GraFIT

A toolbox for fast and accurate frequency response identification in gravitational wave detectors

van Dael, M.; van Haren, M.; Witvoet, G.; Swinkels, B.; Oomen, T.

DOI
10.1063/5.0275060

Publication date
2025

Document Version
Final published version

Published in
Review of Scientific Instruments

Citation (APA)
van Dael, M., van Haren, M., Witvoet, G., Swinkels, B., & Oomen, T. (2025). GraFIT: A toolbox for fast and accurate frequency response identification in gravitational wave detectors. *Review of Scientific Instruments*, *96*(10), Article 104503. https://doi.org/10.1063/5.0275060

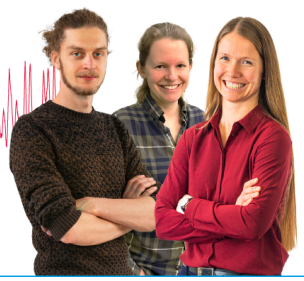
Review of Scientific Instruments

RESEARCH ARTICLE | OCTOBER 06 2025

GraFIT: A toolbox for fast and accurate frequency response identification in gravitational wave detectors

M. van Dael ; M. van Haren ; G. Witvoet ; B. Swinkels ; T. Oomen

Rev. Sci. Instrum. 96, 104503 (2025)
https://doi.org/10.1063/5.0275060



Review of Scientific Instruments ARTICLE pubs.aip.org/aip/rsi

GraFIT: A toolbox for fast and accurate frequency response identification in gravitational wave detectors

Cite as: Rev. Sci. Instrum. 96, 104503 (2025); doi: 10.1063/5.0275060
Submitted: 10 April 2025 • Accepted: 12 September 2025 •
Published Online: 6 October 2025

M. van Dael,[1,2,a] M. van Haren,[1] G. Witvoet,[1,3] B. Swinkels,[2] and T. Oomen[1,4]

AFFILIATIONS

[1] Eindhoven University of Technology, Department of Mechanical Engineering, Control Systems Technology, Eindhoven, The Netherlands
[2] Nikhef, Amsterdam, The Netherlands
[3] TNO, Optomechatronics Department, Delft, The Netherlands
[4] Delft Center for Systems and Control, Delft University of Technology, Delft, The Netherlands

[a] **Author to whom correspondence should be addressed:** m.r.v.dael@tue.nl

ABSTRACT

Frequency Response Function (FRF) measurements are widely used in gravitational wave detectors, e.g., for the design of controllers, calibrating signals, and diagnosing problems with system dynamics. The aim of this paper is to present GraFIT: a toolbox that enables fast, inexpensive, and accurate identification of FRF measurements for gravitational wave detectors compared to the commonly used approaches, including common spectral analysis techniques. The toolbox builds upon recent developments in non-parametric system identification by utilizing a local modeling technique, which is particularly effective in reducing the impact of transient effects. It is furthermore designed to be user-friendly, handling systems of arbitrary input–output dimensions, and systems operating in a closed loop. The toolbox is validated on two experimental case studies of the Virgo detector, illustrating more than a factor 3 reduction in the estimated standard deviation using GraFIT for the same measurement times and comparable estimated standard deviations with up to ten times less data using GraFIT with respect to the commonly used spectral analysis method. As a result, GraFIT can reduce commissioning time and detector downtime due to noise injections, while also improving the overall quality of the measurements.

I. INTRODUCTION

Frequency domain models are essential in Gravitational Wave (GW) detectors, such as Virgo[1] and LIGO,[2] for a variety of purposes, e.g., control design,[3–6] calibration of signals,[7,8] and obtaining noise budgets for the sensitivity of the detector.[9,10] Frequency Response Function (FRF) measurements are commonly used to represent system dynamics in the frequency domain because they are accurate, user-friendly, and inexpensive to obtain (Ref. 11). Identifying an FRF requires the user to choose the perturbation signal and the identification method. Non-periodic perturbation signals (typically filtered white noise) combined with the Spectral Analysis (SA)[11,12] method have been the baseline for FRF identification in almost all application domains, including the GW community. The SA method can provide accurate estimates, but this requires averaging over a sufficient number of data segments, improving the accuracy at the expense of the frequency resolution. Essentially, if sufficient measurement data are taken, sufficient accuracy can be obtained.

In the more recent literature, new methods for FRF identification have been proposed,[13–15] among which is the Local Polynomial Method (LPM).[16] LPM has been developed to better address transient errors, resulting from dynamic transients or leakage effects, by leveraging the observation that such errors exhibit smooth frequency characteristics [Ref. 11, Appendix 6.(b)]. By explicitly estimating this transient effect, the LPM effectively suppresses it in the output signal. The LPM models the transient in the frequency domain using a polynomial function, a method later extended to rational models in the Local Rational Method (LRM).[17] Compared to the SA method

which uses windowing,[18] the LPM has shown significant improvements in reducing leakage effects.[19] The LPM has furthermore been shown to be very effective for systems with large dynamic transients, such as thermal systems[20] and LRM for mechanical systems with lightly damped resonant dynamics.[21] Finally, the local modeling approach has been shown to be very data-efficient,[21,22] allowing for shorter measurements times, which is of particular interest to the GW community, where experimentation time may be scarce and may go at the expense of observation time.

Although the identification of FRFs is standard in GW detectors, the main goal of this paper is to obtain much more data-efficient FRF models that have a better accuracy vs data size ratio. In this paper, the Gravitational waves Advanced Frequency response Identification Toolbox (GraFIT) is therefore presented, which uses the LPM/LRM method to identify the FRFs for GW detectors. The toolbox is validated on two experimental case studies of the Virgo detector and is furthermore designed to be user-friendly as it is tailored for direct use in the GW community, by implementing recent developments from the control community in a GW-relevant context.

This paper is organized as follows: In Sec. II, the application setting is discussed to highlight the importance of FRF identification in the operation of GW detectors. In Sec. III, a recap of the pre-existing approach for FRF identification in GW detectors is given and the newly proposed local modeling method is introduced. Section IV then presents the GraFIT toolbox, and the toolbox is compared to the pre-existing method in Sec. V using two experimental case studies of the Virgo detector. Finally, conclusions on the work are given in Sec. VI.

II. APPLICATION DOMAINS

GW detectors, such as Virgo, are kilometer-scale interferometers consisting of a vast number of optics to detect differences in the arm lengths in the order of 1×10^{-18} m. FRF identification plays an important role in various aspects of GW detectors; three of such application domains are discussed next.

A. Suspension systems for optics

Suspension systems, such as the ones in Virgo,[6,23] are critical to minimizing the motion of the optics due to ground motion. These suspensions use low-stiffness isolators to passively isolate ground motion, reducing, for example, the mirror motion by over 15 orders of magnitude above 10 Hz.[23] To damp the typically lowly damped modes and thus to minimize the Root-Mean-Square (RMS) motion, feedback controllers are used. An FRF of the system is typically sufficient for the design of these controllers, and identification methods for obtaining these FRFs are therefore typically preferred over physical models since they are fast and inexpensive to obtain and less prone to errors compared to physical models.

