Statistical inference approach to time-delay interferometry for gravitational-wave detection

Quentin Baghi,[*] James Ira Thorpe, Jacob Slutsky, and John Baker

Goddard Space Flight Center Mail Code 663 8800 Greenbelt Road, Greenbelt, Maryland 20771, USA

The future space-based gravitational wave observatory laser interferometer space antenna (LISA) will consist of a constellation of three spacecraft in a triangular constellation, connected by laser interferometers with 2.5 million-kilometer arms. Among other challenges, the success of the mission strongly depends on the quality of the cancellation of laser frequency noise, whose power lies 8 orders of magnitude above the gravitational signal. The standard technique to perform noise removal is time-delay interferometry (TDI). TDI constructs linear combinations of delayed phasemeter measurements tailored to cancel laser noise terms. Previous work has demonstrated the relationship between TDI and principal component analysis (PCA). We build on this idea to develop an extension of TDI based on a model likelihood that directly depends on the phasemeter measurements. Assuming stationary Gaussian noise, we decompose the measurement covariance using PCA in the frequency domain. We obtain a comprehensive and compact framework that we call PCI for "principal component interferometry" and show that it provides an optimal description of the LISA data analysis problem.

I. INTRODUCTION

Only five years since the historic first detection of an astrophysical signal, the field of gravitational-wave astronomy continues to rapidly advance through upgrades to ground-based facilities [1], plans for ambitious follow-on terrestrial detectors [2], and a steady increase in sensitivity of pulsar timing arrays [3]. Adding to this in the 2030s will be the space mission known as the Laser Interferometer Space Antenna (LISA) [4]. LISA will open the source-rich frequency window between 0.1 mHz and 1 Hz, enabling the detection of thousands of gravitational-wave emitting systems of varied origin and at distances ranging from our galactic neighborhood (kpc) to cosmological redshift ($z = 15$ and beyond). The types of sources will include compact binary stars, the capture of stellar-remnant black holes by massive black holes, and mergers of (super)massive black hole binaries. LISA will form a triangular constellation of three satellites trailing Earth on heliocentric orbits. Each satellite will house inertial test masses whose trajectories will be monitored through a network of interferometric laser links, connecting the spacecrafts separated from each other by 2.5 million kilometers. Incoming gravitational waves (GW) will affect the space-time across the constellation, introducing a characteristic shift in the light travel time along the six one-way optical links.

Due to the large distance between the inertial references, LISA needs multiple interferometric measurements to operate optimally [5]. In this setup, the phasemeter measurements are sensitive to noise fluctuations of the laser frequencies at a level of 10^{-13} Hz$^{-1/2}$. The typical metric perturbations induced by the target GW sources being about 10^{-21}, the laser noise dominates them by 8 orders of magnitude. Recovery of the GW signal is possible because the correlations between the individual link signals differ for GWs and laser noise. Laser frequency noise is almost entirely canceled by a postprocessing technique called time-delay interferometry (TDI) [6], which constructs linear combinations of delayed phasemeter measurements tailored to cancel the laser frequency noise up to ranging errors. Some TDI variables can be interpreted physically by synthetically retracing the path of light rays traveling in a classical Michelson interferometer. The feasibility of TDI has been studied extensively over the past decade (see, e.g., Ref. [7] for an overview), including the implementation of interpolation filters needed to produce the laser-free data streams [8]. Further analyses have recently been tackled, revealing expected TDI noise artifacts, such as flexing filtering [9] and clock jitters effects [10].

Different generations of TDI variables achieving different accuracy levels can be formed depending on assumptions about the spacecraft motion. TDI generations 1.0, 1.5, and 2.0, respectively, assume a rigid and static constellation, a rigid and rotating configuration, and a flexing, rotating configuration where the arm lengths are linearly varying in time. Recent work also identified new TDI

---

[*]quentin.baghi@protonmail.com

042006-1

combinations by using the explicit dependence of arm delays on the satellites' velocities and accelerations [11].

In addition to its geometric interpretation, TDI can be viewed algebraically as the group of solutions of an equation involving six-tuples polynomials in the delay operators [12], which encodes the cancellation of noise. Beyond laser noise cancellation, the analysis of LISA's measurement can also be interpreted as an inference problem. We observe the same physical effect (the perturbation of the metric due to incoming gravitational waves) through different sensors (interferometric arms) affected by fundamental errors (the laser noise and other stochastic noises). Everything works as if we recorded the same sound using different microphones. The goal of the analysis is generally to estimate parameters of astrophysical sources emitting GWs. Thus, we can infer them by directly writing the likelihood to observe the phasemeter measurements given a prescribed model. The model describes how we expect the GW signal to appear in the various measurements and the relationships between the noises present in each data stream. In Ref. [13], Vallisneri *et al.* recently formalized this idea in the time domain by marginalizing the likelihood with respect to laser noises. In this study, we develop a similar approach in the frequency domain, using an eigendecomposition of the covariance.

Typically, the likelihood depends on the inverse correlation matrix of all measured variables. Applying its inverse to the model residuals yields the weighed, uncorrelated squared errors involved in any optimal parameter estimation scheme. We refer to the process of generating orthogonal variables as principal component analysis (PCA). In a pioneering work [14], Romano and Woan show that we can derive the TDI combinations from the eigenvectors of the single-link covariance matrix through a demonstration based on a simplified time-domain analysis assuming white noise and short time series. Leighton further analyzes the noise covariance matrix structure and its principal components in Ref. [15], extending it to the frequency domain. The aim of the present work is to make this idea readily usable for gravitational-wave data analysis and demonstrate it on an inference problem. For this purpose, we build an analytic formalism implementing PCA, based on a matrix formulation in the frequency domain. We refer to this approach as "principal component interferometry" or PCI for short.

PCI implements the generalized analog of the orthogonal TDI channels A, E, T [16]. While A, E, T's original construction relies on assumptions such as equal arms and uncorrelated acceleration noises, PCI yields the optimal variance in a unified, data-driven formalism. We first decompose the data on the eigenvector basis of the laser-noise covariance matrix. We then estimate the frequency-domain covariance matrix (or spectral matrix) of eigenstreams, assuming that the acceleration noise is stationary. The resulting model likelihood allows us to fit for

noise parameters encoding all types of noise correlations, which we account for by construction in the inference process.

In Sec. II, we present the algebraic formalism used to develop PCI. In Sec. III, we derive the principal components of the covariance in the simple case of rigid arm lengths. In Sec. IV, we demonstrate PCI's performance through simple numerical simulations, with an example of an application where we estimate conjointly the laser light propagation delays (ranging), the noise covariance elements, and the parameters of a compact galactic binary source. We conclude and discuss the further generalization to more complex cases in Sec. V.

II. MODELING SINGLE-LINK MEASUREMENTS USING MATRIX OPERATORS

A. Derivation of the likelihood

In this section, we present the conventions and the formalism that we adopt throughout the study. Using conventions in Fig. 1 of Ref. [9], we consider a spacecraft, labeled i, and the signal, s_i, obtained by comparing the laser light coming from the distant spacecraft $i + 1$ to the local oscillator of optical bench i. The measurement s_i is where gravitational waves imprint their presence and is called the science interferometer signal. The science interferometer measurements $s_i(t) = (\nu_i(t) - \nu_0)/\nu_0$ are expressed as the relative deviation of instantaneous frequency with respect to the carrier frequency ν_0. For the two optical benches onboard spacecraft i, we have, at each time t:

$$s_i = h_{i+2} + \mathcal{D}_{i+2}p_{i'+1} - p_i + n_i;$$
$$s_{i'} = h_{i+1} + \mathcal{D}_{i'+1}p_{i+2} - p_{i'} + n_{i'}, \quad (1)$$

where $\mathcal{D}_i x(t) = x(t - c^{-1}L_i)$ denotes the operator applying the light travel time delay $c^{-1}L_i$ along arm i, h_i is the integrated frequency shift along arm i due to incoming gravitational waves, p_i is the frequency noise contribution of the laser in optical bench i, and n_i gathers all other noises affecting the science measurement on that bench for all $i \in \{1, 2, 3\}$. We adopt the convention of cyclic indexing where p_i actually means $p_{i-3\lfloor(i-1)/3\rfloor}$, and simple indices refer to links and light travel times pointing clockwise, whereas prime indices denote counterclockwise directions. In the following, we assume that the lasers in the two optical benches are identical, so that

$$p_i = p_{i'} \quad \forall i \in \{1, 2, 3\}. \quad (2)$$

This assumption allows us to simplify the analysis but can be adopted without loss of generality. Note that although the laser sources are different in reality, they will be compared through reference interferometer measurements.

