Real-time sinusoidal parameter estimation for damage growth monitoring during ultrasonic very high cycle fatigue tests

Shawn L. Kiser, Marc Rébillat, Mikhail Guskov, Nicolas Ranc

Real-time sinusoidal parameter estimation for damage growth monitoring during ultrasonic very high cycle fatigue tests

Shawn L. Kiser *, Marc Rébillat, Mikhail Guskov, Nicolas Ranc

Laboratoire PIMM, Arts et Métiers Sciences et Technologies, CNRS, CNAM, HESAM Université, Paris, France

ARTICLE INFO

Keywords:
Frequency and amplitude estimation
Ultrasonic fatigue tests
Very high cycle fatigue

ABSTRACT

Ultrasonic fatigue tests (UFT) are used to study the fatigue life behavior of metallic components undergoing a very high number of cycles (typically $\geq 10^7 - 10^9$ cycles) under relatively low mechanical loads. By soliciting fatigue specimens at 20 kHz, ultrasonic fatigue machines are indispensable for monitoring damage growth and fatigue failures in a reasonable amount of time. As fatigue damage accumulates in the specimen, the specimen's free-end exhibits a nonlinear dynamic response. The resulting quasi-stationary, harmonic signals have sinusoidal parameters (frequency and amplitude) which are slowly time-varying with respect to the excitation frequency. The discrete Fourier transform (DFT) is typically used to extract these evolving sinusoidal parameters from a window of finite data of the vibration signal. Alternative spectral estimation methods, specifically line spectra estimators (LSEs), exploit a priori information of the signal via their modeling basis and overcome limitations seen by the DFT. Many LSEs are known to have state-of-the-art results when benchmarked on purely stationary signals with unit amplitudes. However, their performances are unknown in the context of slowly time-varying signals typical of UFT, leading to a widespread use of the DFT. Thus, this paper benchmarks classical and modern LSEs against specific synthetic signals which arise in UFTs. Adequate algorithms are then recommended and made publicly available to process experimental data coming from ultrasonic fatigue tests depending on performance metrics and experimental restraints.

1. Introduction

Engineering design of metallic components according to their classical fatigue limit is no longer sufficient when loaded beyond 10^7 cycles. Mechanical failure due to fatigue takes place during a significant period of service: e.g., most load-bearing metallic components in the transportation industry are loaded in 10^8–10^9 cycles. Characterizing and modeling fatigue behavior in this very high cycle fatigue (VHCF) regime (10^7–10^9 cycles) has only been recently pursued due to experimental data provided by ultrasonic fatigue tests (UFT) [1]. To achieve 10^9 loading cycles, a conventional fatigue machine working at 100 Hz will finish in 115 days whereas an ultrasonic fatigue machine working at 20 kHz will finish in only 14 h. The principle of action of an ultrasonic fatigue machine is the excitation at a working frequency that corresponds to the fatigue specimen's first longitudinal mode. An example of a typical UFT experimental setup can be seen in Fig. 1. The standing wave formed at this frequency allows for periodic tension–compression loadings whose maximal stress occurs at the centroid of the fatigue specimen [2]. The local microstructural evolution at this centroid is of great interest for VHCF researchers interested in characterization with respect to the fatigue life of materials.

* Corresponding author.
E-mail address: shawn_lee.kiser@ensam.com (S.L. Kiser).

Acronyms

AWGN	Additive white Gaussian noise
CFH	Course-to-fine HAQSE
CNN	Convolutional neural network
CRB	Cramér-Rao lower bound
DFT	Discrete Fourier transform
ESPRIT	Estimation of signal parameters via rational invariance techniques
FSR	Frequency success rate
HAQSE	Hybrid Aboutanios and Mulgrew and q-shift estimator
LSE	Line spectral estimator
MAD	Matched amplitude distance
MFD	Matched frequency distance
MSAE	Mean-squared amplitude error
MSFE	Mean-squared frequency error
NOMP	Newtonized orthogonal matching pursuit
NRUS	Non-contact nonlinear resonance ultrasound spectroscopy
RELAX	Relaxation-based estimation method
ReLU	Rectified linear activation unit
SHG	Second harmonic generation
SHM	Structural health monitoring
SNR	Signal-to-noise ratio
UFT	Ultrasonic fatigue tests
VHCF	Very high cycle fatigue

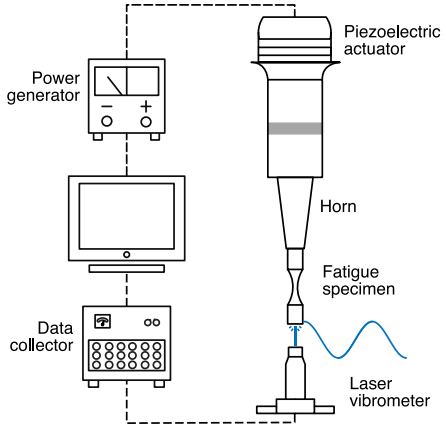

Fig. 1. Experimental setup of an ultrasonic fatigue test.

Multiple techniques currently exist for the assessment of damage and microstructural characteristics in materials during UFTs. They rely on measuring multiple sinusoidal parameters, specifically amplitudes and/or frequencies [3–11]. During UFTs, fatigue specimens are excited at their first longitudinal eigenfrequency, nonlinearities can be observed in the vibration's frequency spectra [12]. The evolution of these parameters can drive a nonlinear model and consequently infer the microstructural changes and damage within the fatigue specimen. These techniques would benefit from in-situ, or real-time sinusoidal parameter estimation since it could correspond to a real-time assessment of microstructural evolution leading to macroscopic damage. Acoustic techniques using ultrasound for damage and microstructural characterization also utilizes amplitude estimation: the second harmonic generation (SHG) technique [13], and non-contact nonlinear resonance ultrasound spectroscopy (NRUS) also utilizes frequency estimation [14]. These techniques are also subject to this study since they both rely on estimates of sinusoidal parameters of harmonics from transmitted tone bursts. More generally, sinusoidal parameters have also been used in fault detection and structural health monitoring (SHM) for signal-based methods [15] and as features for data-driven methods [16].

UFT signals can be observed to have harmonics at integer multiples of the working frequency, corresponding to the fatigue specimen's longitudinal eigenfrequency, see Fig. 2. However, throughout the UFT, damage and other nonlinear phenomena can

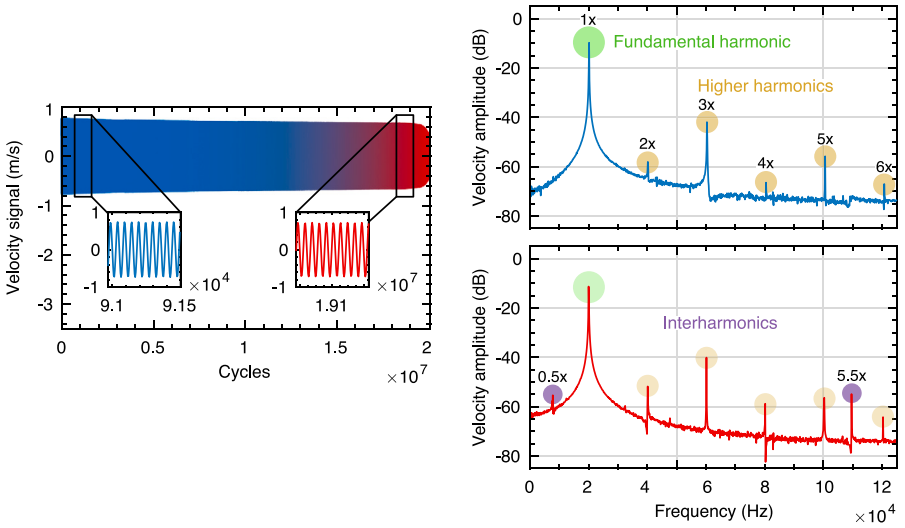

Fig. 2. (Left) Velocity signal from a typical ultrasonic fatigue test with segmented signal windows near the beginning and end of the fatigue test. Frequency spectra corresponding to an undamaged (right-top) and damaged (right-bottom) fatigue specimen are obtained through a DFT. Harmonics correspond to integer multiples of the fundamental harmonic while interharmonics correspond to non-integer multiples of the fundamental harmonic.

introduce interharmonics which exceeds the scope of algorithms that only estimate fundamental frequency [17]. Additionally, UFTs that require a large number of cycles have signals that are nonstationary, but can be adequately modeled as being a locally stationary signal within a finite segment of length N of the data, or window. We refer to the windowed signal as quasi-stationary, which we define as a trend-stationary sequence by the Kwiatkowski–Phillips–Schmidt–Shin test [18]. One main concern is the method of estimating such sinusoidal parameters for a quasi-stationary signal: a choice must be made between choosing algorithms whose basis lies either in spectral analysis or time–frequency analysis.

To date, a sliding window approach has only been coupled with the discrete Fourier transform (DFT) in UFT damage monitoring. When applied per window, the DFT yields a respective frequency spectrum, and a peak picking algorithm extracts sinusoidal parameters, i.e. the frequencies and amplitudes. This is one of the most rudimentary forms of a line spectral estimator (LSE): where sinusoidal parameters correspond to Dirac deltas (lines) in the frequency spectra. However, a finite period (time length) of the data corresponds to the sinc function convolving with the DFT resulting in a loss of frequency resolution, see Fig. 3. This directly poses problems for an UFT signal of a damaged fatigue specimen, where introduced interharmonics can influence the estimation of harmonics. Additional factors can skew the DFT's spectra affecting the accuracy of the peak picking: nonstationary sinusoidal components smearing the spectral envelope, non-periodicity of the data manifesting as spectral leakage, aliasing due to violation of the Nyquist–Shannon sampling theorem, and noise introducing uncertainty into the measured signal. Due to the difficulties seen by the DFT, alternative spectral estimators should be considered, such as nonparametric and parametric spectral estimators.

Nonparametric algorithms estimate the entire frequency spectra which include the DFT-based methods [20,21] and filter-bank methods [22,23]. Filter-bank methods' frequency resolution depends on the spectra length and can achieve super-resolution [24], i.e., resolve closely-spaced frequencies beyond the DFT in Fig. 3. Recently, nonparametric algorithms have been coupled with a peak picking algorithm to become LSEs, but require a method to estimate the number of sinusoids [25] or utilize an amplitude/power threshold [26]. On the other hand, parametric algorithms only estimate parameters for a known number of sinusoids and some are also capable of super-resolution. These parametric estimators are formulated as sinusoidal frequency estimation problems since the estimation of amplitudes and phases become least squares solvable when the frequencies are known [27]. Many of the parametric LSE established in the 1980s and beyond are inspired by Prony's method [28] of converting a non-linear approximation by solving a set of linear equations and a root-finding problem. One of the most popular parametric LSEs are the subspace methods, which include MUSIC [29] and ESPRIT [30]. Subspace methods decompose the finite data into signal and noise subspaces via an eigenvalue decomposition (EVD) and/or singular-value decomposition (SVD). Exploiting the low-rank structure of the signal's covariance matrix allows for these methods to achieve super-resolution. A sliding-window approach [31] is commonly used with a spectral estimator to circumvent the problem of quasi-stationarity, where the signal is uniformly discretized in time. The sinusoidal parameters in each window are obtained by applying a spectral estimator, and the window is shifted in time and the process is repeated.

Other dominant parametric LSEs are derived from the maximum likelihood principle of estimators [32], where estimated sinusoidal parameters are the most likely to explain the finite data. Two classes exist: deterministic maximum likelihood [33] and stochastic maximum likelihood [34] which have criteria derived in the case the sinusoidal parameters are unknown deterministic and stochastic, respectively. Deterministic maximum likelihood has been most often referred to as the nonlinear least squares method in literature [20] since it estimates parameters by minimizing ℓ_2-norm of the difference between the finite data and signal model.

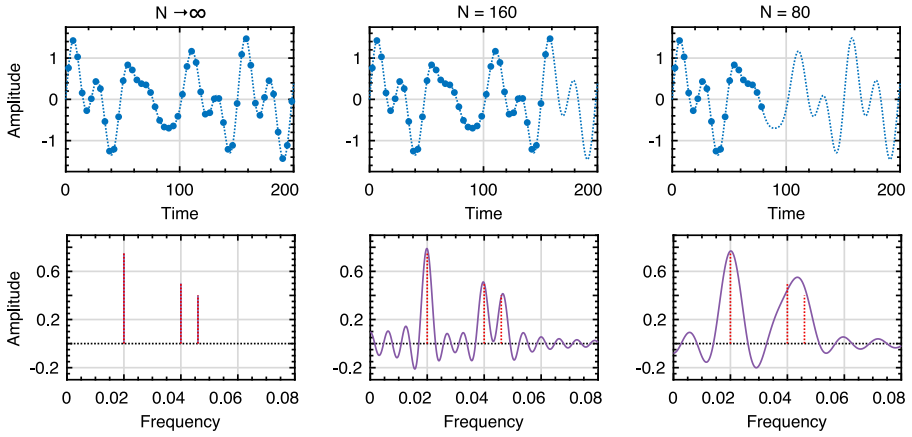

Fig. 3. Demonstration of the frequency resolution of a signal with a fixed sampling frequency, with respect to the Rayleigh limit $1/N$ seen by the DFT, adapted from [19]. A multi-sinusoidal signal (dashed blue line) and its Nyquist–Shannon samples (blue circles) are depicted on the top row. The bottom row shows the DFT's frequency resolution (purple line) and the true line spectra (dashed red line). For a finite N, the DFT's frequency resolution is equal to the convolution between the DFT of the signal and the sinc function.

Some deterministic maximum likelihood methods separate the full multidimensional minimization into iterative searches in lower-dimensional parameter subspaces: these include Expectation–Maximization [35] and RELAX [36]. Stochastic maximum likelihood methods, until most recently, are generally only found in the direction of arrival literature, e.g., [37].

New LSEs have been introduced due to innovations in data science domains, specifically those relating to compressed sensing, sparse regularization, and deep learning. LSEs that use compressed sensing for sparse regularization [38] are sometimes referenced as semi-parametric since they sometimes rely on model order or other signal parameters a priori, such as its noise covariance. One issue arises when resolving the compressed sensing problem onto LSE, since it normally assumes a discrete basis for signal recovery, whereas a sparse signal is continuous in frequency space. Thus grid-based approaches divide the frequency spectra into finite discrete grid points and build bases from these grid points. Popular solutions to this approach can be through convex optimization algorithms like basis pursuit [39] or LASSO [40]. However, the true frequency parameters are not guaranteed to lie within the frequency grid, a problem commonly referred to as basis or grid mismatch [41]. Attempts to remediate grid mismatch includes iterative grid refinement [42], using a dense grid and iteratively optimize sparse solutions [43], generalizing the ℓ_1-norm as continuous in frequency space via an Atomic Norm minimization problem [44,45], or Bayesian approaches which iteratively refine the grid [46,47]. Artificial neural networks for LSE have been used for LSEs in the early 1990s [48], whereas recent deep neural network approaches have provided state-of-the-art performance by adopting convolutional neural network (CNN) [19,49].