B. Relative mirror position control

Feedback control for the relative distances between the optics is essential to achieve the required sensitivity of the detector.[24] Obtaining models for the control design by accurately simulating the system dynamics is difficult since imperfections in the optics are difficult to model and the system exhibits time-varying behavior.[25] Instead, the FRF is measured by having a dedicated experiment, where the optics are perturbed and the subsequent response is measured. If this experiment is shorter than the timescale of the time-varying behavior, the system can be approximated as a linear time-invariant system. Since this experiment requires downtime of the machine, an identification method that requires fewer data, while maintaining similar accuracy to the classical identification method, is preferred to minimize the downtime of the detector.

C. Noise budget

Noise budgets are used to identify the limiting disturbances at each frequency for the detector sensitivity.[9,10] To determine the contribution of a disturbance, both a model of the disturbance and an FRF of the coupling to the sensitivity are required. Identifying the FRF is often preferred over modeling due to the complexity of the system and the time-varying behavior. The quality of the noise projection heavily depends on the quality of the measured FRF, thus requiring accurate FRFs.

III. METHODOLOGY

This section presents the methodology used in GraFIT and the baseline method for obtaining FRFs in GW detectors. First, the identification problem is formalized, after which the baseline approach is briefly revisited. Then, the local modeling approach as used in GraFIT is presented, followed by a brief discussion on the different approaches between the two methods.

A. Identification problem

Consider the basic identification problem as shown in Fig. 1. The objective is to obtain the FRF of the dynamic system G using the input signal $r(n)$ and the output signal $y(n)$, which is perturbed by the unknown disturbance $v(n)$, with $n = 0, 1, \ldots, N-1$, and N being the number of samples.

There are different possibilities for the perturbation signal $r(n)$, but here the standard approach for GW detectors is considered, where filtered white noise as external perturbation for $r(n)$ is used. The noisy output $y(n)$ is measured, and the Discrete Fourier Transform (DFT) of the input and output signals is computed through

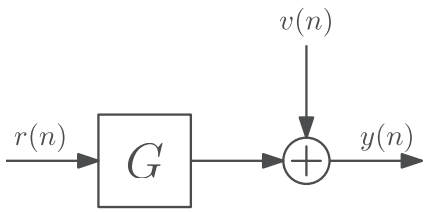


FIG. 1. Standard open-loop identification problem, where the goal is to identify G using the input signal $r(n)$ and the noisy output signal $y(n)$, which is perturbed by a disturbance $v(n)$.

$$M(k) = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} m(n) e^{-j\omega_k n}, \tag{1}$$

with $m = r, y, v$ and $M(k) = R(k), Y(k), V(k)$ and

$$\omega_k = \frac{2\pi k}{N}, \tag{2}$$

where ω_k represents the frequency, with $k \in \mathbb{Z}_{[0,N/2]}$ indexing the discrete frequencies obtained from an N-point DFT of real-valued data. The following input–output relation is then obtained:

$$Y(k) = G(\Omega_k)R(k) + T(\Omega_k) + V(k), \tag{3}$$

where $\Omega_k = e^{j\omega_k T_s}$ represents the angular frequency, $G(\Omega_k)$ represents the dynamic response to a forced input $R(k)$, and $V(k)$ represents the noise.

The additional term $T(\Omega_k)$ represents transient effects resulting from the difference between the initial and final state of the system. An intuitive interpretation of the transient term is given in McKelvey and Guérin[17] and Evers *et al.*,[26] where the state-space representation of a dynamic system is considered,

$$\begin{aligned} x(n+1) &= Ax(n) + Br(n), \\ y(n) &= Cx(n) + Dr(n), \end{aligned} \tag{4}$$

with $x(n) \in \mathbb{R}^{n_x}$ representing the state of the system of dimension n_x. Applying the N-point DFT of (1)–(4) and subsequently eliminating the state $x(n)$ yield the following representations for $G(\Omega_k)$ and $T(\Omega_k)$ in (3):

$$G(\Omega_k) = D + C\left(e^{j\omega_k} I - A\right)^{-1} B \tag{5}$$

and

$$T(\Omega_k) = C\left(e^{j\omega_k} I - A\right)^{-1} (x(0) - x(N)) e^{j\omega_k}, \tag{6}$$

which shows that $T(\Omega_k) \neq 0$ if the initial and final states of the system are not equal. In practice, the two primary sources of a transient term $T(\Omega_k)$ are dynamic transients and leakage. Dynamic transients occur when the system begins from a non-zero initial state, leading to a free (unforced) response of the system, which typically goes to zero. Leakage arises due to the data not being periodic within the finite dataset of length N, which inherently results in $x(0) - x(N) \neq 0$ and thus a non-zero transient term. Note that while both dynamic transients and leakage produce a transient term $T(\Omega_k)$, dynamic transients are a system aspect, which depends on the system dynamics and initial conditions, whereas leakage is an artifact of the data processing method and thus not inherently a system property.

The goal in this section is to obtain an estimate of the FRF $G(\Omega_k)$ using minimal input and output data, $r(n)$ and $y(n)$, respectively, which requires effectively handling both the transient effect $T(\Omega_k)$ and noise term $V(k)$.

B. SA estimator

The most common approach, also in GWs, to obtain an estimate of $G(\Omega_k)$ is the SA method. The method obtains an estimate of $G(\Omega_k)$ by splitting the data into P segments and computing

$$G^{\text{sa}}(\Omega_k) \approx \frac{\sum_{i=1}^{P} Y_i(k)\overline{R}_i(k)}{\sum_{i=1}^{P} R_i(k)\overline{R}_i(k)}, \tag{7}$$

where $Y_i(k)\overline{R}_i(k)$ and $R_i(k)\overline{R}_i(k)$ are the cross- and auto-power spectral densities of the ith segment, respectively, with $\overline{R}_i(k)$ representing the complex conjugate of $R_i(k)$. By averaging over the segments, the variance on the estimate due to the noise is reduced, with more averages leading to a lower variance [see Pintelon and Schoukens (Ref. 11, Sec. 2.6.2) for an expression of the bias and variance of the estimator]. To reduce the effect of the transient term $T(\Omega_k)$, windowing (e.g., Von Hann)[18] or detrending is typically applied.