The measurements s_i in Eq. (1) will be sampled at a cadence of $f_s = 2$ Hz or more over a finite duration T.

The resulting time series can be represented by a column vector s_i of size $N = f_s T$. In this discretized version of the measurement, the delay operators $\mathcal{D}_i$ acting on any variable $x(t)$ can be represented by $N \times N$ matrices depending on the arm lengths L_i. Using consistent notation, all lines in Eq. (1) can be rewritten as

$$y = h + Mp + n, \qquad (3)$$

where $y \in \mathbb{R}^{6N}$ is the column vector stacking the phasemeter measurements for all spacecrafts and all optical benches so that $y((i-1)N + j) = s_i(j/f_s)$ and $y((i+2)N + j) = s_{i'}(j/f_s) \, \forall \, i \in \{1, 2, 3\}, \, \forall \, j \in [0, N-1]$. Vectors h, p, and n, respectively, represent the GW signals, laser noises, and other noises. They have in general the same structure and dimensions as y. Note that under the assumption outlined in Eq. (2), we keep only the first half of p so that $p \in \mathbb{R}^{3N}$.

Matrix M encodes the mixing and delaying of laser noises and can be written as a block matrix:

$$M = \begin{pmatrix} -I_N & D_3 & 0_N \\ 0_N & -I_N & D_1 \\ D_2 & 0_N & -I_N \\ -I_N & 0_N & D_{2'} \\ D_{3'} & -I_N & 0_N \\ 0_N & D_{1'} & -I_N \end{pmatrix}. \qquad (4)$$

Now that the observation equation is written in matrix form, we can derive the corresponding likelihood. Gravitational-wave source parameters θ are usually extracted using Bayesian inference, which estimates their posterior distribution given the data,

$$p(\theta|y) = \frac{p(y|\theta)p(\theta)}{p(y)}, \qquad (5)$$

where $p(y|\theta)$ is the model likelihood, $p(\theta)$ is the prior distribution of the parameters, and $p(y)$ is the evidence, acting as a normalization.

Assuming a zero-mean Gaussian distribution for all noises, the likelihood follows from Eq. (3):

$$p(y|\theta) = \frac{\exp\left\{-\frac{1}{2}(y-h)^\dagger \Sigma^{-1}(y-h)\right\}}{\sqrt{(2\pi)^{6N}|\Sigma|}}, \qquad (6)$$

where $\dagger$ denotes the Hermitian conjugate, and Σ is the $6N \times 6N$ covariance matrix of the observations y, whose expression derives from Eq. (3),

$$\Sigma = M\Sigma_p M^\dagger + \Sigma_n, \qquad (7)$$

where Σ_p and Σ_n are, respectively, the covariance matrices of the laser noises p and of the other noises n, assuming no intrinsic correlation between the two.

In the following, we make the convenient but realistic assumption that all noises are stationary (at least for a relatively short period of time). In that case, their covariance matrices are Toeplitz, and for a sufficiently large N, they are approximately diagonalizable in the discrete Fourier basis W that forms its eigenvectors and writes $W(k, n) = e^{\frac{2\pi jnk}{N}}$, with $j = \sqrt{-1}$ being the complex number. We can therefore rewrite the covariance matrix in Eq. (7) in the Fourier domain as

$$\tilde{\Sigma} = \tilde{M} S_p \tilde{M}^\dagger + S_n, \qquad (8)$$

where S_p and S_n are the covariance matrices of the discrete Fourier-transformed data that we call spectral matrices. Due to stationarity, S_p (respectively, S_n) includes 3×3 (respectively, 6×6) blocks, which are $N \times N$ diagonal matrices, whose diagonal elements are given by the noise cross spectra. Spectral matrices are Hermitian, with real positive diagonal blocks and complex conjugate off-diagonal blocks. If we restrict the analysis to a specific set of N_f frequencies, then each block has size $N_f \times N_f$. In the following, we look for the principal components of the covariance matrix.

III. PRINCIPAL COMPONENT ANALYSIS OF SINGLE-LINK MEASUREMENTS

A. Principle of PCA

PCA aims at finding a transformation of the observations that converts them into uncorrelated variables, ordered according to their variance. The process is often used to reduce the dimension of the problem by discarding the highest variance components. Here, we aim at finding a unitary transformation matrix V, where the covariance matrix can be diagonalized as

$$\Sigma = V\Lambda V^\dagger, \qquad (9)$$

where Λ is a diagonal matrix. Then the log-likelihood can conveniently be rewritten as

$$\log p(y|\theta) = -\frac{1}{2}(y-h)^\dagger V \Lambda^{-1} V^\dagger (y-h) - \frac{1}{2}\log|\Lambda|. \qquad (10)$$

However, finding a full decomposition like Eq. (9) can be tricky, unless we make a few key assumptions, which we do in the following.

B. PCI for equal noises

A way to find a decomposition of the form (9) is to find the eigenvectors of the covariance matrix. To ease their calculation, we make two assumptions.

First, we assume that all delays are constant in time, which implies that the delay operators are commutative; i.e., $D_1 D_2 = D_2 D_1$. In this case, the Fourier basis also provides approximate eigenvectors for the delay operators D_i. Hence, in the Fourier basis, the delay operators are approximately diagonal, and in the limit of large N, we can write their elements as

$$\tilde{D}_{i|k,l} = e^{-2\pi j f_k c^{-1} L_i} \delta_{kl}. \tag{11}$$

Second, we assume for now that all nonlaser noises have the same power spectral density (PSD). Thus, all diagonal blocks $S_{n,i}$ of the spectral matrix S_n are equal: $S_{n,i} = \lambda_n \,\forall\, i$.

Based on these assumptions, from the calculation of the characteristic polynomial of matrix $\tilde{\Sigma}$, we find that there are six eigenvalues per frequency bin, hence, $6N_f$ eigenvalues for the full problem. Half of them (which we label Λ_n) are degenerate and equal to the nonlaser noise PSD values so that $\Lambda_n = \text{diag}(\lambda_n, \lambda_n, \lambda_n)$. We can write the $6N_f \times 3N_f$ matrix gathering their associated eigenvectors $\tilde{V}_n$ analytically as

$$\tilde{V}_n = \begin{pmatrix} \tilde{D}_2^\dagger(\tilde{D}_1^\dagger\tilde{D}_1^\dagger - I) & \tilde{D}_{3'}^\dagger - \tilde{D}_1^\dagger\tilde{D}_2^\dagger & \tilde{D}_2^\dagger\tilde{D}_2^\dagger - I \\ \tilde{D}_{1'}^\dagger - \tilde{D}_2^\dagger\tilde{D}_3^\dagger & \tilde{D}_3^\dagger\tilde{D}_3^\dagger - I & \tilde{D}_3^\dagger(\tilde{D}_2^\dagger\tilde{D}_2^\dagger - I) \\ \tilde{D}_1^\dagger\tilde{D}_1^\dagger - I & \tilde{D}_1^\dagger(\tilde{D}_{3'}^\dagger\tilde{D}_3^\dagger - I) & \tilde{D}_{2'}^\dagger - \tilde{D}_1^\dagger\tilde{D}_3^\dagger \\ 0 & 0 & I - \tilde{D}_1^\dagger\tilde{D}_2^\dagger\tilde{D}_3^\dagger \\ 0 & I - \tilde{D}_1^\dagger\tilde{D}_2^\dagger\tilde{D}_3^\dagger & 0 \\ I - \tilde{D}_1^\dagger\tilde{D}_2^\dagger\tilde{D}_3^\dagger & 0 & 0 \end{pmatrix}. \tag{12}$$

Note that $\tilde{V}_n$ is also a basis for the null space of the laser-noise part of the covariance so that we have $\tilde{\Sigma}\tilde{V}_n = S_n\tilde{V}_n$.

