

Bayesian time delay interferometry for orbiting LISA: Accounting for the time dependence of spacecraft separations

Jessica Page 

*Space Science Department, University of Alabama in Huntsville,
320 Sparkman Drive, Huntsville, Alabama 35899, USA*

Tyson B. Littenberg 

NASA Marshall Space Flight Center, Huntsville, Alabama 35812, USA



(Received 24 May 2023; accepted 10 August 2023; published 28 August 2023)

Previous work demonstrated effective laser frequency noise (LFN) suppression for Laser Interferometer Space Antenna (LISA) data from raw phasemeter measurements using a Markov Chain Monte Carlo (MCMC) algorithm with fractional delay interpolation (FDI) techniques to estimate the spacecraft separation parameters required for time-delay interferometry (TDI) under the assumption of a rigidly rotating LISA configuration. Including TDI parameters in the LISA data model as part of a global fit analysis pipeline enables gravitational wave inferences to be marginalized over uncertainty in the spacecraft separations. Here we extend the algorithm’s capability to perform data-driven TDI on LISA in Keplerian orbits, which introduce a time-dependence in the arm-length parameters and at least $\mathcal{O}(M)$ times greater computational cost since the filter must be applied for every sample in the time series of sample size M . We find feasibility of arm-length estimation on \sim day-long timescales by using a novel Taylor-expanded version of the fractional delay interpolation filter that allows half of the filter computation to be calculated and stored before MCMC iterations and requires shorter filter lengths than previously reported. We demonstrate LFN suppression for orbiting LISA using accurate arm-length estimates parameterized by Keplerian orbital parameters under the assumption of unperturbed analytical Keplerian orbits, and explore the potential extension of these methods to arbitrary numerical orbits.

DOI: [10.1103/PhysRevD.108.044065](https://doi.org/10.1103/PhysRevD.108.044065)

I. INTRODUCTION

Astrophysical sources of gravitational waves (GW) in the 0.1 mHz–0.1 Hz band will be detected through the space-based LISA (Laser Interferometer Space Antenna) mission set to launch in the mid 2030s. LISA combines three spacecraft separated by a mean inter-spacecraft distance of $\sim 2.5 \times 10^6$ km to measure the pathlength change of the free-falling test masses induced by GWs using laser interferometry [1].

Because LISA is in a rotating heliocentric orbit, the interspacecraft distances will differ and vary in time, leaving a fractional frequency fluctuation amplitude of $\frac{\Delta\nu}{\nu} = 10^{-13}$ Hz $^{-1/2}$ remaining in the data. Typical GW signals are found near 10^{-21} Hz $^{-1/2}$, requiring laser frequency noise (LFN) suppression. Time-delay interferometry (TDI) is a well-studied algebraic operation to suppress LFN adequately below other secondary noises such as the test-mass acceleration and optical metrology system noises for GW signal analysis [2]. TDI works by linearly combining the LISA interferometric measurements delayed appropriately by the light travel time between spacecraft to synthesize a digital equal-arm interferometer.

The LISA mission will use pseudorandom noise phase modulation (PRM) encoded in the measurements to obtain the spacecraft separations L (i.e., “arm-lengths,” delays L/c) needed in the TDI combinations.

There are alternative data-driven methods that provide a back-up method to PRM that suppress LFN from the raw LISA measurements such as [3–6]. This paper is an extension of [5] to estimate the time-varying delays needed for second-generation TDI that accounts for the heliocentric motion of the LISA array using a Markov chain Monte Carlo (MCMC) algorithm. The use of MCMC methods to estimate arm length posteriors allows one to marginalize over their uncertainty when computing the LISA response function to GWs and opens the possibility to integrate TDI into the LISA “global fit” as demonstrated in [7]. High signal to noise GW sources were found to require accurate orbital models that account for the arm length time-dependence when computing the GW response function [8], indicating the importance of quantifying the spacecraft separation uncertainty.

Bayesian TDI builds upon time-delay interferometric ranging (TDIR) which estimates the arm lengths from raw interferometer measurements by minimizing the power of

the residual LFN [6] which relies on fractional delay interpolation (FDI) [9] to estimate the required nanosecond precision delays that occur between samples telemetered to Earth. Our previous work used an MCMC to estimate the delays in a rigidly rotating LISA scenario where the clockwise and counterclockwise light travel times differ but remained time-independent. We found that a shorter filter length can be used by limiting the maximum frequency in the likelihood function to reduce data loss in the presence of data gaps. This work extends the ideas of Bayesian data-driven TDI to assess feasibility in the realistic case of LISA heliocentric motion. We review second-generation TDI and describe the relevant LISA measurement details in Sec. II. We introduce a new highly efficient formulation of the Lagrange FDI filter in Sec. III B and examine the computational optimization that would be crucial for Bayesian TDI methods. We explore the MCMC in Sec. III, and a way to parameterize the arm length time dependence in the case of LISA on Keplerian orbits in Sec. IV A. Finally we demonstrate a promising method to estimate arm lengths in the realistic case of arbitrary non-Keplerian orbits that account for all other forces other than solar gravitational motion in Sec. IV B.

II. LISA MEASUREMENT DESCRIPTION AND SECOND-GENERATION TDI

We use the LISA Simulation software [10] for the data simulation to account for the heliocentric motion of the rotating array and therefore follow the LISA labeling conventions used there. The arm length labels in reference to each spacecraft are demonstrated in Fig. 1. The meaning of the first and second indices of a measurement has essentially been alternated from the notation in [5], and we again follow the split interferometry design described in [11]. The science measurement s_{ij} contains the GW signal and is the measurement exchanged between two

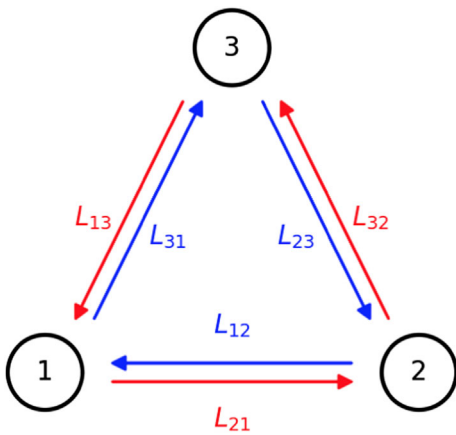


FIG. 1. Arm length labeling convention in relation to spacecraft number. Figure does not reflect that each spacecraft contains two test masses and two optical measurement systems.