C. Local model estimator

An illustration of the local modeling method is shown in Fig. 2. The key concept is to estimate the output DFT $Y(k)$ in a local frequency window $l \in \mathbb{Z}_{[-W,\,W]}$ of size $2W + 1$ around the frequency bin Ω_k, i.e.,

$$\widehat{Y}(k+l, \widehat{\Theta}(k)) = \widehat{G}(\Omega_{k+l}, \widehat{\Theta}(k))R(k+l) + \widehat{T}(\Omega_{k+l}, \widehat{\Theta}(k)), \tag{8}$$

by parameterizing the system dynamics $G(\Omega_k)$ and transient term $T(\Omega_k)$ as rational models of frequency, i.e., $\widehat{G}(\Omega_{k+l}, \widehat{\Theta}(k))$ and $\widehat{T}(\Omega_{k+l}, \widehat{\Theta}(k))$, with $\widehat{\cdot}$ denoting that it is an estimate of the respective variable. The model parameters $\widehat{\Theta}(k)$ are the estimated coefficients of the models in this local frequency window, and they

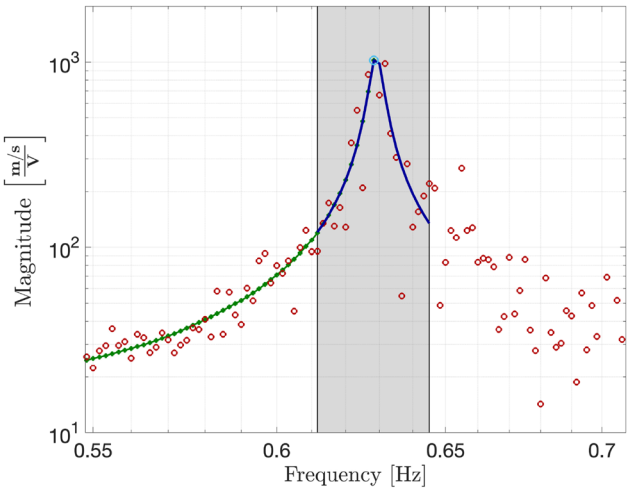


FIG. 2. Illustration of the local modeling method, where (○) represents $Y(k)/R(k)$, (▒) represents the local frequency window in which the model is estimated, (—) represents the fit of the local model $\widehat{G}(\Omega_{k+l}, \widehat{\Theta}(k))$ on the data, (○) represents the estimate of $G(\Omega_k)$ extracted from the fitted model, and (—○—) represents estimates of $G(\Omega_k)$ for previous frequency bins Ω_k. The local frequency window slides over the frequency grid to obtain a local model and subsequent estimate of $G(\Omega_k)$ for each frequency Ω_k.

are obtained by minimizing the following Least-Squares (LS) cost function:

$$J\left(\widehat{\Theta}(k)\right) = \sum_{l=-W}^{W} \left|Y(k+l) - \widehat{Y}\left(k+l,\, \widehat{\Theta}(k)\right)\right|^2. \tag{9}$$

To find the FRF estimate $\widehat{G}(\Omega_k)$, the estimated model $\widehat{G}(\Omega_{k+l}, \widehat{\Theta}(k))$ is evaluated at the center point $l = 0$. The additional advantage of also estimating $\widehat{T}(\Omega_{k+l}, \widehat{\Theta}(k))$ is that this suppresses the transient effects in the output DFT $Y(k)$. This estimation approach is repeated for each Ω_k, resulting in the local frequency window essentially sliding over the frequency grid to obtain a local model and subsequent estimate of $G(\Omega_k)$ for each frequency Ω_k.

To fit a local model in the frequency window, both $\widehat{G}(\Omega_{k+l}, \widehat{\Theta}(k))$ and $\widehat{T}(\Omega_{k+l}, \widehat{\Theta}(k))$ are parameterized as rational models of frequency. The local model for the system dynamics is parameterized as

$$\widehat{G}(\Omega_{k+l}, \widehat{\Theta}(k)) = \frac{A_{k+l}}{D_{k+l}}, \tag{10}$$

where both A_{k+l} and D_{k+l} are polynomial functions of frequency given by

$$A_{k+l} = \widehat{G}(\Omega_k) + \sum_{i=1}^{L_a} a_i(k) l^i \tag{11}$$

and

$$D_{k+l} = 1 + \sum_{i=1}^{L_d} d_i(k) l^i. \tag{12}$$

Similarly, the transient term is parameterized as

$$\widehat{T}(\Omega_{k+l}, \widehat{\Theta}(k)) = \frac{B_{k+l}}{D_{k+l}}, \tag{13}$$

with

$$B_{k+l} = \widehat{T}(\Omega_k) + \sum_{i=1}^{L_b} b_i(k) l^i. \tag{14}$$

Note that $\widehat{G}(\Omega_{k+l})$ and $\widehat{T}(\Omega_{k+l})$ are parameterized to have the same denominator, since both $G(\Omega_k)$ in (5) and $T(\Omega_k)$ in (6) share the same poles and thus denominator of their transfer function when they originate from the same dynamic system. The model parameters are the coefficients of the polynomials, forming the parameter vector $\Theta(k) \in \mathbb{C}^{n_\Theta}$ given by

$$\begin{aligned} \Theta(k) &= [\Theta_A(k) \;\; \Theta_B(k) \;\; \Theta_D(k)], \\ \Theta_A(k) &= \left[\widehat{G}(\Omega_k) \;\; a_1(k) \;\; \ldots \;\; a_{L_a}(k)\right], \\ \Theta_B(k) &= \left[\widehat{T}(\Omega_k) \;\; b_1(k) \;\; \ldots \;\; b_{L_b}(k)\right], \\ \Theta_D(k) &= [d_1(k) \;\; \ldots \;\; d_{L_d}(k)], \end{aligned} \tag{15}$$

with $n_\Theta = L_a + L_b + L_d + 2$ representing the number of parameters in the local model.

The current formulation of $\widehat{Y}(k+l, \widehat{\Theta}(k))$ in (8) yields a parameterization in the cost function in (9), which is not linear in Θ. While there are several approaches to solving this LS problem, see e.g., Voorhoeve *et al.*,[21] the approach used here is the most straightforward and typically leads to satisfactory results. By substituting the modeled output of (8) in (9) and weighting the cost function with the denominator D_{k+l}, a cost function is obtained that is linear in the parameters Θ, i.e.,

$$J\left(\widehat{\Theta}(k)\right) = \sum_{l=-W}^{W} \left|D_{k+l} Y(k+l) - A_{k+l} R(k+l) + B_{k+l}\right|^2. \tag{16}$$

This cost function can be rewritten to the linear LS problem,

$$J\left(\widehat{\Theta}(k)\right) = \sum_{l=-W}^{W} \left|Y(k+l) - \widehat{\Theta}(k) K(k+l)\right|^2, \tag{17}$$

with

$$K(k+l) = \begin{bmatrix} K_a(l) & \otimes R(k+l) \\ K_b(l) & \\ -K_d(l) & \otimes Y(k+l) \end{bmatrix}, \tag{18}$$

with $\otimes$ denoting the Kronecker product and

$$\begin{aligned} K_a(l) &= \left[1 \; l \; \ldots \; l^{L_a}\right]^T, \\ K_b(l) &= \left[1 \; l \; \ldots \; l^{L_b}\right]^T, \\ K_d(l) &= \left[l \; \ldots \; l^{L_d}\right]^T. \end{aligned} \tag{19}$$

The solution to the linear LS in (17) is then given by

$$\widehat{\Theta}(k) = \underset{\Theta(k)}{\operatorname{argmin}} J\left(\widehat{\Theta}(k)\right) = Y_W(k) K_W^H(k) \left(K_W(k) K_W^H(k)\right)^{-1}, \tag{20}$$

with the output DFT $Y_W = \left[Y(k-W) \; \ldots \; Y(k+W)\right]^T$ and regressor matrix $K_W(k) = \left[K(k-W) \; \ldots \; K(k+W)\right]$ evaluated over all frequencies in the window W and $\cdot^H$ denoting the conjugated transpose of a matrix.