The $3N_f$ other eigenvalues, which we label as $\Lambda_p = \text{diag}(\lambda_{p1}, \lambda_{p2}, \lambda_{p3})$, have more complicated expressions but are all proportional to the laser noise PSD S_p. We denote by $\tilde{V}_p$ the associated eigenvector matrix, which has the same dimensions as $\tilde{V}_n$. We plot the laser-noise dominated eigenvalues λ_{pi} in gray as a function of frequency in Fig. 1, along with the degenerate laser-noise free eigenvalues λ_n in blue. This figure confirms that the former are much larger than the latter.

As a result, we can partition the eigenvector matrix as $\tilde{V} = (\tilde{V}_p \tilde{V}_n)$. Let us consider the data transformation $\tilde{e} \equiv \tilde{V}^\dagger \tilde{y}$. As they correspond to different eigenvalues, the eigenvector matrices $\tilde{V}_n$ and $\tilde{V}_p$ are orthogonal. Therefore, the covariance of $\tilde{e}$ is block diagonal,
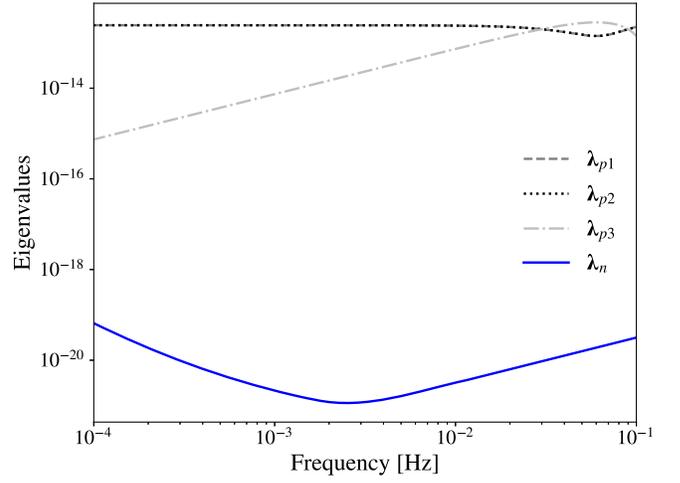

FIG. 1. Laser-noise dominated eigenvalues λ_{pi} (gray) and laser-noise free eigenvalues λ_n (blue) of the phasemeter covariance matrix as a function of frequency.

$$\text{Cov}(\tilde{e}) = \begin{pmatrix} V_p^\dagger \Sigma V_p & 0 \\ 0 & V_n^\dagger S_n V_n \end{pmatrix} \equiv \begin{pmatrix} C_p & 0 \\ 0 & C_n \end{pmatrix}, \tag{13}$$

where we defined the $3N_f \times 3N_f$ matrix C_n (respectively, C_p) as the covariance of the projected data $\tilde{e}_n \equiv \tilde{V}_n^\dagger \tilde{y}$ (respectively, $\tilde{e}_p \equiv \tilde{V}_p^\dagger \tilde{y}$) onto the laser-noise free (respectively, laser-noise dominated) basis.

Then it is possible to separate the laser noise-dominated eigenbasis $\tilde{V}_p$ from the laser noise-free eigenbasis $\tilde{V}_n$ in the calculation of the likelihood,

$$\log p(y|\theta) = -\frac{1}{2}(\tilde{y} - \tilde{h})^\dagger \tilde{V}_p C_p^{-1} \tilde{V}_p^\dagger (\tilde{y} - \tilde{h})$$
$$- \frac{1}{2}(\tilde{y} - \tilde{h})^\dagger \tilde{V}_n C_n^{-1} \tilde{V}_n^\dagger (\tilde{y} - \tilde{h})$$
$$- \frac{1}{2}(\log|C_p| + \log|C_n|). \tag{14}$$

Similarly to Romano and Woan's example in Ref. [14], the laser noise variance being much larger than other noises, the term in the second line in Eq. (14) is almost constant as a function of θ compared to the term in the first line. Therefore, for parameter inference purposes, we can safely approximate the log-likelihood by

$$\log p(y|\theta) \approx -\frac{1}{2}(\tilde{y} - \tilde{h})^\dagger \tilde{V}_n C_n^{-1} \tilde{V}_n^\dagger (\tilde{y} - \tilde{h}) - \frac{1}{2}\log|C_n|. \tag{15}$$

In the next section, we detail how we compute the inverse of C_n.

C. Orthogonalization with respect to nonlaser noise

In the previous section, we saw that the approximate log-likelihood depends on the covariance C of the data

projected onto the eigenbasis associated with null eigenvalues of the laser-noise covariance. If the noise parameters (PSD levels and light travel time delays) are assumed to be known and fixed in the inference scheme, this matrix and its inverse can be computed once for all. However, if we need to update the delays and the noise model along the way, we must compute C_n^{-1} at every parameter update for all frequencies. In this work, we perform this computation by numerically diagonalizing C_n with its eigenvectors $\mathbf{\Phi}$ and eigenvalues Λ_C and then computing $C_n^{-1} = \mathbf{\Phi}\Lambda_C^{-1}\mathbf{\Phi}^{\dagger}$. We use the NumPy library [17,18], which includes efficient algorithms when the number of frequency bins is not large (<1000). It may be more efficient for larger frequency bands to use an analytical formula for 3×3 Hermitian matrices as derived by Ref. [19].

This diagonalization is a generalization of the orthogonalization process that leads to TDI variables A, E, T [7,20]. Indeed, we can apply the formalism developed in this section to any linear transformation that cancels laser frequency noise. For example, TDI transformations can be encoded by some matrix T instead of V_n in Sec. III B. While channels A, E, and T are orthogonal under specific conditions (including equal arm lengths, a nonrotating constellation, identical acceleration noise levels, and uncorrelated noises), the rationale leading to Eq. (15) does not rely on any of these assumptions.

D. PCI with no prior knowledge of noise PSDs

While the projection onto the null space of the laser noise covariance matrix (i.e., the calculation of $\tilde{V}_n$) is independent of the laser noise spectra, the orthogonalization that we outlined in Sec. III C relies on our knowledge of the other noises' spectral matrix S_n. Although we may have a physical model describing acceleration and optical metrology system (OMS) noises, we must expect deviations from the theory when dealing with future LISA data. Therefore, it is necessary to have a formalism that also allows us to estimate C_n robustly. Estimating the full covariance matrix elements is not commonly done in gravitational-wave data analysis, but it can be performed similarly to PSD estimation methods, extending them to off-diagonal terms. This type of problem relates to spectral analysis of co-stationary multivariate time series, for which, several approaches are available, such as estimating the components of the generalized Cholesky decomposition of the spectral matrix or its inverse [21]. Regardless of the model we adopt, one has to ensure that the estimated spectrum is a positive definite matrix and is continuous as a function of frequency. To benefit from fast conditional steps when it comes to posterior sampling, we choose to model the covariance elements themselves with the regression scheme proposed by Ref. [22], which takes advantage of conjugate priors. To this end, let us consider a single frequency f and the 3×3 covariance of the corresponding elements,

$$\check{C}_n(f) \equiv \mathrm{Cov}(\check{e}(f)), \qquad (16)$$

where we labeled as $\check{e}(f_k) \equiv (\tilde{y}_k, \tilde{y}_{N_f+k}, \tilde{y}_{2N_f+k})^T$ the vector of eigenstream elements associated with frequency bin f_k. Thus, we assume that the covariance has the form

$$\check{C}_n(f) = \mathbf{\Psi} + \boldsymbol{B}\boldsymbol{x}(f)\boldsymbol{x}^{\dagger}(f)\boldsymbol{B}^{\dagger}, \qquad (17)$$

where $\mathbf{\Psi}$ is a constant 3×3 Hermitian matrix, $\boldsymbol{x}(f)$ is a $q \times 1$ design matrix depending on frequency, and $\boldsymbol{B}$ is a $3 \times q$ matrix of regression parameters. For example, $\boldsymbol{x}$ can have the form of a polynomial in frequency with elements $\boldsymbol{x} = (1, f, ..., f^{q-1})^T$.

