mathrm{Arg}[E_{\mathrm{res}}] \tag{45}$$

Some care should technically be taken to write the terms outside of the complex argument function Arg as modulo 2π plus some factor to guarantee it is consistent with the branch of complex argument function used for Arg, however since relevant quantities for TTL are derivatives with respect to θ and ϕ, this actually won't matter. The terms before the complex argument function are phase elements common to every HG mode of this particular basis (i.e. background Gouy phase, radius of curvature phase, and $\mathrm{e}^{-\mathrm{i}kz}$ phase, all better detailed in [16, 17, 23]). They contain no information about effects of WFE. The remaining portion E_{res} contains information unique to each mode:

$$\mathrm{Arg}[E_{\mathrm{res}}]=\mathrm{Arg}\left[\sum_{m,n}\frac{a_{m,n}\mathrm{e}^{\mathrm{i}(m+n)\arctan\left(\frac{2L\ \sin\ \theta\ \sin\ \phi}{kw_0^2}\right)}}{\sqrt{2^{m+n}\,m!n!}} \times H_m\left(\frac{L\cos\theta}{w_0\sqrt{1+\left(\frac{2L\ \sin\ \theta\ \sin\ \phi}{kw_0^2}\right)^2}}\right)H_n\left(\frac{L\ \sin\ \theta\ \cos\ \phi}{w_0\sqrt{1+\left(\frac{2L\ \sin\ \theta\ \sin\ \phi}{kw_0^2}\right)^2}}\right)\right] \tag{46}$$

Here $H_j(t)$ is the jth physicist's Hermite polynomial evaluated at t, and all information about the initial WFE is encoded in the $a_{m,n}$ terms.

Evaluation of TTL elements requires derivatives of equation (45) with respect to angular variables θ and ϕ. While we have implemented programs based on a purely analytical expressions for this, we've found it faster to implement algorithms that only utilize analytical expressions for the angular derivatives of terms in equation (45) outside the complex argument function. Then we add to these numerically approximated angular derivatives of the numerically evaluated $\mathrm{Arg}[E_{\mathrm{res}}]$ function at appropriate FF points.

B.2. Fresnel diffraction integrals for field and TTL calculations

The TTL expressions obtained via the HG modal decomposition described above are based on a first order expanding $\mathrm{e}^{\mathrm{i}\Phi_{\mathrm{WFE}}}$ in the outgoing field. Had we not required the expansion of this initial phase, TTL obtained via an infinite modal expansion would be exact for any sufficiently paraxial light (which is certainly the case for the LISA beam at the receiving SC, which has been transmitted 2.5 Gm, for more on paraxial light and the paraxial approximation see [17, 23]).

To estimate the introduced error in TTL estimates, we need another way to obtain the propagated field without requiring this simplification of our initial field. We utilize the Fresnel diffraction integral for a propagated field:

$$E(x,y,z) = \frac{\mathrm{i}\mathrm{e}^{-\mathrm{i}k\Delta z}}{\lambda\Delta z}\iint_S \mathrm{d}x_0\mathrm{d}y_0\left[E_T(x_0,y_0)\mathrm{e}^{\frac{-\mathrm{i}k\left((x-x_0)^2+(y-y_0)^2\right)}{2\Delta z}}\right] \tag{47}$$

where here $E_T(x_0,y_0)$ is the field in the plane of the transmitting telescope exit aperture and Δz is the propagation distance along the optical axis from this aperture to the point of interest, ideally the location of the receiving SC. This simplification of the full wave-equation based Kirchoff diffraction integral produces a paraxial light equivalent. One can show [23] that any paraxial propagation scheme is equivalent, i.e. this yields the same exact field we would get with an infinite HG modal decomposition scheme if the initial field utilized was also $E_T(x_0,y_0)$. The key difference here is that by evaluating this integral numerically we need not expand the initial fields WFE dependence.

