A Wavelet-Based Approach to FRF Identification From Incomplete Data

Dirkx, Nic; Tiels, Koen; Oomen, Tom

DOI
10.1109/TIM.2023.3260271

Publication date
2023

Document Version
Final published version

Published in
IEEE Transactions on Instrumentation and Measurement

A Wavelet-Based Approach to FRF Identification From Incomplete Data

Nic Dirkx, Koen Tiels, *Member, IEEE*, and Tom Oomen, *Senior Member, IEEE*

***Abstract*— Frequency response function (FRF) estimation from measured data is an essential step in the design, control, and analysis of complex dynamical systems, including thermal and motion systems. Especially for systems that require long measurement time, missing samples in the data record, e.g., due to measurement interruptions, often occur. The aim of this article is to achieve accurate identification of nonparametric FRF models of periodically excited systems from noisy output measurements with missing samples. An identification framework is established that exploits a wavelet-based transform to separate the effect of the missing samples in the time domain from the system characteristics in tre frequency domain. The framework encompasses both a time-invariant and a time-varying wavelet-based estimator, which provides different mechanisms to address the missing samples. Experimental results from a thermodynamical system confirm that the estimators enable accurate identification.**

***Index Terms*— Frequency response, linear systems, missing data, system identification, wavelet transforms.**

I. Introduction

ACCURATE nonparametric frequency response function (FRF) identification is a key step toward high-performance control of complex systems. FRF identification, as opposed to first principles modeling, is considered a fast and inexpensive approach to obtain thorough insight into the system dynamics while requiring only mild assumptions on the actual system [1]. FRFs are used in many application domains, including flexible mechanical systems [2] and thermal systems [3], with purposes ranging from controller synthesis [4], diagnostics [5], and parametric identification [6].

The mitigation of leakage errors, arising from the time–frequency domain transformation of finite-length data records, is an important aspect in nonparametric identification. As recognized in, e.g., [7] and [8], leakage errors originate from a transient phenomenon, which has smooth frequency characteristics that depend on the system dynamics. This is exploited in the local polynomial method (LPM) [9], where the transient is approximated by a polynomial function in a local frequency-domain window. Extensions to local rational approximation functions are developed in the local rational method (LRM) [10]. The LPM has shown to be particularly effective for the identification of slow dynamic systems that exhibit significant transient behavior [11].

In many practical applications, and especially for slow dynamic systems that require long measurement time, measurements cannot be performed without interruptions, which leads to missing samples in the data record. Missing samples may also occur due to communication link or sensor failure [12]. Locally missing samples in the time domain introduce global and nonsmooth effects in the frequency domain [13], which complicates the use of standard identification algorithms, including the LPM.

Several techniques have been developed to deal with the missing samples in nonparametric identification. In [14], a signal reconstruction approach is presented in which an additional transient is estimated for each data gap, but the application to FRF estimation is not considered. In [15], the classical LPM is extended to locally estimate the FRF parameters together with additional transient parameters for each data gap. The estimation of additional variables has a negative impact on the locality of the estimation problem, leading to a performance deterioration for an increasing number of data gaps.

Other approaches aim at reconstructing the missing samples. In [16], a nonparametric iterative spectrum reconstruction scheme is presented. In [17], reconstruction is based on a statistical approach. Neither of the approaches are applied for FRF identification. In [13] and [18], the LPM is extended to estimate frequency-domain system parameters along with the missing samples. Due to the global nature of the time–frequency domain transform, errors in the estimates of the missing samples induce errors in the frequency-domain system parameters at all frequencies. This effect is amplified for an increasing number of missing samples due to an increasing correlation between the estimated parameters.

Another line of research has focused on the development of time-dependent frequency analysis techniques for transient signals. In particular, the short-time Fourier transform [19] and the wavelet transform [20], [21] are widely used in, e.g., signal processing [22], finite-element methods [23], and

Manuscript received 6 December 2022; revised 20 February 2023; accepted 4 March 2023. Date of publication 22 March 2023; date of current version 5 April 2023. This work was supported in part by the Research Programme VIDI under Project 15698 and in part by the Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO). The Associate Editor coordinating the review process was Dr. Dong Wang. *(Corresponding author: Nic Dirkx.)*

Nic Dirkx is with ASML, 5504 DR Veldhoven, The Netherlands, and also with the Department of Mechanical Engineering, Control Systems Technology, Eindhoven University of Technology, 5600 MB Eindhoven, The Netherlands (e-mail: nic.dirkx@asml.com).

Koen Tiels is with the Department of Mechanical Engineering, Control Systems Technology, Eindhoven University of Technology, 5600 MB Eindhoven, The Netherlands.

Tom Oomen is with the Department of Mechanical Engineering, Control Systems Technology, Eindhoven University of Technology, 5600 MB Eindhoven, The Netherlands, and also with the Delft Center for Systems and Control, Delft University of Technology, 2628 CD Delft, The Netherlands.

Digital Object Identifier 10.1109/TIM.2023.3260271

image compression [24], but have not been applied for FRF identification from incomplete data.

Although important developments have been made in FRF identification from incomplete data, current approaches involve the estimation of an increasing number of auxiliary variables for an increasing amount of missing data. The aim of this article is to present a nonparametric FRF identification framework for periodically excited systems that avoids the estimation of additional variables and the associated drawbacks.

The main contributions of this article are given as follows.

1) A generalized LPM framework for FRF estimation from incomplete data that avoids the estimation of additional variables by separating the effect of the missing samples in the time-domain system from the frequency-domain system characteristics. The framework encompasses the following:
   a) a time-invariant wavelet-based estimator (TIE);
   b) a time-varying wavelet-based estimator (TVE).
2) An analysis of bias and variance errors for both estimators.
3) A performance comparison based on experimental results from a thermodynamic system.

Preliminary results related to the TIE in contribution 1-a) are presented in [25]. In this current article, the TIE is further developed toward a generalized framework in which the TIE is recovered as a specialization.

This article is organized as follows. In Section II, the problem is formulated. Key principles of wavelet-based convolution and the time–frequency plane are discussed in view of the missing samples problem in Section III. In Section IV, the wavelet-based identification framework is presented, which extends the existing frequency-domain LPM to a time–frequency plane approach. Here, the TIE and TVE are presented as specializations within the general framework, and their bias and variance properties are analyzed. Experimental results are presented in Section V, and conclusions are drawn in Section VI.

Notations and Definitions: The discrete Fourier transform (DFT) of a sequence $x \in \mathbb{R}^N$ at the kth frequency line is

$$X(k) = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} x[n] e^{-1i2\pi nk/N}. \tag{1}$$

The inverse DFT (IDFT) is given by

$$x_N[n] = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} X(k) e^{1i2\pi nk/N} \tag{2}$$

where x_N is the N-periodic extension of x and $x_N[n \pm \alpha N] = x[n]$, $\forall \alpha \in \mathbb{N}, n = [0, \ldots, N-1]$. Operations $X = \mathcal{F}x$ and $x = \mathcal{H}X$ denote the DFT and IDFT, respectively, where $\mathcal{F}$ and $\mathcal{H}$ denote the DFT and IDFT matrices [26, Ch. 2], respectively, satisfying $\mathcal{F}\mathcal{H} = \mathcal{H}\mathcal{F} = I$. Operator $\odot$ denotes the elementwise product, satisfying $Y \odot X = \text{diag}(Y)X$ for a vector Y. Operation $X^\dagger$ denotes the Moore–Penrose inverse, i.e., $X^\dagger = (X^H X)^{-1} X^H$. For notational convenience, zero-based numbering is applied throughout. For $X \in \mathbb{R}^N$, operation $[X]_q$ selects the $q = 0, 1, \ldots, N-1$ element, whereas $[X]_{q,q}$ selects the qth diagonal entry of a matrix $X \in \mathbb{R}^{N \times N}$. The cardinality of a set $\mathcal{A}$ is denoted as $\text{card}(\mathcal{A})$.

II. Problem Formulation

A. Missing Samples Identification Problem

Consider the identification setup in Fig. 1, governed by the discrete-time input–output relation at the nth sample

$$y[n] = \{g_o * u\}[n] + v_y[n]. \tag{3}$$

Here, g_o is the impulse response of the to-be-identified single-input–single-output (SISO) linear time-invariant (LTI) system G_o. Operator $*$ denotes the convolution. Signal $u \in \mathbb{R}^{\bar{N}}$ with $\bar{N} = NP$ is the N-periodic input, for a number of $P \geq 2$ periods with sample size N. Furthermore, $y \in \mathbb{R}^{\bar{N}}$ is the true output signal, corrupted by measurement noise $v_y = \{h_o * e\}$, where h_o is the impulse response of the monic noise dynamics H_o and e is a zero-mean white noise realization with variance λ.

Using the DFT, the exact expression of (3) in the frequency domain is [27]

$$Y(k) = G(\Omega_k)U(k) + T_G(\Omega_k) + N_Y(k) \tag{4}$$

where Y and U denote the DFT of y and u, respectively. The index $k = 0, \ldots, \bar{N} - 1$ points to the frequency $kf_s/(\bar{N})$ with f_s the sampling frequency and $\Omega_k = e^{-1i \cdot 2\pi k/(\bar{N})}$. Term G is the FRF of the system and T_G is the transient term due to the difference between the initial and final system conditions [27]. Term N_Y is the DFT of v_y and is governed by $N_Y(k) = N_{Y_{\text{st}}}(k) + T_H(\Omega_k)$, where $N_{Y_{\text{st}}} = H(\Omega_k)E(k)$ is the stationary part of the noise with $H(\Omega_k)$ the FRF of the noise, E is the DFT of e, and T_H is the transient term due to the difference between the initial and final conditions of the noise dynamics. The noise converges to a circular complex normal distribution $N_{Y_{\text{st}}} \xrightarrow{d} \mathcal{N}(0, C_Y)$ for $\bar{N} \to \infty$ [6].

Suppose that the output signal y is measured with interruptions, leading to a number of n_{m} missing samples in its measurement y^{m}. Assuming that the time indices of these samples are known, the measurement can be expressed as $y^{\text{m}} = W^{\text{m}} y$, where

$$\left[W^{\text{m}}\right]_{n,n} = \begin{cases} 1, & \text{if sample } n \text{ is available} \\ 0, & \text{if sample } n \text{ is missing} \end{cases} \tag{5}$$

with $n = 0, \ldots, \bar{N} - 1$. Also, see Fig. 1. The DFT of y^{m} is expressed as $Y^{\text{m}} = \mathcal{T}^{\text{m}} Y$, where

$$\mathcal{T}^{\text{m}} = \mathcal{F} W^{\text{m}} \mathcal{H}. \tag{6}$$

Missing samples significantly complicate the identification of G. In particular, due to coupling introduced by matrix $\mathcal{T}^{\text{m}}$ when samples are missing, all frequency lines in Y^{m} contain data of all the missing samples. Thus, the locally missing samples in the time domain have a global impact on the frequency domain. Consequently, traditional nonparametric frequency-domain identification methods that rely on the frequency-separation principle [28], including the LPM [27], no longer provide accurate solutions. In this article, a method is presented for accurate FRF identification when samples are missing.

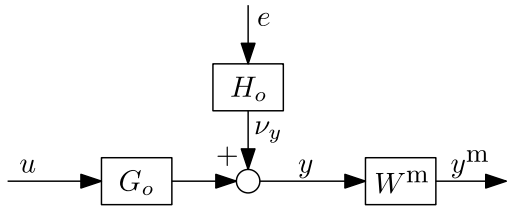


Fig. 1. Identification setup with plant dynamics G_o, noise dynamics H_o, and measurement y^{m} with missing samples, reflected by matrix W^{m}, see (5).

B. Key Idea: Estimation in the Time–Frequency Plane

The main limitation in traditional FRF identification techniques originates from the fact that the DFT is a global transform, making it unsuitable for analysis of signals that exhibit local time-domain phenomena such as missing samples.

The key idea of the identification method presented in this article is to employ a local wavelet-based transform [21] rather than the global DFT transform. Contrary to the DFT, the wavelet-based transform provides frequency information that is time-localized, allowing data to be represented in the 2-D time–frequency plane. Exploiting the extra dimension, the effect of the missing samples can be localized in the time dimension and separated from the system response in the frequency dimension.

In Section III, the wavelet-based transform to the time–frequency plane is introduced, which is exploited for the formulation of wavelet-based estimators in Section IV.

III. Missing Samples in the Time–Frequency Plane

A. Wavelet-Based Transformation

The key step in separating the perturbation from the missing samples from the FRF of G in (4) is the transformation of the data y^{m} to the time–frequency plane. This is realized via convolution with a set of short-time oscillations at the jth frequency line, referred to as the wavelet functions ψ_j.

Definition 1 (Wavelet Convolution): Consider a wavelet function $\psi_j \in \mathbb{C}^{\bar{N}}$ with center frequency j. Let $y^{\mathrm{m}}_{\bar{N}}$ be the $\bar{N}$-periodic extension of signal $y^{\mathrm{m}} \in \mathbb{R}^{\bar{N}}$. The circular convolution $z^{\mathrm{m}}_j \in \mathbb{C}^{\bar{N}}$ of the signal y^{m} with the wavelet ψ_j at sample $n = [0, \ldots, \bar{N}-1]$ is defined as

$$z^{\mathrm{m}}_j[n] := \left\{y^{\mathrm{m}} \circledast \psi_j\right\}[n] := \frac{1}{\sqrt{\bar{N}}} \sum_{m=0}^{\bar{N}-1} \psi_j[m] \cdot y^{\mathrm{m}}_{\bar{N}}[n-m]. \quad (7)$$

By the convolution theorem, the transform z^{m}_j in (7) is equivalently expressed as the frequency-domain operation

$$z^{\mathrm{m}}_j = \mathcal{H}\left(\Psi_j \odot Y^{\mathrm{m}}\right) \quad (8)$$

where Ψ_j is the DFT of ψ_j.

Expression (7) shows that z^{m}_j describes the characteristics of y^{m} in the time domain, whereas (8) shows that Ψ_j acts as a frequency-domain windowing function upon Y^{m}. Hence, z^{m}_j reflects the time-domain behavior of the frequency components of Y^{m} within the selected window, i.e., z^{m}_j provides both time- and frequency-localized information of Y^{m}.

The transform z^{m}_j is affected by the missing samples in y^{m}. The key idea is to confine this effect in z^{m}_j by selecting a suitable wavelet function, which is considered next.

B. Wavelet Selection

The goal of using the wavelet-based transformation in Definition 1 is to obtain a time- and frequency-domain localized representation of the data y^{m}, in which the effect of the missing samples can be confined. As expressed by the Heisenberg–Gabor uncertainty principle, see, e.g., [29], there are fundamental limitations to the localization accuracy that can simultaneously be achieved in the two domains. In particular, it is not possible to have a wavelet with finite support in the time domain that is also band limited in the frequency domain. Wavelet selection, therefore, involves a tradeoff, and its choice is typically problem-specific.