In this model, $\mathbf{\Psi}$ and $\boldsymbol{B}$ are unknown and must be estimated. For a sufficiently short frequency range, we can even approximate the covariance by a constant term across the band, as in Ref. [23]. Under this assumption, Eq. (17) reduces to $\check{C}_n(f) = \mathbf{\Psi}$, and Ref. [22]'s sampling scheme amounts to using the conjugate prior for the Gaussian distribution, i.e., the inverse Wishart prior $\mathrm{IW}(\mathbf{\Psi}_0, \nu_0)$. We adopt this simplification in what follows, where the conditional posterior of $\mathbf{\Psi}$ given the delays L and GW parameters $\boldsymbol{\theta}_{\mathrm{GW}}$ is also inverse Wishart:

$$p(\mathbf{\Psi}|\boldsymbol{y}, \boldsymbol{L}, \boldsymbol{\theta}_{\mathrm{GW}}) = \mathrm{IW}(\mathbf{\Psi}_0 + \hat{\mathbf{\Psi}}, \nu_0 + N_f), \qquad (18)$$

where N_f is the number of frequency bins, and $\hat{\mathbf{\Psi}}$ is the 3×3 sample covariance of the eigenstream residuals,

$$\hat{\mathbf{\Psi}} = \tilde{V}_n^{\dagger}(\tilde{y} - \tilde{h})(\tilde{y} - \tilde{h})^{\dagger}\tilde{V}_n. \qquad (19)$$

We implement this step in Python using statistical packages from the SciPy library [24]. Following Ref. [22]'s suggestions, we set $\nu_0 = d + 2$, where $d = 3$ is the dimension of $\mathbf{\Psi}_0$. We choose $\mathbf{\Psi}_0$ to be the median of the frequency bins' sample covariances after a first run obtaining a rough estimate of the eigenstreams $\tilde{e}_n$.

E. Frequency-domain implementation of delays

Up to now, we assumed that time series have a quasi-infinite length so that the asymptotic frequency-domain formulation of the delay operator in Eq. (11) is valid. In practice, we analyze relatively short measurements for which this approximation breaks. Applying Eq. (11) on Fourier-transformed data leads to large edge effects. To mitigate this behavior, we use a time window that smoothly drops to zero at the time series' edges. However, such an operation usually requires transforming the data back to the time domain, which is computationally expensive compared to the usual cost of one likelihood evaluation. Therefore, we perform the equivalent computation in Fourier space (i.e., a discrete convolution) using a sparse approximation of the convolution kernel, similarly as in

covariance approximation techniques [25]. The delay operation amounts to the following matrix multiplication:

$$\tilde{D}_{\text{tap}} = \tilde{\Omega}\tilde{D}, \tag{20}$$

where $\tilde{D}$ is the asymptotic delay operator as given by Eq. (8), and $\tilde{\Omega}$ is the tapered convolution matrix whose elements are given by

$$\tilde{\Omega}_{k,p} = \begin{cases} \sum_{n=0}^{N-1} w(n)e^{-2\pi j n \frac{k-p}{N}} & \text{if } |k-p| \le p_0; \\ 0 & \text{otherwise.} \end{cases} \tag{21}$$

We denoted by $w(n)$ the time-domain window function and p_0 an integer threshold for the row-column difference, above which, the matrix elements are zero.

IV. CASE STUDY

To demonstrate the developed approach's performance, we consider the simple case where phasemeter data only contain a single GW source buried into stationary Gaussian noise. Unless otherwise stated, we assume arbitrary arm lengths and noise PSDs. The parameters governing the estimation model include:

(i) Laser light travel time delays $L = (L_1, L_2, L_3, L_{1'}, L_{2'}, L_{3'})$;

(ii) GW source parameters that we restrict to intrinsic ones $\theta_{\text{GW}} = (\theta, \phi, f_0, \dot{f}_0)$;

(iii) Nonlaser noise covariance parameters Ψ.

We use the likelihood function in Eq. (15) that we maximize over extrinsic GW amplitudes. In this function, the frequency-domain waveform $\tilde{h}$ depends both on θ_{GW} and on L, while the laser noise covariance eigenvectors V_n depend on L only.

A. Simulation parametrization

1. Noise

We simulate one month of LISA observations that yield single-link time series by implementing Eq. (1) with a Python code. We generate noises at a sampling cadence of 2 Hz. We first applied the delays using time-domain Lagrange interpolation filters, with the same parametrization as in LISANode [9] and checked that the PCI algorithm was successfully canceling laser noise. However, we noted that the applied delays' accuracy was not enough to be unnoticed when recovering delays from data simulated over long periods (one month). In other words, the delay values optimally canceling laser noise were slightly biased compared to injected delay values. This mismatch is understandable, as fractional delay filters have a frequency response that only approximates the ideal delay filter [26]. Therefore, we chose to simulate the data used in this study directly in the frequency domain, relying on noise stationarity, following Ref. [27].

We then filter the data using a Kaiser finite-response filter and down sample it to 0.2 Hz to generate the outputs. We assume a rigid, rotating LISA constellation so that the effective arm lengths do not vary in time and that the light travel time is sensitive to the direction of propagation due to the Sagnac effect.

The noise PSD model includes three components: laser frequency noise, test-mass (TM) acceleration noise, and optical metrology system noise. We assume that the noises affecting two different optical benches are uncorrelated so that matrices S_p and S_n are block diagonal. Matrix S_n has six diagonal blocks S_{n_i} of the form,

$$S_{n_i|k,l} = \alpha_i(S_{\text{TM}}(f_k) + S_{\text{OMS}}(f_k))\delta_{kl}, \tag{22}$$

where α_i is a positive coefficient depending on optical bench i. Expressions for noise PSDs $S_{\text{TM}}(f)$ and $S_{\text{OMS}}(f)$ are given in the Appendix A. Thus, noise spectra have the same shape for every optical benches, up to a coefficient accounting for possible noise level discrepancies.

2. Gravitational-wave signal

We assume that the gravitational signal comes from the loudest verification compact galactic binary known to date, called HM Cnc [28,29]. We simulate single-link gravitational-wave signals sampled at 0.2 Hz in the time domain using the same code as in Ref. [30] that we adapted by removing the TDI transfer function. We also relaxed the low-frequency approximation, using a Fourier series decomposition similar to Cornish and Littenberg's implementation in Ref. [31]. For all parameters, we use uniform priors around the true parameter values. The source's characteristics, along with prior boundaries, are summarized in Table I.

B. Projection onto the null space

The first simulation we consider includes single-link measurements where all noises are generated from the same PSD, as given in Appendix A. It also contains one single GW source, as described in Sec. IV A. For the sake of description, here, we assume that the light travel time

TABLE I. Values of the source parameters used in the simulations, with their uniform prior boundaries.

