Computational de-noising based on deep learning for phase data in digital holographic interferometry

Silvio Montresor, (iD) Marie Tahon, (iD) Antoine Laurent, et al.

COLLECTIONS

Paper published as part of the special topic on Photonics and AI

openaccess

View Online · Export Citation · CrossMark

ARTICLES YOU MAY BE INTERESTED IN

Computational de-noising based on deep learning for phase data in digital holographic interferometry

View Online   Export Citation   CrossMark

Silvio Montresor,[1] Marie Tahon,[2] (iD) Antoine Laurent,[2] (iD) and Pascal Picart[1,3,a] (iD)

AFFILIATIONS

[1] Laboratoire d'Acoustique de l'Université du Mans, LAUM CNRS 6613, Le Mans Université, Avenue Olivier Messiaen, 72085 Le Mans Cedex 09, France
[2] Laboratoire d'Informatique de l'Université du Mans, LIUM, Le Mans Université, Avenue Olivier Messiaen, 72085 Le Mans Cedex 09, France
[3] ENSIM, Ecole Nationale Supérieure d'Ingénieurs du Mans, rue Aristote, 72085 Le Mans, Cedex 09, France

Note: This article is part of the APL Photonics Special Topic on Photonics and AI.
[a] Author to whom correspondence should be addressed: pascal.picart@univ-lemans.fr

ABSTRACT

This paper presents a deep-learning-based algorithm dedicated to the processing of speckle noise in phase measurements in digital holographic interferometry. The deep learning architecture is trained with phase fringe patterns including faithful speckle noise, having non-Gaussian statistics and non-stationary property, and exhibiting spatial correlation length. The performances of the speckle de-noiser are estimated with metrics, and the proposed approach exhibits state-of-the-art results. In order to train the network to de-noise phase fringe patterns, a database is constituted with a set of noise-free and speckled phase data. The algorithm is applied to de-noising experimental data from wide-field digital holographic vibrometry. Comparison with the state-of-the-art algorithm confirms the achieved performance.

I. INTRODUCTION

Digital holography and digital holographic microscopy[1–5] are methods for recording and reconstructing three-dimensional images of objects. Digital holographic interferometry is a very efficient technique for the measurement of deformation fields of object surfaces and for the measurement of surface shapes and contours.[5] As a general rule, the object field is numerically reconstructed using a propagation operator such as the discrete Fresnel transform or the angular spectrum[3] and, from the complex-valued data, the phase of the optical object field can be extracted. By the way, holographic phase imaging measures the optical path length associated with the specimens under interest and translates these data into relevant information. The measured field of interest is wrapped in modulo 2π phase maps, known as the phase fringe pattern.[2,5] In any measurement from digital holographic interferometry, speckle phase noise is induced and de-noising is required.[5] The specificity of the speckle noise in holographic phase data is non-Gaussian, highly non-stationary, has a correlation length, and is amplitude-dependent (noise depends on the fringe density).

Since 10 years, deep learning has emerged as a very efficient approach in signal and image processing. At the heart of this new tool are the convolutional neural networks (CNN). The CNN integrate several fundamental advances of the last few decades: wavelet and multiresolution analysis, shrinkage and thresholding algorithms, sparse representations, block matching, and dictionary learning.[6–9] In recent years, several applications of deep learning in optics have emerged, when processing noise, such as optical coherence tomography,[10,11] hyperspectral imaging,[12] and using multiscale decompositions.[13,14] Deep learning based algorithms applied to the noise reduction in synthetic aperture radar imaging were also proposed.[15,16] The problem of the speckle reduction was approached

with a deep learning based solution when taking into account scattering effects.[17] In the literature, a deep CNN was used as a de-noising algorithm for the speckle noise reduction in the amplitude image[18] and fringe patterns.[19,20] In the latter references, the noise training is carried out on simulated fringes with added Gaussian noise.

However, this does not address the real case of speckle noise. Indeed, the speckle noise is non-Gaussian and non-stationary, and in addition, speckle grains have a correlation length that cannot be reproduced by simply adding Gaussian fluctuations. In this paper, we demonstrate that efficient speckle de-noising can be at the state-of-the-art for phase data in digital holographic interferometry if the network is correctly trained by considering realistic speckle statistics. To the best of our knowledge, this paper proposes the first CNN architecture trained with faithful speckle noise conditions, having non-Gaussian statistics and non-stationary property and exhibiting spatial correlation length, in order to process noisy phase data from holographic interferometry. The obtained results demonstrate that the proposed scheme is state-of-the-art.

This paper is organized as follows: Sec. II describes the basic fundamentals of digital holographic interferometry, Sec. III describes the proposed approach, Sec. IV provides the results, and Sec. V proposes application to experimental data from wide-field digital holographic vibrometry. Section VI draws the conclusions of this paper.