separate spacecraft; s_{ij} is received on spacecraft i from j . The test mass measurement ε_{ij} is the motion of the test mass on spacecraft i that is adjacent to spacecraft j , and the reference measurement on τ_{ij} interferes the light between the two optical benches on a given spacecraft and the i indice is denoted as the receiving. The distance between two spacecraft ($L_{ij}(t)$, i.e. “arm lengths,” “delays”) will differ depending on the direction the laser light is traveling due to the Sagnac effect [$L_{ij}(t) \neq L_{ji}(t)$]. The i indice in the arm length denotes the receiving spacecraft and the j indice denotes the emitting spacecraft. We ignore all secondary noises and GW in the simulated data (Sec. III C) for this demonstration but we show the laser phase noise p_{ij} on the optical bench on spacecraft i that is adjacent to spacecraft j that enters in the three interferometric measurements given in Eqs. (1) and (2) and include our assumption of the test-mass acceleration noise n_{ij}^{TM} and optical metrology system noise n_{21}^{OMS} placement. Cyclic permutation of both indices for both sets of Eqs. (1) and (2) yield the remaining 12 measurements.

$$\begin{aligned} s_{21} &= p_{12,21} - p_{21} + n_{21}^{\text{OMS}} \\ \tau_{21} &= p_{23} - p_{21} \\ \varepsilon_{21} &= p_{23} - p_{21} - 2n_{21}^{\text{TM}} \end{aligned} \quad (1)$$

$$\begin{aligned} s_{23} &= p_{32,23} - p_{23} + n_{23}^{\text{OMS}} \\ \tau_{23} &= p_{21} - p_{23} \\ \varepsilon_{23} &= p_{21} - p_{23} + 2n_{23}^{\text{TM}} \end{aligned} \quad (2)$$

A. TDI 2.0 for LISA in heliocentric orbit

The main purpose of this work is to demonstrate feasibility of Bayesian-estimated arm-lengths for the realistic orbiting LISA scenario. The heliocentric motion of the three spacecraft array introduces a time dependence on $L_{ij}(t)$ with $\dot{L}_{ij} \sim 10$ m/s which leaves fractional frequency fluctuations present on $\mathcal{O}(10)$ times above the secondary noises in the GW signal-rich portion of the LISA band [2,12] if the “first generation” combinations in the previous Bayesian TDI work [5] are used. The “second-generation” (i.e., TDI 2.0) combinations average out the time-dependent effects by linearly approximating each light path ($L_{ij}(t) \approx L_{ij} + \dot{L}_{ij}t$) and synthesizing two equal light paths that now strongly depend on the order that each one-way beam is applied [13] which now requires non-commuting delay operators \mathcal{D}_{ij} in the second-generation combinations. LFN is suppressed sufficiently to first order in spacecraft distance velocities \dot{L}_{ij} . We continue use of the common 16-link Michelson combinations [Eq. (6)] denoted with a “2” subscript [$X_2(t)$, $Y_2(t)$, $Z_2(t)$] since they are second-generation combinations. The $;\mathcal{D}_{ij}\mathcal{D}_{mn}$ notation

indicates the time delay operator $\mathcal{D}_{mn}\mathcal{D}_{ij}$ that acts on a function $f(t)$ as given in Eq. (3), and similarly for more delays; the operator closest to a measurement is applied first and subsequent delays act on the measurement and previously-applied delays. The use of a semicolon indicates the delays do not commute to keep standards in the literature. The first order expansion to the arm length rate of change is shown in Eq. (5), and the comma $_{, \mathcal{D}_{ij}\mathcal{D}_{mn}}$ notation denotes constant valued estimates of commuting operators.

$$\mathcal{D}_{mn}\mathcal{D}_{ij}f(t) = f(t)_{, \mathcal{D}_{ij}\mathcal{D}_{mn}} \quad (3)$$

$$= f\left(t - \frac{L_{ij}(t - L_{mn}(t)/c)}{c} - \frac{L_{mn}(t)}{c}\right) \quad (4)$$

$$\approx f_{, \mathcal{D}_{ij}\mathcal{D}_{mn}} + \dot{f}_{, \mathcal{D}_{ij}\mathcal{D}_{mn}} \dot{L}_{ij}L_{mn} \quad (5)$$

Cyclic permutation of both indices of Eq. (6) yield the $Y_2(t)$ and $Z_2(t)$ combinations for the light paths of the other two spacecraft.

$$\begin{aligned} X_2(t) &= [\mathcal{D}_{12}\mathcal{D}_{21}\mathcal{D}_{13}\mathcal{D}_{31} - \mathcal{I}] \\ &\quad \times [(\eta_{13} + \eta_{31; \mathcal{D}_{13}}) + (\eta_{12} + \eta_{21; \mathcal{D}_{12}})_{; \mathcal{D}_{31}\mathcal{D}_{13}}] \\ &\quad - [\mathcal{D}_{13}\mathcal{D}_{31}\mathcal{D}_{12}\mathcal{D}_{21} - \mathcal{I}] \\ &\quad \times [(\eta_{12} + \eta_{21; \mathcal{D}_{12}}) + (\eta_{13} + \eta_{31; \mathcal{D}_{13}})_{; \mathcal{D}_{21}\mathcal{D}_{12}}] \end{aligned} \quad (6)$$

The intermediary variables η_{ij} in Eq. (6) are combinations of the interferometric measurements that remove 3 of the 6 LFN terms, essentially making the LFN the same on both optical benches on a given spacecraft. Cyclic permutation of both indices twice for both equations of Eq. (7) provide the 6 intermediary variables.

$$\begin{aligned} \eta_{12} &= s_{12} + \frac{\tau_{12} - \varepsilon_{12}}{2} + D_{12} \frac{(\tau_{21} - \varepsilon_{21})}{2} - D_{12} \frac{\tau_{21} - \tau_{23}}{2} \\ \eta_{13} &= s_{13} + \frac{\tau_{13} - \varepsilon_{13}}{2} + D_{13} \frac{(\tau_{31} - \varepsilon_{31})}{2} + \frac{\tau_{12} - \tau_{13}}{2} \end{aligned} \quad (7)$$