The current research is motivated by the tradeoffs between finite window lengths and accuracy when dealing with quasi-stationarity and the Rayleigh limit. Despite the loss of information which affects estimation performance, a small time window is sought for many reasons:

- Smaller window lengths along the entire fatigue vibration signal will be more locally stationary compared to larger windows.
- Parameter estimation algorithms are quicker for smaller data lengths, meaning that a computationally complex algorithm can still be used in real-time.
- A smaller window length about the entire fatigue vibration signal will yield more discrete parameters, which is attractive for data-driven models. While using window overlap [31] can also increase the number of discrete parameter estimates, this would carry over for all spectral estimators. Within the context of ultrasonic fatigue experiments, more discrete parameters would allow for conditional-based rules for stopping a fatigue test before crack failure, and for deterministic models, a prediction of damage evolution.

Thus the merits of potential algorithms have to weigh between window length and parameter accuracy, noise robustness, and computational performance.

Our contribution lies in the study of sinusoidal parameter estimation accuracy on quasi-stationary signals, with a focus on signals found in VHCF vibration. Specifically, we establish two compounding influences that are at play when using a windowed approach to the real-time estimation of sinusoidal parameters: (1) estimators face a frequency and time uncertainty dictated by the Rayleigh limit, and (2) the usage of estimators who assume a stationary signal model being used on a quasi-stationary signal. Minimizing the window length negatively influences the frequency resolution, as mentioned in (1), but increases time resolution and has positive contributions stated in the bullets previously. Additionally, the statistical performance and accuracy of estimators are numerous in literature but generally utilize synthetic signals that are purely stationary, whose results cannot be said for quasi-stationary signals, mentioned in (2). Therefore, we choose to investigate five LSEs, specifically Unitary ESPRIT [50], RELAX [36], CFH [25], NOMP [42], and DeepFreq [19] with respect to the problems aforementioned. These five algorithms were chosen since their theoretical basis are a diverse (subspace-based, maximum likelihood-based, DFT peak interpolation-based, greedy algorithm for

Table 1
Summary of LSE algorithms in implemented.

Algorithms	Main principle	Advantages	Disadvantages	Tunable parameters
Unitary ESPRIT[a] [50]	Rotational invariance of signal's subspaces	Forward-backwards averaging of signal's covariance, faster than ESPRIT	Computation complexity dictated by SVD	Sub-vector length
RELAX [36]	Frequency domain (zero-padding) interpolation, iterative refinement	Conceptually easy to implement	Asymptotic gains in performance tuning large zero-padding	Zero-padding length
CFH [25]	Frequency domain (peak) interpolation, iterative refinement	One of the fastest parametric estimators	Algorithm is not tunable	–
NOMP [42]	Greedy algorithm, iterative refinement	Newton refinements alleviates restriction to DFT basis	Unknown performance in the case of basis mismatch	Number of refinements, zero-padding length[b]
DeepFreq [19]	Deep convolution neural network	Fully-automatic frequency estimates[c]	Expensive offline training	Training data, CNN architecture

[a]The line spectrum can be computed after the estimation of frequencies, using the LS estimate Eq. (21).

[b]In this paper, the zero-padding can be removed to further differentiate NOMP from RELAX.

[c]DeepFreq, as published, features a frequency representation module, which is used in this paper, but also a component counting module which estimates the model order, making the algorithm fully-automatic compared to other LSEs.

grid refinement-based, and deep neural network-based, respectively) representation of the many LSE that exist in the literature, and most are considered to be state-of-the-art LSE for the purely stationary case. We choose not to use fast versions of these algorithms (for example ESPRIT with partial SVDs [51]) since they trade computational complexity for accuracy; this study assumes the original algorithms provide their best asymptotic performance. It should be noted that except for the former two algorithms listed, the original authors provide open access to their algorithms. The remaining part of the paper is organized as follows. Section 2 outlines the theory of each implemented LSE algorithm. The methodology is described in Section 3. Results and discussions are presented in Section 4 and the conclusions are drawn in Section 5.

2. Implemented algorithms

In this section, the mathematical formulation of the implemented LSE algorithms is described. A practical summary of the algorithms used in this work exists in Table 1. The estimation of frequencies and amplitudes for P complex sinusoids of $x(n)$ can be written as the signal model:

$$x(n) = \sum_{i=1}^{P} \beta_i e^{\mathrm{j}2\pi\omega_i n} + \varepsilon(n) \tag{1}$$

where $n \in \mathbb{Z}$ is a discrete index and $\varepsilon(n)$ represents additive white Gaussian noise (AWGN) with variance σ. The sinusoidal parameter β_i is the complex amplitude such that its modulus $|\beta_i|$ is the amplitude, and $\{\omega_i \in [0, 1); \omega_i \neq \omega_j : \forall i \neq j\}$ is the normalized frequency (units of cycles per sample) whose set is distinct. It is assumed that a uniformly sampled noisy signal has been segmented into windows. Each local window $\boldsymbol{x}$ with sample length N can be written in matrix–vector notation:

$$\boldsymbol{x}(n) = \boldsymbol{A}\boldsymbol{\beta}(n) + \boldsymbol{\varepsilon}(n) \tag{2}$$

where $\boldsymbol{x}(n), \boldsymbol{\varepsilon}(n) \in \mathbb{C}^N$:

$$\boldsymbol{x}(n) = \begin{bmatrix} x(n) & \cdots & x(n+N-1) \end{bmatrix}^{\mathrm{T}}$$
$$\boldsymbol{\varepsilon}(n) = \begin{bmatrix} \varepsilon(n) & \cdots & \varepsilon(n+N-1) \end{bmatrix}^{\mathrm{T}}$$

the complex amplitudes for P complex sinusoids $\boldsymbol{\beta}(n) \in \mathbb{C}^P$:

$$\boldsymbol{\beta}(n) = \begin{bmatrix} \beta_1 e^{\mathrm{j}2\pi\omega_1 n} & \cdots & \beta_P e^{\mathrm{j}2\pi\omega_P n} \end{bmatrix}^T$$

and the matrix $\boldsymbol{A} \in \mathbb{C}^{N\times P}$ is a Vandermonde matrix of P complex sinusoids:

$$\begin{aligned} \boldsymbol{A} &= \begin{bmatrix} 1 & \cdots & 1 \\ e^{\mathrm{j}(2\pi)\omega_1} & \cdots & e^{\mathrm{j}(2\pi)\omega_P} \\ e^{\mathrm{j}(2)(2\pi)\omega_1} & \cdots & e^{\mathrm{j}(2)(2\pi)\omega_P} \\ \vdots & \ddots & \vdots \\ e^{\mathrm{j}(N-1)(2\pi)\omega_1} & \cdots & e^{\mathrm{j}(N-1)(2\pi)\omega_P} \end{bmatrix} \\ &= \begin{bmatrix} \boldsymbol{a}(\omega_1) & \cdots & \boldsymbol{a}(\omega_P) \end{bmatrix} \end{aligned} \tag{3}$$

Lastly, it should be reminded real-valued signals can be represented using the complex notation in Eq. (1) through two methods: the use of the (downsampled) analytic signal [52], provided that there is no spectral content in the real-valued signal near zero and Nyquist frequencies [53]; or simply realizing real-valued sinusoidal signals can be represented as complex-valued and applying the algorithms. The latter case is preferred in this work because of its simplicity and to prevent issues seen in computing the analytic signal.

2.1. Unitary ESPRIT

The Unitary ESPRIT algorithm is used since it is one of the most economical and accurate among all ESPRIT implementations [50]. First, Eq. (1) is converted into the necessary matrix notation with overlapping segments such that:

$$\boldsymbol{X}(n) = \begin{bmatrix} x(n) & x(n+1) & \cdots & x(n+M) \\ x(n+1) & x(n+2) & \cdots & x(n+M+1) \\ \vdots & \vdots & \ddots & \vdots \\ x(n+L-1) & x(n+L) & \cdots & x(n+N) \end{bmatrix} \tag{4}$$

where $\boldsymbol{X}(n) \in \mathbb{C}^{M \times L}$ is a Hankel matrix, M is a chosen sub-vector size such that $P < M < N$, and $L = N - M$ is the remaining length. Unitary ESPRIT, as opposed to the standard ESPRIT algorithm, inherently includes forward–backward smoothing [30]. A forward–backward observation matrix $\boldsymbol{Y}'$ is estimated and converted to real-values by the unitary transformation of $\boldsymbol{X}$:

$$\boldsymbol{Y}' = \begin{bmatrix} \Re\left\{\mathbf{Q}_M^{\mathrm{H}} \boldsymbol{X}\right\} & \Im\left\{\mathbf{Q}_M^{\mathrm{H}} \boldsymbol{X}\right\} \end{bmatrix} \in \mathbb{R}^{M \times 2L} \tag{5}$$

where $\square'$ indicates a unitary transformed component, $\square_M$ indicates the square matrix dimension of size $M \times M$, and $\square^{\mathrm{H}}$ indicates the Hermitian (complex conjugate) transpose. This processing steps has been shown to improve parametric and nonparametric estimators that rely on covariance estimates [54]. The even- and odd-dimensioned unitary matrices are given, for an arbitrary D:

$$\begin{aligned} \mathbf{Q}_{2D} &= \frac{1}{\sqrt{2}} \begin{bmatrix} \mathbf{I}_D & \mathrm{j}\mathbf{I}_D \\ \boldsymbol{\Pi}_D & -\mathrm{j}\boldsymbol{\Pi}_D \end{bmatrix} \\ \mathbf{Q}_{2D+1} &= \frac{1}{\sqrt{2}} \begin{bmatrix} \mathbf{I}_D & 0 & \mathrm{j}\mathbf{I}_D \\ 0 & \sqrt{2} & 0 \\ \boldsymbol{\Pi}_D & 0 & -\mathrm{j}\boldsymbol{\Pi}_D \end{bmatrix} \end{aligned} \tag{6}$$

where $\mathbf{I}_D$ and $\boldsymbol{\Pi}_D$ are the identity matrix and the antidiagonal identity matrix (ones along the antidiagonal) of size $D \times D$, respectively. To obtain the signal subspace $\boldsymbol{U}'_{\mathrm{s}}$, a SVD is performed on $\boldsymbol{Y}'$ to yield:

$$\boldsymbol{Y}' = \begin{bmatrix} \boldsymbol{U}'_{\mathrm{s}} & \boldsymbol{U}'_{\mathrm{n}} \end{bmatrix} \begin{bmatrix} \boldsymbol{\Sigma}'_{\mathrm{s}} & 0 \\ 0 & \boldsymbol{\Sigma}'_{\mathrm{n}} \end{bmatrix} \begin{bmatrix} \boldsymbol{V}'^{\mathrm{H}}_{\mathrm{s}} \\ \boldsymbol{V}'^{\mathrm{H}}_{\mathrm{n}} \end{bmatrix} = \boldsymbol{U}' \boldsymbol{\Sigma}' \boldsymbol{V}'^{\mathrm{H}} \tag{7}$$

where $\boldsymbol{\Sigma}'$ corresponds to a diagonal matrix that contains the singular values on the main diagonal in descending order, and $\boldsymbol{V}'$ is an identity matrix orthogonal to $\boldsymbol{U}'$. The signal subspace is extracted via a priori knowledge of the left P column singular vectors:

$$\boldsymbol{U}'_{\mathrm{s}} = \begin{bmatrix} \boldsymbol{u}'(1) & \cdots & \boldsymbol{u}'(P) \end{bmatrix} \tag{8}$$

To exploit the signal subspace, selection matrices are introduced corresponding to a time shift of one sample value:

$$\begin{aligned} \boldsymbol{J}_1 &= \begin{bmatrix} \mathbf{I}_N \mid \mathbf{0}_N \end{bmatrix} \\ \boldsymbol{J}_2 &= \begin{bmatrix} \mathbf{0}_N \mid \mathbf{I}_N \end{bmatrix} \end{aligned} \tag{9}$$

where $\mathbf{0}_N$ is a zero matrix of size $N \times N$. Since the signal model of Eq. (1) has distinct frequencies, the rotational invariance of Eq. (9) allows one to form:

$$\boldsymbol{J}_1 \boldsymbol{A} \boldsymbol{\Phi} = \boldsymbol{J}_2 \boldsymbol{A} \tag{10}$$

where $\boldsymbol{\Phi} = \mathrm{diag}\left(e^{\mathrm{j}2\pi\omega_1}, \ldots, e^{\mathrm{j}2\pi\omega_P}\right)$ is the diagonal matrix of signal poles. The steering matrix $\boldsymbol{A}$ spans the P-dimensional signal subspace $\boldsymbol{U}_{\mathrm{s}}$, such that a transformation matrix $\boldsymbol{T}$ has the property:

$$\boldsymbol{A} = \boldsymbol{U}_{\mathrm{s}} \boldsymbol{T} \tag{11}$$

Eq. (10) can thus be expressed as a function of the eigenvectors of the signal subspace:

$$\boldsymbol{J}_1 \boldsymbol{U}_{\mathrm{s}} \boldsymbol{T} \boldsymbol{\Phi} = \boldsymbol{J}_2 \boldsymbol{U}_{\mathrm{s}} \boldsymbol{T} \tag{12}$$

which is equivalent to:

$$\boldsymbol{J}_1 \boldsymbol{U}_{\mathrm{s}} \boldsymbol{\Psi} = \boldsymbol{J}_2 \boldsymbol{U}_{\mathrm{s}} \quad \text{where} \quad \boldsymbol{\Psi} = \boldsymbol{T} \boldsymbol{\Phi} \boldsymbol{T}^{-1} \tag{13}$$

Eq. (13) is also equivalent to its real-valued counterpart through unitary transformation:

$$\boldsymbol{J}'_1 \boldsymbol{U}'_{\mathrm{s}} \boldsymbol{\Psi}' = \boldsymbol{J}'_2 \boldsymbol{U}'_{\mathrm{s}} \quad \text{where} \quad \boldsymbol{\Psi}' = \boldsymbol{T}' \boldsymbol{\Phi}' \boldsymbol{T}'^{-1} \tag{14}$$

with the respective unitary transformed selection matrices:

$$\begin{aligned} \boldsymbol{J}'_1 &= \Re\left\{\mathbf{Q}_{M-1}^{\mathrm{H}} \boldsymbol{J}_2 \mathbf{Q}_M\right\} \\ \boldsymbol{J}'_2 &= \Im\left\{\mathbf{Q}_{M-1}^{\mathrm{H}} \boldsymbol{J}_2 \mathbf{Q}_M\right\} \end{aligned} \tag{15}$$

This representation results in an overdetermined system of equations for calculating the eigenvalues which correspond to the diagonal elements of the matrix $\boldsymbol{\Phi}$, i.e., the signal poles with the frequency parameters ω_i. The LS approach is taken over the total LS approach [50] to reduce computational complexity with negligible accuracy loss for uniformly sampled signals:

$$\boldsymbol{\Psi}' \approx \left(\boldsymbol{J}_1' \boldsymbol{U}_s'\right)^{-1} \boldsymbol{J}_2' \boldsymbol{U}_s' \tag{16}$$

The right hand of Eq. (14) is the EVD of $\boldsymbol{\Psi}'$ since $\boldsymbol{\Phi}'$ is diagonal, such that:

$$e^{j\omega_i} = \lambda_i(\boldsymbol{\Psi}') \quad \text{for} \quad i = 1, \ldots, P \tag{17}$$

where $\lambda_i(\boldsymbol{\Psi}')$ is the eigenvalue of the ith entry of $\boldsymbol{\Psi}'$. When the signal is consistent with the model of Eq. (1), then the signal poles will lie close to the unit circle. Finally, the unwrapped frequency estimates can be obtained by:

$$\widehat{\omega}_i = \tan^{-1}\left(\lambda_i\right)/\pi \quad \text{for} \quad i = 1, \ldots, P \tag{18}$$

where $\widehat{\square}$ indicates a consistent estimate. The amplitude estimate vector $\widehat{\boldsymbol{\beta}}$ is then LS solvable using the frequency estimate vector $\widehat{\boldsymbol{\omega}}$ through Eq. (21).