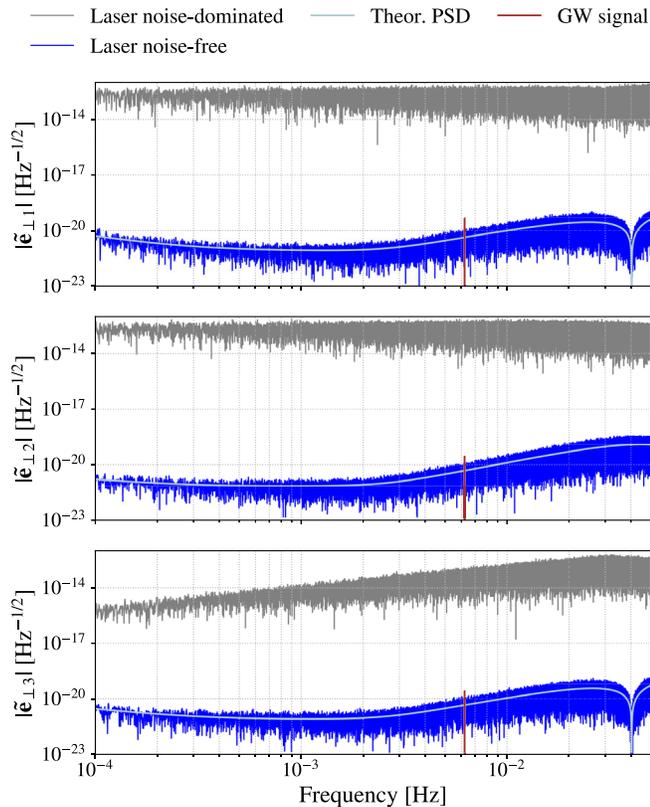| Parameter	Value	Prior range
Frequency [mHz]	6.22	[6.12, 6.32]
Frequency derivative [mHz/s]	7.49×10^{-13}	$[10^{-14}, 10^{-10}]$
Ecliptic latitude [rad]	-0.0821	$[-\pi/2, \pi/2]$
Ecliptic longitude [rad]	2.102	$[-\pi, \pi]$
Amplitude [strain]	6.4×10^{-23}	Marginalized
Inclination [rad]	0.6632	Marginalized
Initial phase [rad]	5.78	Marginalized
Polarization [rad]	3.97	Marginalized

FIG. 2. PCI eigenstreams expressed in relative frequency deviation for the laser-noise free subspace (blue) and the laser-noise dominated subspace (gray) for a month-long simulation. The theoretical PSD function of laser-noise free eigenstreams are plotted in light blue. The red vertical lines denote the gravitational signal from the verification galactic binary HM Cnc that shows up in all channels as a quasimonochromatic signal.

delays L_i/c are known. We compute the null-space eigenvector matrix $\tilde{V}_n$ analytically using Eq. (12) and the other eigenvector matrix $\tilde{V}_p$ numerically. Thanks to these matrices, we apply the PCI transformation to obtain eigenstreams that we orthogonalize as described in Sec. III C. We label as $\tilde{e}_{\perp i}$ these orthogonal streams.

We plot the periodogram of the PCI transformations in Fig. 2, which shows the laser-noise dominated eigenstreams (gray) along with the null space eigenstreams (blue), lying 8 orders of magnitude below. We also plot the GW signal transformed in the null space eigenstreams, which emerges from the noise. We show that this noise is indeed limited by the acceleration and OMS errors by plotting the theoretical PSD (light blue) computed from the diagonal elements of the covariance matrix $\text{Cov}(\tilde{e}_\perp) = \Phi^\dagger \tilde{V}_n^\dagger S_n \tilde{V}_n \Phi$. These plots demonstrate the ability of the frequency-domain algorithm to separate the two orthogonal spaces correctly.

C. Sensitivity analysis

In this section, we investigate the theoretical performance of the PCI approach in two cases: (i) all nonlaser
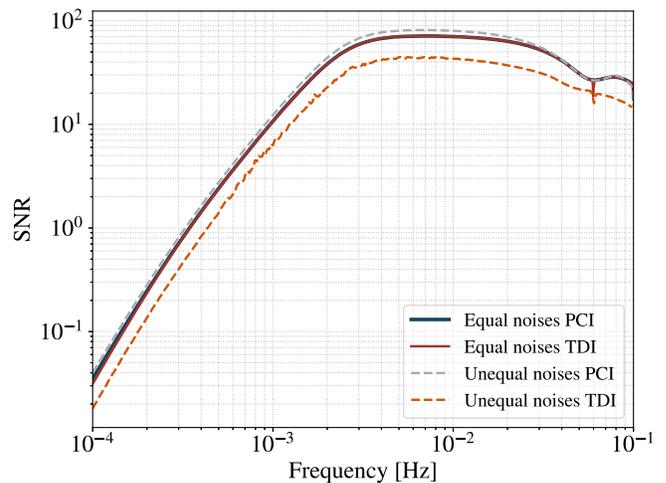

FIG. 3. Sensitivity of laser-noise free channels as a function of frequency obtained with PCI eigenstreams (blue or gray) and unoptimized TDI channels A, E, T (red or orange). The solid lines represent the case of equal noise; dashed lines correspond to unequal noise levels, computed for a source like HM Cnc. PCI and TDI sensitivities are almost identical in the case of equal noises, while unoptimized TDI undergoes a significant SNR loss due to misorthogonalization in the case of unequal noises.

noise levels are the same, i.e., $\alpha_i = 1 \, \forall \, i$ and (ii) nonlaser noise levels are different depending on optical benches, with $\alpha = (4, 0.25, 16, 0.1, 0.4, 1)$. In case (ii), some optical bench noises have larger amplitudes than the baseline, while others have smaller amplitudes. Overall, the mean noise level is larger than in case (i) by a factor 3.6. In Fig. 3, we plot in blue the generalized sensitivity to an ultracompact galactic binary source observed over one year. The source we consider has the same location as HM Cnc, with zero inclination and equal polarization modes, and we allow its frequency to vary. Here, for any frequency f, "sensitivity" refers to the signal-to-noise ratio (SNR) of a source of frequency f measured in its full bandwidth. This calculation takes into account the fact that the covariance matrix is nondiagonal in general, as described in Appendix B. For comparison, we plot in red and orange generalized sensitivities of unoptimized TDI combinations A, E, and T. They are obtained from combining TDI Michelson X, Y, and Z strictly as derived in [16], relying on the assumption of equal noises and equal arm lengths. Thus, we compute associated sensitivities assuming that A, E, T's covariance matrix is perfectly diagonal; hence, we call them "unoptimized."

PCI and unoptimized TDI yield almost the same sensitivity when noises are equal, as shown by the superimposition of the blue and red, solid curves. This similarity confirms that the A, E, T formulation is nearly optimal in this configuration because the assumptions made in their derivation are almost met, except for the equal arm lengths hypothesis, which plays a minor part in the orthogonalization process.

The figure shows that the unequal-noise case (dashed lines) yields a different SNR compared to the case where all noises are equal (solid lines). However, the change in SNR is smaller with the PCI method as it includes the frequency-dependent orthogonalization by design, providing a general extension to classic TDI. For this particular source, the value yielded by the PCI decomposition with unequal noises is even slightly larger than with equal noises. This is because some channels have a larger noise level while others have a smaller one. This comparison illustrates the importance of off-diagonal terms in the covariance $\boldsymbol{C}_n$, which are not equal in the case of heterogeneous noise levels. Thus, not taking this discrepancy into account may result in suboptimal performance.