To evaluate accuracy of an order N expansion in initial phase Φ_{WFE} of the initial field, we then evaluate the difference in TTL estimates found with the simplified initial field:

$$\begin{aligned} E_T(x_0,y_0) &= \mathrm{HG}_{0,0}(x_0,y_0,0;w_{\mathrm{LISA}},0)\sum_{p=0}^{N}\frac{(\mathrm{i}\Phi_{\mathrm{WFE}}(x_0,y_0))^p}{p!} \\ &\equiv \mathrm{HG}_{0,0}(x_0,y_0,0;w_{\mathrm{LISA}},0)F_N(x_0,y_0), \end{aligned} \tag{48}$$

from those found with the actual initial field

$$\begin{aligned} E_T(x_0,y_0) &= \mathrm{HG}_{0,0}(x_0,y_0,0;w_{\mathrm{LISA}},0)\mathrm{e}^{\mathrm{i}\Phi_{\mathrm{WFE}}(x_0,y_0)} \\ &\equiv \mathrm{HG}_{0,0}(x_0,y_0,0;w_{\mathrm{LISA}},0)F_\infty(x_0,y_0), \end{aligned} \tag{49}$$

where F_N is implicitly defined above with $N = \infty$ the unexpanded case. To evaluate these at the proper FF points and obtain TTL expressions, we first rewrite the Fresnel diffraction integral in terms of the proper angular coordinates of appendix A:

$$E(\theta,\phi) = \frac{\mathrm{i}\mathrm{e}^{-\mathrm{i}kL\ \sin\ \theta\ \sin\ \phi}}{\lambda L\ \sin\ \theta\ \sin\ \phi}\iint_S \mathrm{d}x_0\,\mathrm{d}y_0\left[E_T(x_0,y_0)\mathrm{e}^{\frac{-\mathrm{i}k\left(\left(L\ \cos\ \theta - x_0\right)^2+\left(L\ \sin\ \theta\ \cos\ \phi - y_0\right)^2\right)}{2L\ \sin\ \theta\ \sin\ \phi}}\right]. \quad (50)$$

However, at this point numerical evaluation would yield unusable phase information for taking angular derivatives because of the massive $kL\sin\theta\sin\phi$ phase element outside of the integral. In addition, this must be evaluated numerical at each point we want the field. If calculating phase derivatives numerically, this means multiple evaluation points FF location we hope to find the TTL, bogging down computations. To improve computational speed, we instead separated as much phase as possible out of the integrals, then took angular derivatives analytically, yielding derivatives within the Fresnel diffraction integral, and then evaluated these expressions numerically. We first simplify the expressions slightly by expanding out the exponent terms

$$E(\theta,\phi) = \frac{\mathrm{i}\mathrm{e}^{\frac{-\mathrm{i}kL\left(1+\sin^2\theta\sin^2\phi\right)}{2\ \sin\ \theta\ \sin\ \phi}}}{\lambda L\ \sin\ \theta\ \sin\ \phi}\iint_S \mathrm{d}x_0\,\mathrm{d}y_0\left[E_T(x_0,y_0)\mathrm{e}^{\frac{-\mathrm{i}k\left(\left(x_0^2+y_0^2\right)-2L\left(x_0\ \cos\ \theta+y_0\ \sin\ \theta\ \cos\ \phi\right)\right)}{2L\ \sin\ \theta\ \sin\ \phi}}\right] \quad (51)$$

$$= \frac{\mathrm{i}\sqrt{2}\mathrm{e}^{\frac{-\mathrm{i}kL\left(1+\sin^2\theta\sin^2\phi\right)}{2\ \sin\ \theta\ \sin\ \phi}}}{\sqrt{\pi}w_{\mathrm{LISA}}\lambda L\ \sin\ \theta\ \sin\ \phi}\int_0^R \mathrm{d}\rho\left[\rho\mathrm{e}^{-\frac{\rho^2}{w_{\mathrm{LISA}}^2}-\frac{\mathrm{i}k\rho^2}{2L\ \sin\ \theta\ \sin\ \phi}}\int_0^{2\pi}\mathrm{d}\gamma\right.$$
$$\left.\times\left[F_N(\rho\ \cos\ \gamma,\rho\ \sin\ \gamma)\mathrm{e}^{\frac{\mathrm{i}kL\rho(\cos\ \gamma\ \cos\ \theta+\sin\ \gamma\ \sin\ \theta\ \cos\ \phi)}{\sin\ \theta\ \sin\ \phi}}\right]\right] \quad (52)$$