2.2. RELAX

The RELAX algorithm attempts to estimate the maximum likelihood estimate by decoupling the nonlinear least squares problem into iterative one-dimensional minimizations. This idea will be introduced by a description of the nonlinear least squares problem for frequency estimation:

$$\{\widehat{\boldsymbol{\beta}}, \widehat{\boldsymbol{\omega}}\} = \underset{\{\boldsymbol{\beta},\boldsymbol{\omega}\}}{\arg\min}\, C_1(\boldsymbol{\beta}, \boldsymbol{\omega})$$

with:

$$C_1(\boldsymbol{\beta}, \boldsymbol{\omega}) = \|\boldsymbol{x} - \boldsymbol{\beta}\boldsymbol{A}\|_2^2 \tag{19}$$

where $\|\square\|_2$ is the ℓ_2-norm. Eq. (19) minimized with respect to $\boldsymbol{\omega}$ yields the estimate:

$$\widehat{\boldsymbol{\omega}} = \underset{\boldsymbol{\omega}}{\arg\max}\left[\boldsymbol{x}^{\mathrm{H}}\boldsymbol{A}\left(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}\right)^{-1}\boldsymbol{A}^{\mathrm{H}}\boldsymbol{x}\right] \tag{20}$$

and with respect to $\boldsymbol{\beta}$, the least squares error can be minimized through the pseudoinverse:

$$\widehat{\boldsymbol{\beta}} = \left(\boldsymbol{A}^{H}\boldsymbol{A}\right)^{-1}\boldsymbol{A}^{\mathrm{H}}\boldsymbol{x}\Big|_{\boldsymbol{\omega}=\widehat{\boldsymbol{\omega}}} \tag{21}$$

Eq. (20) is a multimodal function and its maximization corresponds to searching for its sharp global maxima. The RELAX algorithm is based on the relaxation of Eq. (19) to minimize the squared error between the observation and the parameterized signal model. A minimization of Eq. (19) with respect to both $\boldsymbol{\omega}$ and $\boldsymbol{\beta}$ through a cyclic minimization approach is detailed [36]. The subsequent minimization can be solved to yield the estimate of $\{\widehat{\beta}_j, \widehat{\omega}_j\}$:

$$\{\widehat{\beta}_j, \widehat{\omega}_j\} = \underset{\{\beta_j,\omega_j\}}{\arg\min}\, C_2(\beta_j, \omega_j)$$

with:

$$C_2(\beta_j, \omega_j) = \left\|\widehat{\boldsymbol{x}}_j - \beta_j \boldsymbol{a}(\omega_j)\right\|_2^2 \tag{22}$$

where $\widehat{\boldsymbol{x}}_j$ is defined in Eq. (25). Eq. (22) minimized similarly to before with respect to ω_j will yield:

$$\begin{aligned}\widehat{\omega}_j &= \underset{\omega_j}{\arg\min}\left\|\left[\mathbf{I} - \frac{\boldsymbol{a}(\omega_j)\boldsymbol{a}^{\mathrm{H}}(\omega_j)}{N}\right]\widehat{\boldsymbol{x}}_j\right\|_2^2 \\ &= \underset{\omega_j}{\arg\max}\left|\frac{\boldsymbol{a}^{\mathrm{H}}(\omega_j)\widehat{\boldsymbol{x}}_j}{N}\right|^2\end{aligned} \tag{23}$$

and with respect to β_j:

$$\widehat{\beta}_j = \left.\frac{\boldsymbol{a}^{\mathrm{H}}(\omega_j)\widehat{\boldsymbol{x}}_j}{N}\right|_{\omega_j=\widehat{\omega}_j} \tag{24}$$

Note Eq. (23) is the definition of the periodogram $\left|\boldsymbol{a}^{\mathrm{H}}(\omega_j)\boldsymbol{x}\right|^2/N$ whose estimate $\widehat{\beta}_j$ corresponds to the maximum peak. This corresponds to a maximum likelihood estimate [55] when the signal only contains a single sinusoid. Eqs. (23) and (24) are interpolated through using a zero-padding [36], appending zeros to the signal:

$$x(n) = \begin{cases} x(n), & n \le N \\ 0, & N < n \le Z_{\mathrm{p}} \end{cases}$$

where Z_{p} is the new signal length such that $\boldsymbol{x}(n) \in \mathbb{C}^{Z_{\mathrm{p}}}$.

The RELAX algorithm starts from the strongest signal component to determine $\{\widehat{\beta}_1, \widehat{\omega}_1\}$ using Eqs. (23) and (24). Then the successive separation of the signal is performed:

$$\widehat{\boldsymbol{x}}_j = \boldsymbol{x} - \sum_{i=1, i\neq j}^{j} \widehat{\beta}_i \boldsymbol{a}(\widehat{\omega}_i) \quad \text{for} \quad j = 1, \ldots, P \tag{25}$$

where Eq. (25) is used to obtain $\{\widehat{\beta}_j, \widehat{\omega}_j\}$ from Eqs. (23) and (24), and $\{\widehat{\beta}_i, \widehat{\omega}_i\}_{i=1}^{j-1}$ are re-estimated in a sub-iteration:

$$\widehat{\boldsymbol{x}}_k = \boldsymbol{x} - \sum_{i=1, i\neq j}^{j} \widehat{\beta}_i \boldsymbol{a}(\widehat{\omega}_i) \quad \text{for} \quad k = 1, \ldots, j-1 \tag{26}$$

until $C_2 < \epsilon$, where ϵ is a tolerance. This cyclic iterative procedure helps remove bias made from initial estimates done before successive separation of the signal.

2.3. CFH

The coarse-to-fine HAQSE (CFH) algorithm also exploits the idea of the maximum likelihood estimator. Many estimators exist which attempt to refine a frequency estimate via interpolation of the DFT grid via zero-padding [21], parabolic fitting [56], and iteratively [57]. However, most have not been extended to the multiple sinusoidal case and thus require a posteriori knowledge of frequencies to interpolate via peak picking. CFH utilizes the HAQSE interpolator [58] which shifts the DFT coefficients by $\pm q$ with:

$$|q| \leq N^{-1/3} \tag{27}$$

where it was shown that it obtains better accuracy compared to parabolic interpolators. CFH begins by finding the dominant signal component:

$$\tilde{\omega}_j = \arg\max_{\omega_j} \left| \frac{\boldsymbol{a}^{\mathrm{H}}(\omega_j)\widehat{\boldsymbol{x}}_j}{N} \right|^2 \tag{28}$$

where $\tilde{\square}$ is a coarse estimate. [57] is used to iterate to the fractional residual components:

$$\delta_1 = \frac{N}{2\pi} \sin^{-1}\left(\sin\left(\frac{\pi}{N}\right) \Re\left\{ \frac{S_{0.5} + S_{-0.5}}{S_{0.5} - S_{-0.5}} \right\} \right) \tag{29}$$

and refined residual component:

$$\delta_2 = \frac{1}{c(q)} \Re\left\{ \frac{S_{+q} - S_{-q}}{S_{+q} + S_{-q}} \right\} + \delta_1 \tag{30}$$

where S are the interpolation functions:

$$\begin{aligned} S_{\pm 0.5} &= \boldsymbol{a}^{\mathrm{H}}(\tilde{\omega}_j \pm 0.5)\widehat{\boldsymbol{x}}_j \\ S_{\pm q} &= \boldsymbol{a}^{\mathrm{H}}(\tilde{\omega}_j + \delta_1 \pm q)\widehat{\boldsymbol{x}}_j \end{aligned} \tag{31}$$

and c is the bias correction:

$$c(q) = \frac{1 - \pi q \cot(\pi q)}{q \cos^2(\pi q)} \tag{32}$$

The coarse frequency estimate is composed of an integer and a residual component, which can be expressed as:

$$\widehat{\omega}_j = \frac{\tilde{\omega}_j + \delta_2}{F_s} N \tag{33}$$

where F_s is the sampling frequency, and the respective complex amplitude:

$$\widehat{\beta}_j = \left. \frac{\boldsymbol{a}^{\mathrm{H}}\left(\omega_j\right)\widehat{\boldsymbol{x}}_j}{N} \right|_{\omega_j = \widehat{\omega}_j} \tag{34}$$

It then uses a similar procedure of RELAX, obtaining $\{\widehat{\beta}_j, \widehat{\omega}_j\}$ from Eqs. (33) and (34), and successfully separates the signal:

$$\widehat{\boldsymbol{x}}_j = \boldsymbol{x} - \sum_{i=1, i\neq j}^{j} \widehat{\beta}_i \boldsymbol{a}(\widehat{\omega}_i) \quad \text{for} \quad j = 1, \ldots, P \tag{35}$$

where Eq. (35) is used to obtain $\{\widehat{\beta}_j, \widehat{\omega}_j\}$ by repeating Eqs. (28)–(34) until the P components are coarsely interpolated. The coarsely estimated signal components are all subtracted from the signal:

$$\widehat{\boldsymbol{x}}_k = \boldsymbol{x} - \sum_{i=1, i\neq k}^{P} \widehat{\beta}_i \boldsymbol{a}(\widehat{\omega}_i) \quad \text{for} \quad \text{for} \quad k = 1, \ldots, P \tag{36}$$

where Eq. (36) is used to obtain $\{\widehat{\beta}_j, \widehat{\omega}_j\}$ by repeating Eqs. (28)–(34) until the P components are finely interpolated. The fine step corrects for selection bias due to estimation from the strongest to the weakest signal component in the coarse step of the algorithm.

2.4. NOMP

Newtonized orthogonal matching pursuit (NOMP) algorithm is built upon orthogonal matching pursuit, which is based on the greedy algorithm that iterates a sparse selection of the best fitting basis of the matrix $\boldsymbol{A}$. An LS optimization is then performed in the subspace spanned by all previously selected bases. The grid mismatch is mitigated through the Newton–Raphson method, which iteratively refines the DFT grid. We describe this algorithm with complete measurements (sensing matrix is an identity matrix) first by rewriting the nonlinear least squares problem Eq. (19) as a maximization:

$$\{\tilde{\beta}_j, \tilde{\omega}_j\} = \underset{\{\beta_j, \omega_j\}}{\arg\max}\, C_3(\beta_j, \omega_j)$$

with:

$$C_3(\beta_j, \omega_j) = 2\Re\left\{\boldsymbol{x}^{\mathrm{H}} \beta_j \boldsymbol{a}(\omega_j)\right\} - |\beta_j|^2 \|\boldsymbol{a}(\omega_j)\|_2^2 \tag{37}$$

Eq. (37) maximized with respect to ω_j will yield:

$$\tilde{\omega}_j = \underset{\omega_j}{\arg\max}\, \frac{\zeta(\omega_j)}{N} \tag{38}$$

where ζ is the power of the spectra:

$$\zeta(\omega_j) = \left|\boldsymbol{a}^{\mathrm{H}}(\omega_j)\hat{\boldsymbol{x}}_j\right|^2 \tag{39}$$

and Eq. (37) maximized with respect to β_j:

$$\tilde{\beta}_j = \left.\frac{\boldsymbol{a}^{\mathrm{H}}(\omega_j)\hat{\boldsymbol{x}}_j}{N}\right|_{\omega_j = \tilde{\omega}_j} \tag{40}$$

The DFT grid of Eqs. (38) and (40) can be densely defined through zero-padding as indicated by [42]. To improve a coarse estimate of $\{\tilde{\beta}_j, \tilde{\omega}_j\}$, the Newton–Raphson method is used for refine the frequency estimate:

$$\widehat{\omega}_j = \tilde{\omega}_j - \frac{\dot{C}_3(\beta_j, \omega_j)}{\ddot{C}_3(\beta_j, \omega_j)} \tag{41}$$

where the derivatives are defined:

$$\dot{C}_3(\beta_j, \omega_j) = \Re\left\{\left(\hat{\boldsymbol{x}}_j - \tilde{\beta}_j \boldsymbol{a}(\tilde{\omega}_j)\right)^{\mathrm{H}} \tilde{\beta}_j \frac{d\boldsymbol{a}(\tilde{\omega}_j)}{d\tilde{\omega}_j}\right\} \tag{42}$$

and:

$$\ddot{C}_3(\beta_j, \omega_j) = \Re\left\{\left(\hat{\boldsymbol{x}}_j - \tilde{\beta}_j \boldsymbol{a}(\tilde{\omega}_j)\right)^{\mathrm{H}} \tilde{\beta}_j \frac{d^2\boldsymbol{a}(\tilde{\omega}_j)}{d\tilde{\omega}_j^2}\right\} - |\tilde{\beta}_j|^2 \left\|\frac{d\boldsymbol{a}(\tilde{\omega}_j)}{d\tilde{\omega}_j}\right\|_2^2 \tag{43}$$

The NOMP algorithm first estimates $\{\tilde{\beta}_j, \tilde{\omega}_j\}$ using Eqs. (38) and (40) and successively separates the signal:

$$\hat{\boldsymbol{x}}_j = \boldsymbol{x} - \sum_{i=1, i\neq j}^{j} \widehat{\beta}_i \boldsymbol{a}(\widehat{\omega}_i) \quad \text{for} \quad j = 1, 2, \ldots \tag{44}$$

until $\zeta(\tilde{\omega}_j) < \tau$, where Eq. (44) is used to obtain $\{\widehat{\beta}_j, \widehat{\omega}_j\}$ from Eqs. (40) and (41) and τ represents a sparsity promoting ℓ_0-norm regularization, which implicitly determines the model order P:

$$\tau = \sigma^2 \log(N) - \sigma^2 \log\log\left(\frac{1}{1-\rho}\right) \tag{45}$$

where σ is the noise variance and ρ is the probability of a false alarm. For the jth estimate, an initial refinement is performed through the Newton step of Eq. (41). The refinement is conditionally accepted only if it serves to maximize C_3, i.e. locally concave $\ddot{C}_3(\beta_j, \mu_j) < 0$, and if the refinement globally reduces the residual energy such that $\zeta(\widehat{\omega}_j) > \zeta(\tilde{\omega}_j)$. After each singular refinement, the sub-iteration of cyclic refinement is performed R_c times to reestimate $\{\widehat{\beta}_j, \widehat{\omega}_j\}$:

$$\hat{\boldsymbol{x}}_k = \boldsymbol{x} - \sum_{i=1, i\neq k}^{k} \widehat{\beta}_i \boldsymbol{a}(\widehat{\omega}_i) \quad \text{for} \quad k = 1, \ldots, j-1 \tag{46}$$

At the end of the jth estimate, the least square procedure of Eq. (40) is performed to refine complex amplitudes $\{\widehat{\beta}_i\}_i^j$. Performing this step just prior to detecting a new sinusoid increases the rate of convergence of NOMP by mirroring arguments used to establish matching pursuit convergence [59].