III. BAYESIAN ESTIMATION OF TIME-DEPENDENT SPACECRAFT SEPARATIONS

A. MCMC

The MCMC algorithm utilizes Bayes' theorem to estimate the posterior distribution $p(\vec{\theta}|\vec{\mathbf{d}})$ of the delay parameters $\vec{\theta} = \vec{\mathcal{D}}_{ij}(t)$ given the data $\vec{\mathbf{d}} = \{\tilde{X}_2(f), \tilde{Y}_2(f), \tilde{Z}_2(f)\}$ which are the Fourier-transformed TDI combinations, and various parameterizations for the $L_{ij}(t)$ time dependence

are described in Sec. IV A. The FDI filter described in Sec. III B is applied at each iteration of the chain to precisely determine the measurement time series delayed by the proposed arm-length. The prior distribution is uniform in all data models tested, but the parameters and parameter ranges vary depending on the data as described in Sec. IV.

The log likelihood function neglecting constant normalization factors is given in Eq. (8).

$$\ln p(\vec{\mathbf{d}}|\vec{\theta}) = \sum_{i=f_{\min}}^{f_{\max}} \left[-\ln(|\mathbf{C}|)_i - \left(\sum_{j,k}^{X,Y,Z} \mathbf{d}_j^\dagger \mathbf{C}_{jk}^{-1} \mathbf{d}_k \right)_i \right], \quad (8)$$

where $f_{\min} = 10^{-4}$ Hz and f_{\max} is restricted to 0.03 Hz instead of 1 Hz which is typically the assumed maximum of the LISA band. Interpolation error increases dramatically at frequencies near the 1 Hz maximum and we previously found in [5] that a 0.1 Hz restriction (just below the transfer frequency $f_* = c/L$) yields better parameter accuracy using a lower sampling rate and significantly shorter filter length. We restrict the likelihood calculation to below 0.03 Hz to reduce the computational cost of time-varying FDI filters and examine LFN suppression to the most relevant parts of the LISA band. Assuming that LFN is white noise in the LISA band $10^{-4} - 1$ Hz with fractional frequency power spectral density (PSD) $S_y^{\text{LFN}} = 10^{-26}$ Hz $^{-1}$, contributions to the likelihood integral above 0.03 Hz are negligible.

The noise covariance matrix $\mathbf{C}(f)$ describes the TM and OMS noises that remain after TDI application in the \mathbf{d}_j^\dagger and \mathbf{d}_k terms. We assume equal and constant arm lengths in the calculation of $\mathbf{C}(f)$ and use the same $\mathbf{C}(f)$ that is applied to the rigidly rotating LISA scenario so that all diagonal elements are given by Eq. (9),

$$\begin{aligned} C_{jj} &= 16 \sin^2(2\pi fL) S_y^{\text{OMS}} + (8 \sin^2(4\pi fL) \\ &\quad + 32 \sin^2(2\pi fL)) S_y^{\text{OMS}} \end{aligned} \quad (9)$$

and all off-diagonal elements are given by Eq. (10)

$$C_{jk} = (4S_y^{\text{TM}} + S_y^{\text{OMS}})(-4 \sin(2\pi fL) \sin(4\pi fL)) \quad (10)$$

since we are focused on arm length estimation in the case of data containing LFN-only. We found no effect on arm length estimation when the delay parameters were included in the unequal-arm noise covariance matrix in [5], but we could similarly derive the second-generation noise covariance matrix elements in future studies that include TM and OMS noises in the data.

We assume the TM and OMS noise PSDs converted to fractional frequency are as given in [1] shown in Eq. (11).

$$\begin{aligned}
 S_y^{\text{TM}}(f) &= (2\pi fc)^{-2} (3 \times 10^{-15})^2 \left(1 + \left(\frac{4 \times 10^{-4}}{f} \right)^2 \right) \left(1 + \left(\frac{f}{8 \times 10^{-3}} \right)^4 \right) \text{Hz}^{-1} \\
 S_y^{\text{OMS}}(f) &= \left(\frac{2\pi f}{c} \right)^2 (1.5 \times 10^{-11})^2 \left(1 + \left(\frac{2 \times 10^{-3}}{f} \right)^4 \right) \text{Hz}^{-1}.
 \end{aligned} \tag{11}$$

B. Optimized Lagrange fractional delay interpolation filter

This work extends upon the ideas of Bayesian data-driven TDI [i.e., time-delay interferometric ranging (TDIR)] [5] to include time-varying delays. TDIR was first introduced in [6] and relies heavily on the ideas of [9] which showed that by interpolating between integer samples to land on delay estimates at nanosecond precision and minimizing the power of the residual LFN, one can perform TDI from the raw interferometric measurements and obtain LFN suppression sufficiently below the secondary noises. We previously used the same Lagrange-windowed fractional delay interpolation (FDI) filter as shown in [9] and adjusted the filter length requirements for the sampling rate and cutoff frequency for the data simulated in [5]. Now that the delays are time-varying, re-use of the same Lagrange filter would require a new N th order filter for every M th sample where M is the number of samples in the data. The previously used Lagrange filter is too computationally costly for updating the filter each sample when working with stochastic sampling algorithms such as an MCMC, so we adapted the Lagrange-windowed sinc filter based off of [14] to remove half of the computation out of every MCMC iteration which dramatically reduces total cost and is described as follows. Equation (12) is the FDI process for a constant delay that is applied to every sample of the time series.

$$\begin{aligned}
 f(n-D) &= f(n) * w_L(D) \text{sinc}(n-D) \\
 &= \sum_{k=-(N-1)/2}^{(N-1)/2} f(n+k) w_L(k, D) \text{sinc}(D-k) \tag{12}
 \end{aligned}$$

The Lagrange window is given in Eq. (7) of [9]. We used the convolution theorem and computed the Fourier transforms of the LISA measurements before beginning MCMC iterations which significantly reduced the computation time in the constant delay case. Now, the overall delay applied to a measurement is a function of time so the interpolation filter is different for each sample in the time series [Eq. (13)].