D. Parameter inference scheme

We further assess the performance of the developed method by using it to recover injected parameters from the numerical simulations described in Sec. IV A. Estimated parameters include light travel time delays, source parameters, and noise covariance. We restrict the inference to a portion of the frequency band between 6.20 and 6.24 mHz around the binary's frequency. We sample the posterior distribution of delays and GW parameters through parallel-tempered Markov Chain Monte Carlo (PTMCMC) sampling, using the PTEMCEE algorithm [32], a parallel-tempered version of the affine-invariant ensemble sampler EMCEE [33].

We modify the sampling algorithm to include noise covariance parameters in the inference, using a two-step Blocked Gibbs sampling scheme where noise parameters are sampled conditionally to delays and GW parameters:

$$\text{Step 1}: \boldsymbol{L}, \boldsymbol{\theta}_{\text{GW}} \sim p(\boldsymbol{L}, \boldsymbol{\theta}_{\text{GW}} | \boldsymbol{y}, \boldsymbol{\Psi}, \boldsymbol{B});$$
$$\text{Step 2}: \boldsymbol{\Psi} \sim p(\boldsymbol{\Psi} | \boldsymbol{y}, \boldsymbol{L}, \boldsymbol{\theta}_{\text{GW}}). \tag{23}$$

While Step 1 is still based on PTMCMC, Step 2 uses direct sampling as described in Sec. III D. We describe sampling results in the next section.

E. Inference results

We present the results of the inference of delays, GW parameters, and covariance parameters applied to the synthetic data described in Sec. IV A with the sampling scheme presented in Sec. IV D. To obtain sufficient precision, we extend the simulation duration to one year, sampled at 20 mHz.

First, we use two data sets: one corresponding to the equal noise case (i) described in Sec. IV C, the other for the unequal noise case (ii). We run 40 chains in parallel with 10 different temperatures, and we retain 4×10^5 samples after chains have reached convergence.

1. Delays

We plot in Fig. 4 the delays posteriors marginalized over other parameters in the case of equal (dashed lines)
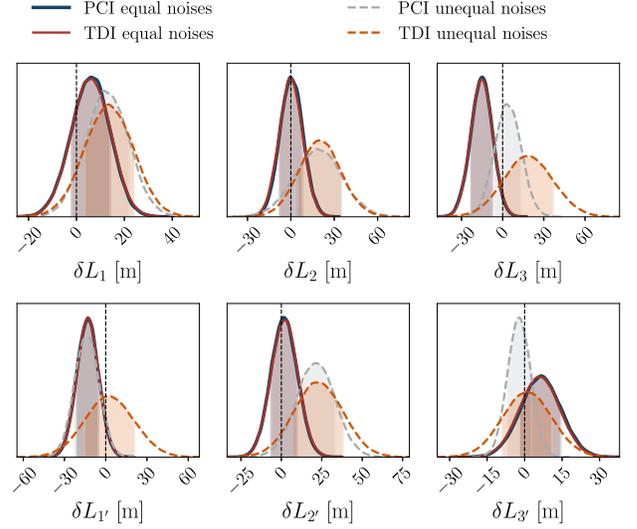

FIG. 4. Posterior distribution of the six light-travel time delays expressed in equivalent arm lengths. Solid lines correspond to the case of equal acceleration noises, whereas dashed lines correspond to unequal noises. Posteriors obtained with PCI are in blue or gray, and posteriors obtained with unoptimized TDI are in red or orange. Thin vertical black dashed lines represent true values, and shaded areas under the curves cover the 1σ region.

and unequal (solid lines) acceleration noises, using the chainConsumer package [34]. We express delays in equivalent interspacecraft distances. In the case of equal noise levels, posteriors obtained from PCI and unoptimized TDI are almost equal to each other, confirming the result found in Fig. 3.

Delays distributions are broader in the case of unequal noises because, in this example, the overall noise power is larger than in the case of equal noises. However, for most delays, posteriors have a larger variance with classic TDI than with the PCI analysis, because PCI accounts for the change of off-diagonal covariance terms, maintaining orthogonalization. This result is consistent with the significant SNR loss shown in Fig. (3) when using unoptimized TDI combinations.

We remark that the uncertainty in estimating equivalent arm lengths is of order 40 m, which is enough to cancel laser noise in this particular case. Should we wish to, we could obtain a better precision in using the entire frequency data instead of restricting it to a narrow band.

2. GW parameters

Then, in Fig. 5, we examine the posterior of the GW source's frequency and frequency derivative, marginalized over all other parameters. Here, we focus on PCI results only, comparing equal (solid blue lines) and unequal (gray dashed lines) noises. The figure shows that the frequency is accurately recovered by the PCI analysis (within about $1 \, \mu\text{Hz}$), even when performed simultaneously with the estimation of laser light delays. We observe a minor
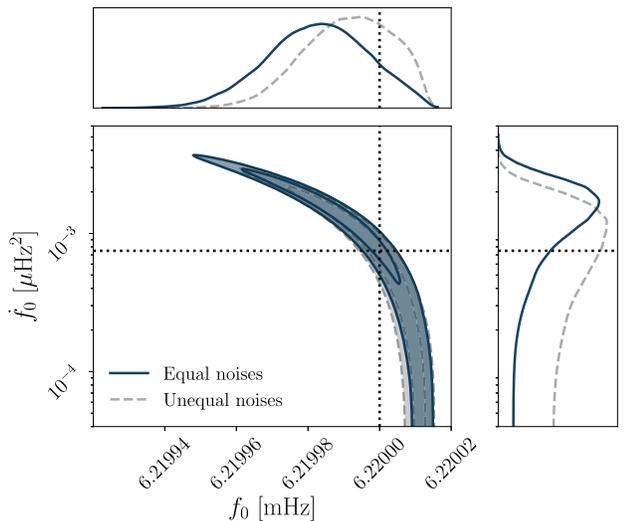
FIG. 5. Joint posterior distribution of GW source's frequency f_0 and frequency derivative $\dot{f}_0$, obtained with PCI in the case of equal (solid blue) and unequal (dashed gray) noises applied to a one-month-long simulation of phasemeter measurements. Contours correspond to 1σ and 2σ regions.

difference between equal and unequal noise cases, showing that the adaptive orthogonalization built in the PCI process minimizes the impact of the noise heterogeneity.

The GW source sky location posteriors that we plot in Fig. 6 exhibit the same behavior as for the frequency parameters. The joint PCI analysis accurately spots the source's location in the sky (dashed black lines), as shown by the maximum *a posteriori* estimate (MAP) represented by the blue (equal noises) and gray (unequal noises)
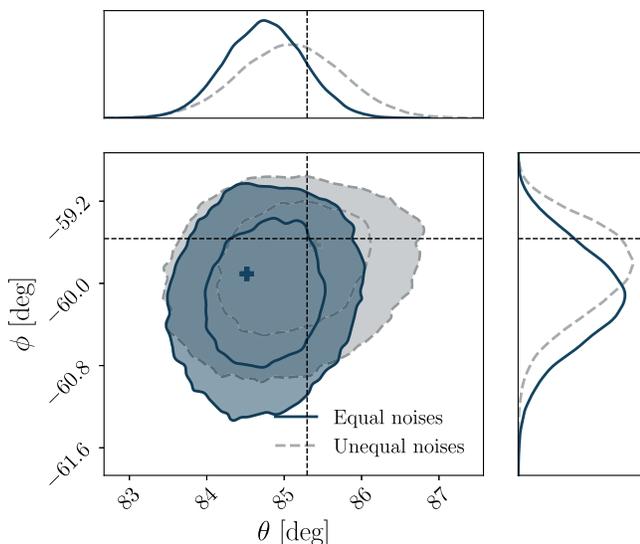

FIG. 6. Sky localization posterior distribution of the GW source in celestial coordinates, obtained with PCI in the case of equal (solid blue) and unequal (dashed gray) noises. The dashed lines represent the true location.

crosses. Hence, the MAP estimate remains close to the actual sky location even in the case of unequal noises.