D. SA vs LRM estimator

The SA and LRM estimators differ significantly in their approach toward minimizing the effect of the additional noise and transient term on the estimate of $G(\Omega_k)$.

Both the SA and LRM estimators use an averaging approach to reduce the effect of the noise term $V(k)$. The SA method does this by splitting the data into segments, essentially obtaining multiple estimates per frequency bin and averaging over them. The LRM method instead uses a local frequency window of size $2W + 1$ and estimates the set of n_Θ parameters by having at least as many frequency bins in the window as parameters, where a larger window than the number of parameters, i.e., $2W + 1 > n_\Theta$, essentially allows a form of averaging. While both methods effectively handle the noise term, the downside of the SA method is that to obtain more segments, either longer datasets have to be measured or the data have to be split into more segments, which reduces the frequency resolution proportionally by the number of segments.

The approach toward minimizing the transient effect $T(\Omega_k)$ is also fundamentally different. To minimize leakage effects, the SA method uses windowing, while the LRM method explicitly estimates $T(\Omega_k)$, which has been shown to be a more effective approach toward handling leakage effects.[19] Large dynamic transients are also more effectively handled by the LRM method,[20] as it captures the transient behavior in the estimated $T(\Omega_k)$ term, while the SA method ignores this effect. In essence, the SA method requires approaches, such as windowing and detrending, which effectively alter the data to reduce the transient effect, while the LRM method explicitly estimates the transient effect and thus can distinguish transient contributions from the system dynamics in the output.

IV. GraFIT

This section presents the GraFIT toolbox, which implements the LRM method for FRF identification in GW detectors. First, the GraFIT toolbox and its use are presented. Then, some extensions in the toolbox are presented, such as indirect identification and uncertainty estimation, and finally, a brief guide on parameter selection is given.

A. GraFIT

GraFIT consists of a single function, written in both MATLAB and Python, which performs FRF identification for systems operating in both open- and closed-loops and for arbitrary input and output dimensions. The toolbox is specifically developed for the identification approach commonly used in Virgo, where filtered white noise is consecutively injected into each degree of freedom of the system. The identification inside the toolbox is then performed in a Single-Input Multiple-Output (SIMO) setting, where the estimate of $G(\Omega_k)$ is directly computed from the input $r(n)$ to the n_y output signals $y(n)$ of the system. When performing SIMO experiments, the total number of parameters estimated is given by

$$n_\Theta = L_a + L_b + L_d \cdot n_y + 2, \tag{21}$$

since n_y denominator terms D_{k+l} are estimated. This SIMO approach is repeated for the n_r input signals $r(n)$ of the system to find the Multiple-Input Multiple-Output (MIMO) estimate of $G(\Omega_k)$. In Virgo, the noise is typically injected on the level of the error signal when the system is operating in a closed-loop, but the toolbox can also be used when the noise is injected on the level of the correction signal, i.e., the signal going into $G(\Omega_k)$.

The example code here is all in MATLAB and the usage of the function is shown in Listing 1.

LISTING 1. GraFIT usage.

```
[G,G_var,Gry,Gru,Gry_var,Gru_var,Y_contr] = ...
    GraFIT(r,y,u,W,freqIdentBand,Fs,L)
```

For systems operating in an open-loop, the variable r is the perturbation signal and y is the measured output, with the variable u left blank. An example code for open-loop identification for the first case study of a Virgo system is shown in Listing 2. Note here that the data must be stored as a three-dimensional matrix, with the data points in the third dimension.

LISTING 2. Basic open-loop identification.

```
% Load data
% r = 1 x nr x N matrix of the input signals
% y = ny x nr x N matrix of the output signals

% Frequencies to identify
freqIdentBand = [0.04 10];
W = 12; % Size of frequency window is 2W+1
Fs = 1e4; % Sampling frequency
La = 2; % Order of plant numerator
Lb = 2; % Order of transient numerator
Ld = 4; % Order of plant & transient denominator

[G,G_var] = ...
    GraFIT(r,y,[],W,freqIdentBand,Fs,[La Lb Ld]);
```

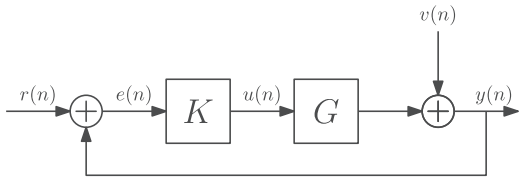


FIG. 3. Standard closed-loop identification problem, where the goal is to identify G using the input signal $r(n)$ and noisy output signals $y(n)$ and $u(n)$.

For systems operating in a closed-loop, r is again the perturbation signal and u and y are, respectively, chosen as the input and output of the dynamic system $G(\Omega_k)$ to be identified, with n_u representing the number of input signals to $G(\Omega_k)$; see Fig. 3. An example code for indirect identification is shown in Listing 3.

LISTING 3. Basic indirect identification.

```
% Load data
% r = 1 x nr x N matrix of the injected noise
% y = ny x nr x N matrix of the output of G
% u = nu x nr x N matrix of the input of G

% Frequencies to identify
freqIdentBand = [10 100];
W = 18; % Size of frequency window is 2W+1
Fs = 1e4; % Sampling frequency
La = 2; % Order of plant numerator
Lb = 2; % Order of transient numerator
Ld = 2; % Order of plant & transient denominator

[G,G_var,Gry,Gru,Gry_var,Gru_var] = ...
    GraFIT(r,y,u,W,freqIdentBand,Fs,[La Lb Ld]);
```

The argument freqIdentBand must be specified by the user, as it restricts the identification to a given frequency band, thereby reducing unnecessary computational effort. In addition, the user has to specify W and L, which affect the quality of the estimate.

B. Indirect identification

This section revisits the standard identification procedure for systems operating in a closed-loop, as this is the approach implemented in the toolbox. A more detailed overview of identifying systems operating in a closed-loop is given by Evers *et al.*[27]

In Fig. 3, the block diagram for a system operating in a closed-loop is shown, and the goal is to obtain the FRF of $G(\Omega_k)$ using the perturbation signal $r(n)$ and the noisy output signals $y(n)$ and $u(n)$. Direct identification of $G(\Omega_k)$ through $y(n)$ and $u(n)$ leads to a biased estimate,[27] since $u(n)$ and $y(n)$ are correlated through the feedback loop by $v(n)$.