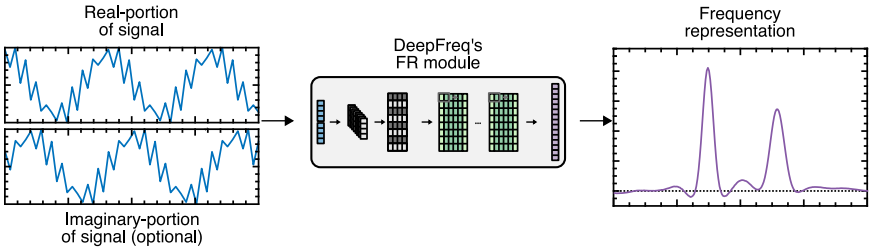

Fig. 4. Algorithm flowchart of DeepFreq's frequency representation module. The assumed complex input of the DeepFreq is split into its real and imaginary parts before input. For real-valued signals, the imaginary-valued portion is supplied with zeros of equal length. The frequency representation generates a pseudospectrum (purple line) subject to a peak-picking algorithm.

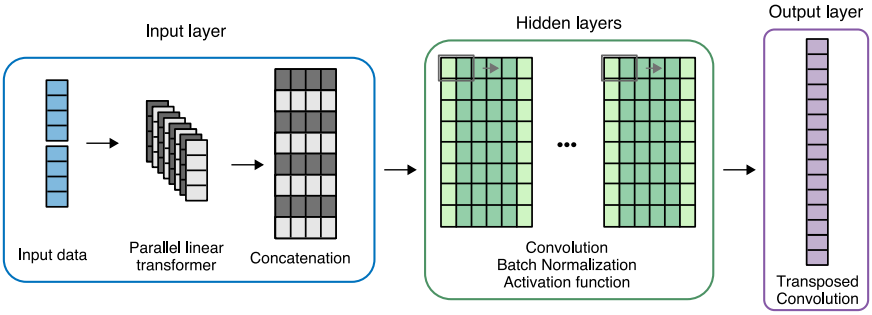

Fig. 5. Architecture of the DeepFreq frequency representation module.

2.5. DeepFreq

A fully connected neural network trained on the signal model of Eq. (1) is unlikely to converge to the maximum likelihood solution of the cost function Eq. (19). In general, the nonlinear least squares problem of Eq. (19) has many local minima, and training generally iterates to some local minimum (or a point near a local minimum). This is opposed to converging to the sought global minimum, which can depend on the initialization conditions and loss function optimization. DeepFreq differs from this direct problem in that it is schematized into different deep neural networks. A pseudospectrum is generated in a data-driven manner, training the frequency representation neural network to produce superimposed Gaussian kernels directly from the measurements. The representation is fed into a second frequency counting neural network that estimates the number of sinusoids. Frequency estimation is then carried out through the selection of the Pth most prominent peaks of the frequency representation, see Fig. 4.

While this yields a fast, fully-automatic method for frequency and model estimation, we are interested only in the frequency representation module. The pseudospectrum FR is defined as the convolution of the P frequencies of the signal with Gaussian kernels $K : \mathbb{R} \to \mathbb{R}$:

$$\mathrm{FR}(n) \triangleq \sum_{i=1}^{P} K\left(n - \omega_i\right) \tag{47}$$

This frequency representation is continuously differentiable and features smooth peaks at the true frequencies for the signal model of Eq. (1). Despite not including amplitude and/or phase information into the CNN, this pseudo-spectra serves as the regression that the CNN weights are tasked to learn from a noisy and finite signal. I.e., this is achieved by the loss function that penalizes the ℓ_2-norm between the output of the CNN and the true pseudo-spectra for a large number of training data. Fig. 5 shows the proposed architecture for the frequency representation neural network.

First, the input layer maps the complex signal to a real-valued feature space. Then, the features are processed in the hidden layers that contain a series of convolutional layers with localized filters of length three and batch normalization [22], interleaved with rectified linear activation unit (ReLUs). The dimension of the input is preserved using circular padding. No pooling layers are used since frequency sensitivity is prioritized over invariance. Finally, the output utilizes a decoder that produces the estimate FR applying a transposed convolution. The original network described in [19] is slightly modified to accept signals $N > 50$. This and all technical details are detailed in Table 2.

3. Experimental methodology

The performance of the implemented LSEs was analyzed using five benchmarks, each using relevant performance metrics defined below. All chosen algorithms were subjected to the same signal model shown in Eq. (1). Additionally, estimation of the number of

Table 2
Layer summary of frequency representation module with respect to input size N, where layers notated by * have no bias. The set of convolution, batch normalization, and ReLU layers are repeated for a depth of 20.

Layer		Features	Filter	Stride	Padding	Output
Input		–	–	–	–	$1 \times 2N$
Linear*		–	–	–	–	1×8000
$20 \times$	Conv*	64	3	–	1, Circ	$1 \times 64 \times 125$
	BatchNorm	64	–	–	–	$1 \times 64 \times 125$
	ReLU	–	–	–	–	$1 \times 64 \times 125$
ConvTrans*		64	17+Stride	$2^{\log_2(N)-3}$	9	$1 \times \frac{125}{8} N$

sinusoidal components P is given a priori. This is because LSEs are parametric and the algorithms typically used for estimation of the number of sinusoidal components within a signal is a separate problem, see [20] for a brief introduction on model order estimation. The normalized frequencies are wrapped on the bounds $[0, 1)$. The AWGN variance σ is defined such that a desired signal-to-noise ratio (SNR) can be obtained, $\text{SNR} \triangleq \|\boldsymbol{A}(\boldsymbol{\omega})\boldsymbol{\beta}\|_2^2/\sigma^2$. A Raspberry Pi 4b, which is an affordable option for control of ultrasonic fatigue tests, is utilized for all benchmarks using Python 3.8.8 via the Anaconda distribution. We readily make available our algorithms and synthetic data [60].

This section aims to define multiple benchmarks for LSEs which are lacking in the literature. Specifically, the first test aims to demonstrate the degeneration of LSEs when exposed to quasi-stationarity within each window: this corresponds to the effects of the Rayleigh limit (if the signal was stationary) compounded with signal noise. The following three tests aim to benchmark in the stationary signal case when LSEs are subject to less favorable conditions, e.g., non-unit amplitudes, interharmonics, and order mismatch. The last test looks at the asymptotic computational costs and the algorithm runtimes for a Raspberry Pi 4b.

3.1. Test signals

For usage in the first two tests, a synthetic real-valued signal is generated from an experimental UFT signal. The discrete UFT velocity signal has a sampling frequency $F_s = 10^6$ Hz which was performed on a pure copper fatigue specimen in a setup similar to Fig. 1. The time-evolving amplitude and frequencies are extracted from the signal per window via the DFT for the first three harmonics. Then cubic polynomial fits are evaluated on a quasi-stationary basis for each ith window along the entire experimental signal seen in Fig. 6. These averaged sinusoidal parameters can be quantified as:

$$\bar{a}_i = \frac{1}{N}\int_0^N a_i(n)\,dn$$

and:

$$\bar{\omega}_i = \frac{1}{NF_s}\int_0^N f_i(n)\,dn$$

where $\bar{a}_i$ and $\bar{\omega}_i$ are quantized (window-averaged) amplitude and normalized frequencies, respectively. The frequencies in Section 3 are normalized to cycles per sample corresponding to the wrapped bounds $[0, 1)$. This achieves a similar signal model defined in Eq. (1) with a model order of $P = 6$ complex sinusoids but has a symmetrical frequency component about $\omega = 0.5$. The real-valued synthetic signal is created with a sampling frequency $F_s = 250$ kHz which is compared with the experimentally obtained signal in Fig. 6. The time evolving sinusoidal parameters cubic polynomial fits can be found in Appendix A. For all tests except the first, the synthetic signal excludes the last 10 s of the data, where the formation of a large crack manifests large nonstationary components unsuitable for the LSEs.

For the next 3 tests, we test the sensitivity to amplitude variability, interharmonics, and order mismatch for randomly generated stationary test signals. The set of true normalized frequencies representing harmonics is generated in these signals in two steps. First, a fundamental harmonic is generated randomly, such that $\{\omega_1 \in [2/N, 1/P - 2/N)\}$. Second, the remaining higher harmonics ($i > 1$) are simply integer multiples of ω_1 creating a set $\{\omega_i \in [0, 1) : |\omega_i - \omega_j| > 2/N, \forall j \neq i\}$. Note, a minimum frequency separation below the Rayleigh limit of $1/N$ makes the estimation problem ill-posed [24]. The complex amplitudes of the signal components in $\boldsymbol{\beta}$ are generated with independent and identically distributed uniformly random phases from $[0, 2\pi)$. The complex amplitudes' magnitudes $|\beta_i|$ in are generated in ratio descending manner. The fundamental harmonic $i = 1$ has a unit complex magnitude, with subsequent harmonics ($i > 1$) being scaled such that $\{|\beta_i| \in (0, 1] : |\beta_{i+1}| = |\beta_i|/\alpha, \forall i\}$.

3.2. Configuration of algorithms

We compare the five algorithms Unitary ESPRIT [50], RELAX [36], CFH [25], NOMP [42], and DeepFreq [19] with configurations:

- Unitary ESPRIT requires an estimate of the measurement matrix and the model order. The former is estimated from the signal vector with size $M = N/2$ since this sub-vector length was shown to be optimal by [61]. After obtaining the frequency estimates $\widehat{\boldsymbol{\omega}}$, the complex amplitude estimates $\widehat{\boldsymbol{\beta}}$ are LS solvable using Eq. (21).

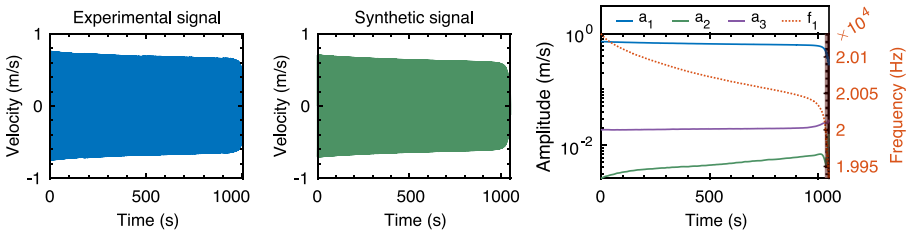

Fig. 6. Copper specimen ultrasonic fatigue test excited at 20 kHz with vibration signal sampled at 200 kHz (left). A synthetic signal (center) is created from the first 3 harmonics' amplitudes and frequencies fits (right). The frequency f_1 is only shown, where it should be understood that higher harmonic frequencies occur at integer multiples, e.g. $f_2 = 2f_1$. Near the end of the ultrasonic fatigue test, the signal becomes increasingly nonstationary, where the last 20 s (shaded) are excluded from evaluation.

- RELAX utilizes a zero-padded DFT and iterates over the model order. We select this zero-padding length such that its grid size is interpolated over $Z_\mathrm{p} = 4N$. This was chosen since it corresponded to a computational speed similar to other iterative LSEs, see Fig. 15.
- CFH requires only a model order and a minimum separation parameter which is already set by default to the Rayleigh limit and is considered optimal by [25].
- NOMP requires a cyclic refinement parameter R_c and the sparsity promoting parameter τ which implicitly determines the model order. We set $R_c = 3$ since it was shown by [42] that beyond number of cycles of Newton refinement has diminishing returns in estimation accuracy. We modify the NOMP algorithm to terminate on model order P, bypassing the regularization Eq. (45) to put it on par with other LSEs, especially when trying to test model order robustness.
- DeepFreq is inherently input data length-dependent so that separate frequency representation neural networks are trained per input length. We specifically use Table 2 and a standard deviation of the Gaussian kernel to $0.3/N$. It was found that for all input sizes, DeepFreq was able to generalize for all SNR when its frequency representation neural network was trained on AWGN with SNR = 1 dB. The training data consists of the generating signals with the following sets on uniformly distributed bounds: $\{P \in [1, 10]\}$, $\{\boldsymbol{\omega} \in [0, 1) : |\omega_i - \omega_j| > 2/N, \forall j \neq i\}$, and $\{|\boldsymbol{\beta}| \in [0.001, 1]\}$. For other training data parameters, we used the defaults described in the original paper [19]. The amplitude estimates are obtained from its frequency representation through peak-picking via amplitude prominence with a minimum separation of $2/N$, when appropriate.

We also include the DFT, utilizing the Fast Fourier Transform algorithm, as a baseline which determines the P sinusoidal parameters through peak-picking via amplitude prominence. with a minimum separation of $2/N$, when appropriate.

3.3. Performance metrics

Five performance metrics are used: matched frequency distance (MFD), matched amplitude distance (MAD), frequency success rate (FSR), mean-squared frequency error (MSFE), and the mean-squared amplitude error (MSAE). All metrics are averaged over each window (discretized) of the entire synthetic signal for the first test, and the number of Monte Carlo trials for the latter 3 tests. To determine the accuracy of the recovered sinusoidal parameters, the MFD is defined as:

$$\mathrm{MFD} \triangleq \sum_{\widehat{\omega}_i \in \widehat{\boldsymbol{\omega}}} \left(\min_{\overline{\omega}_j \in \overline{\boldsymbol{\omega}}} \left| \widehat{\omega}_i - \overline{\omega}_j \right| \right) + \sum_{\overline{\omega}_j \in \overline{\boldsymbol{\omega}}} \left(\min_{\widehat{\omega}_i \in \widehat{\boldsymbol{\omega}}} \left| \overline{\omega}_j - \widehat{\omega}_i \right| \right) \tag{48}$$

The MFD aims to match each estimated frequency with its closest window-averaged frequency, and vice versa, and record a chamfered error. The MAD follows a similar logic, seen in Eq. (49).

$$\begin{aligned} \mathrm{MAD} &\triangleq \widehat{\sum} + \overline{\sum} \\ \text{with } \widehat{\sum} &= \sum_{\hat{\omega}_i \in \hat{\boldsymbol{\omega}}} \frac{\left| \hat{a}_i(\hat{\omega}_i) - \bar{a}_j(\bar{\omega}_j) \right|}{\bar{a}_j(\bar{\omega}_j)}, \quad \bar{\omega}_j = \underset{\bar{\omega}_j \in \bar{\boldsymbol{\omega}}, \widehat{\omega}_i \in \widehat{\boldsymbol{\omega}}}{\arg\min} \left| \widehat{\omega}_i - \bar{\omega}_j \right| \\ \overline{\sum} &= \sum_{\bar{\omega}_j \in \bar{\boldsymbol{\omega}}} \frac{\left| \hat{a}_i(\hat{\omega}_i) - \bar{a}_j(\bar{\omega}_j) \right|}{\bar{a}_j(\bar{\omega}_j)}, \quad \widehat{\omega}_i = \underset{\bar{\omega}_j \in \bar{\boldsymbol{\omega}}, \widehat{\omega}_i \in \widehat{\boldsymbol{\omega}}}{\arg\min} \left| \widehat{\omega}_i - \bar{\omega}_j \right| \end{aligned} \tag{49}$$

Note, the MAD is normalized by the window-averaged amplitude to remove the scaling bias from amplitudes with larger magnitudes.