3. Covariance

We collect the covariance parameter samples computed in the Gibbs steps in Fig. 7. We use them to compute the chains of covariance values. We plot the estimated posterior of the diagonal terms in light green in Fig. 7 as a 3σ interval around the mean, and we compare it to the true value represented by the solid line. The real value is located within the 3σ interval, demonstrating the covariance estimate's accuracy. We obtain similar-looking plots for off-diagonal covariance elements (real and imaginary parts), confirming the accurate characterization of nonlaser noise and the proper orthogonalization of the eigenstreams.

We also plot the periodograms of the three eisgenstreams using the delays' actual values (solid blue line), along with the 3σ interval of the posterior samples. The posterior closely encompasses the target value, showing that the ranging estimates' variability is acceptable.

Finally, we plot the 3σ interval (red shaded area) of the GW waveform samples against the true value of the signal
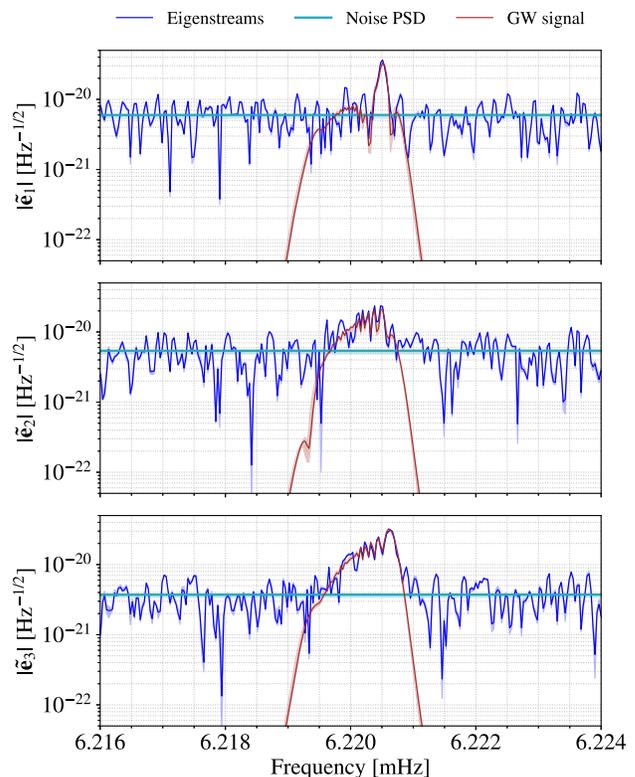

FIG. 7. True value (solid lines) and estimated 3σ posterior interval (light shaded areas) of the eigenstream periodograms (blue), the GW signal (red), and the noise PSDs (diagonal covariance entries) expressed in relative frequency deviation. Posterior intervals are computed using 1000 MCMC samples. The data shown here corresponds to the case of unequal nonlaser noises.

(solid red curve), showing the right consistency between the two. As a result, Fig. 7 provides a summary of the multiparameter inference enabled by the PCI framework.

V. DISCUSSION AND PROSPECTS

We revisited space-based GW data modeling by writing the model likelihood directly as a function of phasemeter measurements. Based on Romano and Woan's idea in Ref. [14], we restricted the likelihood to its lowest variance terms using principal component analysis. This work's main contribution is to formalize the PCA approach in the frequency domain, using an asymptotic formulation of the delay operators with sparse matrices. This formalization provides a framework that is readily applicable to parameter inference, yielding optimal precision. We show that it allows us to handle the TDI transfer function of both signal and noise in a single compact, matrix-based formalism.

With a simplified example of simulated LISA data, we show that the method allows us to consistently and simultaneously fit for interspacecraft phase delays, an ultracompact galactic binary source's parameters, and noise covariance parameters. We show that a numerical and data-driven diagonalization of the covariance yields an optimal sensitivity to gravitational waves and minimal source parameter uncertainty.

This work lays the foundation for a more robust analysis of LISA data. First, it provides a way to derive sensitivities from instrumental noises, tracking all correlations straightforwardly. Second, it generalizes the concept of orthogonal TDI variables to arbitrary arm lengths and noise correlations. However, to apply to in-orbit data, the approach must be extended to time-varying interspacecraft distances. Future work will focus on formalizing this time dependence in the frequency domain to maintain computational efficiency. Furthermore, we can extend the developed framework to a complete set of LISA measurements by dropping the assumption of identical spacecraft's laser noise and include reference interferometer measurements. Preliminary work has already demonstrated the successful decomposition into large and low variance components in this configuration, that we will present in followup studies.

Finally, our work highlights the importance of L0 data including raw phasemeter measurements in LISA's science analysis. The availability of such data will ensure that alternative processing methods complementary to the standard TDI pipeline are possible. This diversity of approaches is an essential tool for cross-checking and validating the data in an off-line interferometry step that is crucial for the precise characterization of astrophysical sources.

APPENDIX A: EXPRESSIONS OF NOISE PSD MODEL

The noise PSD models used in this study are given in fractional frequency per Hertz (Hz^{-1}) as

$$S_{\text{laser}}(f) = \left(\frac{a_0}{\nu_0}\right)^2;$$

$$S_{\text{TM}}(f) = \left(\frac{a_{\text{TM}}}{2\pi f c}\right)^2 \left(1 + \left(\frac{f_1}{f}\right)^2\right)\left(1 + \left(\frac{f}{f_2}\right)^4\right);$$

$$S_{\text{OMS}}(f) = a_{\text{OMS}}^2 \left(\frac{2\pi f}{c}\right)^2 \left(1 + \left(\frac{f_3}{f}\right)^4\right). \tag{A1}$$

We indicate the values of the noise model parameters in Table II.

APPENDIX B: COMPUTATION OF SENSITIVITY TO MONOCHROMATIC BINARIES

In this section, we derive the expression for the generalized SNR plotted in Fig. 3, which extends the classic SNR calculation to the case where the noise covariance used for the estimation is different than the correct one.

Let us consider some linear transformation of the single-link measurements, encoded by a $6N_f \times 3N_f$ matrix $\boldsymbol{W}$ that can represent any TDI or PCI transformation in the frequency domain, over a bandwidth of N_f frequency bins. Applying $\boldsymbol{W}$ to the phasemeter measurements vector $\tilde{\boldsymbol{y}}$ of size $6N_f$ yields the vector $\tilde{\boldsymbol{e}}$ of size $3N_f$.

We define the generalized SNR $\rho_{\boldsymbol{W}}(f_0)$ of a monochromatic GW of frequency f_0 obtained with the data transformation $\boldsymbol{W}$ as the ratio between the absolute value of its gravitational-wave strain amplitude h_0 and the standard deviation σ_0 of its maximum likelihood estimate,

TABLE II. Values of the noise parameters used in the simulations.

Noise type	Parameter	Value
Laser	a_0	$28.2 \ \text{Hz} \cdot \text{Hz}^{-1/2}$
	ν_0	$281759 \ \text{GHz}$
Test mass	a_{TM}	$3 \ \text{fm s}^{-2} \, \text{Hz}^{-1/2}$
	f_1	$0.4 \ \text{mHz}$
	f_2	$8 \ \text{mHz}$
OMS	a_{OMS}	$15 \ \text{pm Hz}^{-1/2}$
	f_3	$2 \ \text{mHz}$

$$\rho_W(f_0) \equiv \frac{|h_0|}{\sigma(f_0)}. \tag{B1}$$

We assume that the GW signal appears in the single-link measurements $\tilde{y}$ as

$$\tilde{y} = \tilde{A}h_0, \tag{B2}$$

where $\tilde{A}$ is a $6N_f \times 1$ design matrix. From Eq. (3), this transformation maps the GW amplitude as

$$\tilde{e} = W^\dagger \tilde{A}h_0 + \tilde{\epsilon}, \tag{B3}$$

where $\tilde{\epsilon}$ is the noise contribution to the data stream vector $\tilde{e}$. The effective $3N_f \times 1$ transfer function matrix is thus $H \equiv W^\dagger \tilde{A}$.