$$f(n - D_n) = f(n) * w_L(D_n) \text{sinc}(n - D_n) \tag{13}$$

We implement a Taylor series expansion of the delayed signal based on [14] in terms of the difference operator

$$\Delta f[n] = f[n] - f[n-1],$$

powers of the LISA measurements $f[n]$ and factorial polynomials of the delays

$$D^{[k]} = D(D+1)\dots(D+k-1)$$

to result in Eq. (14).

$$f(n - D_n) = \sum_{k=0}^N \Delta^{[k]} f[n] \frac{(-D_n)^{[k]}}{k!} \tag{14}$$

The $\Delta^{[k]} f[n]$ powers of the measurements are computed and stored for a given N th order filter before running the MCMC iterations and $\Delta^{[k]} f[n] = \Delta \Delta^{[k-1]} f[n]$. Paper [14] quotes $3N - 2$ additions and $3N - 1$ multiplications for an N th order filter per output sample. Storing the $\Delta^{[k]} f[n]$ factors of Eq. (14) beforehand further reduces the new filter's additions to N per output sample. This results in approximately 30 times greater computational speed improvement for one MCMC iteration on one hour of data. Following standard guidelines set by the LISA Consortium for interpolating filters, the filter is shifted so that the n th sample to delay is centered in the middle of the filter and both even and odd filter lengths can be used.

This new Lagrange filter formulation also requires shorter filter lengths than the standard Lagrange-windowed sinc or Lagrange polynomials. The previous Bayesian TDIR work that used Eq. (12) found that the minimum filter length N required was ≈ 23 for a sampling rate $f_s = 4$ Hz and a likelihood function maximum frequency of 0.1 Hz. The new Taylor series representation requires $N = 7$ for the same sampling specifications resulting in 4 seconds of data saved each time the delay estimation process must be recomputed to account for gaps in mission data.

Additionally, the code implements an algebraically-reduced and optimized version of the TDI combinations by substituting Eq. (7) into Eq. (6) and combining terms with the same delay operator combination so that each unique delay sequence need only be applied once which significantly reduces run-time. We implement the $\dot{f}_{, \mathcal{D}_{ij}}$ term in Eq. (5) by taking the NumPy numerical gradient of the delayed signal with second order central differences of the interior points and first order forward and backwards differences for the end samples.

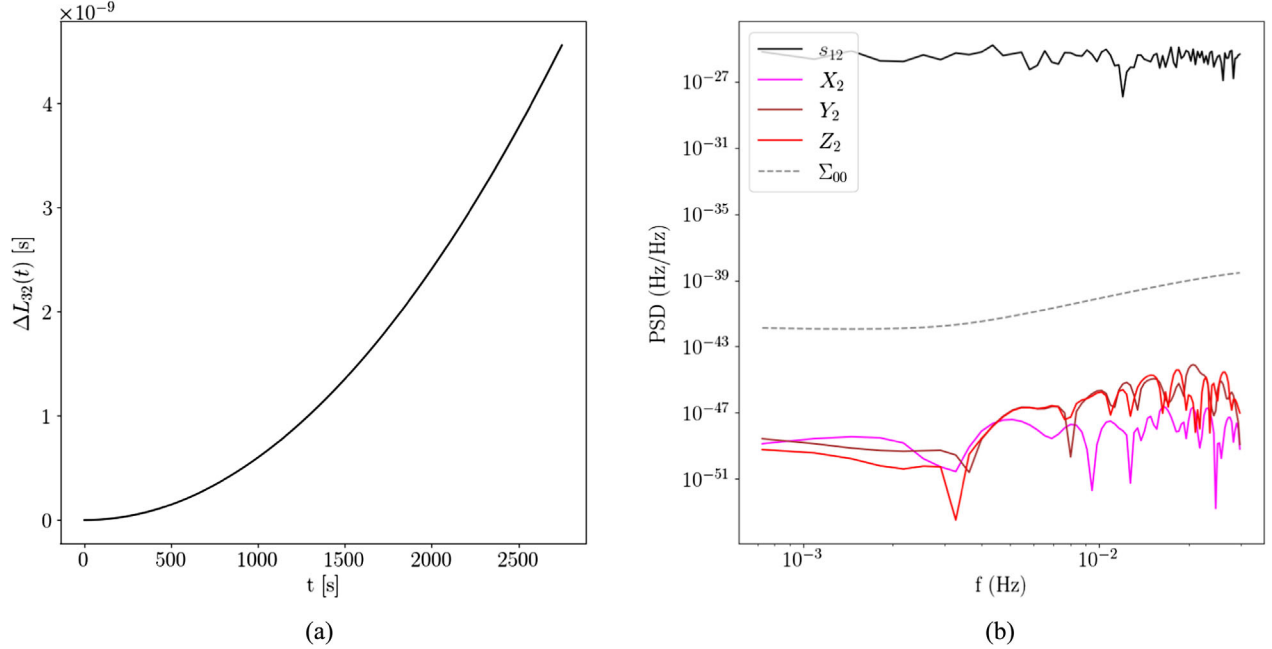


FIG. 2. (a) Linear approximation—true time dependence ($\Delta L_{32}(t) = L_{32}^{\text{LINEAR}} - L_{32}^{\text{TRUTH}}$) for arm length $L_{32}(t)$ for LISA in Keplerian orbit on ~ 1 hour of data. (b) TDI residuals $X_2(f)$ (magenta), $Y_2(f)$ (brown), $Z_2(f)$ (red) using linear approximation compared to the secondary noise reference curve (dashed gray) given by Eq. (9). Although the TDI residual suppresses LFN noise, the linear time dependence is not sufficient for estimating L_{ij} posteriors.