The FSR is defined as follows:

$$\mathrm{FSR} \triangleq \frac{\sum_{\widehat{\omega}_i \in \widehat{\boldsymbol{\omega}}} \mathcal{S}\left(\widehat{\omega}_i, \boldsymbol{\omega}\right) + \sum_{\omega_j \in \boldsymbol{\omega}} \mathcal{S}\left(\omega_j, \widehat{\boldsymbol{\omega}}\right)}{2P} \tag{50}$$

with the success function defined as:

$$\mathcal{S}(i, \boldsymbol{j}) \triangleq \mathbb{1}\left[\min_{j_k \in \boldsymbol{j}} |i - j_k| < \frac{1}{2N} \right] \tag{51}$$

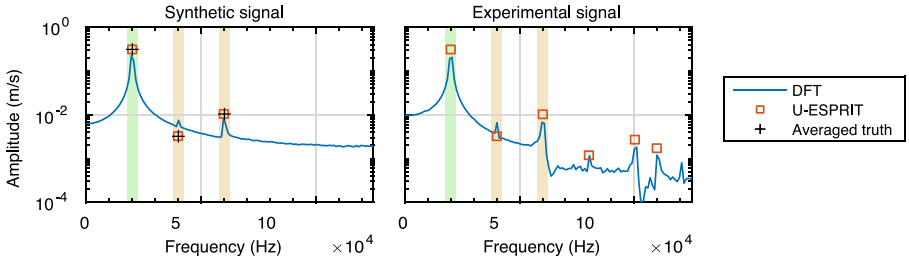

Fig. 7. Demonstration of the DFT spectra and Unitary ESPRIT line spectra for the synthetic signal (left) and experimental signal (right). To extract sinusoidal parameters of the first three harmonics: first, a maximum amplitude of the fundamental is searched for, represented by the green shaded area; second, integer multiples (two and three times) of the green shaded area are created and represented by the yellow shaded area; lastly, the peak is searched within the yellow shaded area, within a $\pm 2/N$ tolerance.

where $\mathbb{1}[\square]$ denotes the indicator function. An FSR of 1 is achieved if, and only if, all estimated frequencies are in the vicinity of one or more true frequencies and all true frequencies are in the vicinity of one or more estimated frequencies. The MSFE and MSAE are also introduced as statistical measures:

$$\text{MSFE} \triangleq \frac{1}{P}\sum_{j=1}^{P}\left(\min_{\widehat{\omega}_i\in\widehat{\boldsymbol{\omega}}}\left|\widehat{\omega}_i-\omega_j\right|\right)^2 \tag{52}$$

and:

$$\text{MSAE} \triangleq \frac{1}{P}\sum_{j=1}^{P}\left|\widehat{a}_i(\widehat{\omega}_i)-a_j(\omega_j)\right|^2 \quad \text{with} \quad \omega_i = \underset{\omega_j\in\boldsymbol{\omega},\widehat{\omega}_i\in\widehat{\boldsymbol{\omega}}}{\arg\min}\left|\widehat{\omega}_i-\omega_j\right| \tag{53}$$

The MSFE is juxtaposed with the averaged Cramér–Rao lower bound (CRB) of Eq. (B.6), which characterizes the asymptotic performance at large N and/or high SNR, see Appendix B. The MSAE metric is compared with an "Oracle" estimator, which is the LS estimate of the noisy signal given the true frequencies. Lastly, it is to be understood that all metrics are subject to the constraint $\{\omega_k \backslash \boldsymbol{\omega} : \forall\omega_{k+1} \in \boldsymbol{\omega}\}$, i.e. no components are repeated for evaluation.

4. Results

4.1. Parameter estimation for synthetic and experimental signals

The differences between an estimator's ability to extract the first three harmonics of a synthetic and experimental signal are demonstrated in this test. The vibration signal is distinguished by a large amplitude variability between the first harmonic and higher harmonics, seen in Fig. 6. Harmonics are extracted through all estimators using the following methodology: the maximum prominent amplitude is assumed to be the fundamental harmonic, and the following peak at integer multiples of the fundamental harmonic are searched for within $\pm 2/N$ and extracted. These details are visualized in Fig. 7 for windows of both the synthetic and experimental signals. The SNR for the experimental signal is estimated to be 50 dB through a periodogram-based method [21]; the synthetic signal has AWGN applied per window such that its SNR matches the experimental signal. A signal length of $N = 2^8$ is empirically chosen to trade off computational time for the LSEs and still allow for peak resolution in the DFT frequency spectra. For the LSEs, the model order $P = 6$ and $P = 12$ are imposed for the synthetic and experimental signals, respectively.

In Fig. 8, the estimation of harmonics is plotted against the synthetic signal. For Unitary ESPRIT and NOMP, the estimation of the amplitudes match most closely to the averaged truth; their frequency estimates are slightly outperformed by DeepFreq. However, DeepFreq demonstrates that it fails to estimate accurate amplitudes, even for a strong fundamental harmonic. Due to the small signal length, it can be seen that the DFT struggles to accurately capture even the fundamental harmonic's amplitude. It should be reminded that even though zero-padding the DFT would improve the amplitude results (to a certain extent), its frequency estimation ability would remain the same in resolution. CFH outperforms RELAX for fundamental harmonic amplitude estimation, despite its fundamental harmonic frequency estimation being offset more than RELAX. However, both CFH and RELAX struggle to steadily estimate the second harmonic's amplitude due to a difference of magnitudes with respect to the fundamental and third harmonic. In general, Unitary ESPRIT and NOMP show the most stability in parameter estimation for the synthetic signal. DeepFreq is unique in that it completely fails to provide accurate amplitude estimates for any part of the synthetic signal, however, it is able to only track the fundamental frequency with success. While not shown, it is unable to extract the frequencies of the higher harmonics.

In contrast with the synthetic signal, the estimation of harmonics is performed for the experimental signal. A glance of Fig. 9 reveals a qualitative difference in the behavior of all estimators. For the fundamental harmonic's amplitude, Unitary ESPRIT, NOMP, and CFH all closely estimate a similar amplitude evolution. For the fundamental harmonic's frequency, CFH follows more closely to RELAX, while Unitary ESPRIT and NOMP overlap one another. All LSEs struggle with stability more than compared with the synthetic signal, e.g. Unitary ESPRIT and NOMP demonstrate a large oscillation at the beginning of the experimental signal before stabilizing for the second harmonic's amplitude. This might be because of the presence of more harmonics within the real experimental signal,

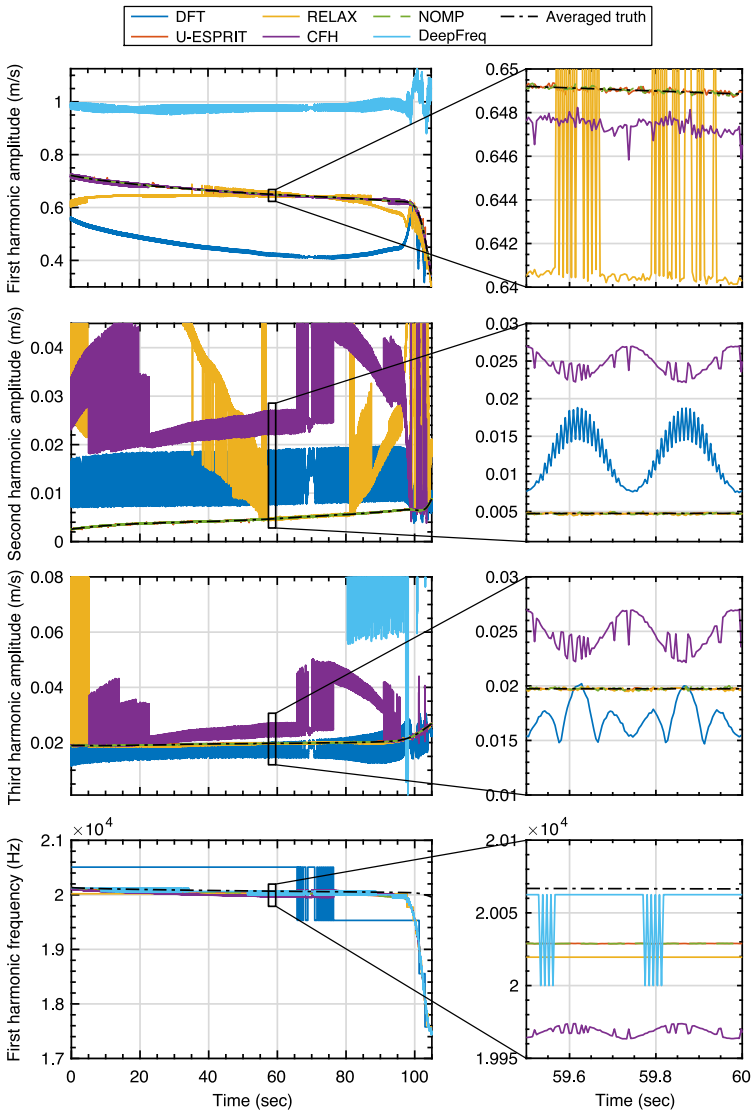

Fig. 8. Sinusoidal parameter estimation of the three harmonics' amplitudes and fundamental harmonic's frequency for synthetic signal. The legend shown above applies to all plots.

which presents a more difficult estimation problem for the LSEs. Similar to the synthetic test, DeepFreq has a subpar performance for extraction of amplitudes but can capture the fundamental harmonic only. If the LSEs' behavior on the synthetic signal test is to be indicative of its behavior on the experimental signal, one may conclude that the general results of the LSEs can be carried over. For example, the widely spread and oscillatory nature of amplitude estimation should be understood to be an uncertainty of the estimator instead of a detectable very small vibratory oscillation. Lastly, it can be noted that during the nonstationary evolution of sinusoidal parameters near fatigue failure, i.e., the end of the synthetic signal, the LSEs provide frequency estimates which diverge from truth, but still maintain accurate amplitude estimates.

4.2. Tolerance to noise, sensitivity to signal length

In this test, we opt to study the quasi-stationary synthetic signal of three well-separated sinusoids, with large amplitude variability between the first harmonic and higher harmonics, as seen in Fig. 6. Grid-based simulations with respect to (windowed) signal lengths and AWGN levels are applied per window along the entire signal. For purely stationary signals, it is well known that an increase in signal length and/or a decrease in noise variance increases the performance of LSEs. A specific quality of the synthetic signal is the unknown interaction between frequency resolution and non-stationarity with respect to signal length.

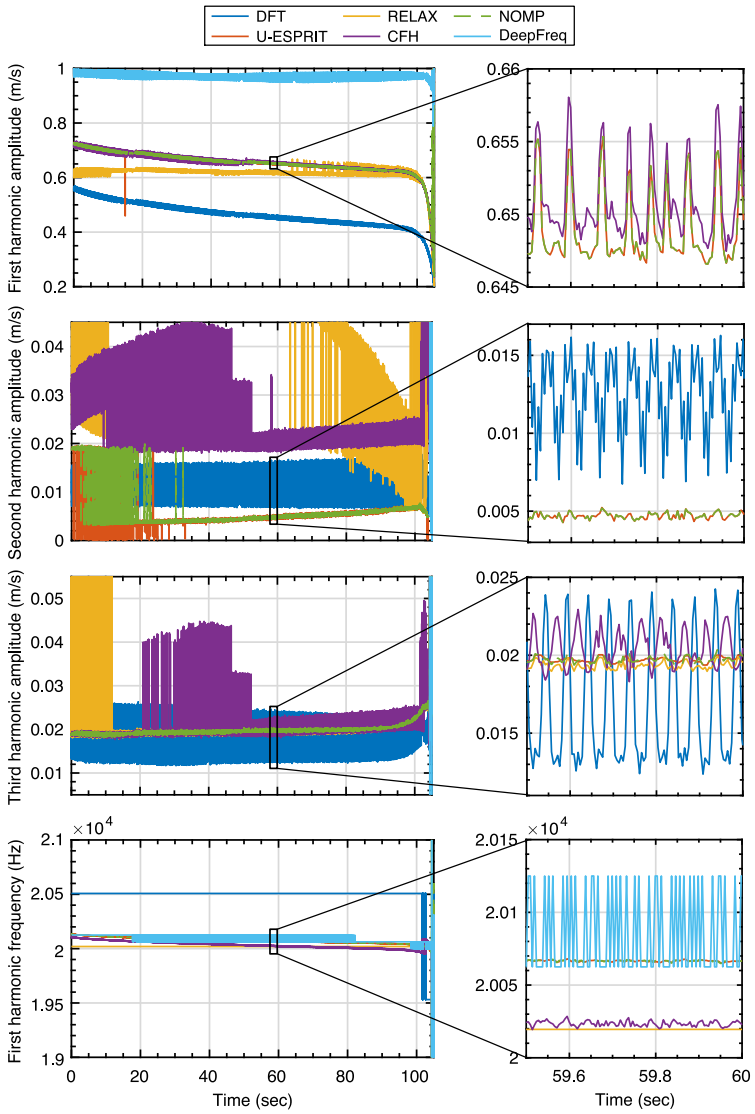

Fig. 9. Sinusoidal parameter estimation of the three harmonics' amplitudes and fundamental harmonic's frequency for experimental signal. The legend shown above applies to all plots.

In Fig. 10, the mean MFD, with outliers outside of five standard deviations removed, is plotted as a heatmap for specific LSEs with respect to signal length and AWGN. Most notably, Unitary ESPRIT has the largest area of "best" frequency estimation performance (in dark blue) compared to other estimators. Additionally, similarities between estimators that explicitly rely on the DFT are also seen, despite their attempts to refine frequency estimation: the DFT has an area of "best" frequency estimation that roughly intersects the CFH, RELAX, and NOMP's areas of "best" frequency estimation. As mentioned previously, CFH, RELAX, and NOMP all utilize DFT operations in their algorithm, but their refinements can be characterized as peak parabolic interpolation, zero-padding interpolation, and Newton–Raphson refinements of the DFT grid, respectively. This is likely due to the limitations of the DFT operation with respect to the quasi-stationarity in the sinusoidal parameters, whose expression can be seen in the fits in Appendix A. The other LSEs' expanded area in Fig. 10 compared to the DFT represents their ability to mitigate the quasi-stationarity and AWGN in each signal segment.