Instead, an unbiased estimate of $G(\Omega_k)$ is obtained by identifying $\widehat{G}_{ru}$, given by

$$U(k) = \underbrace{K(\Omega_k)(I + G(\Omega_k)K(\Omega_k))^{-1}}_{=\widehat{G}_{ru}(\Omega_k)} R(k), \tag{22}$$

and $\widehat{G}_{ry}$, which is given by

$$Y(k) = \underbrace{G(\Omega_k)K(\Omega_k)(I + G(\Omega_k)K(\Omega_k))^{-1}}_{=\widehat{G}_{ry}(\Omega_k)} R(k), \tag{23}$$

and then computing

$$\widehat{G}(\Omega_k) = \widehat{G}_{ry}(\Omega_k)\widehat{G}_{ru}^{-1}(\Omega_k). \tag{24}$$

The estimation of $\widehat{G}_{ry}$ and $\widehat{G}_{ru}$ can be considered as an open-loop identification problem of the closed-loop response, i.e., the standard identification procedure between a perturbation signal and output signal as presented in Sec. III can be applied, since the perturbation signal $r(n)$ is uncorrelated with the noise $v(n)$ when (filtered) white noise is used. The same derivation holds for MIMO systems, where care has to be taken that matrix inversion and multiplication are properly applied in (24) to avoid a biased estimate.[27] The variance of the FRF is furthermore directly dependent on the signal-to-noise ratio of the perturbation signal $r(n)$ with respect to the noise $v(n)$ in the signals $u(n)$ and $y(n)$, so an FRF with a low variance can only be obtained when a sufficiently high signal-to-noise ratio is present in $u(n)$ and $y(n)$.

C. Uncertainty estimation

To assess the quality of the estimated FRF, the variance of the estimate is approximated by computing the residuals of the local model estimate,

$$\widehat{V}(k) = Y_w(k) - \widehat{\Theta}(k)K_w(k), \tag{25}$$

and the estimated covariance of G is given by [Ref. 11, Eq. (7.21)]

$$\mathrm{cov}(G(\Omega_k)) = S^H S \otimes \mathrm{cov}(\widehat{V}(k)), \tag{26}$$

with

$$S = K_W^H(K_W K_W^H)^{-1}\begin{bmatrix} I_{n_r} \\ 0 \end{bmatrix}, \tag{27}$$

with I_{n_r} representing the identity matrix of size n_r. To obtain the variance of G using the indirect approach, the in-loop signals $Y(k)$ and $U(k)$ are combined into a single signal $Z(k) = [Y(k)\ U(k)]^T$ and the estimated covariance of G is then obtained by [Ref. 11, Eq. (7.50)]

$$\mathrm{cov}(G) = \left(\widehat{G}_{ru}^{-T} \otimes \left[I_{n_y} - \widehat{G}\right]\right)\mathrm{cov}(\widehat{G}_{rz})\left(\widehat{G}_{ru}^{-T} \otimes \left[I_{n_y} - \widehat{G}\right]\right)^H, \tag{28}$$

where the frequency operator was left out for brevity purposes and with I_{n_y} representing the identity matrix of size n_y. The covariance matrix $\mathrm{cov}(\widehat{G}_{rz})$ is obtained from Eq. (26).

D. Parameter selection

The choice of the polynomial orders L depends on the system dynamics. Based on current experience with multiple sets of GW data, choosing the numerator orders $L_a = L_b = 2$ typically leads to satisfactory results for GW detectors. For the denominator order, the choice depends on whether the system has lightly damped resonant dynamics, in which case a higher order up to $L_d = 4$ is recommended. Without such dynamics, choosing $L_d = 0$ is often sufficient but choosing a non-zero L_d may yield marginally better results. For W, the total frequency window size $2W + 1$ needs to be larger than the number of parameters to be estimated. Any additional increase in W will typically lead to better parameter estimates and thus a lower variance since there will be more data points to average over. However, the critical consideration is the bias-variance trade-off when selecting these parameters, which is inherent to the LRM approach (Ref. 11, Sec. 7.2.6), as a higher W will typically lead to a lower variance at the expense of an increased bias in the estimate. In practice, lower model orders, such as the ones proposed here, and a frequency window size, which yields a few additional points to average over, typically already lead to satisfactory results. Using this estimate as a baseline, then increasing the frequency window size W, and comparing it to the original estimate can be helpful in determining when significant bias starts to occur.

V. EXPERIMENTAL CASE STUDIES

In this section, GraFIT is compared to the SA method in two of the three application domains from Sec. II to illustrate the effectiveness of the toolbox. The metric for comparison is first introduced, after which experimental results from two of the application domains are presented.

A. Metric for comparison

For the experimental case studies, the standard deviations on the estimates of the SA method and LRM are compared to assess the quality of the estimators. The estimated standard deviation for the LRM is computed using (26) for the open-loop case and using (28) for the closed-loop case. The coherence function $c \in \mathbb{R}_{[0,1]}$ [Ref. 11, Eq. (2.47)] is commonly used for the SA method to assess the FRF quality. Using Pintelon and Schoukens [Ref. 11, Eq. (2.50)], the estimated standard deviation using c is obtained for the SA method.

To ensure a fair comparison, the same q is used for both methods, where q represents the number of data points relative to the number of parameters to be estimated, since more data points per parameter leads to a lower standard deviation. The parameter $q_{\mathrm{SA}} = P$ for the SA estimator, since only a single parameter per frequency bin is estimated using P averages. For the LRM, q_{LRM} is given by [Ref. 11, Eq. (7.14)]

$$q = 2W + 1 - n_\Theta, \tag{29}$$

since there are $2W + 1$ data points in the window and n_Θ parameters to be estimated for the local model in the window, with n_Θ defined

in (21). The parameters W for the LRM and P for the SA method are therefore selected such that $q_{\mathrm{LRM}} = q_{\mathrm{SA}}$.

B. Direct identification for a suspension system

The first application domain is the suspension system (see Sec. II A), and the MultiSAS suspension[6] is used as an example system. This system is used to suspend the auxiliary optics of Virgo, such as photodiodes, for which the residual motion requirements are much less stringent, but it uses the same working principle as the suspensions for the main optics. Active control is applied on the first isolation stage of the system to damp the suspension modes and reduce the RMS motion. Three actuators and sensors are therefore located between the first isolation stage and the ground to damp the modes in the two horizontal translational directions (x and z) and the modes around the vertical axis (t_y).