C. Data

We use the LISA Instrument software to simulate the interferometric measurements s_{ij} , τ_{ij} , and ϵ_{ij} . They are given in units of fractional frequency fluctuations for a given sampling rate, data duration, data content, reference frame and orbit model using LISA Orbits. We assume a 4 Hz sampling rate with varying duration depending on the orbit model used. We are demonstrating feasibility of Bayesian TDIR for second-generation TDI, so we consider LFN only for now, which excludes the optical metrology system (OMS), test mass (TM), clock and all other optional noises as well as GWs. We previously showed the effect of the presence of OMS and TM noises on arm length estimation in Ref. [5], and found that the MCMC produced adequate posterior estimates and LFN suppression for the rigidly rotating LISA configuration assumption. We assume no up-sampled on-board physics and therefore no anti-aliasing filter, and keep the remaining arguments to LISA Instrument at their default settings. Since we are interested in examining Bayesian estimates of time-varying arm lengths, we explore both the Keplerian and the numerically computed ESA (European Space Agency) orbit models in LISA Orbits as input to LISA Instrument. The Keplerian orbit model assumes the center of mass of the LISA constellation is on a heliocentric orbit. A more realistic numerical orbit model ESA Orbits accounts for gravitational forces from the Earth and other planetary bodies, as well as other non-gravitational forces [15]. The LISA Instrument output is in units of frequency fluctuations Hz and we convert to

fractional frequency fluctuations HzHz^{-1} so we implement the suggestion given in [16] to mitigate for the Doppler shift caused by the armlength variation in Eq. (6) by using the Doppler delay operator in place of delay operators.

Since the second-generation combinations suppress LFN to first order in the spacecraft distance velocities, we analyzed a linear parametrization of the time-dependence by estimating all 6 L_{ij} and 3 \dot{L}_{ij} (since $\dot{L}_{ij} \approx \dot{L}_{ji}$). While we do see LFN suppression in this parametrization [Fig. 2(b)], a comparison of the true $L_{ij}(t)$ in the Keplerian orbit model to the linear expansion is shown in Fig. 2(a). Even on data segments with a duration of $\mathcal{O}(1)$ hour, the log-likelihood difference between the use of linear approximated $L_{ij}(t)$ and true $L_{ij}(t)$ was greater than expected for a 9-parameter model. We additionally tried a quadratic approximation which added three more \ddot{L}_{ij} parameters but all expansion-based $L_{ij}(t)$ parameterization methods result in biased posterior estimates, so a more accurate parametrization is required.

IV. PERFORMANCE TESTING

A. Keplerian parametrization

The light travel time $L_{ij}(t)$ is simply the time difference between reception on spacecraft i from spacecraft j and we follow the procedure written in the LISA Orbits documentation based on a post-Minkowskian expansion given that the emission time is unknown upon reception which ultimately depends on the positions [Eq. (15)], velocities

and accelerations of the three spacecraft in the solar-system barycentric (SSB) reference frame at the time of reception. These quantities are derived from the orbital motion of each spacecraft using the 6 Keplerian orbital elements $\{a, e, \iota, M_0, \Omega, \omega\}$.

$$\begin{aligned} x_i(t) &= r_{c_i}(t) [\cos \nu_i(t) (\cos \omega \cos \Omega_i - \sin \omega \cos \iota \sin \Omega_i) \\ &\quad - \sin \nu_i(t) (\sin \omega \cos \Omega_i + \cos \omega \cos \iota \sin \Omega_i)] \\ y_i(t) &= r_{c_i}(t) [\cos \nu_i(t) (\cos \omega \sin \Omega_i + \sin \omega \cos \iota \cos \Omega_i) \\ &\quad + \sin \nu_i(t) (\cos \omega \cos \iota \cos \Omega_i - \sin \omega \sin \Omega_i)] \\ z_i(t) &= r_{c_i}(t) [\cos \nu_i(t) \sin \omega \sin \iota + \sin \nu_i(t) \cos \omega \sin \iota], \quad (15) \end{aligned}$$

where $r_{c_i}(t) = a(1 - e \cos E_i(t))$ is the distance to the solar system barycenter and $E_i(t)$ is the eccentric anomaly related to the true anomaly $\nu_i(t)$ by $\nu_i(t) = \arctan \frac{\sqrt{1+e}}{\sqrt{1-e}} \tan \frac{E_i(t)}{2}$. The velocities are given by taking the analytical derivative of Eq. (15), and the accelerations \ddot{r}_i are

$$\ddot{r}_i = \mu_\odot \frac{\vec{r}_i}{|\vec{r}_i|^3},$$

where μ_\odot is the solar gravitational parameter $\mu_\odot = GM_\odot$.

In the case of optimized Keplerian orbits for LISA, all 3 spacecraft share the same argument of perihelion ω and semi-major axis a which defines the eccentricity e and inclination ι given a tilt angle factor of $\delta = 5/8$ to minimize the constellation breathing [17]. The eccentric anomaly $E_i(t)$ is related to the initial mean anomaly parameter M_0 by Kepler's equation $M(t) = M_0 + (t - t_0) \frac{\mu_\odot}{a^3} = E(t) - e \sin E(t)$ and $M_{i_0} = M_0 - 2(i - 1)\pi/3$ for spacecraft $i \in \{1, 2, 3\}$. The longitude of the ascending node for spacecraft i is $\Omega_i = \Omega_1 + 2(i - 1)\pi/3$. The longitude of the ascending node is unconstrained due to the LISA orbit. When the inclination $\iota = 0$, Ω is undefined, and our orbital elements are described in ecliptic coordinates in the barycentric reference frame. The LISA guiding center will be in an Earth-leading or Earth-trailing heliocentric orbit, making the inclination $i \sim 0$. Therefore the only parameters required to describe all 6 armlengths $L_{ij}(t)$ in the Keplerian model are

$$\vec{\theta} = \{a, M_0, \omega\} \rightarrow L_{ij}(t).$$

While these assumptions do not address the realistic nature of the future LISA spacecraft orbits, this model provides a demonstration for MCMC-estimated time-dependent delays for data simulations commonly used in current LISA mock data challenges. Realistic LISA Orbits are considered in Sec. IV B.