In this article, a class of wavelets is considered that is particularly suitable for the confinement of missing samples.

Definition 2 (Wavelet Selection): Let $j \in \{\mathbb{N} : j \in [0, \bar{N}-1]\}$, and let $\delta \in \{\mathbb{R} : N/(2\delta) \in \mathbb{N}, \delta P \in \mathbb{N}\}$ be a design variable. Then, the wavelet $\psi^{[p]}_j \in \mathbb{C}^{\bar{N}}$, $p \in \mathbb{N}$ is defined as

$$\psi^{[p]}_j[m] = c^{[p]} h^{[p]}[m] \cdot e^{1i \cdot 2\pi \frac{j}{\bar{N}} m}, \quad m = 0, \ldots, \bar{N}-1.$$

Here, $h^{[p]} = \mathcal{H}(\mathcal{F}h)^p$, with superscript p the elementwise exponent. Function $h^{[p]}$ constitutes the pth order cardinal B-spline [30] when $p > 0$, with h the characteristic function

$$h[m] = \begin{cases} 1, & \text{if } m \le N/\delta \\ 0, & \text{otherwise.} \end{cases}$$

The parameter $c^{[p]} = (\delta(P/N)^{1/2})^p$ is a scaling constant.

The wavelet function $\psi^{[p]}_j$ in Definition 2 is composed of a complex exponential multiplied by an envelope $h^{[p]}$ based on a pth order cardinal B-spline. The p-parameter controls the smoothness of the envelope and hereby the wavelet's time- and frequency-domain localization properties. In the frequency domain, the DFT $\Psi^{[p]}_j$ is characterized by a Dirichlet kernel, also referred to as aliased sinc function [31]. This gives rise to the key properties in the following.

Lemma 1 (Wavelet Properties): The wavelet function $\psi^{[p]}_j$ in Definition 2 satisfies the following properties.

1) $\Psi^{[p]}_j$ equals 1 at the jth frequency line.
2) $\Psi^{[p]}_j$ equals zero at frequency lines $j \pm \alpha\delta P$, $\alpha, p \in \mathbb{N}_+$.
3) $\psi^{[p]}_j$ has finite time-domain support of length $l^{[p]}_\psi = pN/\delta - p + 1$, i.e., $\psi^{[p]}_j[n] = 0$ for $n \ge l^{[p]}_\psi$.
4) $|\Psi^{[p+1]}_j| < |\Psi^{[p]}_j|$ at all but the jth and $j \pm \alpha\delta$th frequency lines.

The proof is presented in the Appendix. The wavelet properties are further illustrated in Example 1.

Example 1: An example of $\psi^{[2]}_{20}$ and $\psi^{[3]}_{20}$ is shown in Fig. 2, where $N = 500$, $P = 1$, and $\delta = 5$. The wavelets enable approximately extracting local frequency information around the jth frequency line within a local time window of $l^{[2]}_\psi = 199$ and $l^{[3]}_\psi = 298$ samples, see properties 1) and 3) in Definition 2. Increasing p improves the frequency-domain localization accuracy, at the cost of time-domain accuracy, see properties 3) and 4).

Remark 1: The transform (7) for $\psi^{[p]}_j$ as defined in Definition 2 can be reformulated in terms of the discrete-time short-time Fourier transform [21]. The explicit expression in

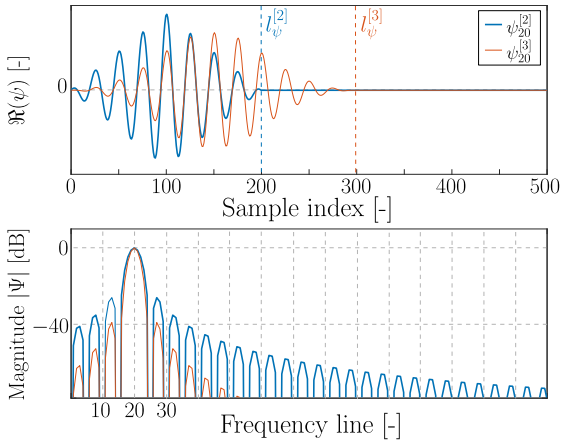


Fig. 2. Wavelets $\psi_{20}^{[2]}$ (—) and $\psi_{20}^{[3]}$ (—) in the time (top) and frequency domains (bottom). The wavelets extract frequency information around the frequency line 20, within the finite wavelet duration $l_\psi^{[2]} = 199, l_\psi^{[3]} = 298$.

terms of a wavelet function serves the interpretation in view of the missing samples problem.

The compactness of the support, as expressed by property 3) in Lemma 1, is crucial for confining the effect of the missing samples, which is illustrated next.

C. Missing Samples in the Time–Frequency Plane

Since the wavelets in Definition 2 have compact support in the time domain, the effect of missing samples in the data y^{m} can be confined in the transform z_j^{m} in (7). Let $w_{\mathrm{m}} \in \mathbb{N}^{n_{\mathrm{m}}}$ contain the indices of the n_{m} missing samples. Then, the samples v in the transform $z_j^{\mathrm{m}}[v]$ that are unaffected by the missing samples are given by $v \in \mathcal{D}^{[p]}$, where $\mathcal{D}^{[p]} = \bigcap_{i=1}^{n_{\mathrm{m}}} \mathcal{D}_i^{[p]}$ and

$$\mathcal{D}_i^{[p]} = \Big\{ v \in \mathbb{N} : \max\Big(w_{\mathrm{m}}[i] - \bar{N} + l_\psi^{[p]}, 0\Big) \le v < w_{\mathrm{m}}[i] \quad \text{or} \quad \min\Big(w_{\mathrm{m}}[i] + l_\psi^{[p]}, \bar{N}\Big) \le v < \bar{N} \Big\}. \tag{9}$$

Likewise, the indices v of the samples $z_j^{\mathrm{m}}[v]$ that are affected by the missing samples are contained in the set

$$\mathcal{A}^{[p]} = \Big\{ v \in \big[0, \bar{N} - 1\big] \setminus \mathcal{D}^{[p]} \Big\}. \tag{10}$$

The set $\mathcal{D}^{[p]}$ in (9) depends on the length of the wavelet support $l_\psi^{[p]}$ and on the number and pattern of the missing samples. The implications on the estimator are further discussed in Section IV.

Clearly, the sets $\mathcal{A}^{[p]}$ and $\mathcal{D}^{[p]}$ are disjoint sets, which enables separating the affected samples from the unaffected samples. This concept is illustrated in Example 2.

Example 2: Consider the system G with FRF, as shown in (—) in Fig. 3. The system is periodically excited with $P = 5$ and $N = 1200$ at frequency lines $10, 20, \ldots, 2990$. The excitation spectral magnitude has a uniform frequency distribution. A noiseless measurement is performed, where samples $1200, \ldots, 1800$ and $3900, \ldots, 4000$ are missing. The result of wavelet convolution z_j^{m} with wavelets $\psi_j^{[2]}$, $j = 10, 30, \ldots, 2990$ is shown in Fig. 3.

The main result is that the transforms z_j^{m} shown in Fig. 3(b) provide a time–frequency representation of the combination of the system G and the transient T_G, in which the part of z_j^{m} associated with the set $\mathcal{D}^{[p]}$ (in ■) is fully unaffected by the missing samples. In Section IV, a wavelet-based estimator is presented, which exploits these unaffected data to accurately identify the two individual components G and T_G. Furthermore, an estimator is presented, which additionally extracts information from the affected samples in $\mathcal{A}^{[p]}$ to reduce variance errors. These estimators constitute contributions 1-a) and 1-b).

IV. Wavelet-Based LPM

In this section, the wavelet-based estimation algorithms are presented for the accurate FRF identification from measurements with missing samples. The methods extend the existing frequency-domain LPM [27], [32] to estimators in the time–frequency plane. First, a recapitulation of the classical LPM approach is provided in Section IV-A, which provides the mathematical notation and foundation for the extended estimator presented in Section IV-B. This extension encompasses both a TIE and a TVE, which are presented in Sections IV-C and IV-D, respectively. The two estimators are unified in a single framework, as presented in Section IV-E. All proofs are presented in the Appendix.

A. Classical LPM and Its Limitations

The classical LPM [32] provides an effective method for FRF identification for complete data but does not produce accurate estimates when samples are missing. In this section, the classical LPM for periodic input signals is recapitulated and its limitations are discussed. The LPM forms the basis for the wavelet-based estimators presented in this article, which address these limitations.

Consider input–output relation (4). For periodic inputs $u \in \mathbb{R}^{\bar{N}}$, the DFT $U \in \mathbb{C}^{\bar{N}}$ contains energy only at the excited frequency (EF) lines $\bar{k} = [\bar{k}_0, \ldots, \bar{k}_{\bar{N}-1}]$, where $\bar{k}_j = j\gamma P$, $j = 0, \ldots, \bar{N} - 1$, and $\bar{N} = \lfloor N/\gamma \rfloor$. User-defined parameter $\gamma \in \mathbb{N}_+$ selects the sparsity of the EF lines in the frequency grid. Due to the periodicity of u, the system response $G(\Omega_k)U(k)$ is zero at the nonexcited frequency (NEF) lines $\bar{k}+q$ with $q = 1, 2, \ldots, \gamma P - 1$. This is visualized in Fig. 4. Thus, the output (4) at the NEF lines reduces to

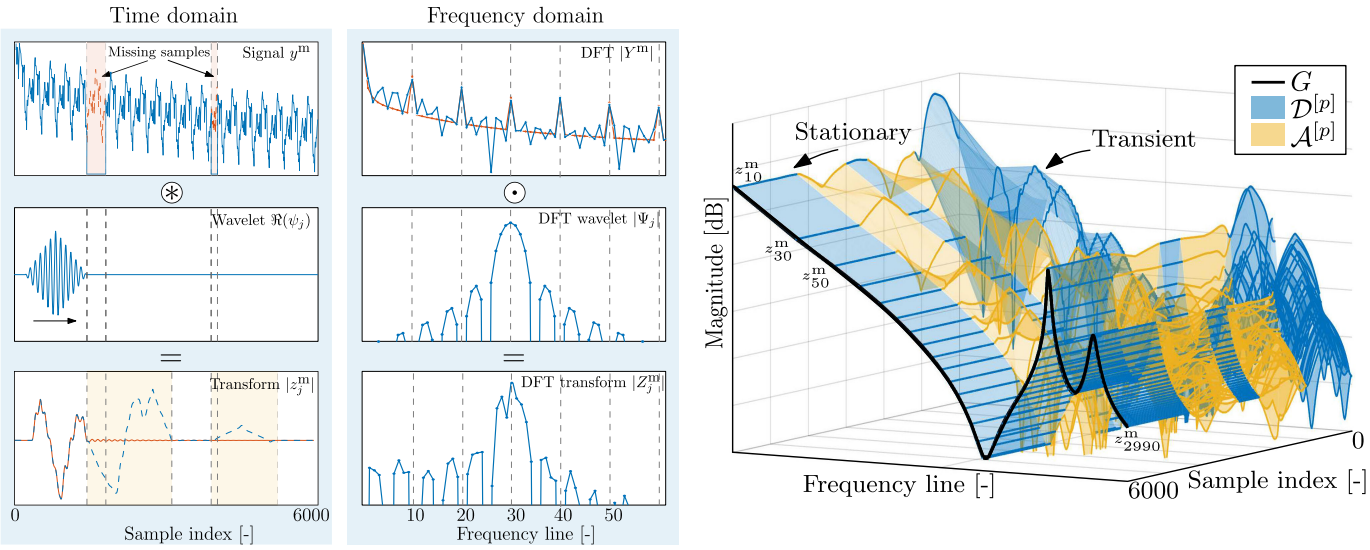
$$Y\big(\bar{k}_j + q\big) = T\Big(\Omega_{\bar{k}_j + q}\Big) + N_{Y_{\mathrm{st}}}\big(\bar{k}_j + q\big) \tag{11}$$

where $T = T_G + T_H$ combines the transient of the system and the noise dynamics. The transient term T is a smooth function of the frequency. It can be locally approximated by the $(R+1)$th order polynomials

$$T\Big(\Omega_{\bar{k}_j + q}\Big) = T\Big(\Omega_{\bar{k}_j}\Big) + \textstyle\sum_{s=1}^{R} t_s\big(\bar{k}_j\big) q^s + \mathcal{O}_{\bar{k}_j}(q) \tag{12}$$

where $\mathcal{O}_{\bar{k}_j}(q) = (1/\sqrt{\bar{N}})\mathcal{O}((q/\bar{N})^{R+1})$ is the remainder of an $(R+1)$th order Taylor series expansion around $T(\Omega_{\bar{k}_j})$, see [27]. The polynomial approximation (12) is shown in Fig. 4. At the EF lines, the output (4) is expressed as

$$Y\big(\bar{k}_j\big) = G\Big(\Omega_{\bar{k}_j}\Big) U\big(\bar{k}\big) + T\Big(\Omega_{\bar{k}_j}\Big) + N_{Y_{\mathrm{st}}}\big(\bar{k}_j\big). \tag{13}$$



(a) (b)

Fig. 3. (a) Wavelet-based convolution of signal y^{m} with a wavelet $\psi_j^{[2]}$, $j = 30$, represented in the time domain (left column) and the frequency domain (right column). Left top: signal y^{m} (—) with locally missing samples (- -). Left center: convolution is visualized as a translation over time of the wavelet ψ_j with respect to the signal y^{m}, see (7). Left bottom: transforms $z_j^{\mathrm{m}} = \{y^{\mathrm{m}} \circledast \psi_j\}$ (- -) and $z_j = \{y \circledast \psi_j\}$ in case of complete data (—). The effect of the missing samples in z_j^{m} is confined within the yellow boxes, which is associated with the set $\mathcal{A}^{[p]}$, see (10). The samples outside the boxes, associated with set $\mathcal{D}^{[p]}$, see (9), provide information of the system and the transient that is unaffected by the missing samples. Right top: DFT Y^{m} (—) and the DFT Y (—). The locally missing samples in the time domain affect the DFT Y^{m} at all frequency lines, showing the global impact on missing samples in the frequency domain. Right center: DFT magnitude $|\Psi_j|$ is largest at its central frequency line $j = 30$, enabling approximate frequency-domain localization around this line. Right bottom: DFT Z_j^{m} is dominated by the frequency content around the jth frequency line. (b) Transforms z_j^{m}, $j = 10, 30, \ldots, 2990$ provide a time–frequency representation of the stationary response of the system G and the transient T_G, in which the unaffected samples in $\mathcal{D}^{[p]}$ (■) are separated from the affected samples in $\mathcal{A}^{[p]}$ (■) on the time axis, see (9) and (10).