To assess the coupling between the Degrees of Freedom (DoF) and to design a controller for each DoF, the MIMO FRF of the system is measured in an open-loop by consecutively injecting bandpass filtered white noise in each DoF. The system is identified with both the standard SA method and GraFIT using 300 s of data to guarantee enough frequency resolution. For the SA method, $P = 7$ segments [see (7)] are used to have some averaging while maintaining enough frequency resolution. For GraFIT, a frequency window size of $W = 12$ is used, and the orders of the polynomials are chosen as $L_a = L_b = 2$ and $L_d = 4$ due to the largely undamped modes. This results in $n_\Theta = 18$ and $q_{\mathrm{LRM}} = 7$, which is equal to the number of segments P in the SA estimate.

The resulting estimates for both methods and their estimated standard deviations are shown in Fig. 4. The methods find almost identical estimates, indicating similar bias levels, but GraFIT almost consistently achieves roughly a factor 2–3 lower standard deviation on the estimate with even larger standard deviation differences at the resonances. The reduced estimation error at the resonances provides a better estimate of the eigenfrequency and damping of the modes, which may be desirable for control design or estimating modal models.[22] An advantage of GraFIT is that the frequency window size W could be significantly increased to reduce the standard deviation while maintaining the same frequency resolution, while for the SA method, increasing the number of segments P would reduce the frequency resolution proportionally to the number of segments. However, the bias-variance trade-off as discussed in Sec. IV D has to be considered when increasing W.

C. Indirect identification for an optical system

The second application domain is the longitudinal control of the mirrors in the Advanced Virgo+ detector (see Sec. II B). The control system consists of five longitudinal DoFs, but only the DARM, MICH, and SRCL DoFs[9] will be considered here, where DARM contains the GW signal and MICH and SRCL are two of the most strongly coupled DoFs. The control loop uses error signals derived from the photodiodes to actuate on the mirror positions, and the system must operate in a closed-loop to identify the FRF. In each DoF, bandpass filtered white noise is consecutively injected for 120 s. For the SA method, $P = 25$ segments are used, a detrend of the signals is performed by subtracting the DC component, and a Von Hann window is applied to handle the leakage effects. These two methods form the basic approach for handling leakage effects in the GW community, allowing a fair comparison to the LRM approach. For GraFIT, a frequency window size of $W = 18$ is used and the orders of the polynomials are chosen as $L_a = L_b = L_d = 2$ such that $n_\Theta = 12$ and $q_{\mathrm{LRM}} = 25$. The indirect identification approach from Sec. IV B is used to obtain an estimate of G.

In Fig. 5, the estimate of G for the SA method with 120 s of data and GraFIT with 120 and 12 s of data is shown. The standard deviations for both GraFIT estimates are also shown but not for the

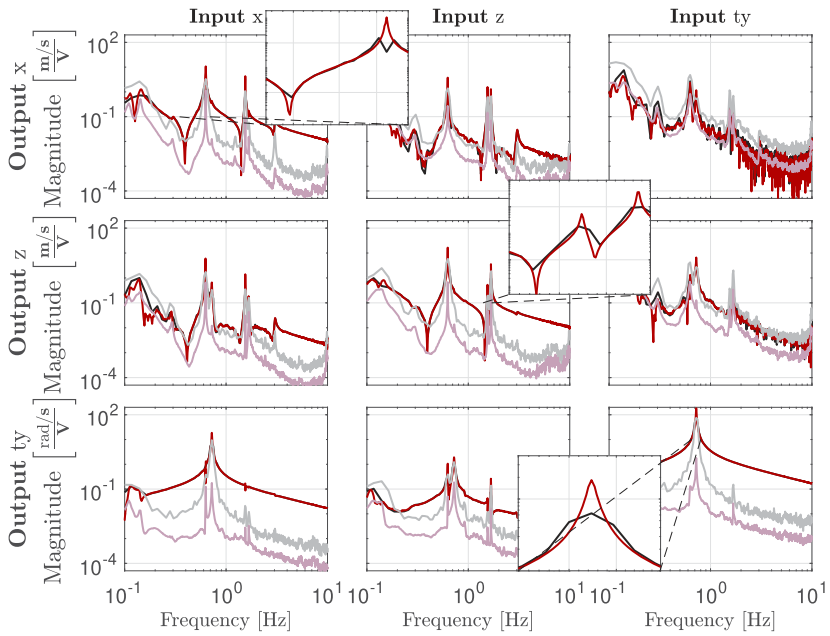


FIG. 4. FRF $\widehat{G}(\Omega_k)$ for the x, z, and t_y degrees of freedom of the top stage of the MultiSAS using the SA estimation (—) and its standard deviation (—) and GraFIT (—) and its standard deviation (—). GraFIT obtains almost consistently a factor 2 to 3 lower estimated standard deviation with even lower standard deviations at the resonance peaks, while also having a factor $P = 6$ higher frequency resolution.

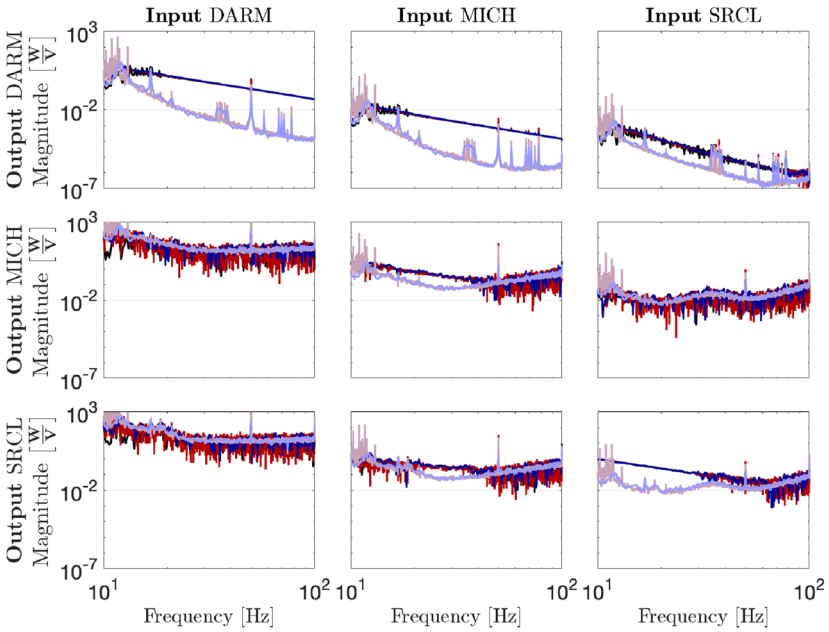


FIG. 5. FRF $\widehat{G}(\Omega_k)$ for the three longitudinal DoFs using the SA estimation (—) for 120 s of data, GraFIT (—) and its estimated standard deviation (—) for 120 s of data, and GraFIT (—) and its estimated standard deviation (—) for the first 12 s of the same dataset. The standard deviation for SA is not available. Just 12 s of data using GraFIT is sufficient to get a good quality FRF.

SA method since they are not directly available. All three estimates are roughly identical, indicating that they have similar bias levels. Since the standard deviations are computed based on the residuals of the local model estimate $\widehat{V}(k)$ in (25) and since both the 120 and 12 s datasets use the same number of neighboring frequency bins W to estimate the same number of parameters n_Θ, the standard deviations are almost identical.