We use the EMCEE affine-invariant MCMC ensemble sampler [18] on a 1-day duration of data simulated with the Keplerian orbit model using the ‘‘stretch move’’ proposal

jumps. We use a uniform prior over each parameter with $p(a) = \mathcal{U}[1 \text{ AU} - 2 \times 10^8 \text{ m}, 1 \text{ AU} + 1 \times 10^8 \text{ m}]$ which spans the linear range over a 10 year LISA mission as estimated in the numerical model given in [15] centered at 1 AU. The remaining angular orbital element priors are $p(M_0) = \mathcal{U}[-\pi, \pi]$ to avoid incomplete distributions at the prior boundaries and $p(\omega) = \mathcal{U}[-\pi, 0]$ under the assumption of clockwise rotation of the spacecraft. To reduce computational cost even further for a day's worth of data with a time-varying FDI filter, we reduce the sampling rate to 2 Hz and set the maximum frequency in the likelihood function to $f_{\max} = 0.03 \text{ Hz}$ since we are primarily concerned with LFN suppression while maintaining a significant portion of the LISA signal band in the analysis. The new FDI filter (Sec. III B) to delay the measurements for each MCMC iteration now requires a filter length of only $N = 7$ which saves on run-time and significantly reduces the amount of data loss.

Figure 3(a) shows the $\Delta a, M_0$, and $\Delta \omega$ posterior distributions and the 90% credible intervals centered around the truth values used in the data simulation. The semimajor axis and the initial mean anomaly of spacecraft 1 are represented by differences from their truth values that were input into the data simulation for figure clarity. The 1σ (dark blue) and 2σ (blue) shaded contours are shown. Figure 3(b) show the TDI residuals using the median posterior estimates for $\{a, M_0, \omega\}$ compared to one of the science measurements containing LFN to demonstrate sufficient noise suppression. The boundaries of the 90% credible intervals also produce TDI residuals well below the secondary noises.

B. A proposed Keplerian model for realistic numerical LISA Orbits

The more realistic orbit model used in the data simulation accounts for the other forces that act upon the LISA spacecraft such as the gravitational force of other solar system bodies, especially the Earth's. The numerical model described in [15] and used in LISA Orbits points out that the Earth's gravity leads to a linear drift in the semi-major axis and is the most significant perturbation to analytical Keplerian orbits. We tested whether our sampler could estimate the time-varying spacecraft separations using an adiabatic Keplerian model, fitting Keplerian orbital elements to the spacecraft orbits on short time segments. Now the spacecraft motion cannot be described by 6 orbital elements, but rather an osculating orbit with time-varying orbital parameters for all three spacecraft. If we can sample the orbital elements on a sufficiently short duration of data such that the time-variance of the orbital parameters is negligible for calculation of \vec{L}_{ij} and sufficient LFN suppression, the posterior results can be interpolated across a chosen duration of LISA data to estimate the time-varying Keplerian orbital elements on any chosen scale.

We began by looking at the position and velocity for each spacecraft in the SSB time frame for 1 hour of

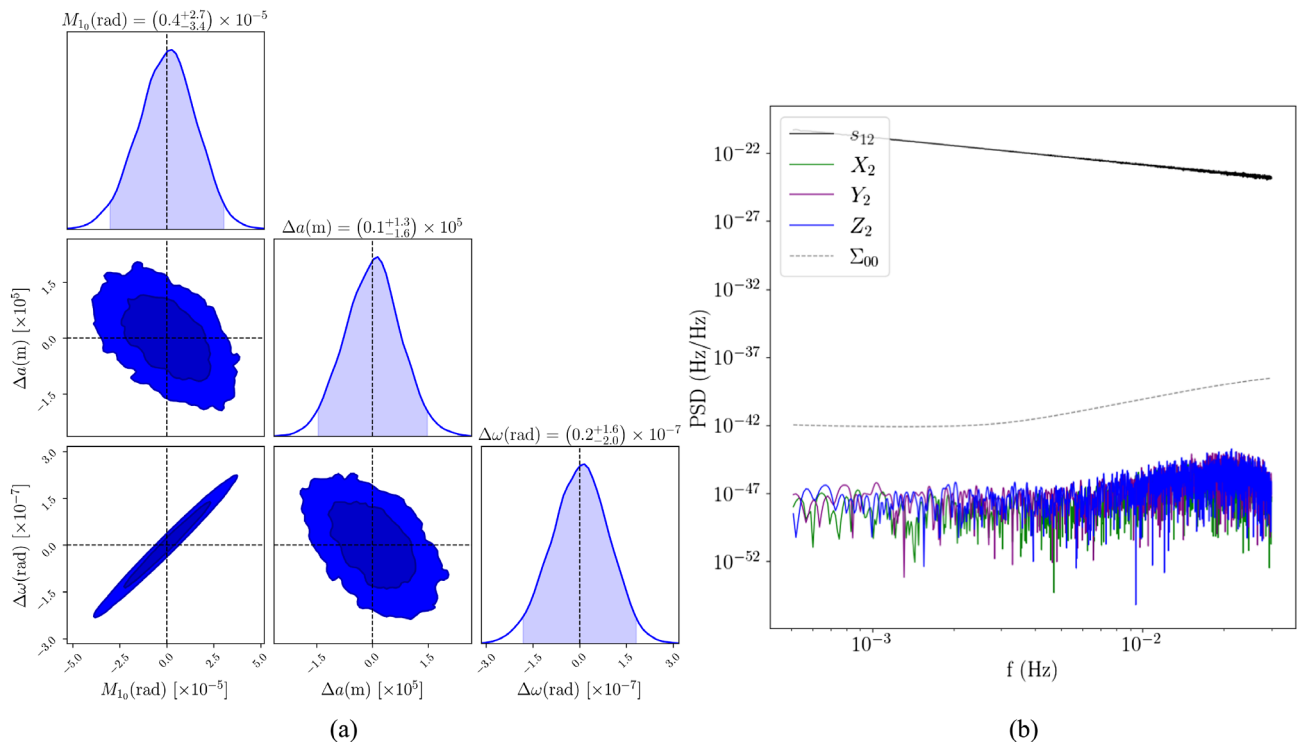


FIG. 3. (a) Posteriors for $\vec{\theta} = \{a, M_{0_i}, \omega\} \vec{L}_{ij}$ parameterization on Keplerian orbital data. The semi-major axis and argument of perihelion are shown as differences from truth values used in the simulation (Δa and $\Delta \omega$). 90% credible interval quantities are quoted and 1σ (dark blue) and 2σ (blue) shaded contours are shown. (b) TDI residuals $X_2(f)$ (green), $Y_2(f)$ (purple), $Z_2(f)$ (blue) using median values from posterior distribution compared to the secondary noise reference curve (dashed grey) given by Eq. (9).