Lastly, it should be mentioned that DeepFreq has errors larger than any of the other LSEs presented; the frequency representation produced by DeepFreq is unable to compensate for any quasi-stationarity in the signal and creates a problem for a peak picking algorithm. This is a realistic reminder that this particular data-driven approach cannot be deployed so easily even if the data is quasi-stationary. Similarly, the average MAD is shown in Fig. 11, with outliers outside of five standard deviations removed, LSEs show a similar pattern with the MFD. As seen in the previous test, the LSEs have a better amplitude estimation ability compared to their frequency estimation, indicating amplitude estimation is more "lenient". The exception is the large difference of amplitude

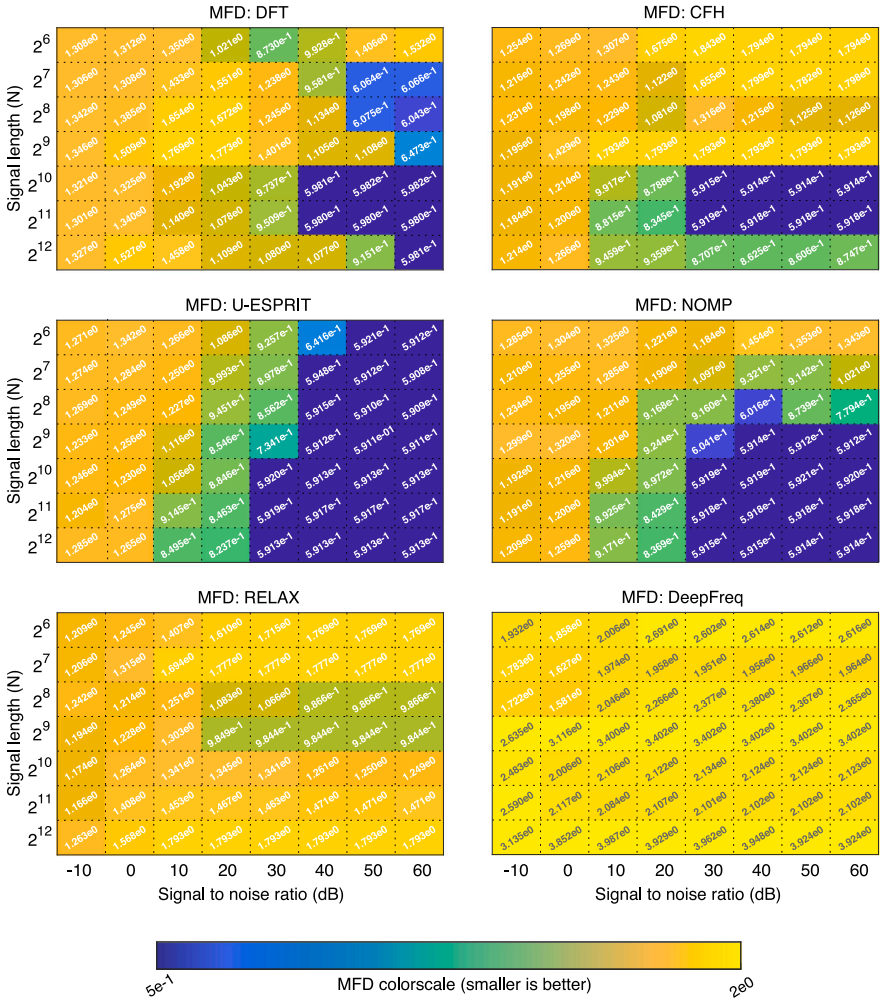

Fig. 10. Simulation results as heatmaps of the MFD metric for each LSE, whose colorscales are plotted logarithmically.

estimation performance for the DFT: the corresponding ranges at which the DFT provides excellent frequency estimations (dark blue) provide poor amplitude estimates with a MAD of approximately one magnitude higher than other LSEs. For all LSEs for this particular synthetic signal, in the low SNR region (below 20 dB) provides unsatisfactory results for even the best-performing LSEs.

4.3. Sensitivity to amplitude variability

For randomly generated well-spaced sinusoids, the components whose descending magnitudes vary by the amplitude magnitude scaling parameter α correspond to a proportionally difficult problem as α increases. We remind the reader that the complex amplitudes are generated in a ratio descending manner, such that $|\beta_1| = 1$ and subsequent harmonics ($i > 1$) are scaled such that $\left\{|\beta_i| \in (0, 1] : |\beta_{i+1}| = |\beta_i|/\alpha, \forall i\right\}$. We perform 5000 Monte Carlo simulations on five complex sinusoids for a signal length $N = 2^7$ and an SNR = 40 dB. In Fig. 12, the statistical performance is shown. At first glance, Unitary ESPRIT performs best or on par with all other LSEs for the three metrics. For the FSR, the DFT can be seen to be most unsuccessful in estimating frequency components within the defined vicinity of $0.5/N$. RELAX, which is bound by its zero-padding operation, is also expected to plateau in performance. For all other LSEs at unit amplitude magnitudes ($\alpha = 1$), their FSR parallel results typically seen in the literature.

In the plots of MSFE for $\alpha > 4$, Unitary ESPRIT and NOMP begin to descend towards the CRB. NOMP characteristically shows an upturn after $\alpha = 4$ which is not seen by other LSEs. After $\alpha = 2$, CFH overtakes Unitary ESPRIT in terms of distance from the CRB. The DFT, RELAX, and DeepFreq can be seen to plateau approaching unit amplitudes. LSEs in the MSAE plot do not follow the same pattern as their MSFE counterparts. Essentially, ESPRIT is the first to approach the lower bound at $\alpha = 3$, and CFH achieve the lower bound at $\alpha = 1.5$. The Oracle estimator (Oracle knows the true frequencies and performs an LS fit) shows a lower, but realistic bound tied to the accuracy of the frequency estimation.

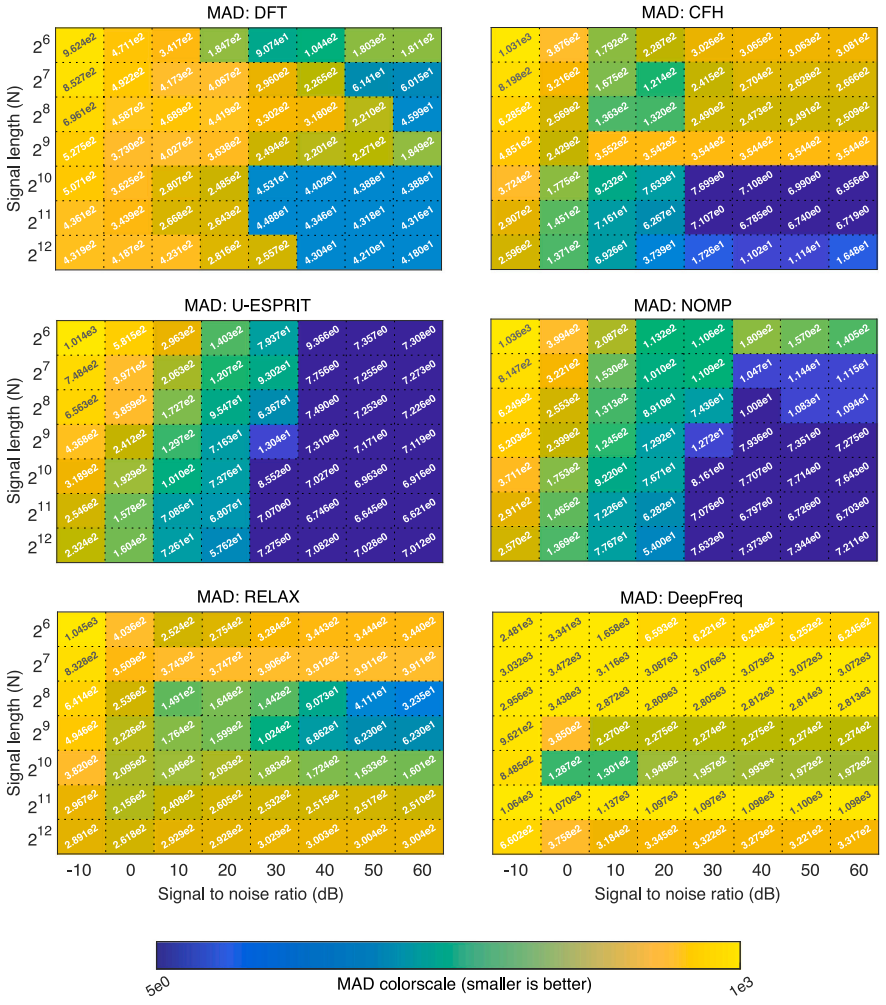

Fig. 11. Simulation results as heatmaps of the MAD metric for each LSE, whose colorscales are plotted logarithmically.

4.4. Influence of interharmonics (super-resolution)

When maximizing for the smallest signal length, interharmonics close to harmonics with respect to the Rayleigh limit can influence measurement accuracy. Interharmonics correspond to non-integer multiples of the fundamental harmonic where an example can be seen in Fig. 2. With signal parameters SNR = 40 dB, $N = 2^7$, a similar test with three well-spaced complex sinusoids of magnitude ratios dictated by $\alpha = 1.5$, is performed for 5000 Monte Carlo simulations. A single interharmonic is uniformly randomly generated such that its distance is d from at least one harmonic. In Fig. 13, the FSR of the DFT and RELAX can be seen to already inconsistently predict the well-spaced harmonics, with an average success rate 63% and 16% respectively. At $d = 0.75/N$, only the Unitary ESPRIT and CFH can still achieve a high FSR. Below $d = 0.75/N$, only Unitary ESPRIT is capable of obtaining a consistent FSR until $d = 0.1/N$ when its frequency resolution begins to taper. The performance dip seen by some LSEs in the FSR is simply the conflict between the LSE to predict the interharmonic with another harmonic. The gradual increase despite still decreasing d indicates that the interharmonic is simply masked by its harmonic.

For the MSFE, all LSE follow a similar trend, where the order of best performance is Unitary ESPRIT, followed by CFH, DeepFreq, RELAX, NOMP, and the DFT. Interestingly, DeepFreq is capable of generalizing below $d = 1/N$ until $d = 0.4$ despite being trained on randomly generated signals with $d = 1/N$ minimum separation. The MSAE shows a comparable pattern of performance, except that DeepFreq provides very poor amplitude estimates, similarly seen in the previous simulation plot of Fig. 12.

4.5. Robustness against order mismatch

While it is not the main focus of this paper, robustness against model order mismatch is an attractive feature since algorithms used for order estimation can under/overestimate. Generally, the most commonly used model order estimation methods are the Akaike

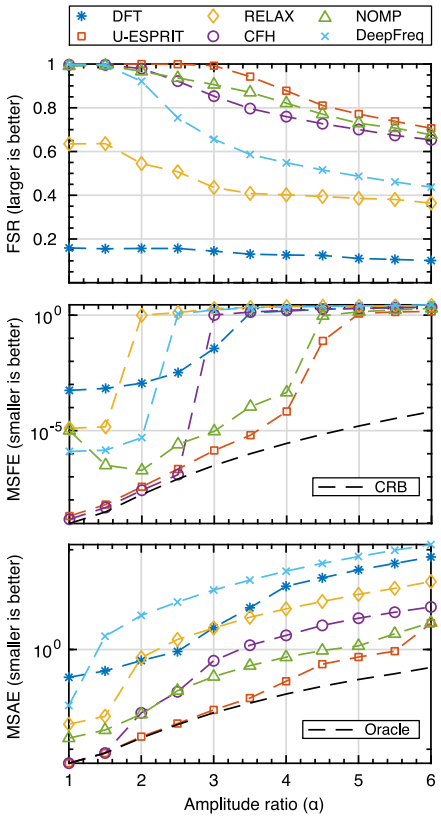

Fig. 12. Simulation results for varying amplitude magnitude scaling parameter α. The signal length $N = 2^7$, SNR = 40 dB, and model order $P = 5$ are pre-determined. The legend shown above applies to all plots, with sub-legends within the plots applying only to sub-plots.

information criterion and the minimum description length criterion [20]. In short, they balance a measure between a goodness of fit with the data while penalizing a higher model order. With signal parameters SNR = 40 dB, $N = 2^7$, 5000 Monte Carlo simulations with $P = 5$ frequencies drawn uniformly random in the set $\{\boldsymbol{\omega} \in [0, 1) : |\omega_i - \omega_j| > 2/N, \forall j \neq i\}$ with amplitude magnitude ratios dictated by $\alpha = 1.5$ are performed. The LSEs are given a varying $\widehat{P}$ instead of the true P. In Fig. 14, Unitary ESPRIT can be seen to provide the best results when the imposed model order $\widehat{P} > P$ for MSFE and MSAE. At $\widehat{P} \leq P$, CFH provides the lowest MSFE and MSAE, and benefit the most when $\widehat{P} = P$. CFH can be seen to perform best as $\widehat{P} \to P$, whereas Unitary ESPRIT is very robust when $\widehat{P} > P$.

4.6. Computational effort

The computational time plays a large factor when considering LSE algorithms for real-time usage. In Fig. 15, averaged computation times are shown with respect to signal lengths. It can be seen that the DFT always achieves the best runtime, this is due to the usage of the Fast Fourier Transform algorithm, which has the relatively lowest asymptotic complexity in Table 3. RELAX, CFH, and NOMP can be seen to have similar runtimes as the signal length varies. ESPRIT begins to diverge at $N = 10$, having the greatest runtime analogous to its large asymptotic complexity. It and the other LSE's asymptotic complexities in Big-$\mathcal{O}$ notation can be found in Table 3. Unique to DeepFreq, offline training requires a considerable amount of computational effort and training data that is a good representation of experimental data. The GPU training times can be seen at the bottom of Fig. 15 on a specialized computer for offline training with four virtual Intel Xeon CPUs clocked at 2.00 GHz, 32 Gb of memory, and an NVIDIA Tesla P100 GPU. It should be noted that Python (versions 3.8 and prior) is known to have a poor just-in-time compilation, meaning that LSEs that utilize for-loops, namely CFH, RELAX, and NOMP, suffer with respect to their theoretical computational complexities. One can expect the LSEs to perform closer to their asymptotic complexities when programmed in a statically compiled language. A few options exist to leverage static typing of existing Python code, e.g., using Pythran [62].