Combining (11)–(13) enables expressing the output Y in the local frequency window $\bar{k}_j + r$, with

$$r = \left[-\gamma P + 1, -\gamma P + 2, \ldots, \gamma P - 1\right]^T \tag{14}$$

as the regression model

$$Y\left(\bar{k}_j + r\right) = K_{\bar{k}_j}(r)\Theta_{\bar{k}_j} + \mathcal{O}_{\bar{k}_j}(r) + N_{Y_{\mathrm{st}}}\left(\bar{k}_j + r\right). \tag{15}$$

Here, $\Theta_{\bar{k}_j}$ collects the $n_\Theta = R+2$ unknown plant and transient parameters around the EF line $\bar{k}_j$

$$\Theta_{\bar{k}_j} = \left[\, G\left(\Omega_{\bar{k}_j}\right) \quad T\left(\Omega_{\bar{k}_j}\right) \quad t_1\left(\bar{k}_j\right) \quad \cdots \quad t_R\left(\bar{k}_j\right) \,\right]^T \tag{16}$$

and the regressor $K_{\bar{k}_j}(r)$ is composed as

$$K_{\bar{k}_j}(r) = \left[\, U\left(\bar{k}_j + r\right) \quad r^0 \quad r^1 \quad \cdots \quad r^R \,\right]. \tag{17}$$

Choosing the polynomial order R such that $n_r > n_\Theta$ enables estimating the parameters $\Theta_{\bar{k}_j}$ from the $n_r = 2(\gamma P - 1) + 1$ frequency lines around the EF line $\bar{k}_j$ via the linear least-squares minimization problem

$$\hat{\Theta}_{\bar{k}_j} = \underset{\Theta_{\bar{k}_j} \in \mathbb{C}^{n_\Theta}}{\arg\min} \left\| Y\left(\bar{k}_j + r\right) - K_{\bar{k}_j}(r)\Theta_{\bar{k}_j} \right\|_2^2. \tag{18}$$

Asymmetric frequency windows r may be chosen near the boundaries, see [27].

The limitations of the LPM (18) in view of the missing samples problem become apparent by observing that the method, by exploiting local behavior, hinges on the frequency-separation principle. In addition, through (12), it hinges on smooth frequency-domain characteristics of the transient. These properties no longer apply when missing samples are present. In Section IV-B, the LPM is extended toward a wavelet-based framework that addresses these limitations, which constitutes contribution 1.

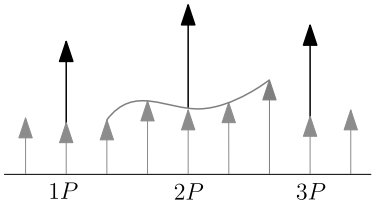


Fig. 4. Frequency-domain representation of response Y in (4) for $N_Y = 0$, $P = 3$, and $\gamma = 1$. At the EF lines $[1P, 2P, 3P]$, the response consists of both the transient (gray arrows) and the periodic response (black arrows). At the NEF lines, the response contains only transient contributions. The gray line represents the local transient estimate based on the n_r frequency lines around the $2P$th EF line.

B. Wavelet-Based LPM for Handling Missing Data

In this section, the wavelet-based LPM estimator is formulated for FRF identification in the presence of missing samples, as part of contribution 1. The extension upon the traditional LPM (18) is twofold. First, the local criteria (18) are combined into a single global criterion to facilitate handling the global effect of the missing samples. Second, a global weighting matrix $\bar{M}$ is introduced, which embeds the wavelet-based transformation for a bank of wavelets.

Exploiting the local parametrization (15) of the classical LPM, the global frequency-domain output Y is parametrized exactly by the regression model

$$Y = \bar{K}\bar{\Theta} + \bar{\mathcal{O}} + N_{Y_{\mathrm{st}}} + Y_\Delta. \tag{19}$$

Here, the $n_{\bar{\Theta}} = 2\tilde{N}n_{\Theta}$ parameters are given by

$$\bar{\Theta} = \begin{bmatrix} \Re\left(\begin{bmatrix} \Theta_{\tilde{k}_0}^T & \cdots & \Theta_{\tilde{k}_{\tilde{N}-1}}^T \end{bmatrix}^T\right) \\ \Im\left(\begin{bmatrix} \Theta_{\tilde{k}_0}^T & \cdots & \Theta_{\tilde{k}_{\tilde{N}-1}}^T \end{bmatrix}^T\right) \end{bmatrix} \in \mathbb{R}^{n_{\bar{\Theta}}}$$

where $\tilde{k} = [\tilde{k}_0, \ldots, \tilde{k}_{\tilde{N}-1}] \subset \bar{k}$ is a yet to be chosen vector $\tilde{k}$ that selects the desired frequency lines and $\Theta_{\tilde{k}_j}$ is of the form (16). The regressor $\bar{K}$ is defined as

$$\bar{K} = \begin{bmatrix} 0_{1\times\frac{1}{2}n_{\bar{\Theta}}} & 0_{1\times\frac{1}{2}n_{\bar{\Theta}}} \\ \tilde{K} & i\cdot\tilde{K} \\ 0_{1\times\frac{1}{2}n_{\bar{\Theta}}} & 0_{1\times\frac{1}{2}n_{\bar{\Theta}}} \\ \mathcal{I}\mathrm{conj}(\tilde{K}) & -i\cdot\mathcal{I}\mathrm{conj}(\tilde{K}) \end{bmatrix} \in \mathbb{C}^{\bar{N}\times n_{\bar{\Theta}}}. \quad (20)$$

Here, the zero rows coincide with the zero and Nyquist frequencies, matrix $\mathcal{I}$ is a $((1/2)\bar{N}-1)$-sized exchange matrix, and $\tilde{K} \in \mathbb{C}^{(1/2)\bar{N}-1\times(1/2)n_{\bar{\Theta}}}$ is composed of the local matrices $K_{\tilde{k}_j}(r_j)$ of the form (17) as

$$\left[\tilde{K}\right]_{\tilde{k}_j+r_j-1,c_j} = K_{\tilde{k}_j}(r_j) \quad (21)$$

with column index $c_j = [jn_{\Theta}, jn_{\Theta}+1, \ldots, (j+1)n_{\Theta}-1]$. The term $\bar{\mathcal{O}} \in \mathbb{C}^{\bar{N}}$ in (19) consists of the local remainders $\mathcal{O}$ in (12) and is vertically stacked accordingly to (20) and (21). The term $Y_{\Delta} \in \mathbb{C}^{\bar{N}}$ in (19) consists of the deterministic contributions in Y that lie outside the support of $\bar{K}$ and $\bar{\mathcal{O}}$ such that $[Y_{\Delta}]_{\tilde{k}_j+r_j} = 0,\ j = 0, 1, \ldots, \bar{N}-1$.

Similar to (19), the frequency-domain output Y^{m} with missing samples is parameterized as the regression model

$$Y^{\mathrm{m}} = \bar{K}\bar{\Theta} + \bar{\mathcal{O}} + N_{Y_{\mathrm{st}}} + Y_{\Delta} + \left(Y^{\mathrm{m}} - Y\right). \quad (22)$$

The term $(Y^{\mathrm{m}} - Y)$ forms an omitted variable in the inference of the parameters $\bar{\Theta}$ from the incomplete data (22). The key point of the wavelet-based LPM is the use of the weighting matrix $\bar{M}$ in the regression problem

$$\hat{\bar{\Theta}} = \underset{\bar{\Theta}\in\mathbb{R}^{n_{\bar{\Theta}}}}{\arg\min} \left\|\bar{M}\left(Y^{\mathrm{m}} - \bar{K}\bar{\Theta}\right)\right\|_2^2 \quad (23)$$

in which $\bar{M}$ incorporates a bank of wavelets to eliminate the perturbation term $(Y^{\mathrm{m}} - Y)$. Specifically, $\bar{M}$ is selected as

$$\bar{M} = \begin{bmatrix} M_{\tilde{q}_0}^T & \cdots & M_{\tilde{q}_{\bar{N}_q-1}}^T \end{bmatrix}^T \quad (24)$$

where

$$M_{\tilde{q}_j} = \mathcal{W}\mathcal{H}\mathrm{diag}\left(\Psi_{\tilde{q}_j}^{[p]}\right)\mathcal{F}\mathcal{R}\mathcal{H} \quad (25)$$

with $\tilde{q} = [\tilde{q}_0, \ldots, \tilde{q}_{\bar{N}_q-1}] \subset \bar{k}$. Here, $\mathcal{W}$ is a diagonal weighting matrix, $\Psi_{\tilde{q}_j}$ represents the DFT of the wavelet function $\psi_{\tilde{q}_j}$ defined in Definition 2, and $\mathcal{R}$ is a to-be-chosen matrix. The specific selection and interpretation of these parameters are presented later.

The considered estimator associated with (23) is given by

$$\hat{\bar{\Theta}} = \left(\Re\left(\bar{K}^H\bar{M}^H\bar{M}\bar{K}\right)\right)^{-1}\Re\left(\bar{K}^H\bar{M}^H\bar{M}Y^{\mathrm{m}}\right). \quad (26)$$

The following result provides the conditions for $\bar{M}$ under which the estimate $\hat{\bar{\Theta}}$ is exact.

Theorem 1 (Exact Estimation): Assume that the noise term $N_{Y_{\mathrm{st}}}$ and the remainder $\bar{\mathcal{O}}$ in (22) are zero. Then, the estimate $\hat{\bar{\Theta}}$ to (26) is exact, i.e., $\hat{\bar{\Theta}} = \bar{\Theta}$, if for $j = 0, 1, \ldots, \bar{N}-1$.

1) $M_{\tilde{q}_j}(Y - Y^{\mathrm{m}}) = 0$.
2) $\mathrm{rank}(\bar{M}\bar{K}) = n_{\bar{\Theta}}$.
3) $M_{\tilde{q}_j}Y_{\Delta} = 0$.

Conditions 1)–3) in Theorem 1 formulate the essential requirements for establishing an exact estimator in the sense of Theorem 1. Conditions 1) and 3) impose that the solution is invariant to the effects of the missing samples and the contribution Y_{Δ}, respectively, i.e., $\bar{M}$ may be regarded as acting as an instrumental variable [33, Ch. 8] to these variables. Condition 2) ensures that the solution $\hat{\bar{\Theta}}$ is unique.

The conditions in Theorem 1 impose specific demands on the selection of the parameters $\{\mathcal{W}, \Psi, \mathcal{R}\}$. In addition, this parameter selection is crucial in the tradeoff between variance and bias errors. In Sections IV-C and IV-D, it is shown that the estimator (26) encompasses both a TIE and a TVE, through the choice of the matrix $\bar{M}$. The TIE and TVE estimators provide different mechanisms to achieve conditions 1)–3) in Theorem 1 and to deal with bias and variance errors.

C. Time-Invariant Wavelet-Based Estimator

The main idea of the TIE is to identify the FRF in the time–frequency plane from the data associated with set $\mathcal{D}^{[p]}$ in (9). In this section, the method is formalized, and its properties in view of bias and variance errors are analyzed. This section constitutes contributions 1-a) and 2.

1) Estimator Formulation: The TIE arises from the generalized estimator (26) by setting $\mathcal{R} = I$ in (25), by which $M_{\tilde{q}_j}$ reduces to

$$M_{\tilde{q}_j} = \mathcal{W}\mathcal{H}\mathrm{diag}\left(\Psi_{\tilde{q}_j}^{[p]}\right). \quad (27)$$

Note that $M_{\tilde{q}_j}Y^{\mathrm{m}}$ reflects the time–frequency plane transformation (8) with the additional time weight $\mathcal{W}$.

To obtain an exact estimator (23) in the sense of Theorem 1, the parameters $\mathcal{W}$ and $\Psi_{\tilde{q}_j}$ in (27) must be selected such that conditions 1)–3) in Theorem 1 are satisfied. To satisfy condition 1), the weighting matrix $\mathcal{W}$ is chosen such that the samples of $z_{\tilde{q}_j}^{\mathrm{m}}$ that are affected by the missing samples in y^{m} are discarded, to retain only the unaffected samples collected in set $\mathcal{D}^{[p]}$ in (9), i.e., $\mathcal{W} = W^{\mathcal{D}}$, where

$$\left[W^{\mathcal{D}}\right]_{v,v} = \begin{cases} 1, & \text{if } v \in \mathcal{D}^{[p]} \\ 0, & \text{otherwise} \end{cases} \quad (28)$$

for $v = 0, \ldots, \bar{N}-1$.

By virtue of (28), the cardinality of the set $\mathcal{D}^{[p]}$ upper bounds the rank of the weighted regressor, i.e., $\mathrm{rank}(\bar{M}\bar{K}) \leq \mathrm{card}(\mathcal{D}^{[p]})$. Hence, $\mathrm{card}(\mathcal{D}^{[p]})$ plays an important role in view of the rank condition 2) in Theorem 1. Since $\mathrm{card}(\mathcal{D}^{[p]})$ depends on the support of the wavelet functions $\psi_{\tilde{q}_j}$, see (9), sensible selection of the wavelet parameters is crucial. In the next section, a selection of wavelet and LPM parameters is presented, which addresses conditions 2) and 3) in Theorem 1.

TABLE I
SELECTION OF PARAMETERS FOR IDENTIFICATION AT THE ODD FREQUENCY LINES

	Par.	Value		Description
LPM	$\tilde{k}$	$[1, 3, \ldots, \frac{N}{2\gamma} - 1]$		LPM window center freq.
	r_j	r		LPM frequency window
	R	≥ 0		Transient fit order
Wavelet Ψ	$\tilde{q}$	$[1, 3, \ldots, \frac{N}{2\gamma} - 1]$		Wavelet center frequency
	δ	γ		Wavelet zero frequencies
	p	≥ 0		Trade-off time-freq. accuracy
Estimator		TIE	TVE	
	$\mathcal{W}$	$W^{\mathcal{D}}$	I	Data discardment
	$\mathcal{R}$	I	see (37)	Modified wavelet properties
	A^{f}	–	$\mathcal{F}_v$	Frequency-domain constraints
	A^{t}	–	see (40)	Time-domain constraints

2) Selection of LPM Window and Wavelet Parameters: In this section, the selection of the LPM parameters $\{\tilde{k}, r\}$ in (21) and wavelet parameters $\{\tilde{q}, \delta, p\}$ in (25) is presented, which leads to satisfaction of condition 3) in Theorem 1 while maximizing the cardinality of the set $\mathcal{D}^{[p]}$ in favor of condition 2). An overview of the parameters is provided in Table I. The specific selection divides the identification problem into two independent subproblems.