where we have plugged in explicitly the form of the unaberrated $\mathrm{HG}_{0,0}$ mode for LISA as well as explicitly written integral bounds to the telescope exit aperture radius R (γ was used as the polar angle to avoid confusion with FF angular variables θ/ϕ). Again, here we use $N = \infty$ for the actual TTL to compare finite N values to. This allows us to ignore unnecessary constants and write the FF phase as

$$\Phi_{\mathrm{FF}}(\theta,\phi) = \frac{\pi}{2} - \frac{kL\left(1+\sin^2\theta\sin^2\phi\right)}{2\ \sin\ \theta\ \sin\ \phi} + \mathrm{Arg}[D_N(\theta,\phi)] \quad (53)$$

where

$$D_N(\theta,\phi) = \int_0^R \mathrm{d}\rho\left[\rho\mathrm{e}^{-\frac{\rho^2}{w_{\mathrm{LISA}}^2}-\frac{\mathrm{i}k\rho^2}{2L\ \sin\ \theta\ \sin\ \phi}}\int_0^{2\pi}\mathrm{d}\gamma\right.$$
$$\left.\times\left[F_N(\rho\ \cos\ \gamma,\rho\ \sin\ \gamma)\mathrm{e}^{\frac{\mathrm{i}kL\rho(\cos\ \gamma\ \cos\ \theta+\sin\ \gamma\ \sin\ \theta\ \cos\ \phi)}{\sin\ \theta\ \sin\ \phi}}\right]\right] \quad (54)$$

where as in the last section, the terms outside the argument function carry information independent of the WFE, while WFE information is carried explicitly in the complex argument of the D_N function. Using that

$$\frac{\partial}{\partial \mathrm{Arg}[f(x)]}x = \frac{\partial \frac{\ln(f(x))-\ln(f^*(x))}{2i}}{\partial x} = \frac{\frac{\partial f}{\partial x}}{2if(x)} - \frac{\frac{\partial f^*}{\partial x}}{2if^*(x)} = \mathrm{Im}\left[\frac{\frac{\partial f}{\partial x}}{f}\right] \tag{55}$$

at any point where $f(x) \neq 0$, where Im is the imaginary part of a complex number, we then wrote

$$\frac{\partial \Phi_{\mathrm{FF}}(\theta,\phi)}{\partial\theta/\phi} = -\frac{kL}{2}\frac{\partial}{\partial\theta/\phi}\left[\frac{1+\sin^2\theta\sin^2\phi}{\sin\,\theta\,\sin\,\phi}\right] + \mathrm{Im}\left[\frac{\frac{\partial D_N(\theta,\phi)}{\partial\theta/\phi}}{D_N(\theta,\phi)}\right]. \tag{56}$$

In evaluating this the first term was differentiated purely analytically. For the second term, the integrand in $D_N(\theta,\phi)$ was first differentiated analytically within the integral (the portion dependent on θ/ϕ is independent of WFE and thus will remain the same for all WFE), and the resulting expression numerically integrated using Matlab's 2D integration function.