Let us assume that we analyze the data with the following likelihood model:

$$p(\tilde{e}|h_0) = \frac{\exp\left\{-\frac{1}{2}(\tilde{e} - Hh_0)^\dagger C^{-1}(\tilde{e} - Hh_0)\right\}}{\sqrt{(2\pi)^{6N}|C|}}, \tag{B4}$$

where C is some model of the covariance of $\tilde{e}$. Assuming that C is fixed, the maximum likelihood estimator error on h_0 from data in the bandwidth of size N_f is then given by the inverse Fisher matrix, which is defined as

$$\sigma_0^2 = \mathrm{E}\left[\left(\frac{\partial \log p}{\partial h_0}\right)^2\right]^{-1}. \tag{B5}$$

Inserting Eq. (B4) into Eq. (B5) yields

$$\sigma_0^2 = \frac{H^\dagger C^{-1} \Sigma_e C^{-1} H}{(H^\dagger C^{-1} H)^2}, \tag{B6}$$

where Σ_e is the true variance of data streams $\tilde{e}$, and is given by

$$\Sigma_e = W^\dagger \tilde{\Sigma} W, \tag{B7}$$

where $\tilde{\Sigma}$ is the covariance of $\check{y}$. We can remark two properties:

  (i) If W cancels laser noise, then $\Sigma_e = W^\dagger S_n W$.
 (ii) If W is an orthogonal transformation with respect to the noise, then Σ_e is diagonal by construction.
(iii) If $C = \tilde{\Sigma}_e$ (i.e., the model covariance is equal to the true one), then $\rho_W(f_0) = |h_0|\sqrt{H^\dagger \Sigma_e^{-1} H}$, which is exactly equivalent to the standard SNR formula used in the gravitational-wave literature.

[1] J. Aasi, J. Abadie, B. P. Abbott, R. Abbott, T. D. Abbott *et al.*, Enhanced sensitivity of the LIGO gravitational wave detector by using squeezed states of light, Nat. Photonics **7,** 613 (2013).

[2] B. P. Abbott, R. Abbott, T. D. Abbott, M. R. Abernathy, K. Ackley *et al.*, Exploring the sensitivity of next generation gravitational wave detectors, Classical Quantum Gravity **34,** 044001 (2017).

[3] P. K. Dahal, Review of pulsar timing array for gravitational wave research, J. Astrophys. Astron **41,** 1 (2020).

[4] P. Amaro-Seoane, H. Audley, S. Babak, J. Baker, E. Barausse *et al.*, Laser Interferometer Space Antenna, *European Space Agency Proposal* (2017), pp. 1–41.

[5] 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).

[6] J. W. Armstrong, F. B. Estabrook, and M. Tinto, Time-delay interferometry for spacebased gravitational wave searches, Astrophys. J. **527,** 814 (1999).

[7] M. Tinto and S. V. Dhurandhar, Time-delay interferometry, Living Rev. Relativity **17,** 6 (2014).

[8] D. A. Shaddock, B. Ware, R. E. Spero, and M. Vallisneri, Postprocessed time-delay interferometry for LISA, Phys. Rev. D **70,** 081101 (2004).

[9] J.-B. Bayle, M. Lilley, A. Petiteau, and H. Halloin, Effect of filters on the time-delay interferometry residual laser noise for LISA, Phys. Rev. D **99,** 084023 (2019).

[10] O. Hartwig and J.-B. Bayle, Clock-jitter reduction in LISA time-delay interferometry combinations, arXiv:2005.02430.

[11] M. Muratore, D. Vetrugno, and S. Vitale, Revisitation of time delay interferometry combinations that suppress laser noise in LISA, arXiv:2001.11221.

[12] K. R. Nayak and J. Y. Vinet, Algebraic approach to time-delay data analysis for orbiting LISA, Phys. Rev. D **70,** 16 (2004).

[13] M. Vallisneri, J.-B. Bayle, S. Babak, and A. Petiteau, TDI-infinity: Time-delay interferometry without delays, arXiv:2008.12343.

[14] J. D. Romano and G. Woan, Principal component analysis for LISA: The time delay interferometry connection, Phys. Rev. D **73,** 102001 (2006).

[15] M. D. Leighton, A principal component approach to space-based gravitational wave astronomy, Technical Report, Glasgow U., 2016.

[16] T. A. Prince, J. W. Armstrong, M. Tinto, and S. L. Larson, LISA optimal sensitivity, Phys. Rev. D **66,** 122002 (2002).

[17] T. E. Oliphant, *A Guide to NumPy* (Trelgol Publishing, USA, 2006).

[18] S. Van Der Walt, S. C. Colbert, and G. Varoquaux, The NumPy array: A structure for efficient numerical computation, Comput. Sci. Eng. **13,** 22 (2011).

[19] J. Kopp, Efficient numerical diagonalization of hermitian 3×3 matrices Int. J. Mod. Phys. C **19,** 523 (2008).

[20] F. B. Estabrook, M. Tinto, and J. W. Armstrong, Time-delay analysis of LISA gravitational wave data: Elimination of spacecraft motion effects, Phys. Rev. D **62,** 042002 (2000).

[21] M. Dai and W. Guo, Multivariate spectral analysis using Cholesky decomposition, Technical Report 3, 2004.

[22] P. D. Hoff and X. Niu, A covariance regression model, Statistica Sinica **22,** 729 (2012).

[23] T. Littenberg, N. Cornish, K. Lackeos, and T. Robson, Global Analysis of the Gravitational Wave Signal from Galactic Binaries, Phys. Rev. D **101,** 123021 (2020).

[24] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy *et al.*, SciPy 1.0: Fundamental algorithms for scientific computing in Python, Nat. Methods **17,** 261 (2020).

[25] R. Furrer, M. G. Genton, and D. Nychka, Covariance tapering for interpolation of large spatial datasets, J. Comput. Graph. Stat. **15,** 502 (2006).

[26] C. C. Tseng and S. L. Lee, Design of fractional delay filter using discrete Fourier transform interpolation method, Signal Process. **90,** 1313 (2010).

[27] J. Timmer and M. Koenig, On generating power law noise, Astron. Astrophys. **3,** 707 1995.

[28] A. Stroeer and A. Vecchio, The LISA verification binaries, Classical Quantum Gravity **23,** S809 (2006).

[29] T. Kupfer, V. Korol, S. Shah, G. Nelemans, T. R. Marsh, G. Ramsay, P. J. Groot, D. T. H. Steeghs, and E. M. Rossi, LISA verification binaries with updated distances from Gaia Data Release 2, Mon. Not. R. Astron. Soc. **480,** 302 (2018).

[30] Q. Baghi, J. I. Thorpe, J. Slutsky, J. Baker, T. D. Canton, N. Korsakova, and N. Karnesis, Gravitational-wave parameter estimation with gaps in LISA: A Bayesian data augmentation method, Phys. Rev. D **100,** 022003 (2019).

[31] N. J. Cornish and T. B. Littenberg, Tests of Bayesian model selection techniques for gravitational wave astronomy, Phys. Rev. D **76,** 083006 (2007).

[32] W. Vousden, W. M. Farr, and I. Mandel, Dynamic temperature selection for parallel-tempering in Markov chain Monte Carlo simulations Mon. Not. R. Astron. Soc. **455,** 1919 (2015)..

[33] D. Foreman-Mackey, D. W. Hogg, D. Lang, and J. Goodman, emcee: The MCMC Hammer, Publ. Astron. Soc. Pac. **125,** 925 (2012).

[34] S. Hinton, Chain-Consumer, J. Open Source Software **1,** 45 (2016).