The first subproblem is aimed at estimating $\hat{G}$ at the odd EF lines $\bar{k}_{\mathrm{odd}} = \gamma k_{\mathrm{odd}} P$ with $k_{\mathrm{odd}} = [1, 3, \ldots, (N/2\gamma) - 1]$, as shown in Fig. 5. For notational convenience, suppose that $N/(4\gamma) \in \mathbb{N}$ in the following. The LPM windows are centered at the odd EF lines by setting $\tilde{k} = \bar{k}_{\mathrm{odd}}$, and $r_j = r$ for $j = 0, \ldots, \tilde{N}-1$, with r defined in (14). Likewise, the wavelets are centered at $\tilde{q} = \bar{k}_{\mathrm{odd}}$. This enables extracting local information around these lines while suppressing the contributions at other frequency lines. In addition, by selecting $\delta = \gamma$, the zeros of $\Psi_{\tilde{q}}$ coincide with all EF lines outside the current local window and with the zero and Nyquist frequencies (see Fig. 5).

The result of this selection is twofold. First, the regressor $\bar{K}$ in (20) is zero at the even lines $[0, 2\gamma P, 4\gamma P, \ldots, (N - \gamma)P]$. Hence, Y_Δ in Theorem 1 can attain nonzero values at these lines only. Since the wavelet DFT magnitude is zero at precisely these even lines, condition 3) in Theorem 1 is satisfied. Second, separating the estimation at the odd lines from the even lines allows for a larger δ parameter, which reduces the time-domain support l_ψ, see property 3) in Lemma 1, and thus enlarges $\mathrm{card}(\mathcal{D}^{[p]})$ in favor of condition 2) in Theorem 1.

The second subproblem is aimed at estimating $\hat{G}$ at the even EF lines $\bar{k}_{\mathrm{even}} = \gamma k_{\mathrm{even}} P$ with $k_{\mathrm{even}} = [2, 4, \ldots, (N/2\gamma) - 2]$. The selection is analogous to the first subproblem.

Next, the role of the parameter p and the number of missing samples on noise-induced variance and remainder-induced bias is analyzed.

3) Variance Analysis: The variance error is generally a complicated function of the number and the specific pattern in which the missing samples occur. The following result presents an upper bound on the achievable variance.

Theorem 2: Assume that N_Y in (4) is the DFT of a zero-mean white noise realization, i.e., $H = I$ and $T_H = 0$. Let the wavelets $\psi_{\tilde{q}}^{[p]}$ be defined by Definition 2, with parameters $\{\tilde{q}, \delta\}$ selected as in Section IV-C2. Then, the variance of the

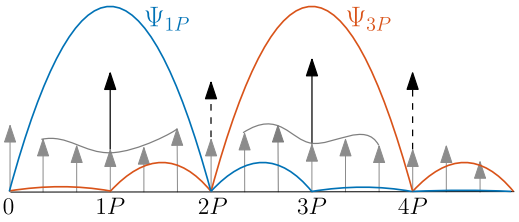


Fig. 5. Frequency-domain representation of Y in (4) for $P = 3$, $\gamma = 1$, and $\tilde{k} = \gamma k_{\mathrm{odd}} P$. The LPM windows K_{1P} and K_{3P} (—) and the wavelets Ψ_{1P} (—) and Ψ_{3P} (—) are centered at the odd EF lines $[1P, 3P]$ (solid arrows), and have zero magnitude at the even lines $[0, 2P, 4P]$ (dashed arrows).

estimates $[\hat{\bar{\Theta}}]_m$, $m = 1, \ldots, n_{\bar{\Theta}}$ is upper bounded by

$$\mathrm{var}\left([\hat{\bar{\Theta}}]_m\right) \leq \left[\left(P_{\mathrm{TIE}}^{[p]}\right)^{-1}\right]_{m,m} \leq \left[\left(P_{\mathrm{TIE}}^{[p+1]}\right)^{-1}\right]_{m,m} \tag{29}$$

with $p \in \mathbb{N}$, $\bar{B} = \bar{M}^H \bar{M}$, and

$$P_{\mathrm{TIE}}^{[p]} = \frac{2}{\lambda} \Re \sum_{i \in \mathcal{D}^{[p]}} \sum_{j=0}^{\tilde{N}_q - 1} \left(\mathcal{H}_i \mathrm{diag}\left(\Psi_{\tilde{q}_j}^{[p]}\right) \bar{K}\right)^H \left(\mathcal{H}_i \mathrm{diag}\left(\Psi_{\tilde{q}_j}^{[p]}\right) \bar{K}\right) \tag{30}$$

where $\mathcal{H}_i$ selects the ith sample from the IDFT operation.

The first inequality in (29) originates from a general variance upper bound in weighted least-squares regression [34, Sec. 9.5], in conjunction with the specific properties of the considered class of wavelets. The second inequality in (29) indicates that using wavelets with good frequency localization properties (large p) leads to an increase in variance errors. This is due to property 3) in Lemma 1, which states that the support of the wavelet function becomes larger for increasing p, which in turn reduces the cardinality of the set $\mathcal{D}^{[p]}$, leading to a smaller $P_{\mathrm{TIE}}^{[p]}$ in (30). In addition, expression (30) implies that the variance increases for an increasing number of missing samples via the reduction of the cardinality of $\mathcal{D}^{[p]}$.

Minimal variance is achieved for $p = 0$, as stated in the following result.

Corollary 1: Assume that N_Y in (4) is the DFT of a zero-mean white noise realization. The estimator (23) with $M_{\tilde{q}_j}$ in (27) achieves minimal variance for $p = 0$, which yields $\Psi^{[0]} = 1_{\tilde{N}}$. The variance is computed by $\mathrm{var}([\hat{\bar{\Theta}}]_m) = [P_{\mathrm{TIE}}^{-1}]_{m,m}$, where $P_{\mathrm{TIE}} = (2/\lambda)\Re \sum_{i \in \mathcal{D}^{[p]}} (\mathcal{H}_i \bar{K})^H (\mathcal{H}_i \bar{K})$. Note that for $p = 0$, set $\mathcal{D}^{[p]}$ contains all available samples.

Corollary 1 recovers the situation without wavelet function, by which the transform (8) reduces to the IDFT. Theorem 2 and Corollary 1 show that frequency-domain localizing wavelets negatively affect variance errors. However, frequency-domain localization provides an important instrument to attenuate bias errors, as shown next.

4) Bias Analysis: Bias in the parameter estimates due to the remainders $\mathcal{O}$ in (12) plays a particularly important role in the presence of missing samples since these samples have an effect on all the frequencies. Provided that the matrices $\bar{M}$ and $\bar{K}$ are selected to comply with conditions 1)–3) in Theorem 1 and assuming that $N_Y = 0$, this is apparent from the parameter estimates

$$\hat{\bar{\Theta}} = \bar{\Theta} + \left(\Re\left(\bar{K}^H \bar{B} \bar{K}\right)\right)^{-1} \Re\left(\bar{K}^H \bar{B} \bar{\mathcal{O}}\right) \tag{31}$$

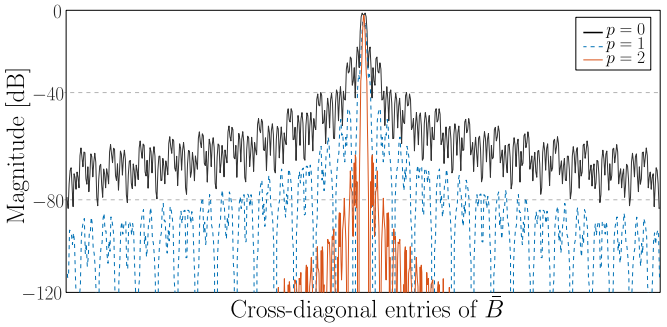


Fig. 6. Cross-diagonal entries of $\bar{B}$ for $p = 0$ (—), $p = 1$ (—), and $p = 2$ (—), as indicator for coupling. For increasing p, the magnitude of $\bar{B}$ at all entries other than the center entry reduces significantly, which indicates a significant reduction of coupling.

where $\bar{B} = \bar{M}^H \bar{M}$. Expression (31) shows that nondiagonality of the matrix $\bar{B}$ introduces global coupling between the local estimates $\Theta_{\tilde{k}_j}$, even when the structure of $\tilde{K}$ in (21) is decoupled (block diagonal). Thus, matrix $\bar{B}$ is an indicator of coupling-induced bias errors in the parameter estimates. Expressing $\bar{B}$ as

$$\bar{B} = \bar{M}^H \bar{M} = \sum_{j=0}^{\tilde{N}_q - 1} \text{diag}\left(\Psi_{\tilde{q}_j}^{[p]}\right)^H \mathcal{F}\mathcal{W}\mathcal{H}\text{diag}\left(\Psi_{\tilde{q}_j}^{[p]}\right) \quad (32)$$

shows that $\Psi_{\tilde{q}}^{[p]}$ forms the crucial design parameters to achieve approximate diagonality of the matrix $\bar{B}$ and hence to mitigate the coupling-induced bias errors. The following example demonstrates the role of the wavelet $\Psi^{[p]}$, and in particular of its design parameter p, in the diagonalization of $\bar{B}$.

Example 3: Consider the scenario of Example 2, for wavelets $\psi_{\tilde{q}}^{[p]}$ for $p = 0, 1$, and 2. To visualize the achieved approximate diagonality of the matrix $\bar{B}$ in (32), its cross-diagonal entries are shown in Fig. 6. For increasing values of p, the cross-diagonal entries of $\bar{B}$ reduce significantly.

Example 3 and (32) indicate that achieving approximate diagonality of $\bar{B}$ requires wavelets with good frequency-domain localization properties and thus large values of p. This observation, together with Corollary 1, shows that the minimization of variance on the one hand and of bias on the other hand require the opposite wavelets in terms of parameter p. This once more illustrates the intrinsic tradeoff between time- and frequency-domain accuracy in wavelet selection, as discussed in Section III-B.

5) Computational Aspects: In contrast to existing approaches such as the LPM for missing samples [13] and the LPM for concatentated datasets [15], the wavelet-based approach aims at eliminating the effects of the missing samples, rather than reconstructing the missing parts. As a result, the total number of estimated parameters does not increase with the number of missing data samples (n_m) or data gaps (n_g). This is supported by the results in the second column of Table II, which compares the total number of estimation variables for different approaches. Here, $n_G = 2\tilde{N}$ represents the total number of real-valued plant parameters for all EF lines, while $n_T = 2\tilde{N}(R + 1)$ represents the total number of real-valued transient parameters.

TABLE II
COMPARISON OF COMPUTATIONAL ASPECTS FOR DIFFERENT METHODS

Method	#Params.	Complexity in n_m, n_g
Classic LPM (Sec. IV-A)	$n_G + n_T$	$\mathcal{O}(1)$
LPM Miss. samples [13]	$n_G + n_T + n_m$	$\mathcal{O}(n_m^3) + \mathcal{O}(\tilde{N} n_\Theta n_m^2)$
LPM Concat. data [15]	$n_G + n_g n_T$	$\mathcal{O}\left(\tilde{N} n_\Theta^3 n_g^3\right)$
Wavelet-based LPM	$n_G + n_T$	$\mathcal{O}(1)$

In addition, the third column of Table II compares the complexity of the n_m- and n_g-dependent computations. The dominant computation in the missing data approach [13] is in the QR-factorization of a matrix whose size depends linearly on n_m, resulting in a cubic complexity in terms of n_m. In the concatenated data approach [15], a number of $\tilde{N}$ systems of equations with dimensionality $n_g n_\Theta$ are solved, resulting in a cubic complexity in terms of n_g. For the wavelet-based LPM approach, the pattern of the missing data appears in the form of the weighting matrix $\mathcal{W}$ in (32), by which the computational complexity is independent of the number of missing samples. The wavelet-based LPM requires the computation of the DFT of $\tilde{N}$ wavelet functions ψ_j. This computation is performed once and is done efficiently using the fast Fourier transform (FFT), resulting in a complexity of $\tilde{N}\bar{N}\log(\bar{N})$.

6) Discussion: The TIE forms an effective approach to obtain low-bias FRF estimates, making the method well-suited as a nonparametric preprocessing step prior to identifying a parametric model. Furthermore, avoiding reconstruction of the missing data eliminates the need for introducing priors, e.g., smoothness assumptions, associated with the reconstruction procedure, as in [13].

A disadvantage of the TIE is that the cardinality of set $\mathcal{D}^{[p]}$ becomes small when (many) missing samples occur in a scattered pattern in the signal y^{m}, which leads to a large variance induced uncertainty in the parameters. In a worst case situation set, $\mathcal{D}^{[p]}$ is empty, in which case the TIE does not provide a viable method. In Section IV-D, the TVE is presented, which addresses the limitation of the TIE by eliminating the need for data discardment.

D. Time-Varying Wavelet-Based Estimator

The main idea of the TVE is to avoid data discardment by adjusting the wavelet, at each sample of the convolution, to the pattern of the missing samples. As such, the TVE addresses the main limitation of the TIE. In this section, the time-varying wavelet-based transformation and the resulting TVE are presented, and its bias and variance properties are analyzed. This section constitutes contributions 1-b) and 2.

1) Time-Varying Wavelet-Based Transformation: The time-varying wavelet-based transformation lies at the basis of the TVE and is defined as follows.

Definition 3 (Time-Varying Wavelet Convolution): Consider the wavelet function $\tilde{\psi}_j^{\langle n \rangle} \in \mathbb{C}^{\bar{N}}$ that is employed at instance n. Let $y_{\bar{N}}$ be the $\bar{N}$-periodic extension of $y \in \mathbb{R}^{\bar{N}}$. The circular convolution $z_j \in \mathbb{C}^{\bar{N}}$ of the signal y with the

wavelet $\tilde{\psi}_j^{\langle n\rangle}$ at samples $n = 0, \ldots, \bar{N}-1$ is defined as

$$z_j[n] = \left\{ y \circledast \tilde{\psi}_j^{\langle n\rangle} \right\}[n] = \frac{1}{\sqrt{\bar{N}}} \sum_{m=0}^{\bar{N}-1} \tilde{\psi}_j^{\langle n\rangle}[m] \cdot y_{\bar{N}}[n-m]. \quad (33)$$

The key point of the time-varying wavelet convolution in (33) is that it enables obtaining a transformation that is invariant to the missing samples, i.e.,

$$z_j^{\mathrm{m}}[n] = \left\{ y^{\mathrm{m}} \circledast \tilde{\psi}_j^{\langle n\rangle} \right\}[n] = \left\{ y \circledast \tilde{\psi}_j^{\langle n\rangle} \right\}[n] \quad (34)$$

provided that $\tilde{\psi}^{\langle n\rangle}$ is suitably selected.

Specifically, (34) is achieved by enforcing structural zeros at the entries of $\tilde{\psi}_j^{\langle n\rangle}$ that coincide with the missing samples in y^{m} during the convolution, as shown in Fig. 7. A similar principle is applied in [24] in image compression. Such wavelet modification enables avoiding the discardment of samples compared to the TIE in Section IV-C, which is beneficial in view of the rank condition 2) in Theorem 1 and for the mitigation of variance errors.