numerical data and converted them to time-varying Keplerian orbital elements. Figure 4 demonstrates adequate LFN suppression using the first time sample of the

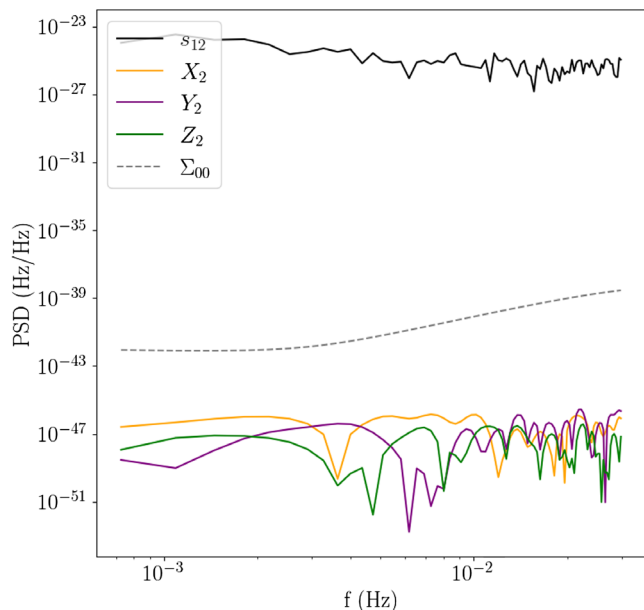


FIG. 4. TDI residuals $X_2(f)$ (orange), $Y_2(f)$ (purple), $Z_2(f)$ (green) on 1 hour of arbitrary numerical orbit data using constant-valued Keplerian orbital elements compared to the secondary noise reference curve (dashed grey) given by Eq. (9).

converted orbital elements as input to the position equations in Eq. (15), now with an “i” subscript after every orbital element since they all differ for each spacecraft. Use of the first sample for the M_{0_i} parameter is crucial for accurate spacecraft separations, but one could explore using median values or other constant valued representations for the other parameters. We then examined sampling this 18-parameter model with

$$\vec{\theta}_i = \{a_i, e_i, i_i, M_{0_i}, \omega_i, \Omega_i\}, i \in \{1, 2, 3\}$$

for each spacecraft i on the hour-long data segment using MCMC methods since constant values of the time-varying parameters suppress LFN. This is a challenging parameter model to sample, and work is still in progress to find the best sampling techniques to gain efficiently processed posteriors. We noticed it performed exceptionally better when the uniform prior ranges for the semimajor axis were narrowed from the 10-year expected range to a smaller range indicating they may need adjustment depending on the start time and duration of the segment. It is a realistic expectation that some mission data will provide more informative priors (e.g. narrowed uniform distributions) to speed up sampling convergence.

Figure 5 compares the TDI residuals using the constant-valued orbital elements to calculate the arm lengths versus the true positions and velocities from the orbit simulation

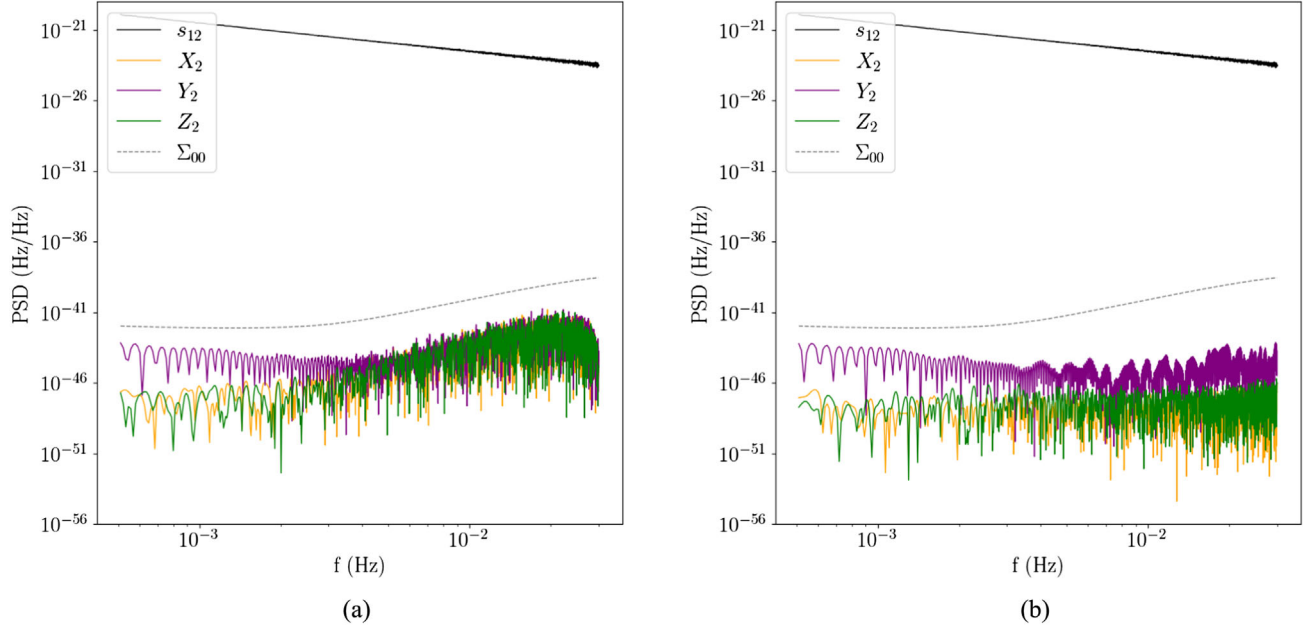


FIG. 5. TDI residuals $X_2(f)$ (orange), $Y_2(f)$ (purple), $Z_2(f)$ (green) on 1 day of arbitrary numerical orbit data using constant-valued Keplerian orbital elements compared to the secondary noise reference curve (dashed gray) given by Eq. (9) for: (a) Initial values in time for the Keplerian orbital elements of each S/C. (b) True positions and velocities from orbit data simulation.