Appendix C. Coefficient map recovery

A general LS based method minimizes the error of a linear combination of basis functions in approximating a known set of data. For example, if we have some function we wish to approximate F, known only at a set of N data point vectors $\{\vec{x}_i\}_{i=1}^N$, with a set of known functions $\{f_j\}_{j=1}^K$, the least squared solution is the set of coefficients $\{a_j\}_{j=1}^K$ such that the expression

$$F_{\mathrm{approx}}(\vec{x}) = \sum_{j=1}^{K} a_j F_j(\vec{x}) \tag{57}$$

minimizes the sum of squared errors from known data values $F(\vec{x}_i)$ over all known values of F. This squared error becomes

$$E^2 = \sum_{i=1}^{N}\left(F(\vec{x}_i) - \sum_{j=1}^{K} a_j F_j(\vec{x}_i)\right)^2 \tag{58}$$

Minimizing this by setting derivatives with respect to $\{a_i\}$ to zero yields the set of linear equations

$$\langle F_i F\rangle = \sum_{j=1}^{K}\langle F_i F_j\rangle a_j \tag{59}$$

Where this is the traditional expectation value over the known data set i.e. for function $g(\vec{x})$:

$$\langle g\rangle = \frac{1}{N}\sum_{i=1}^{N} g(\vec{x}_i) \tag{60}$$

Defining matrix $\hat{\mathbf{M}}$ by $(\hat{\mathbf{M}})_{ij} \equiv \langle F_i F_j\rangle$, if this matrix is invertible (which depends on the values of the basis functions over the data set) we can then write out minimizing choice of $\{a_i\}$:

$$a_i = \sum_{j=1}^{K} (\hat{\mathbf{M}}^{-1})_{ij}\langle F_j F\rangle \tag{61}$$

If we have more basis functions than known data points, we see this becomes degenerate as we will necessarily have $F_m = \sum_{j\neq m} c_j F_j$ for some m at all known points $\{\vec{x}_i\}_{i=1}^{N}$. More specifically we can show the functions must be linearly independent over the chosen data set for this to be successful.

In our case of approximating of TTL components, our F function is now the corresponding TTL component value at a specific point in the FF. Thus we use this for each relevant FF point we wish to approximate TTL at, meaning the a_j are now maps of values at each relevant FF point. For our situation $\vec{x}_i = \vec{\text{WFE}}_i = (c_4, c_5, \ldots, c_{36})_i$ where here the c_m are the heights in nm of the mth contributing Zernike in the ith generating WFE used. The a_j we are finding are the coefficient maps such as $\text{TTL}_{\theta/\phi,0}$. As mentioned above, we need at least as many data points as we have representative functions to approximate the TTL. These include the constant function (with corresponding a_i the $\text{TTL}_{\theta/\phi,0}$ map) and the linear projection functions $\Pi_m\left(\vec{\text{WFE}}\right) = c_m$ for $4 \leqslant m \leqslant 36$ the mth Zernike amplitude (with corresponding a_i the $A_{\theta/\phi,m}$ map), giving a total of 34 representative functions (33 projections plus constant) in our linear approximations. The second order case has additional representative functions of the form $\Pi_m\Pi_n$, so that we have an additional $33 + \frac{33*32}{2} = 561$ terms (33 quadratic and 528 cross). This means for a first order LS based method we need at least 34 generating WFE while for second order we require 595 generating WFE. Just having this number is not enough to guarantee non-degenerate $\hat{\mathbf{M}}$ matrices. Utilizing more may guarantee a better representative expression over a larger range of WFE, however if restricting to only linear or second order regimes additional generating WFE will contribute little to changing optimal coefficient map values.

For example, in linear recreations we utilized a generating WFE set composed of:

$$\vec{\text{WFE}}_1 = \vec{0} \tag{62}$$

$$\Pi_m\left(\vec{\text{WFE}}_{1+n}\right) = 2\delta_{m,n+3} \tag{63}$$

so that besides the first $\vec{0}$ WFE vector the $n+1$ generating vector has an RMS height of 2 nm contributed by Z_{n+3}, i.e. $\vec{\text{WFE}}_2$ is just 2 nm of Z_4 or defocus. This set gave nearly identical coefficient maps to a set found by using several hundred randomly generated WFE all having RMS heights below 10 nm.