Adjusting the wavelet function to the missing samples affects its frequency-domain characteristics and hence must be done carefully. An algorithm to compute time-varying wavelets $\tilde{\psi}^{\langle n\rangle}$ that achieve (34) and additionally adopt certain properties of a so-called template wavelet, i.e., a user-selected fixed wavelet, is presented next.

Definition 4 (Time-Varying Wavelet): Consider a user-selected template wavelet ψ_j of the form in Definition 2. The time-varying wavelet $\tilde{\psi}_j^{\langle n\rangle}$ at sample $n = 0, \ldots, \bar{N}-1$ is defined as

$$\tilde{\psi}_j^{\langle n\rangle} = \underset{\tilde{\psi}_j^{\langle n\rangle}}{\arg\min} \left\| \psi_j - \tilde{\psi}_j^{\langle n\rangle} \right\|_2^2 \quad \text{s.t. } A^{\langle n\rangle} \tilde{\psi}_j^{\langle n\rangle} = b_j^{\langle n\rangle}$$

where

$$A^{\langle n\rangle} = \underbrace{\begin{bmatrix} A^{\mathrm{m}} \\ A^f \\ A^t \end{bmatrix}}_{:=A} \mathcal{P}^{\langle n\rangle}, \quad b_j^{\langle n\rangle} = \begin{bmatrix} 0 \\ A^f \\ A^t \end{bmatrix} \mathcal{P}^{\langle n\rangle} \psi_j. \quad (35)$$

Matrix $A^{\mathrm{m}} = I_{\bar{N}} - W^{\mathrm{m}}$ incorporates the time-domain pattern of the missing samples in the wavelet function $\tilde{\psi}_j^{\langle n\rangle}$, with W^{m} in (5). The matrix A^f imposes frequency-domain constraints on $\tilde{\psi}_j$ and is chosen as $A^f = \mathcal{F}_v$, where user-defined parameter v selects the vth frequency line in $\mathcal{F}$. The matrix A^t imposes time-domain constraints on $\tilde{\psi}_j$ and is selected as $A^t = [\varphi_1^T, \ldots, \varphi_{n_{\mathrm{tc}}}^T]^T$, where φ_i, $i = 1, \ldots, n_{\mathrm{tc}}$ represent user-defined basis-functions. The time-varying matrix $\mathcal{P}^{\langle n\rangle}$ incorporates the effect of the translation of the wavelet with respect to the signal y^{m} and is given by $\mathcal{P}^{\langle n\rangle} = \left[\begin{smallmatrix} \mathcal{I}_{n+1} & 0 \\ 0 & \mathcal{I}_{\bar{N}-(n+1)} \end{smallmatrix}\right]$, where $\mathcal{I}_m$ is an $(m \times m)$-dimensional exchange matrix.

The rationale of Definition 4 is that the adjusted wavelet $\tilde{\psi}_j^{\langle n\rangle}$ remains close to the template wavelet ψ_j, in 2-norm, as to approximately preserve its properties. In addition, the constraint matrices A^t and A^f in (35) enable adopting specific time- and frequency-domain properties of the template wavelet, as selected by the user. The selection of these constraints is discussed later. First, consider the following result, which enables incorporating the transform (33) in the estimator (26).

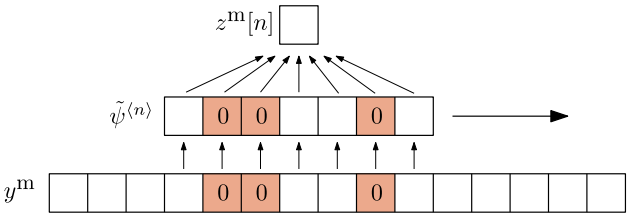


Fig. 7. Visualization of time-varying wavelet-based convolution at sample n. The entries of the wavelet function $\tilde{\psi}_j^{\langle n\rangle}$ that coincide with the missing samples in y^{m} are set to 0.

Lemma 2: Let $\tilde{\psi}_j^{\langle n\rangle}$ be defined in Definition 4. Then, the transform in (33) is equivalently expressed as

$$z_j[n] = \mathcal{H}_n \mathrm{diag}(\Psi_j) \mathcal{F}\mathcal{R}\mathcal{H} Y \quad (36)$$

where $\mathcal{H}_n$ selects the nth sample in $\mathcal{H}$ and

$$\mathcal{R} = \left(I - \begin{bmatrix} A^{\mathrm{m}} & 0 & 0 \end{bmatrix} (A^T)^{\dagger} \right) W^{\mathrm{m}} \quad (37)$$

with A given in (35) and W^{m} in (5).

The achieved result of Lemma 2 is that the convolution with a time-varying wavelet in (33) is posed as the frequency-domain operation (36) involving time-invariant matrices. In addition, expression (36) separates the template wavelet Ψ_j from the pattern of the missing samples in A^{m} and W^{m} in (37) and from the selected constraints in matrices A^f and A^t. Next, the transform (36) is embedded in the generalized LPM estimator.

2) Estimator Formulation: The TVE is obtained from the general estimator (26) by setting $\mathcal{W} = I$ in (25) and setting $\mathcal{R}$ as in (37). This yields the weighting matrix

$$M_{\tilde{q}_j} = \mathcal{H} \mathrm{diag}\left(\Psi_{\tilde{q}_j}^{[p]} \right) \mathcal{F}\mathcal{R}\mathcal{H} \quad (38)$$

which embeds the transformation (36) into (23). The template wavelets $\Psi_{\tilde{q}_j}^{[p]}$ are selected as discussed in Section IV-C.

The key advantage of the resulting TVE compared to the TIE in Section IV-C is that it enables satisfying condition 1) in Theorem 1 without the discardment of samples. By avoiding discardment, the cardinality of the set $\mathcal{D}^{[p]}$ in (9) no longer limits the rank condition 2) in Theorem 1. Instead, for the TVE, it is obtained that $\mathrm{rank}(\bar{M}\bar{K}) \leq \bar{N} - n_m$. Comparing this bound to that of the TIE, where $\mathrm{rank}(\bar{M}\bar{K}) \leq \mathrm{card}(\mathcal{D}^{[p]})$, and observing that $\bar{N} - n_m > \mathrm{card}(\mathcal{D}^{[p]})$ when $p > 0$, shows that the TVE is capable of dealing with a larger number of missing samples than the TIE.

The remaining design freedom in the weighting matrix $M_{\tilde{q}_j}$ in (38) lies in the choice of constraints that determine $\mathcal{R}$ in (37). A constraint selection that leads to satisfaction of condition 3) in Theorem 1 is presented next.

3) Selection of Constraint Parameters: The constraints in (35) provide an important instrument to ensure that the time-varying wavelet $\tilde{\psi}^{\langle n\rangle}$ preserves, as much as possible, the favorable properties of the template wavelet.

In particular, to obtain an exact estimator in the sense of Theorem 1, it is required that condition 3) in Theorem 1 remains met for the time-varying wavelet. As discussed in Section IV-C2, the key mechanism to achieve this condition

is the specific placement of the zeros in the DFT Ψ. For the TVE, it is therefore crucial that the zeros in the DFT Ψ of the template wavelet are adopted in the DFT $\tilde{\Psi}^{\langle n \rangle}$ of the modified wavelet. This is achieved via the selection of the constraint matrix A^f, as shown in the following result.

Theorem 3: Suppose that $\text{rank}(\bar{M}\bar{K}) = n_{\bar{\Theta}}$. Furthermore, let the LPM window and wavelet center frequencies be selected as $\tilde{k} = \tilde{q} = \bar{k}_{\text{odd}}$, and let $p > 0$ and $\delta = \gamma$. Then, the TVE is an exact estimator in the sense of Theorem 1 if the matrix A^f is selected as $A^f = \mathcal{F}_v$ with v the even frequency lines

$$v = \big[0, 2\gamma P, 4\gamma P, \ldots, (N - 2\gamma)P\big]. \tag{39}$$

A similar result applies to the estimation of $\hat{G}$ at the even frequencies. An example of a modified wavelet $\tilde{\psi}_j$ that adopts the zeros of the DFT of template wavelet Ψ_j is given in the following.

Example 4: Let $N = 500, P = 1$, and $\gamma = 5$. Consider the template wavelet $\psi_{15}^{[2]}$ (—) in Fig. 8. For samples $150, \ldots, 180$ missing, the corresponding modified wavelet $\tilde{\psi}_{15}^{[2]}$ is shown in (—). It is constructed by selecting the frequency-domain constraints as in Theorem 3. The wavelet $\tilde{\psi}_{15}^{[2]}$ has zero magnitude in time domain at the missing samples. In addition, the zeros in the DFT $\tilde{\Psi}_{15}^{[2]}$ at the even frequency lines $v = [0, 10, 20, \ldots, 490]$ coincide with those of the template wavelet. In addition, the modified wavelet well-preserves the global shape of the template wavelet.

To further shape the modified wavelet $\tilde{\psi}^{\langle n \rangle}$, additional constraints may be imposed. For example, the time-domain constraint matrix A^t may be selected as the orthonormal basis of $[\varphi_1^T, \ldots, \varphi_{n_{\text{tc}}}^T]^T$, where

$$\varphi_i = \left[-\frac{1}{2}\bar{N}, -\frac{1}{2}\bar{N} + 1, \ldots, \frac{1}{2}\bar{N} - 1\right]^i \tag{40}$$

represents real-valued polynomials of degree $i = 1, \ldots, n_{\text{tc}}$. The basis functions (40) emphasize the preservation of the low-frequency characteristics of the template wavelet, which is typically desirable for systems with first-order transient behavior. Other basis functions may be chosen, depending on the application. The TVE parameters are summarized in Table I. Variance and bias properties of the TVE are analyzed in the following.

4) Variance Analysis: The use of the time-varying wavelet-based transformation is beneficial for mitigating variance errors. The following result shows that the TVE admits a tighter upper bound on the variance in the estimated parameters than the TIE.

Theorem 4: Assume that N_Y in (4) is the DFT of a zero-mean white noise realization and assume that $A^t, A^f = 0$. Let the wavelets $\psi_{\tilde{q}}^{[p]}$ be defined by Definition 2, with parameters $\{\tilde{q}, \delta\}$ selected as in Section IV-C2. The variance of $[\hat{\bar{\Theta}}]_m$, $m = 1, \ldots, n_{\bar{\Theta}}$ obtained from the TVE is upper bounded by

$$\text{var}\big([\hat{\bar{\Theta}}]_m\big) \leq \big[P_{\text{TVE}}^{-1}\big]_{m,m} \leq \big[P_{\text{TIE}}^{-1}\big]_{m,m} \tag{41}$$

where $\bar{B} = \bar{M}^H\bar{M}$ with $\bar{M}$ given by (24) and $M_{\tilde{q}_j}$ by (38). Furthermore, P_Θ is the matrix in (30) related to the variance

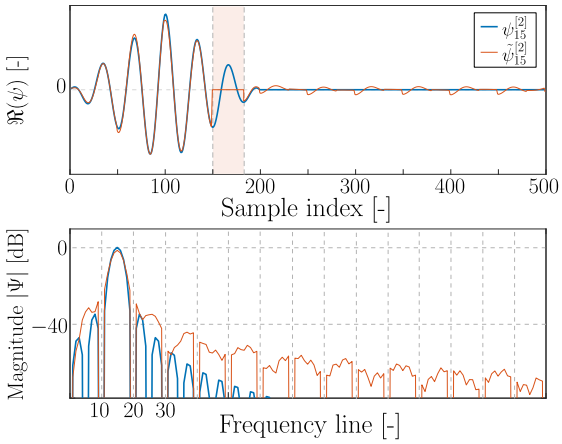


Fig. 8. Template wavelet $\psi_{15}^{[2]}$ (—) and modified wavelet $\tilde{\psi}_{15}^{[2]}$ (—). The modified wavelet $\tilde{\psi}_{15}^{[2]}$ has structural zeros at the missing samples $150, \ldots, 180$, and its DFT preserves the zeros at the even frequency lines $v = [0, 10, 20, \ldots, 490]$.

in the TIE, and P_{TVE} is given by

$$P_{\text{TVE}} = \frac{2}{\lambda}\Re\ \bar{K}^H\big(\mathcal{T}^{\mathcal{R}}\big)^H \sum_{j=0}^{\tilde{N}_q - 1} \text{diag}\left(\left|\Psi_{\tilde{q}_j}^{[p]}\right|^2\right)\mathcal{T}^{\mathcal{R}}\bar{K} \tag{42}$$

where $\mathcal{T}^{\mathcal{R}} = \mathcal{F}\mathcal{R}\mathcal{H}$.

The inequalities in (41) show that the TVE is a more efficient estimator than the TIE. Comparing P_{TVE} in (42) with P_{TIE} in (30) shows that the TVE is indeed less affected by the choice of p than the TIE since P_{TVE} does not depend on the size of the set $\mathcal{D}^{[p]}$, in contrast to P_{TIE}. Still, it can straightforwardly be shown that increasing p enlarges the upper bound on the variance, due to the frequency-domain characteristics of $\Psi_{\tilde{q}_j}^{[p]}$ in (42). Consider the following result.

Corollary 2: Under the assumptions in Theorem 4 and for $p = 0$, the TVE recovers the minimal variance estimator in Corollary 1.

5) Bias Analysis: The wavelet plays a different role in the mitigation of bias errors for the TVE than for the TIE. Recall that when $N_Y = 0$, the parameter estimates are expressed by (31), in which matrix $\bar{B}$ forms an indicator of coupling-induced bias errors. For the TVE, $\bar{B}$ is expressed as

$$\bar{B} = \bar{M}^H\bar{M} = \big(\mathcal{T}^{\mathcal{R}}\big)^H \sum_{j=0}^{\tilde{N}_q - 1} \text{diag}\left(\left|\Psi_{\tilde{q}_j}^{[p]}\right|^2\right)\mathcal{T}^{\mathcal{R}}. \tag{43}$$

This shows that for the TVE the matrix, $\bar{B}$ cannot be diagonalized via the choice of $\Psi_{\tilde{q}}^{[p]}$. Hence, the TVE provides fewer opportunities to reduce coupling-induced bias than the TIE. Still, since increasing the value of p leads to a tighter frequency window around the $\tilde{q}_j$th frequency line, choosing larger values of p generally reduces bias in the estimates.

In summary, this illustrates that the TVE generally yields low-variance estimates, at the price of bias errors, whereas the TIE produces low-bias estimates with larger variance. In Section IV-E, the two estimators are unified via a joint criterion.