used as input to the data simulation for a 1-day duration. Figure 5(a) represents an approximation to the arm length time dependence. It is the initial value of the time-dependent Keplerian orbital elements, the same quantity the MCMC would take as parameters. In reality, the orbital elements themselves are time-dependent, which would ultimately result in the true $x_i(t)$, $y_i(t)$, and $z_i(t)$ Cartesian positions for each spacecraft i that were taken from the data simulation used in Fig. 5(b). Constant values still suppress LFN, but one needs to examine whether the resulting $\log \mathcal{L}$ values are comparable to the true positions for accurate posterior distributions. For a sufficiently sampled posterior chain, the $\Delta \log \mathcal{L} \propto N_D/2$ where N_D is the model dimension. Therefore we would want the $\log \mathcal{L}$ difference to be $\lesssim 9$ for an 18-parameter Keplerian model. We find that a duration of 9 hours of numerical LISA data sampled at 4 Hz gives a $\Delta \log \mathcal{L} = 7.4$ using a filter length $N = 7$ and maximum frequency $f_{\max} = 0.03$ Hz. One could sample 9 hour segments at a time of LISA data and interpolate point estimates and credible intervals of the resulting posterior distribution to estimate the LISA Orbits independent of any mission ephemeris data as a back-up and compliment to on-board phase modulating pseudorange once future work demonstrates a successful Bayesian sampling of the 18 parameter Keplerian model.

V. CONCLUSIONS AND FUTURE WORK

The goal of Bayesian data-driven methods for estimating the TDI delays is acquiring posterior estimates on the arm

length parameters to be marginalized over when computing the LISA response function to GWs. We demonstrate the challenges and ways to mitigate them in the realistic LISA scenario in heliocentric orbit. We use a Taylor-expanded version of the FDI filter to compute and store approximately half of the time-varying filter operations prior to running MCMC iterations to significantly improve computational efficiency. The time-dependent delays are estimated using Keplerian orbital elements for the Keplerian orbit model and provide a foundation for this parametrization in the realistic arbitrary orbit scenario. Future efforts to build upon this data driven TDI strategy will move forward with MCMC sampling algorithms to fully characterize the LISA spacecraft orbits by individual estimation of approximately 9 hour segments of data using a 4 Hz sampling rate and interpolating segment posteriors. Arm length posteriors can then be incorporated into the LISA response function to GWs, and marginalization over arm-length uncertainty will remove potential sources of bias during GW signal parameter estimation. Future work could also include test mass and optical metrology system noises in the simulated data to examine their effect on the more realistic LISA orbital assumptions explored here to make comparisons with previous studies on the rigidly rotating LISA configuration.

ACKNOWLEDGMENTS

This project was supported by the NASA LISA Study Office.

- [1] P. Amaro-Seoane *et al.*, Laser interferometer space antenna, [arXiv:1702.00786](#).
- [2] M. Tinto and S. V. Dhurandhar, Time-delay interferometry, *Living Rev. Relativity* **24**, 1 (2021).
- [3] Q. Baghi, J. G. Baker, J. Slutsky, and J. I. Thorpe, Fully data-driven time-delay interferometry with time-varying delays, *Ann. Phys. (Berlin)* **2023**, 2200447 (2022).
- [4] Q. Baghi, J. I. Thorpe, J. Slutsky, and J. Baker, Statistical inference approach to time-delay interferometry for gravitational-wave detection, *Phys. Rev. D* **103**, 042006 (2021).
- [5] J. Page and T. B. Littenberg, Bayesian time delay interferometry, *Phys. Rev. D* **104**, 084037 (2021).
- [6] M. Tinto, M. Vallisneri, and J. W. Armstrong, Time-delay interferometric ranging for space-borne gravitational-wave detectors, *Phys. Rev. D* **71**, 041101 (2005).
- [7] T. B. Littenberg and N. J. Cornish, Prototype global analysis of LISA data with multiple source types, *Phys. Rev. D* **107**, 063004 (2023).
- [8] M. L. Katz, J.-B. Bayle, A. J. K. Chua, and M. Vallisneri, Assessing the data-analysis impact of LISA orbit approximations using a GPU-accelerated response model, *Phys. Rev. D* **106**, 103001 (2022).
- [9] D. A. Shaddock, B. Ware, R. E. Spero, and M. Vallisneri, Postprocessed time-delay interferometry for LISA, *Phys. Rev. D* **70**, 081101 (2004).
- [10] J.-B. Bayle and O. Hartwig, Unified model for the LISA measurements and instrument simulations, *Phys. Rev. D* **107**, 083019 (2023).
- [11] M. Otto, G. Heinzel, and K. Danzmann, TDI and clock noise removal for the split interferometry configuration of LISA, *Classical Quantum Gravity* **29**, 205003 (2012).
- [12] M. Tinto, F. B. Estabrook, and J. W. Armstrong, Time delay interferometry with moving spacecraft arrays, *Phys. Rev. D* **69**, 082001 (2004).
- [13] D. A. Shaddock, M. Tinto, F. B. Estabrook, and J. W. Armstrong, Data combinations accounting for LISA spacecraft motion, *Phys. Rev. D* **68**, 061303 (2003).
- [14] A. Candan, An efficient filtering structure for Lagrange interpolation, *IEEE Signal Process. Lett.* **14**, 17 (2007).
- [15] W. Martens and E. Joffre, Trajectory design for the ESA LISA mission, *J. Astronaut. Sci.* **68**, 402 (2021).
- [16] J.-B. Bayle, O. Hartwig, and M. Staab, Adapting time-delay interferometry for LISA data in frequency, *Phys. Rev. D* **104**, 023006 (2021).
- [17] K. R. Nayak, S. Koshti, S. V. Dhurandhar, and J. Y. Vinet, On the minimum flexing of LISA's arms, *Classical Quantum Gravity* **23**, 1763 (2006).
- [18] D. Foreman-Mackey, D. W. Hogg, D. Lang, and J. Goodman, EMCEE: The MCMC Hammer, *Publ. Astron. Soc. Pac.* **125**, 306 (2013).