5. Conclusions

In the field of fatigue research, VHCF of metallic components is a microstructural-dominated phenomenon that relies on UFTs for experimental study. Due to the recent abundance of experimental data in the literature, new modeling attempts will seek

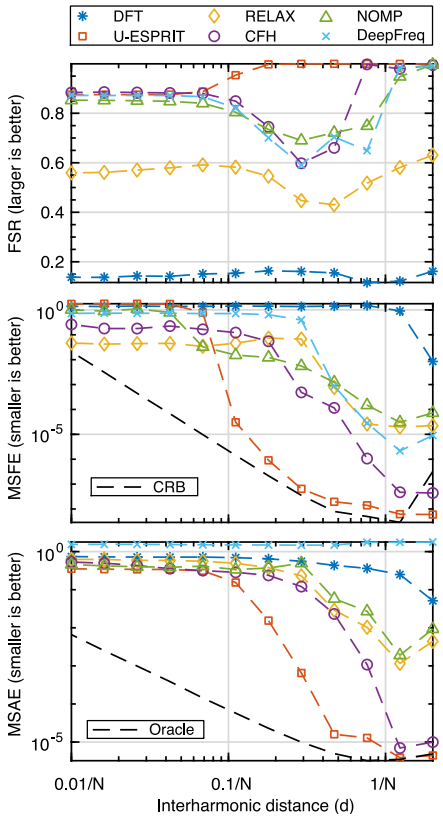

Fig. 13. Simulation results for varying interharmonic distance parameter d. The signal length $N = 2^7$, SNR = 40 dB, amplitude magnitude ratio $\alpha = 1.5$, and model order $P = 3 + 1$ are pre-determined. The legend shown above applies to all plots, with sub-legends within the plots applying only to sub-plots.

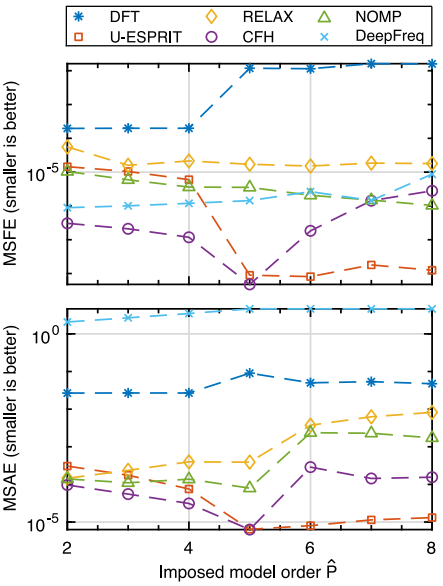

Fig. 14. Simulation results for varying imposed model orders $\hat{P}$. The signal length $N = 2^7$, SNR = 40 dB, amplitude magnitude ratio $\alpha = 1.5$, and model order $P = 5$ are pre-determined. The legend shown above applies to all plots.

to determine the microstructural state from the vibration signal measured from a fatigue specimen. The real-time estimation of these sinusoidal parameters, specifically frequencies and amplitudes, are of utmost importance. Similar domains which rely on the

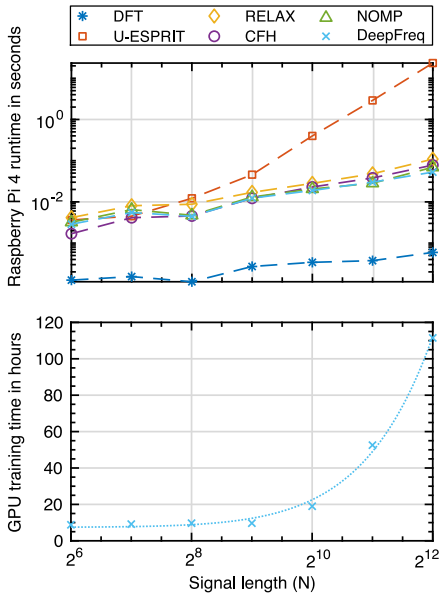

Fig. 15. Averaged runtimes of LSEs on a Raspberry Pi 4 (top) and offline Tesla P100 GPU training time required for frequency representation module of DeepFreq (bottom) with respect to signal length N. The legend shown above applies to all plots.

Table 3
Comparison of asymptotic computational complexities of implemented line spectra estimators in big-$\mathcal{O}$ notation.

LSE	Asymptotic complexity
DFT	$\mathcal{O}(N\log(N))$
Unitary ESPRIT	$\mathcal{O}((N/2)^3)$
RELAX	$\mathcal{O}(PZ_{\mathrm{p}}\log(Z_{\mathrm{p}}))$
CFH	$\mathcal{O}(PN\log(N))$
NOMP	$\mathcal{O}(PR_{\mathrm{c}}N\log(N))$
DeepFreq[a]	$\mathcal{O}(\mathcal{L}\mathcal{K}\mathcal{N}\mathcal{R}^2+N)$

[a]The notation is defined where $\mathcal{L}$ is the layer depth, $\mathcal{N}$ is the sequence length, $\mathcal{R}$ is the representation dimension, $\mathcal{K}$ is the kernel size of convolutions.

estimation of the damage state can be found in SHM, but also in nonlinear acoustics, namely SHG and NRUS. Unique to these ultrasonic vibration signals is the quasi-stationarity of the sinusoidal parameters, which slowly evolve with respect to many periods of the signal. When real-time usage is sought, the dichotomy between signal length, i.e., the Rayleigh limit $1/N$, and the quasi-stationarity of sinusoidal components can pose difficulties for estimators whose original formulation relies on the assumption of stationary signals.

This motivates an experimental methodology comparing alternatives to the DFT. LSEs provide direct estimates of sinusoidal parameters instead of estimating the entire frequency spectra and can overcome shortcomings of the DFT given a signal model and model order a priori. Notably, many LSEs are considered to have state-of-the-art results for purely stationary signals with unit amplitudes, which cannot be said for quasi-stationary signals with non-unit amplitudes.

The benchmarks on a synthetic UFT signal and randomly generated signals with unique estimation challenges seen in ultrasonic vibration, show the capabilities of the various estimators in terms of their adaptability to quasi-stationary, amplitude variability, small interharmonic distance (super-resolution), and robustness to model order mismatch (number of sinusoidal components). In general, Unitary ESPRIT and then NOMP can be seen to offer the best performance with respect to the quasi-stationarity of a synthetic UFT signal. In the subsequent benchmarks, their statistical performance on stationary signals further validates their general usage for sinusoidal parameter estimation. A drawback of this study and current literature, however, is the lack of theory that precisely dictates when a signal is "too nonstationary", thus prohibiting the use of LSEs.

In the practice of UFT, three algorithms can be recommended depending on usage criteria. When accuracy and precision are dominating criteria, Unitary ESPRIT is recommended for all cases when the signal length is $2^6 \leq N < 2^9$. Beyond this signal length, we recommend NOMP for its computational simplicity. This is justified by their area of "best" performance when dealing with the nonstationarity of sinusoidal parameters and signal length dichotomy for the synthetic signal. When computational complexity becomes a constraint, e.g., if using a low-powered Arduino, CFH becomes extremely attractive since it offers the lowest computational complexity of the five algorithms. However, evident in Figs. 10 and 11, it suffers from an "island" of good performance

which is bounded by the nonstationarity of components and the Rayleigh limit (signal length). In terms of flexibility beyond the metrics used in the paper, NOMP is advised to be used simply because it offers more tunable parameters (specifically zero-padding and refinement parameters) than Unitary ESPRIT (sub-vector length) and CFH (which has none). Tuning the algorithmic parameters seen in Table 1 can be done heuristically on synthetic ultrasonic vibration signals representative of a particular fatigue cycle regime. Lastly, it should be noted that all algorithms presented require a priori knowledge of the signal: e.g., the number of sinusoids and/or statistical properties of the noise. However, model order estimation is a separate problem; instead, the test of Section 4.5 examines when the LSEs are given an incorrect model order estimate. A simple solution is to overestimate the number of sinusoidal parameters and extract the harmonics only, given the robustness of the parametric algorithms seen in Fig. 14.

Data-driven CNNs remain a hopeful and interesting prospect for LSEs and vibration signals, despite it requiring a large offline training investment. CNNs have not been limited to parameter estimation, but have also been used for LSEs in denoising stationary signals [63], demonstrating a substantial improvement in estimation accuracy. The poor estimation performance of DeepFreq for the tested synthetic UFT signal is a notable reminder that a data-dependent algorithm cannot necessarily generalize beyond its test data in all facets. Interestingly our implementation of DeepFreq generalized in terms of super-resolution estimation: the training data used in Fig. 13 had an imposed minimum separation of $1/N$, but DeepFreq was able to estimate frequencies with good results up to $0.4/N$. Beyond the scope of the paper, we noticed DeepFreq tended to fail even for stationary signals when amplitudes of the harmonics differed beyond a magnitude of $\alpha \approx 10^2$. For more practical usage in vibration signals, the use of quasi-stationary signals instead of stationary signals for training data and the modification of the optimization function to account for the average frequencies can be a possible remedy to nonstationarity. The poor amplitude estimates created by the frequency representation can be remedied in one of two ways: including amplitude information into the architecture of the frequency representation neural network, or simply performing LS fits on its estimated frequencies into the algorithm. It is suspected DeepFreq can be improved concerning these benchmarks with the former points explored, in addition to modifying its CNN architecture e.g., increasing layer depth in conjunction with using residual nets [64].

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Synthetic UFT sinusoidal parameters

In this Appendix, we detail the cubic fits of the time-varying sinusoidal parameters seen in right of Fig. 6, specifically $a_1(t)$, $a_2(t)$, $a_3(t)$ and $f_1(t)$. The original signal comes from a pure copper fatigue specimen subject to a UFT. It is reminded that $f_2(t)$ and $f_3(t)$ are integer multiples of f_1(t), since they correspond to higher harmonics used for the first test in Section 3. The cubic fits for the amplitudes in meters per second are:

$$a_1(t) = -(7.564\times10^{-10})t^3 + (1.038\times10^{-6})t^2 - (4.818\times10^{-4})t + 7.378\times10^{-1}$$

$$a_2(t) = (6.094\times10^{-12})t^3 - (7.455\times10^{-9})t^2 + (5.791\times10^{-6})t + 2.751\times10^{-3}$$

$$a_3(t) = (2.346\times10^{-11})t^3 - (2.904\times10^{-8})t^2 + (1.09\times10^{-5})t + 1.823\times10^{-2}$$

The cubic fits for the frequencies in cycles per second are:

$$f_1(t) = -(2.859\times10^{-7})t^3 + (4.286\times10^{-4})t^2 - (2.628\times10^{-1})t + 2.013\times10^{4}$$

Appendix B. Averaged Cramér–Rao bound

In this Appendix, the CRB is reviewed for an unbiased estimator from statistical theory, and then resolved to the line spectra problem [20]. For a parametric model $\boldsymbol{a} \in \mathbb{R}$, with data vector $\boldsymbol{x}$ and parameter vector $\boldsymbol{\theta}$:

Theorem 1. *If $\boldsymbol{a}^{\mathrm{T}}\hat{\boldsymbol{\theta}}(\boldsymbol{x})$ is an unbiased estimator of $\boldsymbol{a}^{\mathrm{T}}\boldsymbol{\theta}$, than the variance of the estimator given by:*

$$\mathbb{E}_{\boldsymbol{x}|\theta}\left\{\left(\boldsymbol{a}^{\mathrm{T}}\hat{\boldsymbol{\theta}}(\boldsymbol{x}) - \boldsymbol{a}^{\mathrm{T}}\boldsymbol{\theta}\right)^2\right\} \tag{B.1}$$

has an asymptotic lower bound defined by:

$$\boldsymbol{a}^{\mathrm{T}}\boldsymbol{F}^{-1}(\boldsymbol{\theta})\boldsymbol{a} \tag{B.2}$$

where the (i,j)th component of the Fisher Information Matrix [65] $\boldsymbol{F}$ is:

$$F_{i,j}(\boldsymbol{\theta}) = \mathbb{E}_{\boldsymbol{x}|\theta}\left\{\frac{\partial \ln p(\boldsymbol{x}|\boldsymbol{\theta})}{\partial\theta_i}\frac{\partial \ln p(\boldsymbol{x}|\theta)}{\partial\theta_j}\right\} \tag{B.3}$$

Thus in the case of parametric estimators for line spectra in AWGN, Eq. (B.3) is simply:

$$F_{i,j}(\boldsymbol{\theta}) = \frac{2}{\sigma^2}\Re\left\{\left(\frac{\partial \boldsymbol{s}(\boldsymbol{\theta})}{\partial \theta_i}\right)^{\mathrm{H}} \frac{\partial \boldsymbol{s}(\boldsymbol{\theta})}{\partial \theta_j}\right\} \tag{B.4}$$

where $\boldsymbol{s}(\boldsymbol{\theta}) = \boldsymbol{A}(\boldsymbol{\omega})\boldsymbol{\beta}$ is the (parametric) signal (model). The corresponding parameter vector used for this measurement is:

$$\boldsymbol{\theta} = \left[\omega_1 \dots \omega_P, \beta_1 \dots \beta_P, \sigma^2\right] \tag{B.5}$$

$\boldsymbol{F}(\boldsymbol{\theta})$ is utilized for the signal model Eq. (1) and the diagonal elements of $\boldsymbol{F}^{-1}(\boldsymbol{\theta})$ give the CRB values corresponding to complex amplitudes $\{\beta_i : i = 1, \dots, P\}$ and frequencies $\{\omega_i : i = 1, \dots, P\}$. The averaged CRB metric used for frequency metrics in this study is:

$$\mathrm{CRB}_{\boldsymbol{\omega}} = \frac{1}{P}\sum_{i=1}^{P} \mathrm{diag}\{\boldsymbol{F}^{-1}(\theta_i)\} \tag{B.6}$$

References

[1] H. Mayer, Recent developments in ultrasonic fatigue: recent developments in ultrasonic fatigue, Fatigue Fract. Eng. Mater. Struct. 39 (1) (2016) 3–29, http://dx.doi.org/10.1111/ffe.12365.

[2] V. Jacquemain, N. Ranc, C. Cheuleu, V. Michel, V. Favier, O. Castelnau, D. Vinci, D. Thiaudiere, C. Mocuta, Estimation of stress in specimens loaded with ultrasonic fatigue machines, Int. J. Fatigue 153 (2021) 106474, http://dx.doi.org/10.1016/j.ijfatigue.2021.106474.

[3] A. Kumar, C.J. Torbet, J.W. Jones, T.M. Pollock, Nonlinear ultrasonics for in situ damage detection during high frequency fatigue, J. Appl. Phys. 106 (2) (2009) 024904, http://dx.doi.org/10.1063/1.3169520.

[4] A. Kumar, C.J. Torbet, T.M. Pollock, J. Wayne Jones, In situ characterization of fatigue damage evolution in a cast Al alloy via nonlinear ultrasonic measurements, Acta Mater. 58 (6) (2010) 2143–2154, http://dx.doi.org/10.1016/j.actamat.2009.11.055.

[5] A. Kumar, R.R. Adharapurapu, J.W. Jones, T.M. Pollock, In situ damage assessment in a cast magnesium alloy during very high cycle fatigue, Scr. Mater. 64 (1) (2011) 65–68, http://dx.doi.org/10.1016/j.scriptamat.2010.09.008.