E. Unification

The TIE and TVE employ different approaches to deal with missing data, yet the two estimators are closely related. Define the signal $y^r = \mathcal{R}y^{\mathrm{m}}$ and observe that, by virtue of the definition of $\mathcal{R}$ in (37), signal y^r is equal to y^{m} except at the missing samples. Since the weighting matrix $M_{\tilde{q}_j}$ in (27) used in the TIE discards the data that are affected by precisely these missing samples, it holds that

$$z_j^{\mathrm{m}}[n] = \mathcal{H}_n \mathrm{diag}\left(\Psi_{\tilde{q}_j}^{[p]}\right) Y^{\mathrm{m}} = \mathcal{H}_n \mathrm{diag}\left(\Psi_{\tilde{q}_j}^{[p]}\right) \mathcal{F}\mathcal{R}\mathcal{H} Y^{\mathrm{m}}$$

for $n \in \mathcal{D}^{[p]}$ and when $p > 0$. Thus, the TIE and TVE act identically upon the samples in the set $\mathcal{D}^{[p]}$. This enables unifying the two estimators via the joint weighting matrix

$$M_{\tilde{k}_j} = \left((1-\beta)W^{\mathcal{D}} + \beta I\right)\mathcal{H}\mathrm{diag}\left(\Psi_{\tilde{q}_j}^{[p]}\right)\mathcal{F}\mathcal{R}\mathcal{H} \qquad (44)$$

where $\beta \in [0, 1]$. The TIE is recovered when $\beta = 0$ and the TVE is recovered when $\beta = 1$.

V. Experimental Results

In this section, the TIE and TVE are experimentally validated and compared on a thermodynamic system, in support of the theoretical claims made in Section IV.

A. Experiment Setup

1) System Description: The experimental setup is shown in Fig. 9. It consists of two aluminum cylinders separated by a piece of polyoxymethylene (POM). The setup represents an industrial mixed-material thermodynamic system, in which the aluminum parts reflect a thermal conductor and the POM reflects a thermal resistance. The system is equipped with a heater and temperature sensor at the indicated locations, which form the system input u and output y, respectively.

The system dynamics are characterized by slow thermodynamic behavior, which gives rise to significant transient behavior with a duration of several hours. When there are no missing samples, the classical LPM provides an effective approach to FRF identification from transient data, as illustrated in [35]. In this section, identification in the presence of missing samples is considered.

2) Experiment Description: The system is excited with a periodic excitation signal u consisting of $P = 12$ periods with period length $N = 720$. The sample time $T_s = 5$ s, leading to a period duration of 1 h. The spectral magnitudes of the excitation signal u are uniformly distributed over the EF lines $\bar{k} = \gamma k P$ with $k = 0, \ldots, \lfloor N/\gamma \rfloor - 1$, with sparsity parameter $\gamma = 2$, recall Section IV-A. The spectral phase is selected from a uniform random distribution on $[0, 2\pi)$. Since the heat input is intrinsically constrained to be positive, the input spectrum has a positive-valued contribution at the 0-frequency, resulting in an offset in the heater input.

The measured output response y is shown in Fig. 10(a). The signal contains a significant transient response, e.g., due to the heater offset. To mimic the situation of a data record y^{m} with missing samples, a number of $n_{\mathrm{m}} = 1105$ out of $\bar{N} = 8640$ samples is set to zero via the selection matrix W^{m} in (5). The missing samples form five data gaps, as shown in Fig. 10(a). These data y^{m} form the input to the estimators. The corresponding frequency-domain representations Y and Y^{m} are shown in Fig. 10(b). The global frequency-domain impact of the locally missing samples in the time domain is evident.

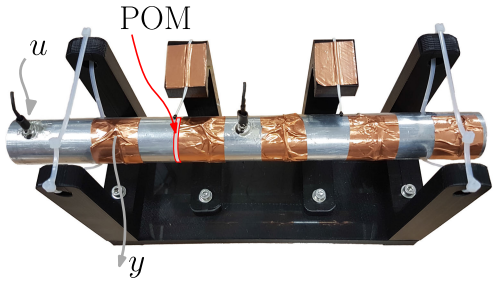


Fig. 9. Experimental setup. The aluminum cylinder has a heater input u and temperature measurement y.

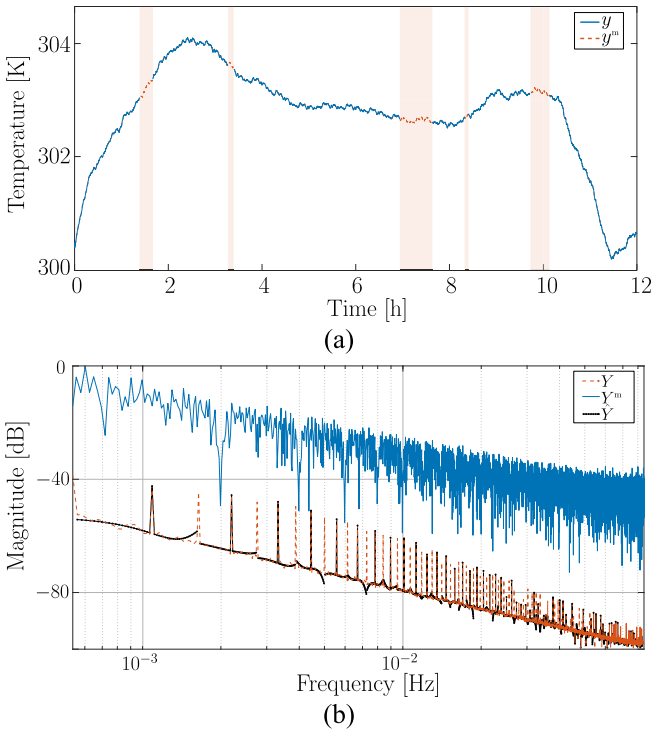


Fig. 10. Time- and frequency-domain representations of the responses of the experimental setup with and without missing samples. (a) Temperature response y (- -) and the mimicked incomplete signal y^{m} (—). Note that the signals largely overlap. The missing data parts are highlighted in red. (b) DFT Y (- -) of the complete data record and DFT Y^{m} (—) of the incomplete record. The locally missing samples in the time domain have a global effect in the frequency domain. Using the TIE with $p = 2$, the estimate $\hat{Y} = \bar{K}\bar{\Theta}$ (—) accurately approximates Y (- -) at and around the odd EF lines while having access to Y^{m} only (—).

3) Estimator Selection: In the following, the performance of the TIE and TVE is analyzed for wavelet parameter values $p = 1, 2, 3, 4$. The polynomial order R in (12) for the transient estimation is selected as $R = 3$, based on the LPM guidelines in [27]. For the TVE, a number of $n_{t_c} = 10$ basis functions φ is selected, see (40), to emphasize on low-frequency accuracy in the computation of the time-varying wavelet functions $\tilde{\psi}$.

In addition, for the purpose of error analysis, a reference model G^{ref} is identified from the complete data y using the classical LPM estimator described in Section IV-A. The resulting reference model is shown in black in Figs. 11 and 12.

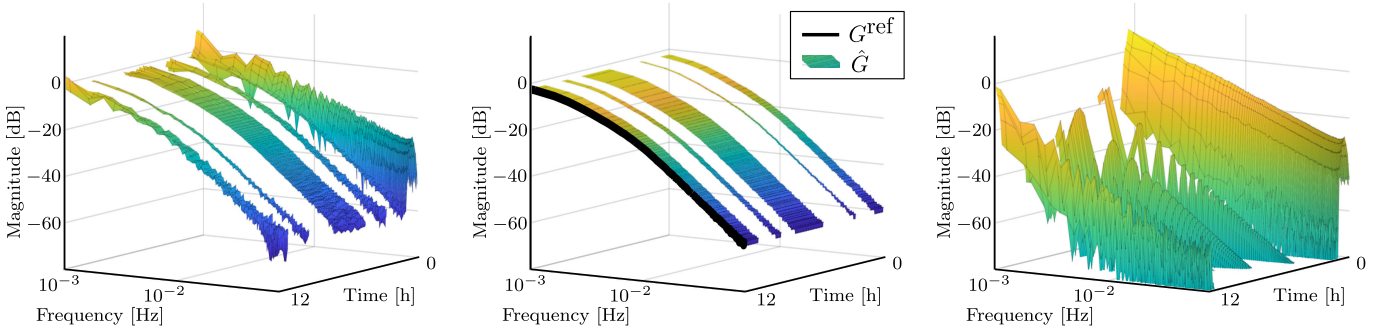


Fig. 11. Frequency–time plane results using the TIE for $p = 2$. Left: transforms $z^{\text{m}}[v]$ for samples $v \in \mathcal{D}$. Center: estimated FRF $\hat{G}$ (surface) at samples v and the FRF of the reference model G^{ref} (—). Right: estimated transient $\hat{T}$ at samples v.

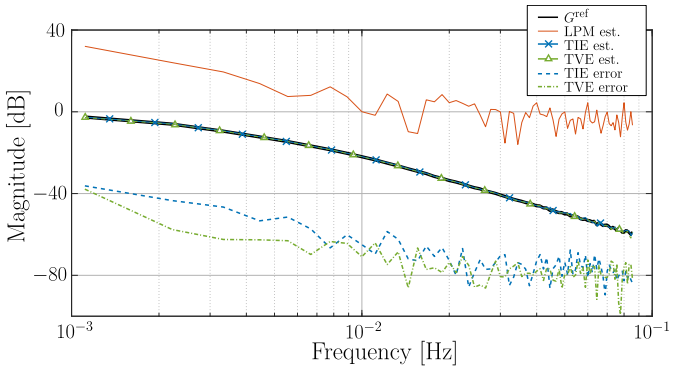


Fig. 12. Bode magnitude plot of the reference model G^{ref} (—), and the estimates obtained from the classical LPM (- -), the TIE (⋆), and the TVE (△) for $p = 2$. In addition, the model errors of the TIE (- -) and TVE are shown (- -). Both the TIE and TVE achieve accurate identification from incomplete data, in contrast to the classical LPM.

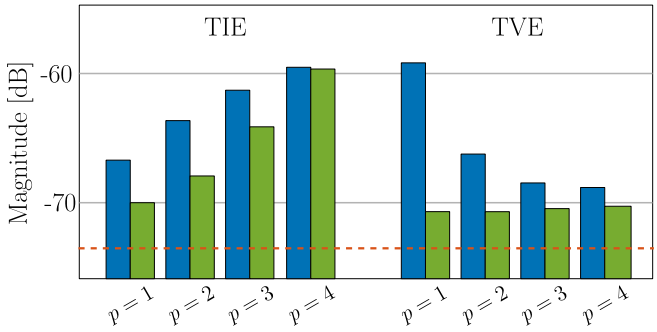


Fig. 13. Model errors $\varepsilon_{\text{TIE}}^{[p]}$ and $\varepsilon_{\text{TVE}}^{[p]}$ (■) and the estimated contributions due to stochastic noise (■).

B. Results

1) Time–Frequency Plane Results: The transforms $z_{\tilde{q}}^{\text{m}}[v]$ according to (7) are shown in the time–frequency plane for $v \in \mathcal{D}$ on the left in Fig. 11, for the wavelet parameter $p = 2$. Both stationary and transient contributions can be recognized. These contributions are distinguished by the TIE, resulting in the estimated FRF $\hat{G}$ and the transient $\hat{T}$ shown in the center and right plots, respectively. The estimated FRF $\hat{G}$ obtained from the incomplete data accurately matches the reference model in (—). Similar observations are made for the TIE for different values of p, but these results are not displayed to conserve space.

2) Frequency-Domain Results: The estimated DFT $\hat{Y} = \bar{K}\hat{\bar{\Theta}}$ using the TIE for $p = 2$ and $\bar{k} = \gamma k_{\text{odd}} P$ is shown in (—) in Fig. 10(b). Although the available incomplete data Y^{m} (—) significantly differs from the complete data Y (- -), the estimate $\hat{Y}$ (—) provides an accurate estimate of Y (- -) at and around the odd EF lines.

In Fig. 12, the models estimated by the TIE (⋆) and the TVE (△) for $p = 2$ are shown. In addition, the respective model errors with respect to the reference model (—) are shown in (- -) and (- -). Both estimators achieve accurate FRF identification. The classical LPM (- -) fails to estimate an accurate model. An error analysis for different values of p is presented next.

3) Bias and Variance Analysis: In this section, the bias and variance contributions in the model error are compared for the TIE and the TVE for different values of p.

As a global indicator of estimation performance, consider the frequency-averaged model error defined as $\varepsilon_X^{[p]} = (1/\tilde{N}) \sum_{\tilde{k}} |G^{\text{ref}}(\tilde{k}) - \hat{G}_X^{[p]}(\tilde{k})|$, where $X \in \{\text{TIE}, \text{TVE}\}$. These indicators are shown in (■) in Fig. 13. In addition, the estimated contribution of the stochastic noise to the model error is shown in (■), which is estimated from an auxiliary measurement. The lower bound on the noise contribution based on the minimal variance estimator in Corollary 1 is shown in (- -). For low values of p, the TIE and TVE both achieve variance errors close to the lower bound.

For the TIE, the model error is increasingly dominated by stochastic noise contributions for larger values of p, resulting in a model error increase. In turn, this implies that the contribution of systematic errors is reduced for larger p values. These observations are in line with the results in Section IV-C.

For the TVE, the noise contribution is lower overall and increases only slightly for increasing p. Observing that the total model error decreases with increasing p implies that the model error is dominated by systematic errors for small p, which are reduced for larger values of p. These observations are in line with the results in Section IV-D.

Overall, the smallest model error is achieved by the TVE for $p = 4$. This outcome depends on the (application-) specific balance between the level of stochastic and systematic contributions. In general, the TIE produces low-bias models, yet with a larger stochastic uncertainty. Such models may be preferred as an intermediate model toward a parametric model. The TVE, on the other hand, can provide more accurate models when the measurement noise is large or when the number of missing samples is large.

VI. Conclusion

The presented methods achieve accurate identification of nonparametric FRF models from measured data records with many missing samples. This is realized by generalizing the existing frequency-domain LPM to a wavelet-based LPM that enables isolating the effect of the missing samples in the time–frequency plane. The generalization encompasses a time-invariant estimator and a time-varying estimator, which have different performances in terms of bias and variance errors. Contrary to most existing algorithms for dealing with missing samples, a benefit of the presented approach is that it does not require the explicit estimation or reconstruction of the missing data parts, and as such, the number of estimation variables does not increase with the amount of missing data. Confrontation with a thermodynamical experimental setup confirms that accurate FRF identification is achieved.

Appendix

Proof of Lemma 1: First, properties 1) and 2) will be proven for $p = 1$. Using (1), the DFT of $\psi_j^{[p]}$ is expressed as

$$\Psi_j^{[p]}(k) = \frac{c^{[p]}}{\sqrt{\tilde{N}}} \sum_{m=0}^{\tilde{N}-1} h^{[p]}[m] \cdot e^{-1i \cdot 2\pi \frac{k-j}{\tilde{N}} m}. \tag{45}$$

For $p = 1$, expression (45) describes the DFT of a rectangular function, which is a Dirichlet kernel [31] obeying

$$\Psi_j^{[1]}(q+j) = \frac{c^{[1]}}{\sqrt{\tilde{N}}} \exp\left(\frac{i\pi q}{\tilde{N}}\left(\frac{N}{\delta} - 1\right)\right) \cdot \frac{\sin\left(\frac{\pi q}{\delta P}\right)}{\sin\left(\frac{\pi q}{\tilde{N}}\right)} \tag{46}$$

where variable substitution $k = q + j$ is applied. Properties 1) and 2) for $p = 1$ then follow from the properties of the Dirichlet kernel [31]. In particular, $\Psi_j^{[1]}(q+j)|_{q\to 0} = 1$ is obtained via L'Hopital's rule, which proves property 1). Property 2) is immediate since $\sin((\pi q/\delta P)) = 0$ for $q = \alpha\delta P,\ \forall \alpha \mathbb{N}_+$.