To have a more direct assessment of GraFIT compared to SA, the estimate of G_{ru} is shown in Fig. 6 (similar results for G_{ry} have been observed, and this plot has therefore been left out for brevity), together with the estimated standard deviations for the three estimates. GraFIT using both 120 and 12 s of data obtains almost consistently a full order of magnitude lower estimated standard deviation compared to the SA method with 120 s of data, while the bias

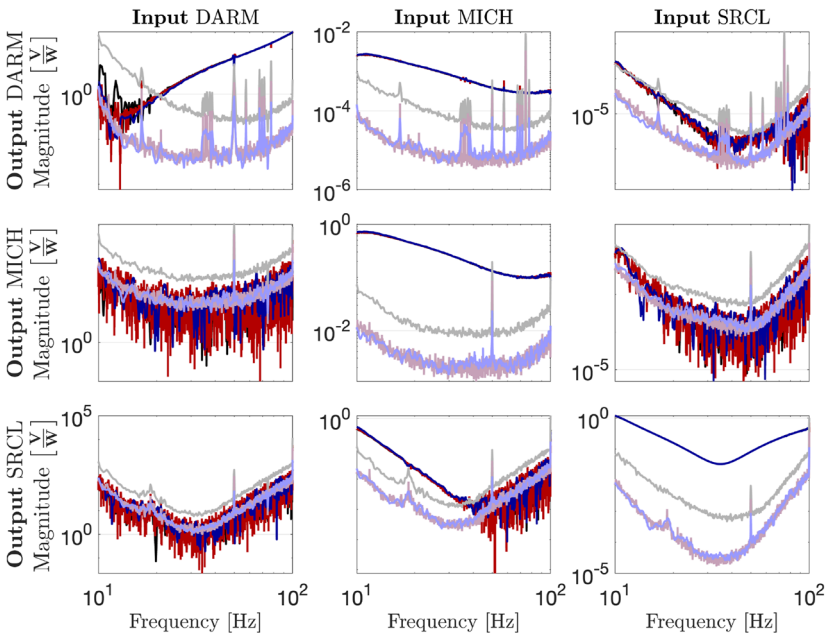


FIG. 6. FRF $\widehat{G}_{ru}(\Omega_k)$ for the three longitudinal DoFs using the SA estimation (—) and its estimated standard deviation (—) for 120 s of data, GraFIT (—) and its estimated standard deviation (—) for 120 s of data, and GraFIT (—) and its estimated standard deviation (—) for the first 12 s of the same dataset. GraFIT is able to obtain a full order of magnitude lower standard deviation with just a tenth of the data compared to the SA, while having almost identical estimates, indicating similar bias levels.

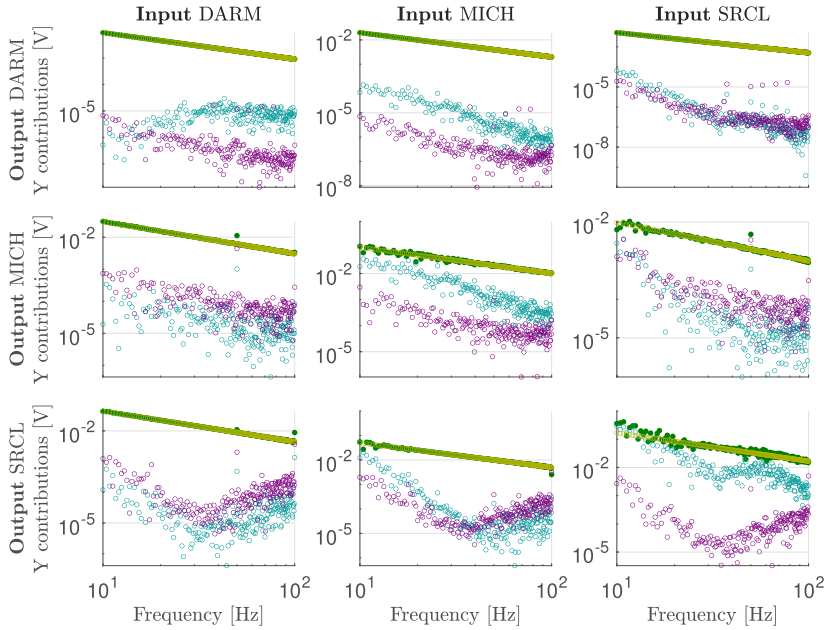


FIG. 7. Output DFT $U(k)$ (●), input contribution $G_{ru}(\Omega_k)R(k)$ (○), transient contribution $T(\Omega_k)$ (○), and noise $V(k)$ (○). The DFTs are downsampled in the frequency domain for clarity of presentation. The transient contribution dominates the output DFT for almost all entries.

is similar for all three estimates, illustrating that GraFIT achieves superior estimation performance compared to the SA method. Furthermore, it is also much more effective with fewer data, with GraFIT obtaining significantly lower standard deviations with ten times fewer data compared to the SA method. Next to the main benefit of requiring less data, this potentially also has the additional benefit of suffering less from non-stationary noise effects due to the reduced measurement time.

One of the reasons for the substantially better estimate using GraFIT is the handling of leakage effects, which are significant in the DFT of $U(k)$. In Fig. 7, the contributions of the input, transient, and noise term to the DFT of u are shown, which have all been estimated using GraFIT. The contribution of the transient term, which identifies the leakage effect, is in almost every entry at least an order higher than the input contribution. This leakage stems from the specific structure of the signal u, the amplitude of which is orders higher below 10 Hz and contains many pure harmonics, resulting in a significant amount of leakage in the 10–100 Hz band due to the difference between the initial and end states in (6). By directly estimating the transient effect in GraFIT, rather than using windowing as in the SA method, the leakage effects are substantially reduced, leading to between a factor 3 and 10 lower estimated standard deviations on the estimate of G_{ru} using GraFIT compared to the SA method and therefore also an improved estimate of G.

VI. CONCLUSIONS

The main requirements for frequency response identification in GW detectors are accuracy and minimal experimentation time. In a representative suspension system of Virgo, GraFIT achieves at least a factor 2–3 lower standard deviation across the identified frequency range, with even lower standard deviations at the resonance peaks, which provides better estimates for the eigenfrequency and damping of the modes. In the longitudinal control system of Advanced Virgo+, GraFIT obtains almost a full order of magnitude lower standard deviation across the entire frequency range with 10 times less data than SA, illustrating that GraFIT is very effective in using minimal data for the estimation. In addition, GraFIT accommodates the identification of systems with arbitrary input–output dimensions and systems operating in a closed-loop, which significantly simplifies the identification procedure, at the expense of increased computation time by a factor of a few. Since experimentation time on a GW detector is scarce and expensive, GraFIT provides a valuable tool to be more efficient with experimentation time.