For second order recreation, we added additional WFE having contributions of -2 nm from each individual Zernike, as well a single WFE for each pair of Zernike terms having both having heights of 2 nm.

ORCID iDs

Alexander Joseph Weaver https://orcid.org/0000-0001-5439-6811

References

[1] Danzmann K *et al* 2017 Laser interferometer space antenna *A Proposal in Response to the ESA Call for L3 Missions* (LISA Consortium)
[2] The LIGO Scientific Collaboration *et al* 2015 Advanced LIGO *Class. Quantum Grav.* **32** 074001
[3] Moore C J, Cole R H and Berry C P L 2014 Gravitational-wave sensitivity curves *Class. Quantum Grav.* **32** 015014
[4] Accadia T *et al* 2012 Virgo: a laser interferometer to detect gravitational waves *J. Instrum.* **7** P03012
[5] Acernese F *et al* 2015 Advanced Virgo: a second-generation interferometric gravitational wave detector *Class. Quantum Grav.* **32** 024001
[6] Gregory H M 2012 Second generation gravitational wave detectors *Marcel Grossman Meeting on General Relativity The 12th Marcel Grossmann Meeting* pp 628–44
[7] Abbott B P *et al* 2020 Prospects for observing and localizing gravitational-wave transients with Advanced LIGO, Advanced Virgo and KAGRA *Living Rev. Relativ.* **23** 3
[8] Akutsu T *et al* 2021 Overview of KAGRA: detector design and construction history *Prog. Theor. Exp. Phys.* **2021** 05A101
[9] Hild S *et al* 2011 Sensitivity studies for third-generation gravitational wave observatories *Class. Quantum Grav.* **28** 094013
[10] Maggiore M *et al* 2020 Science case for the Einstein telescope *J. Cosmol. Astropart. Phys.* JCAP03(2020)050
[11] Abbott B P *et al* 2017 Exploring the sensitivity of next generation gravitational wave detectors *Class. Quantum Grav.* **34** 044001
[12] Dwyer S *et al* 2015 Gravitational wave detector with cosmological reach *Phys. Rev.* D **91** 082001
[13] Mueller G 2005 Beam jitter coupling in advanced LIGO *Opt. Express* **13** 7118–32
[14] LISA Mission Science Office 2007 LISA: probing the universe with gravitational waves LISA-LIST-RP-436
[15] Noll R J 1976 Zernike polynomials and atmospheric turbulence *J. Opt. Soc. Am.* **66** 207–11
[16] Weaver A, Fulda P and Mueller G 2020 Analytic HG-mode propagation through circular apertures with Zernike phase offset *OSA Continuum* **3** 1891–916
[17] Siegman A E 1986 *Lasers* (Mill Valley, CA: University Science Books) ch 16 pp 626–48
[18] Sasso C P, Mana G and Mottini S 2018 Coupling of wavefront errors and jitter in the LISA interferometer: far-field propagation *Class. Quantum Grav.* **35** 185013
[19] Kenny F and Devaney N 2020 Beam propagation simulations for LISA in the presence of telescope aberrations *Class. Quantum Grav.* **38** 035010
[20] Faulks Henry *et al* 2000 *LISA Study of the Laser Interferometer Space Antenna Final Technical Report* LISA final technical report https://lisa.nasa.gov/archive2011/Documentation/FTR.pdf 7–3
[21] Hasselmann N *et al* 2021 LISA optical metrology: tilt-to-pathlength coupling effects on the picometer scale *Int. Conf. on Space Optics 2020* vol 11852 (International Society for Optics and Photonics) pp 1734–44
[22] Caldwell M, McNamara P and Glennmar A 1998 Optical engineering requirements for the LISA wavefront error budget *AIP Conf. Proc.* **456** 156–64
[23] Weaver A 2021 Investigating limits on gravitational wave detection by laser interferometry using Hermite–Gauss mode representations of paraxial light propagation *PhD Dissertation* (University of Florida)