Next, properties 1) and 2) will be proven for $p \in \mathbb{N}$. Let $H^{[p]}$ be the DFT of $h^{[p]}$ and observe from Definition 2 that $H^{[p]} = (H^{[1]})^p$ and that $c^{[p]} = (c^{[1]})^p$. Using (45), the DFT of $\psi_j^{[p]}$ is expressed as

$$\begin{aligned}\Psi_j^{[p]}(q+j) &= c^{[p]} H^{[p]}(q) = \left(c^{[1]} H^{[1]}(q)\right)^p \\ &= \left(\Psi_j^{[1]}(q+j)\right)^p. \end{aligned} \tag{47}$$

Thus, properties 1) and 2) are immediate. Property 4) follows from (47) and the fact that the Dirichlet kernel attains its maximum magnitude of 1 at $q = 0$.

Finally, property 3) is proven. By Definition 2, for $p = 1$, the support $l_\psi^{[1]} = N/\delta = pN/\delta - p + 1$. For $p = 0$, the function $h^{[p]}$ reduces to the discrete unit impulse, and hence, $l_\psi^{[0]} = 1 = pN/\delta - p + 1$. For $p \geq 2$, it is observed that $h^{[p]}[m] = \{h^{[p-1]} \circledast h^{[1]}\}[m]$. Exploiting the standard result that the length of convolution of two sequences with lengths N and M equals $N + M - 1$, the support of $h^{[p]},\ p \geq 2$ is $l_\psi^{[p]} = l_\psi^{[p-1]} + l_\psi^{[1]} - 1 = p l_\psi^{[1]} - p + 1 = pN/\delta - p + 1$.

Proof of Theorem 1: By condition 1), $\bar{M} Y^{\mathrm{m}} = \bar{M} Y$. Furthermore, the assumptions impose that $Y = \bar{K}\bar{\Theta} + Y_\Delta$. Substituting this result into (26), it follows that $\hat{\bar{\Theta}} = \bar{\Theta}$ when condition 3) is satisfied and, in addition, when $\Re(\bar{K}^H \bar{M}^H \bar{M} \bar{K})$ is invertible. The latter always holds under condition 2).

Proof of Theorem 2: First, the first inequality in (29) is proven. For ease of notation, the superscript $[p]$ is omitted at first. Exploiting the closed-form solution (26), the property $\bar{M} Y^{\mathrm{m}} = \bar{M} Y$, and the circular complex normal distribution of $N_{Y_{\mathrm{st}}}$, the covariance matrix $\mathrm{cov}(\hat{\bar{\Theta}}) = \mathbb{E}\{\hat{\bar{\Theta}}\hat{\bar{\Theta}}^H\}$ is given by

$$\mathrm{cov}\left(\hat{\bar{\Theta}}\right) = \frac{2}{\lambda} P_{\mathrm{TIE}}^{-1} \Re\left(\bar{K}^H \bar{B} \bar{B} \bar{K}\right) P_{\mathrm{TIE}}^{-1} \tag{48}$$

where $\bar{B} = \bar{M}^H \bar{M}$ and P_{TIE} is given in (30), also see [34, Sec. 9.5]. By the definition of $\bar{M}$ in (24), matrix $\bar{B}$ is expressed as

$$\begin{aligned}\bar{B} = \bar{M}^H \bar{M} &= \sum_{j=0}^{\tilde{N}_q - 1} \mathrm{diag}\left(\Psi_{\tilde{q}_j}^{[p]}\right)^H \mathcal{F}\mathcal{W}\mathcal{H}\mathrm{diag}\left(\Psi_{\tilde{q}_j}^{[p]}\right) \\ &= \sum_{i \in \mathcal{D}^{[p]}} \sum_{j=0}^{\tilde{N}_q - 1} \left(\mathcal{H}_i \mathrm{diag}\left(\Psi_{\tilde{q}_j}^{[p]}\right)\right)^H \left(\mathcal{H}_i \mathrm{diag}\left(\Psi_{\tilde{q}_j}^{[p]}\right)\right)\end{aligned}$$

and hence, P_{TIE} satisfies the expression (30).

The maximum singular value $\bar{\sigma}$ of $\bar{B}$ is upper bounded by

$$\begin{aligned}\bar{\sigma}\left(\bar{B}\right) &= \bar{\sigma}\left(\sum_{i \in \mathcal{D}^{[p]}} \sum_{j=0}^{\tilde{N}_q - 1} \left(\mathcal{H}_i \mathrm{diag}\left(\Psi_{\tilde{q}_j}^{[p]}\right)\right)^H \left(\mathcal{H}_i \mathrm{diag}\left(\Psi_{\tilde{q}_j}^{[p]}\right)\right)\right) \\ &\leq \bar{\sigma}\left(\sum_{i=0}^{\tilde{N}-1} \sum_{j=0}^{\tilde{N}_q - 1} \left(\mathcal{H}_i \mathrm{diag}\left(\Psi_{\tilde{q}_j}^{[p]}\right)\right)^H \left(\mathcal{H}_i \mathrm{diag}\left(\Psi_{\tilde{q}_j}^{[p]}\right)\right)\right) \\ &= \bar{\sigma}\left(\sum_{j=0}^{\tilde{N}_q - 1} \mathrm{diag}\left(\left|\Psi_{\tilde{q}_j}^{[p]}\right|^2\right)\right) = 1 \end{aligned} \tag{49}$$

where the first inequality in (49) is due to Weyl's inequality [36] and the latter equality follows from the construction of $\Psi_{\tilde{q}_j}^{[p]}$, as defined in Definition 2. Since $\bar{B} \in \mathbb{H}_+$, the singular value decomposition yields $\bar{B} = \mathcal{U}\Sigma\mathcal{U}^H$, and thus, $\bar{B}\bar{B} = \mathcal{U}\Sigma^2\mathcal{U}^H$. From (48),

$$\begin{aligned}\mathrm{var}\left(\left[\hat{\bar{\Theta}}\right]_m\right) &= \frac{2}{\lambda}\left[P_{\mathrm{TIE}}^{-1} \Re\left(\bar{K}^H \bar{B} \bar{B} \bar{K}\right) P_{\mathrm{TIE}}^{-1}\right]_{m,m} \\ &\leq \bar{\sigma}\left(\bar{B}\right) \frac{2}{\lambda}\left[P_{\mathrm{TIE}}^{-1} \Re\left(\bar{K}^H \bar{B} \bar{K}\right) P_{\mathrm{TIE}}^{-1}\right]_{m,m} \\ &= \bar{\sigma}\left(\bar{B}\right)\left[P_{\mathrm{TIE}}^{-1}\right]_{m,m} \leq \left[P_{\mathrm{TIE}}^{-1}\right]_{m,m} \end{aligned} \tag{50}$$

where the latter inequality is due to (49). This proofs (29). Next, the second inequality in (29) is proven. Define $Q_i^{[p]} = (2/(\bar{\sigma}(C_Y)))\Re\sum_{j=0}^{\tilde{N}_q - 1}(\mathcal{H}_i \mathrm{diag}(\Psi_{\tilde{q}_j}^{[p]} \bar{K}))^H (\mathcal{H}_i \mathrm{diag}(\Psi_{\tilde{q}_j}^{[p]}) \bar{K})$ such that $P_{\mathrm{TIE}}^{[p]} = \sum_{i \in \mathcal{D}^{[p]}} Q_i^{[p]}$. Define the difference set $\mathcal{Z}^{[p+1]} = \mathcal{D}^{[p]} \backslash \mathcal{D}^{[p+1]}$. By definition of $\psi_{\tilde{q}_j}^{[p]}$ in Definition 2, it holds that $\Psi_{\tilde{q}_j}^{[p+1]} = \Psi_{\tilde{q}_j}^{[p]} \odot \Psi_{\tilde{q}_j}^{[1]}$. Since $Q_i^{[p]}$ is rank-one and $|\Psi_{\tilde{q}_j}^{[1]}| \leq 1$, it can be written as

$$Q_i^{[p]} = Q_i^{[p+1]} + \Delta_i^{[p+1]} \tag{51}$$

where $\Delta_i^{[p+1]}$ is symmetric positive semidefinite (PSD) and is defined by (51) itself. Thus, it can be expressed as

$$P_{\mathrm{TIE}}^{[p]} = P_{\mathrm{TIE}}^{[p+1]} + \sum_{i \in \mathcal{D}^{[p+1]}} \Delta_i^{[p+1]} + \sum_{i \in \mathcal{Z}^{[p+1]}} Q_i^{[p]}. \tag{52}$$

Since both Δ_i and Q_i are PSD, the second inequality in (29) follows by applying the Woodbury matrix identity, which completes the proof.

Proof of Lemma 2: Define $\psi_j^{\mathcal{P},\langle n\rangle} := \mathcal{P}^{\langle n\rangle}\psi_j$ and $\tilde{\psi}_j^{\mathcal{P},\langle n\rangle} := \mathcal{P}^{\langle n\rangle}\tilde{\psi}_j^{\langle n\rangle}$. Then, the transform in (33) is also expressed as

$$z_j[n] = \frac{1}{\sqrt{\bar{N}}} y^T \tilde{\psi}_j^{\mathcal{P},\langle n\rangle} = \frac{1}{\sqrt{\bar{N}}} y^T W^{\mathrm{m}} \tilde{\psi}_j^{\mathcal{P},\langle n\rangle} \tag{53}$$

where the latter equality is due to the constraint matrix A^{m} in (35). The closed-form solution to $\tilde{\psi}_j^{\mathcal{P},\langle n\rangle}$ in Definition 4 is

$$\tilde{\psi}_j^{\mathcal{P},\langle n\rangle} = (I - D)\psi_j^{\mathcal{P},\langle n\rangle} = W^{\mathrm{m}}(I - D)\psi_j^{\mathcal{P},\langle n\rangle} \tag{54}$$

where $D = A^{\dagger}[\,(A^{\mathrm{m}})^T \; 0 \; 0\,]^T$. Substituting (54) into (53) yields

$$z_j[n] = \frac{1}{\sqrt{\bar{N}}} v^T \psi_j^{\mathcal{P},\langle n\rangle} \tag{55}$$

where $v := (I - D^T)W^{\mathrm{m}} y = \mathcal{R}y$, in which $\mathcal{R}$ is given by (37). By Definition 1, transform $z_j[n]$ in (55) is expressed as

$$z_j[n] = \frac{1}{\sqrt{\bar{N}}} v^T \psi_j^{\mathcal{P},\langle n\rangle} = \left\{v \circledast \psi_j\right\}[n] = \mathcal{H}_n\left(\Psi_j \odot V\right)$$

where V is the DFT of v, satisfying $V = \mathcal{F}\mathcal{R}\mathcal{H}Y$. This completes the proof.

Proof of Theorem 3: Satisfaction of condition 1) in Theorem 1 is immediate by observing that $\mathcal{R}\mathcal{H}Y = \mathcal{R}y = \mathcal{R}y^{\mathrm{m}}$, due to the matrix W^{m} in (37).

Condition 2) is satisfied by assumption.

Next, condition 3) is proven. Observe that the constraints (35) impose that $\mathcal{F}_v\tilde{\psi}^{\mathcal{P},\langle n\rangle} = \mathcal{F}_v\psi^{\mathcal{P},\langle n\rangle} = \mathcal{F}_v\mathcal{R}^T\psi^{\mathcal{P},\langle n\rangle}$ where the latter equality is due to (54). Thus,

$$\left[\mathcal{F}\mathcal{R}^T\mathcal{H}\Psi^{\mathcal{P},\langle n\rangle}\right]_v = \left[\Psi^{\mathcal{P},\langle n\rangle}\right]_v. \tag{56}$$

Define the complementary indices $w \in [0, 1, \ldots, \bar{N} - 1]\backslash v$. Since (56) holds for any $\Psi^{\mathcal{P},\langle n\rangle}$, it is obtained that the submatrix $[\mathcal{F}\mathcal{R}^T\mathcal{H}]_{v,v} = I$ and the submatrix $[\mathcal{F}\mathcal{R}^T\mathcal{H}]_{v,w} = 0$. Since by selection of v in (39) matrix A^f is composed of conjugate pairs, it can be shown that $\mathcal{R}$ in (37) is real-valued. Thus, taking the Hermitian transpose of the submatrices above yields $[\mathcal{F}\mathcal{R}\mathcal{H}]_{v,v} = I$ and $[\mathcal{F}\mathcal{R}\mathcal{H}]_{w,v} = 0$. By virtue of property 2) in Lemma 1, $[\Psi]_v = 0$ when $\delta = \gamma$ and $p > 0$, and hence, $[\mathrm{diag}(\Psi)\mathcal{F}\mathcal{R}\mathcal{H}]_{:,v} = 0$. By observing that Y_Δ in Theorem 1 can attain nonzero values at the frequency lines v only, it is immediate that $\mathrm{diag}(\Psi)\mathcal{F}\mathcal{R}\mathcal{H}Y^{\mathrm{m}} = 0$, and hence, $M_{\tilde{q}_j}Y_\Delta = 0$ with $M_{\tilde{q}_j}$ in (38). This completes the proof.

Proof of Theorem 4: First, the first inequality in (41) is proven. Similar to the proof of Theorem 2, the parameter covariance for the TVE is expressed as

$$\mathrm{cov}\left(\hat{\bar{\Theta}}\right) = \frac{2}{\lambda} P_{\mathrm{TVE}}^{-1} \Re\left(\bar{K}^H \bar{B} \bar{B} \bar{K}\right) P_{\mathrm{TVE}}^{-1} \tag{57}$$

where $P_{\mathrm{TVE}} = (2/\lambda)\Re(\bar{K}^H\bar{B}\bar{K})$, and for the TVE, the matrix $\bar{B} = \bar{M}^H\bar{M} = (\mathcal{T}^{\mathcal{R}})^H \sum_{j=0}^{\tilde{N}_q-1} \mathrm{diag}(|\Psi_{\tilde{q}_j}|^2)\mathcal{T}^{\mathcal{R}}$. Hence, expression (42) is immediate.