The toolbox is freely available under the MIT license at https://github.com/MathynVanD/GraFIT.

ACKNOWLEDGMENTS

This work was funded by the Netherlands Organisation for Scientific Research (NWO) under Grant No. 680.92.18.02. The authors gratefully acknowledge the contributions made by Luuk van Vliet for his testing on GW data. The authors also would like to acknowledge the contributions made by Rick van der Maas, Annemiek van Rietschoten, Enzo Evers, Robbert Voorhoeve, and Paul Tacx in developing the toolbox. The authors also gratefully acknowledge the contributions made by the ISC team at Advanced Virgo+ for helping to perform the experiments on the interferometer and providing the necessary information and support to conduct this work. The authors also gratefully acknowledge the Italian Istituto Nazionale di Fisica Nucleare (INFN), the French Centre National de la Recherche Scientifique (CNRS), and the Netherlands Organisation for Scientific Research for the construction and operation of the Virgo detector and the creation and support of the EGO consortium. The authors also gratefully acknowledge research support

from these agencies and by the Spanish Agencia Estatal de Investigación, the Consellera d'Innovació, Universitats, Ciència i Societat Digital de la Generalitat Valenciana, the CERCA Programme Generalitat de Catalunya, Spain, the National Science Centre of Poland, the Foundation for Polish Science (FNP), the European Commission, the Hungarian Scientific Research Fund (OTKA), the French Lyon Institute of Origins (LIO), the Belgian Fonds de la Recherche Scientifique (FRS-FNRS), Actions de Recherche Concertées (ARC), and Fonds Wetenschappelijk Onderzoek–Vlaanderen (FWO), Belgium.

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

M. van Dael: Conceptualization (lead); Data curation (lead); Formal analysis (lead); Investigation (lead); Software (lead); Validation (lead); Visualization (lead); Writing – original draft (lead); Writing – review & editing (equal). **M. van Haren**: Software (supporting); Validation (supporting); Writing – review & editing (equal). **G. Witvoet**: Conceptualization (supporting); Investigation (supporting); Project administration (equal); Writing – review & editing (equal). **B. Swinkels**: Conceptualization (supporting); Investigation (supporting); Software (supporting); Validation (supporting); Writing – review & editing (equal). **T. Oomen**: Conceptualization (supporting); Project administration (equal); Writing – review & editing (equal).

DATA AVAILABILITY

The data that support the findings of this study are available from the European Gravitational Observatory (EGO) and Nikhef. Restrictions apply to the availability of these data, which were used under license for this study. The data are available from the authors upon reasonable request and with the permission of the European Gravitational Observatory (EGO) and Nikhef.

REFERENCES

[1] F. Acernese *et al.*, J. Phys.: Conf. Ser. **2429**, 012039 (2023).
[2] J. Aasi *et al.*, Classical Quantum Gravity **32**, 074001 (2015).
[3] G. D. Meadors, K. Kawabe, and K. Riles, Classical Quantum Gravity **31**, 105014 (2014).
[4] A. Allocca, D. Bersanetti, J. Casanueva Diaz, C. De Rossi, M. Mantovani, A. Masserot, L. Rolland, P. Ruggi, B. Swinkels, E. N. Tapia San Martin, M. Vardaro, and M. Was, Galaxies **8**, 85 (2020).
[5] S. M. Aston *et al.*, Classical Quantum Gravity **29**, 235004 (2012).
[6] J. V. van Heijningen, A. Bertolini, E. Hennes, M. Beker, M. Doets, H. Bulten, K. Agatsuma, T. Sekiguchi, and J. F. van den Brand, Classical Quantum Gravity **36**, 075007 (2019).
[7] F. Acernese *et al.*, Classical Quantum Gravity **39**, 045006 (2022).
[8] A. D. Viets, M. Wade, A. L. Urban, S. Kandhasamy, J. Betzwieser, D. A. Brown, J. Burguet-Castell, C. Cahillane, E. Goetz, K. Izumi, S. Karki, J. S. Kissel, G. Mendell, R. L. Savage, X. Siemens, D. Tuyenbayev, and A. J. Weinstein, Classical Quantum Gravity **35**, 095015 (2018).
[9] D. Bersanetti, B. Patricelli, O. J. Piccinni, F. Piergiovanni, F. Salemi, and V. Sequino, Universe **7**, 322 (2021).
[10] A. Buikema *et al.*, Phys. Rev. D **102**, 062003 (2020).
[11] R. Pintelon and J. Schoukens, *System Identification: A Frequency Domain Approach*, 2nd ed. (John Wiley, 2012).
[12] J. S. Bendat and A. G. Piersol, *Engineering Applications of Correlation and Spectral Analysis* (Wiley, New York, 1980).
[13] J. Schoukens, R. Pintelon, T. Dobrowiecki, and Y. Rolain, Automatica **41**, 491 (2005), a part of Special Issue: Data-Based Modelling and System Identification.
[14] P. Hägg, J. Schoukens, M. Gevers, and H. Hjalmarsson, Automatica **68**, 314 (2016).
[15] J. Lataire and T. Chen, Automatica **72**, 217 (2016).
[16] J. Schoukens, G. Vandersteen, K. Barbé, and R. Pintelon, in *2009 European Control Conference (ECC)* (IEEE, 2009), pp. 1–14.
[17] T. McKelvey and G. Guérin, in *16th IFAC Symposium on System Identification* (IFAC Proceedings Volumes, 2012), Vol. 45, p. 49.
[18] J. Schoukens, Y. Rolain, and R. Pintelon, Automatica **42**, 27 (2006).
[19] M. Gevers, R. Pintelon, and J. Schoukens, in *2011 50th IEEE Conference on Decision and Control and European Control Conference* (IEEE, 2011), pp. 4302–4307.
[20] E. Evers, N. van Tuijl, R. Lamers, B. de Jager, and T. Oomen, Mechatronics **70**, 102401 (2020).
[21] R. Voorhoeve, A. van der Maas, and T. Oomen, Mech. Syst. Signal Process. **105**, 129 (2018).
[22] P. Tacx, R. Habraken, G. Witvoet, M. Heertjes, and T. Oomen, Mechatronics **99**, 103158 (2024).
[23] S. Braccini *et al.*, Astropart. Phys. **23**, 557 (2005).
[24] D. Bersanetti, M. Boldrini, J. C. Diaz, A. Freise, R. Maggiore, M. Mantovani, and M. Valentini, Galaxies **10**, 115 (2022).
[25] M. van Dael *et al.*, Classical Quantum Gravity **41**, 215008 (2024).
[26] E. Evers, B. de Jager, and T. Oomen, IEEE Trans. Control Syst. Technol. **30**, 142 (2022).
[27] E. Evers, R. Voorhoeve, and T. Oomen, in *2020 IEEE 16th International Workshop on Advanced Motion Control (AMC)* (IEEE, 2020), pp. 1–6.