[6] H. Mayer, M. Fitzka, R. Schuller, Constant and variable amplitude ultrasonic fatigue of 2024-T351 aluminium alloy at different load ratios, Ultrasonics 53 (8) (2013) 1425–1432, http://dx.doi.org/10.1016/j.ultras.2013.02.012.

[7] M. Fitzka, H. Mayer, R. Schuller, S.E. Stanzl-Tschegg, T. Przeorski, P. Krug, Variable amplitude loading of spray-formed hypereutectic aluminium silicon alloy DISPAL® S232 in the VHCF regime, Fatigue Fract. Eng. Mater. Struct. 37 (9) (2014) 945–957, http://dx.doi.org/10.1111/ffe.12178.

[8] W. Li, C. Engler-Pinto Jr., H. Cui, W. Wen, X. Su, In situ characterization of humidity effect on the fatigue damage evolution of a cast aluminium alloy, Fatigue Fract. Eng. Mater. Struct. 39 (10) (2016) 1263–1271, http://dx.doi.org/10.1111/ffe.12441.

[9] M.V. Bannikov, V.A. Oborin, S.V. Uvarov, O.B. Naimark, Investigation of damage accumulation in a prestrained aluminum-magnesium alloy under gigacycle fatigue, AIP Conf. Proc. 2053 (1) (2018) 030003, http://dx.doi.org/10.1063/1.5084364.

[10] A. Messager, A. Junet, T. Palin-Luc, J.-Y. Buffiere, N. Saintier, N. Ranc, M. El May, Y. Gaillard, A. King, A. Bonnin, Y. Nadot, In situ synchrotron ultrasonic fatigue testing device for 3D characterisation of internal crack initiation and growth, Fatigue Fract. Eng. Mater. Struct. 43 (3) (2020) 558–567, http://dx.doi.org/10.1111/ffe.13140.

[11] M. Fitzka, U. Karr, M. Granzner, T.s. Melichar, M. Rödhammer, A. Strauss, H. Mayer, Ultrasonic fatigue testing of concrete, Ultrasonics 116 (2021) 106521, http://dx.doi.org/10.1016/j.ultras.2021.106521.

[12] S.L. Kiser, G. Mikhail, N. Ranc, M. Rébillat, Harmonic balance framework for ultrasonic fatigue vibration, in: Proceedings of the 2021 Eighth International Conference on Very High Cycle Fatigue, Hokkaido University, Sapporo, Hokkaido, Japan, 2021, pp. 215–220.

[13] K.H. Matlack, J.-Y. Kim, L.J. Jacobs, J. Qu, Review of second harmonic generation measurement techniques for material state determination in metals, J. Nondestruct. Eval. 34 (1) (2015) 273, http://dx.doi.org/10.1007/s10921-014-0273-5.

[14] S. Maier, J.-Y. Kim, M. Forstenhäusler, J.J. Wall, L.J. Jacobs, Noncontact nonlinear resonance ultrasound spectroscopy (NRUS) for small metallic specimens, NDT E Int. 98 (2018) 37–44, http://dx.doi.org/10.1016/j.ndteint.2018.04.003.

[15] Z. Gao, C. Cecati, S.X. Ding, A survey of fault diagnosis and fault-tolerant techniques—Part I: fault diagnosis with model-based and signal-based approaches, IEEE Trans. Ind. Electron. 62 (6) (2015) 3757–3767, http://dx.doi.org/10.1109/tie.2015.2417501.

[16] Z. Gao, C. Cecati, S.X. Ding, A survey of fault diagnosis and fault-tolerant techniques part II: fault diagnosis with knowledge-based and hybrid/Active Approaches, IEEE Trans. Ind. Electron. (2015) 1, http://dx.doi.org/10.1109/tie.2015.2419013.

[17] W. Hess, Pitch Determination of Speech Signals, in: Springer Series in Information Sciences, vol. 3, Springer Berlin Heidelberg, Berlin, Heidelberg, 1983, http://dx.doi.org/10.1007/978-3-642-81926-1.

[18] B. Everitt, A. Skrondal, The Cambridge Dictionary of Statistics, fourth ed., Cambridge University Press, Cambridge, UK ; New York, 2010, p. 468.

[19] G. Izacard, S. Mohan, C. Fernandez-Granda, Data-driven estimation of sinusoid frequencies, 2021, arXiv:1906.00823.

[20] P. Stoica, R.L. Moses, Spectral Analysis of Signals, Pearson/Prentice Hall, Upper Saddle River, N.J, 2005, p. 452.

[21] J.O. Smith, Spectral Audio Signal Processing, Stanford University and Center for Computer Research in Music and Acoustics, W3K, Stanford, Calif., 2011.

[22] J. Capon, High-resolution frequency-wavenumber spectrum analysis, Proc. IEEE 57 (8) (1969) 1408–1418, http://dx.doi.org/10.1109/proc.1969.7278.

[23] P. Stoica, Hongbin Li, Jian Li, A new derivation of the APES filter, IEEE Signal Process. Lett. 6 (8) (1999) 205–206, http://dx.doi.org/10.1109/97.774866.

[24] E.J. Candès, C. Fernandez-Granda, Towards a mathematical theory of super-resolution, Comm. Pure Appl. Math. 67 (6) (2014) 906–956, http://dx.doi.org/10.1002/cpa.21455.

[25] A. Serbes, K. Qaraqe, A fast method for estimating frequencies of multiple sinusoidals, IEEE Signal Process. Lett. 27 (2020) 386–390, http://dx.doi.org/10.1109/lsp.2020.2970837.

[26] J. Selva, ML estimation and detection of multiple frequencies through periodogram estimate refinement, IEEE Signal Process. Lett. 24 (3) (2017) 249–253, http://dx.doi.org/10.1109/lsp.2016.2645283.

[27] P. Stoica, Hongbin Li, Jian Li, Amplitude estimation of sinusoidal signals: Survey, new results, and an application, IEEE Trans. Signal Process. 48 (2) (2000) 338–352, http://dx.doi.org/10.1109/78.823962.

[28] G.R. de Prony, Essai expérimental et analytique: Sur les lois de la dilatabilité de fluides elastiques et sur celles de la force expansive de la vapeur de l'eau et de la vapeur de l'alkool, à différentes temperatures, J. l'Ecole Polytech. 1 (1795) 24–76.

[29] R. Schmidt, Multiple emitter location and signal parameter estimation, IEEE Trans. Antennas and Propagation 34 (3) (1986) 276–280, http://dx.doi.org/10.1109/tap.1986.1143830.

[30] R. Roy, T. Kailath, ESPRIT-estimation of signal parameters via rotational invariance techniques, IEEE Trans. Acoust. Speech Signal Process. 37 (7) (1989) 984–995, http://dx.doi.org/10.1109/29.32276.

[31] M. Bastiaans, On the sliding-window representation in digital signal processing, IEEE Trans. Acoust. Speech Signal Process. 33 (4) (1985) 868–873, http://dx.doi.org/10.1109/tassp.1985.1164653.

[32] Y. Bresler, A. Macovski, Exact maximum likelihood parameter estimation of superimposed exponential signals in noise, IEEE Trans. Acoust. Speech Signal Process. 34 (5) (1986) 1081–1089, http://dx.doi.org/10.1109/tassp.1986.1164949.

[33] E.J. Hannan, Non-linear time series regression, J. Appl. Probab. 8 (4) (1971) 767–780, http://dx.doi.org/10.2307/3212240.

[34] K.R. Parthasarathy, On the estimation of the spectrum of a stationary stochastic process, Ann. Math. Stat. 31 (3) (1960) 568–573, http://dx.doi.org/10.1214/aoms/1177705784.

[35] M. Feder, E. Weinstein, Parameter estimation of superimposed signals using the EM algorithm, IEEE Trans. Acoust. Speech Signal Process. 36 (4) (1988) 477–489, http://dx.doi.org/10.1109/29.1552.

[36] Jian Li, P. Stoica, Efficient mixed-spectrum estimation with applications to target feature extraction, IEEE Trans. Signal Process. 44 (2) (1996) 281–295, http://dx.doi.org/10.1109/78.485924.

[37] B. Ottersten, M. Viberg, T. Kailath, Analysis of subspace fitting and ML techniques for parameter estimation from sensor array data, IEEE Trans. Signal Process. 40 (3) (1992) 590–600, http://dx.doi.org/10.1109/78.120802.

[38] D. Donoho, Compressed sensing, IEEE Trans. Inform. Theory 52 (4) (2006) 1289–1306, http://dx.doi.org/10.1109/tit.2006.871582.

[39] S. Chen, D. Donoho, Application of basis pursuit in spectrum estimation, in: Proceedings of the 1998 IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP '98 (Cat. No.98CH36181), Vol. 3, IEEE, Seattle, WA, USA, 1998, pp. 1865–1868, http://dx.doi.org/10.1109/icassp.1998.681827.

[40] S. Wright, R. Nowak, M. Figueiredo, Sparse reconstruction by separable approximation, IEEE Trans. Signal Process. 57 (7) (2009) 2479–2493, http://dx.doi.org/10.1109/tsp.2009.2016892.

[41] D.H. Chae, P. Sadeghi, R.A. Kennedy, Effects of basis-mismatch in compressive sampling of continuous sinusoidal signals, in: 2010 2nd International Conference on Future Computer and Communication, IEEE, Wuhan, China, 2010, pp. V2–739–V2–743, http://dx.doi.org/10.1109/icfcc.2010.5497605.

[42] B. Mamandipoor, D. Ramasamy, U. Madhow, Newtonized orthogonal matching pursuit: frequency estimation over the continuum, IEEE Trans. Signal Process. 64 (19) (2016) 5066–5081, http://dx.doi.org/10.1109/tsp.2016.2580523.

[43] M.F. Duarte, R.G. Baraniuk, Spectral compressive sensing, Appl. Comput. Harmon. Anal. 35 (1) (2013) 111–129, http://dx.doi.org/10.1016/j.acha.2012.08.003.

[44] B.N. Bhaskar, G. Tang, B. Recht, Atomic norm denoising with applications to line spectral estimation, IEEE Trans. Signal Process. 61 (23) (2013) 5987–5999, http://dx.doi.org/10.1109/tsp.2013.2273443.

[45] Z. Yang, L. Xie, On gridless sparse methods for line spectral estimation from complete and incomplete data, IEEE Trans. Signal Process. 63 (12) (2015) 3139–3153, http://dx.doi.org/10.1109/tsp.2015.2420541.

[46] T.L. Hansen, M.A. Badiu, B.H. Fleury, B.D. Rao, A sparse Bayesian learning algorithm with dictionary parameter estimation, in: 2014 IEEE 8th Sensor Array and Multichannel Signal Processing Workshop (SAM), IEEE, A Coruna, Spain, 2014, pp. 385–388, http://dx.doi.org/10.1109/sam.2014.6882422.

[47] T.L. Hansen, B.H. Fleury, B.D. Rao, Superfast line spectral estimation, IEEE Trans. Signal Process. 66 (10) (2018) 2511–2526, http://dx.doi.org/10.1109/tsp.2018.2807417.

[48] L. Han, S.K. Biswas, Neural networks for sinusoidal frequency estimation, J. Franklin Inst. B 334 (1) (1997) 1–18, http://dx.doi.org/10.1016/s0016-0032(96)00079-8.

[49] G. Izacard, B. Bernstein, C. Fernandez-Granda, A learning-based framework for line-spectra super-resolution, in: ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), IEEE, Brighton, United Kingdom, 2019, pp. 3632–3636, http://dx.doi.org/10.1109/icassp.2019.8682882.

[50] M. Haardt, J. Nossek, Unitary ESPRIT: How to obtain increased estimation accuracy with a reduced computational burden, IEEE Trans. Signal Process. 43 (5) (1995) 1232–1242, http://dx.doi.org/10.1109/78.382406.

[51] D. Potts, M. Tasche, Fast ESPRIT algorithms based on partial singular value decompositions, Appl. Numer. Math. 88 (2015) 31–45, http://dx.doi.org/10.1016/j.apnum.2014.10.003.

[52] A. Reilly, G. Frazer, B. Boashash, Analytic signal generation-tips and traps, IEEE Trans. Signal Process. 42 (11) (1994) 3241–3245, http://dx.doi.org/10.1109/78.330385.

[53] L. Marple, Computing the discrete-time "analytic" signal via FFT, IEEE Trans. Signal Process. 47 (9) (1999) 2600–2603, http://dx.doi.org/10.1109/78.782222.

[54] M. Jansson, P. Stoica, Forward-only and forward-backward sample covariances – a comparative study, Signal Process. 77 (3) (1999) 235–245, http://dx.doi.org/10.1016/s0165-1684(99)00037-7.

[55] D. Rife, R. Boorstyn, Single tone parameter estimation from discrete-time observations, IEEE Trans. Inform. Theory 20 (5) (1974) 591–598, http://dx.doi.org/10.1109/tit.1974.1055282.

[56] C. Candan, A method for fine resolution frequency estimation from three DFT samples, IEEE Signal Process. Lett. 18 (6) (2011) 351–354, http://dx.doi.org/10.1109/lsp.2011.2136378.

[57] E. Aboutanios, B. Mulgrew, Iterative frequency estimation by interpolation on Fourier coefficients, IEEE Trans. Signal Process. 53 (4) (2005) 1237–1242, http://dx.doi.org/10.1109/tsp.2005.843719.

[58] A. Serbes, Fast and efficient sinusoidal frequency estimation by using the DFT coefficients, IEEE Trans. Commun. 67 (3) (2019) 2333–2342, http://dx.doi.org/10.1109/tcomm.2018.2886355.

[59] A.R. Barron, A. Cohen, W. Dahmen, R.A. DeVore, Approximation and learning by greedy algorithms, Ann. Statist. 36 (1) (2008) 64–94, http://dx.doi.org/10.1214/009053607000000631.

[60] https://github.com/slkiser/lineSpectraVibration.

[61] N. Golyandina, On the choice of parameters in singular spectrum analysis and related subspace-based methods, 2011, arXiv:1005.4374.

[62] S. Guelton, P. Brunet, M. Amini, A. Merlini, X. Corbillon, A. Raynaud, Pythran: Enabling static optimization of scientific python programs, Comput. Sci. Discov. 8 (1) (2015) 014001, http://dx.doi.org/10.1088/1749-4680/8/1/014001.

[63] Y. Jiang, H. Li, M. Rangaswamy, Deep learning denoising based line spectral estimation, IEEE Signal Process. Lett. 26 (11) (2019) 1573–1577, http://dx.doi.org/10.1109/lsp.2019.2939049.

[64] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, 2015, arXiv:1512.03385.

[65] R.A. Fisher, On the mathematical foundations of theoretical statistics, Philos. Trans. R. Soc. Lond. Ser. A 222 (594–604) (1922) 309–368, http://dx.doi.org/10.1098/rsta.1922.0009.