By the assumption that $A^f = 0$, $A^t = 0$, matrix $\mathcal{R}$ in (37) reduces to $\mathcal{R} = W^{\mathrm{m}}$, and consequently, $\mathcal{T}^{\mathcal{R}} = \mathcal{T}^{\mathrm{m}}$ with $\mathcal{T}^{\mathrm{m}}$ in (6). Since $\bar{\sigma}(\mathcal{T}^{\mathrm{m}}) = 1$ and $\bar{\sigma}(\sum_{j=0}^{\tilde{N}_q-1} \mathrm{diag}(|\Psi_{\tilde{q}_j}|^2)) = 1$ by construction, it follows that $\bar{\sigma}(\bar{B}) \leq 1$. Then, along the lines of (50), it is obtained that $\mathrm{var}([\hat{\bar{\Theta}}]_m) \leq [P_{\mathrm{TVE}}^{-1}]_{m,m}$.

Second, the second inequality in (41) is proven. Define $M_{\tilde{q}_j}[i]$ as the ith row of the matrix $M_{\tilde{q}_j}$ in (38). Then, the matrix P_{TVE} in (42) can be decomposed as

$$P_{\mathrm{TVE}} = P_{\mathrm{TVE}}^{\mathcal{D}} + P_{\mathrm{TVE}}^{\mathcal{A}} \tag{58}$$

with

$$P_{\mathrm{TVE}}^{X} = \frac{2}{\lambda}\Re \sum_{i\in X} \sum_{j=0}^{\tilde{N}_q-1} \bar{K}^H M_{\tilde{q}_j}^H[i] M_{\tilde{q}_j}[i]\bar{K} \tag{59}$$

where X denotes either set $\mathcal{D}$ or $\mathcal{A}$, see (9) and (10).

Consider the transform $z[i] = M_{\tilde{q}_j}[i]Y$, where Y is the DFT of any $y \in \mathbb{R}^{\bar{N}}$, and consider $y^{\mathrm{m}} = W^{\mathrm{m}}y$ with DFT $Y^{\mathrm{m}} = \mathcal{T}^{\mathrm{m}}Y$. Using the definition of $M_{\tilde{q}_j}$ in (38) and $\mathcal{T}^{\mathcal{R}} = \mathcal{T}^{\mathrm{m}}$, this transform $z[i]$ is equivalently expressed as $z[i] = \mathcal{H}_i\mathrm{diag}(\Psi_{\tilde{k}_j})\mathcal{T}^{\mathrm{m}}Y = \mathcal{H}_i\mathrm{diag}(\Psi_{\tilde{k}_j})Y^{\mathrm{m}}$. By definition of $\mathcal{D}$, the transform $z[i]$ is unaffected by the missing samples for $i \in \mathcal{D}$, and thus, $z[i] = \mathcal{H}_i\mathrm{diag}(\Psi_{\tilde{k}_j})\mathcal{T}^{\mathrm{m}}Y = \mathcal{H}_i\mathrm{diag}(\Psi_{\tilde{k}_j})Y$ for $i \in \mathcal{D}$. This applies to any Y, and hence, $M_{\tilde{q}_j}[i] = \mathcal{H}_i\mathrm{diag}(\Psi_{\tilde{q}_j})\mathcal{T}^{\mathrm{m}} = \mathcal{H}_i\mathrm{diag}(\Psi_{\tilde{q}_j})$ for $i \in \mathcal{D}$. So, $P_{\mathrm{TVE}}^{\mathcal{D}}$ in (59) is equal to P_{TIE} in (30).

Finally, applying the Woodbury matrix identity to (58) gives

$$P_{\mathrm{TVE}}^{-1} = P_{\mathrm{TIE}}^{-1} - \left(P_{\mathrm{TIE}}\left(P_{\mathrm{TVE}}^{\mathcal{A}}\right)^{-1}P_{\mathrm{TIE}} + P_{\mathrm{TIE}}\right)^{-1}. \tag{60}$$

Then, the second inequality in (41) follows since the most right term is a symmetric PSD matrix, which completes the proof.

Proof of Corollary 2: Under the assumptions in Theorem 4, the matrix $\mathcal{R}$ in (37) reduces to $\mathcal{R} = W^{\mathrm{m}}$. Observing that $\Psi_{\tilde{q}_j}^{[p]} = 1_{\bar{N}}$ when $p = 0$, the matrix $M_{\tilde{q}_j}$ in (38) used in the TVE simplifies to $M_{\tilde{q}_j} = \mathcal{H}\mathcal{F}W^{\mathrm{m}}\mathcal{H} = W^{\mathrm{m}}\mathcal{H}$, which is identical to the matrix $M_{\tilde{q}_j}$ in (27) in the TIE for $p = 0$.

ACKNOWLEDGMENT

The authors would like to thank Enzo Evers for his contribution.

REFERENCES

[1] L. Ljung et al., *System Identification: Theory for the User*. Upper Saddle River, NJ, USA: Prentice-Hall, 1987.

[2] E. Geerardyn, M. L. D. Lumori, and J. Lataire, "FRF smoothing to improve initial estimates for transfer function identification," *IEEE Trans. Instrum. Meas.*, vol. 64, no. 10, pp. 2838–2847, Oct. 2015.

[3] G. Monteyne, G. Vandersteen, R. Pintelon, and D. Ugryumova, "Transient suppression in FRF measurement: Comparison and analysis of four state-of-the-art methods," *Measurement*, vol. 46, no. 7, pp. 2210–2222, Aug. 2013.

[4] A. Karimi and C. Kammer, "A data-driven approach to robust control of multivariable systems by convex optimization," *Automatica*, vol. 85, pp. 227–233, Nov. 2017.

[5] R. van der Maas, A. van der Maas, R. Voorhoeve, and T. Oomen, "Accurate FRF identification of LPV systems: ND-LPM with application to a medical X-ray system," *IEEE Trans. Control Syst. Technol.*, vol. 25, no. 5, pp. 1724–1735, Sep. 2017.

[6] R. Pintelon and J. Schoukens, *System Identification: A Frequency Domain Approach*. Hoboken, NJ, USA: Wiley, 2012.

[7] J. L. Douce and L. Balmer, "Transient effects in spectrum estimation," *IEE Proc. D Control Theory Appl.*, vol. 132, no. 1, pp. 25–29, Jan. 1985.

[8] J. Schoukens, Y. Rolain, and R. Pintelon, "Improved frequency response function measurements for random noise excitations," in *Proc. IEEE Instrum. Meas. Technol. Conf. Sens., Process., Netw. (IMTC)*, vol. 1, May 1997, pp. 749–753.

[9] J. Schoukens, G. Vandersteen, K. Barbe, and R. Pintelon, "Nonparametric preprocessing in system identification: A powerful tool," in *Proc. Eur. Control Conf. (ECC)*, Aug. 2009, pp. 1–14.

[10] T. McKelvey and G. Guérin, "Non-parametric frequency response estimation using a local rational model," *IFAC Proc. Volumes*, vol. 45, no. 16, pp. 49–54, 2012.

[11] E. Evers, B. de Jager, and T. Oomen, "Incorporating prior knowledge in local parametric modeling for frequency response measurements: Applied to thermal/mechanical systems," *IEEE Trans. Control Syst. Technol.*, vol. 30, no. 1, pp. 142–152, Jan. 2022.

[12] S. Kar and J. M. F. Moura, "Distributed consensus algorithms in sensor networks: Quantized data and random link failures," *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1383–1400, Mar. 2010.

[13] D. Ugryumova, R. Pintelon, and G. Vandersteen, "Frequency response function estimation in the presence of missing output data," *IEEE Trans. Instrum. Meas.*, vol. 64, no. 2, pp. 541–553, Feb. 2015.

[14] K. Barbé, W. Van Moer, L. Lauwers, and N. Björsell, "A simple nonparametric preprocessing technique to correct for nonstationary effects in measured data," *IEEE Trans. Instrum. Meas.*, vol. 61, no. 8, pp. 2085–2094, Aug. 2012.

[15] J. Schoukens, G. Vandersteen, Y. Rolain, and R. Pintelon, "Frequency response function measurements using concatenated subrecords with arbitrary length," *IEEE Trans. Instrum. Meas.*, vol. 61, no. 10, pp. 2682–2688, Oct. 2012.

[16] P. Stoica, J. Li, J. Ling, and Y. Cheng, "Missing data recovery via a nonparametric iterative adaptive approach," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, Apr. 2009, pp. 3369–3372.

[17] P. M. T. Broersen, S. de Waele, and R. Bos, "Autoregressive spectral analysis when observations are missing," *Automatica*, vol. 40, no. 9, pp. 1495–1504, Sep. 2004.

[18] D. Ugryumova, R. Pintelon, and G. Vandersteen, "Frequency response matrix estimation from partially missing data—For periodic inputs," *IEEE Trans. Instrum. Meas.*, vol. 64, no. 12, pp. 3615–3628, Dec. 2015.

[19] Z. Cvetkovic, "On discrete short-time Fourier analysis," *IEEE Trans. Signal Process.*, vol. 48, no. 9, pp. 2628–2640, Sep. 2000.

[20] I. Daubechies, *Ten Lectures Wavelets*. Philadelphia, PA, USA: SIAM, 1992.

[21] I. Daubechies, "The wavelet transform, time-frequency localization and signal analysis," *IEEE Trans. Inf. Theory*, vol. 36, no. 5, pp. 961–1005, Sep. 1990.

[22] M. Jansen, M. Malfait, and A. Bultheel, "Generalized cross validation for wavelet thresholding," *Signal Process.*, vol. 56, no. 1, pp. 33–44, Jan. 1997.

[23] X. Chen, S. Yang, J. Ma, and Z. He, "The construction of wavelet finite element and its application," *Finite Elements Anal. Des.*, vol. 40, nos. 5–6, pp. 541–554, Mar. 2004.

[24] P. Y. Simard and H. S. Malvar, "A wavelet coder for masked images," in *Proc. Data Compress. Conf.*, Mar. 2001, pp. 93–102.

[25] N. Dirkx, K. Tiels, and T. Oomen, "Frequency response function identification from incomplete data: A wavelet-based approach," *IFAC-PapersOnLine*, vol. 55, no. 37, pp. 439–444, 2022.

[26] K. R. Rao and P. C. Yip, *The Transform and Data Compression Handbook*. Boca Raton, FL, USA: CRC Press, 2018.

[27] R. Pintelon, J. Schoukens, G. Vandersteen, and K. Barbé, "Estimation of nonparametric noise and FRF models for multivariable systems—Part I: Theory," *Mech. Syst. Signal Process.*, vol. 24, no. 3, pp. 573–595, Apr. 2010.

[28] J. Zundert and T. Oomen, "Beyond equidistant sampling for performance and cost: A loop-shaping approach applied to a motion system," *Int. J. Robust Nonlinear Control*, vol. 29, no. 2, pp. 408–432, Jan. 2019.

[29] K. Gröchenig, *Foundations of Time-Frequency Analysis*. Cham, Switzerland: Springer, 2001.

[30] G. V. Milovanović and Z. Udovičić, "Calculation of coefficients of a cardinal B-spline," *Appl. Math. Lett.*, vol. 23, no. 11, pp. 1346–1350, Nov. 2010.

[31] H. Levi, "A geometric construction of the Dirichlet kernel," *Trans. New York Acad. Sci.*, vol. 36, no. 7, pp. 640–643, 1974.

[32] R. Pintelon, G. Vandersteen, J. Schoukens, and Y. Rolain, "Improved (non-)parametric identification of dynamic systems excited by periodic signals—The multivariate case," *Mech. Syst. Signal Process.*, vol. 25, no. 8, pp. 2892–2922, Nov. 2011.

[33] T. Söderström and P. Stoica, *System Identification*. Hemel Hempstead, U.K.: Prentice-Hall, 1989.

[34] R. Isermann and M. Münchhof, *Identification of Dynamic Systems: An Introduction With Applications*, vol. 85. Cham, Switzerland: Springer, 2011.

[35] E. Evers, N. van Tuijl, R. Lamers, B. de Jager, and T. Oomen, "Fast and accurate identification of thermal dynamics for precision motion control: Exploiting transient data and additional disturbance inputs," *Mechatronics*, vol. 70, Oct. 2020, Art. no. 102401.

[36] J. N. Franklin, *Matrix Theory*. Chelmsford, MA, USA: Courier Corporation, 2012.

Nic Dirkx received the M.Sc. degree from the Eindhoven University of Technology, Eindhoven, The Netherlands, in 2011, where he is currently pursuing the Ph.D. degree with the Control Systems Technology Group, Department of Mechanical Engineering.

Since 2012, he has been a Mechatronics Design Engineer at ASML, Veldhoven, The Netherlands. His research interests are in the fields of system identification, precision motion control, and mechatronic systems.

Koen Tiels (Member, IEEE) received the master's degree in electromechanical engineering and the Ph.D. degree from the Vrije Universiteit Brussel (VUB), Brussels, Belgium, in July 2010 and March 2015, respectively.

He was a Post-Doctoral Researcher with VUB from 2015 to 2018. From February 2018 to January 2020, he was with the Department of Information Technology, Division of Systems and Control, Uppsala University, Uppsala, Sweden, as a Post-Doctoral Researcher. From February 2020 to December 2022, he was an Assistant Professor with the Control Systems Technology (CST) Group, Eindhoven University of Technology (TU/e), Eindhoven, The Netherlands, where he is currently a Docent at the Department of Mechanical Engineering. His main research interests are in the field of nonlinear system identification.

Tom Oomen (Senior Member, IEEE) received the M.Sc. (cum laude) and Ph.D. degrees from the Eindhoven University of Technology, Eindhoven, The Netherlands, in 2005 and 2010, respectively.

He held visiting positions at KTH Royal Institute of Technology, Stockholm, Sweden, and The University of Newcastle, Callaghan, NSW, Australia. He is currently a Full Professor with the Department of Mechanical Engineering, Eindhoven University of Technology. He is also a part-time Full Professor with the Delft University of Technology, Delft, The Netherlands. His research interests are in the fields of data-driven modeling, learning, and control, with applications in precision mechatronics.

Dr. Oomen has been a member of the Eindhoven Young Academy of Engineering and the Vice-Chair of IFAC TC 4.2. He was a recipient of the 7th Grand Nagamori Award, the Corus Young Talent Graduation Award, the IFAC 2019 TC 4.2 Mechatronics Young Research Award, the 2015 IEEE TRANSACTIONS ON CONTROL SYSTEMS TECHNOLOGY Outstanding Paper Award, the 2017 IFAC Mechatronics Best Paper Award, the 2019 IEEJ Journal of Industry Applications Best Paper Award, and the Veni and Vidi personal grant. He is a Senior Editor of IEEE CONTROL SYSTEMS LETTERS (L-CSS) and an Associate Editor of *IFAC Mechatronics*. He has served on the Editorial Board of the IEEE CONTROL SYSTEMS LETTERS (L-CSS) and IEEE TRANSACTIONS ON CONTROL SYSTEMS TECHNOLOGY.