II. DIGITAL HOLOGRAPHIC INTERFEROMETRY

Basically, digital holograms are obtained by recording, with an image sensor organized as a matrix of pixels, the coherent mixing of the diffracted optical wave from the object surface, and a known reference wave. Since the complex-valued optical field is recorded in any digital holograms, the optical phase, and then the optical path difference, can be retrieved and may yield the measurement of any kind of perturbation at the surface of the object (vibration, heating, pneumatic changes, mechanical loads, etc.). Let O be the wave front from the illuminated object and R be the wave front from the reference wave, then the digital hologram can be expressed by the following equation:[21]

$$H = |R|^2 + |O|^2 + R^*O + RO^*. \qquad (1)$$

The object surface producing the object wave front is generally at distance d_0 from the recording sensor, which is used without any imaging lens (lens-less Fresnel configuration). Then, the Fresnel approximation provides the object field at the sensor plane[21] ($i = \sqrt{-1}$),

$$O(x,y,d_0) = -\frac{i}{\lambda d_0}\exp\left(\frac{2i\pi d_0}{\lambda}\right)\exp\left(\frac{i\pi d_0}{\lambda}(x^2+y^2)\right)$$
$$\times \iint A(X,Y)\exp\left(\frac{i\pi}{\lambda d_0}(X^2+Y^2)\right)$$
$$\times \exp\left(\left(-\frac{2i\pi}{\lambda d_0}\right)(xX+yY)\right)dXdY. \qquad (2)$$

In Eq. (2), the object wave front at the object plane is $A(X, Y)$ = $A_0(X, Y)\exp[i\psi_0(X, Y)]$, λ is the wavelength of light, A_0 is related to the object reflectance, and ψ_0 is the optical phase related to the object surface profile and roughness. It follows that the wave front O is speckled due to roughness. From digitally recorded holograms, the discrete Fresnel transform permits us to reconstruct the object field at any distance d_r from the recording plane. The numerically reconstructed complex-valued image is obtained from the following equation:

$$A_r = h_F \times FFT[H \times h_F]. \qquad (3)$$

In Eq. (3), FFT means two-dimensional fast Fourier transform and h_F is the kernel defined by the following equation:

$$h_F(x,y) = \frac{1}{\sqrt{\lambda d_r}}\exp\left(i\pi\frac{d_r}{\lambda} - i\frac{\pi}{4}\right)\exp\left[\frac{i\pi}{\lambda d_r}(x^2+y^2)\right]. \qquad (4)$$

From the complex-valued data computed from Eq. (3), the amplitude and phase of the diffracted field A_r can be evaluated. When considering two different states at the surface of the object and recording two holograms, the phase change can be computed by calculating the difference between the two phases from the two holograms. This phase change, equivalent to a Doppler phase, is related to the displacement field $\mathbf{U}$ at the object surface and is a result of the object solicitation. The relation $\Delta\varphi = 2\pi\mathbf{U}\cdot(\mathbf{K_e} - \mathbf{K_o})/\lambda$ holds,[5] where $\mathbf{K_e}$ is the normalized illumination vector from the light source to the object and $\mathbf{K_o}$ is the observation vector (also normalized) from the object to the sensor, both defined in a set of reference axes ($\mathbf{i}, \mathbf{j}$, and $\mathbf{k}$) attached to the object surface, with $\mathbf{k}$ being perpendicular to the surface. Generally, the Doppler phase is obtained by modulo 2π and requires phase unwrapping to yield physically sound data.[22] Due to the speckle nature of the recorded holograms, the Doppler phase includes speckle decorrelation noise and noise reduction techniques have to be involved. The specificity of the speckle noise in the Doppler phase is that it is non-Gaussian and non-stationary.[23] In addition, it exhibits a correlation length related to the speckle grain size and is amplitude-dependent. The noise amount in the phase image is then related to the local fringe density, and the probability density function of the phase noise is not shaped as a Gaussian curve (in particular, it is bounded by $[-\pi, +\pi]$). In order to yield quantitative data with reduced noise, the raw modulo 2π phase has to be low-pass filtered. Figure 1 shows the illustrations of digital holographic interferometry: panel (a) depicts the basic holographic configuration in the Fresnel arrangement, panel (b) shows a raw phase difference obtained in the case of stroboscopic holography applied to a loud speaker,[24] panel (c) exhibits the noise from panel (b), and panel (d) shows the probability density function of the phase noise. Figures 1(c) and 1(d) show that the speckle noise in phase data from digital holographic interferometry is non-Gaussian and non-stationary. In Fig. 1(d), the red line is the theoretical probability density that fits the experimental one.[23] In this example, the standard deviation of noise is found to be almost $\sigma = 0.86$ rad and the correlation factor of the speckle phases is around $|\mu| = 0.81$.

In the literature, there does exist a large diversity of approaches to cope with noise in images.[23] The best filtering scheme for de-noising phase data from digital holographic metrology was found to be the two-dimensional windowed Fourier transform (WFT2F).[25] The main drawback of this algorithm is the computation time. For example, on a laptop personal computer (PC) Intel core i5 at 2.30 GHz with MATLAB, the processing of the image sized 1024 × 1024 pixels with the WFT2F requires 63 s, 230 s, and 1600 s
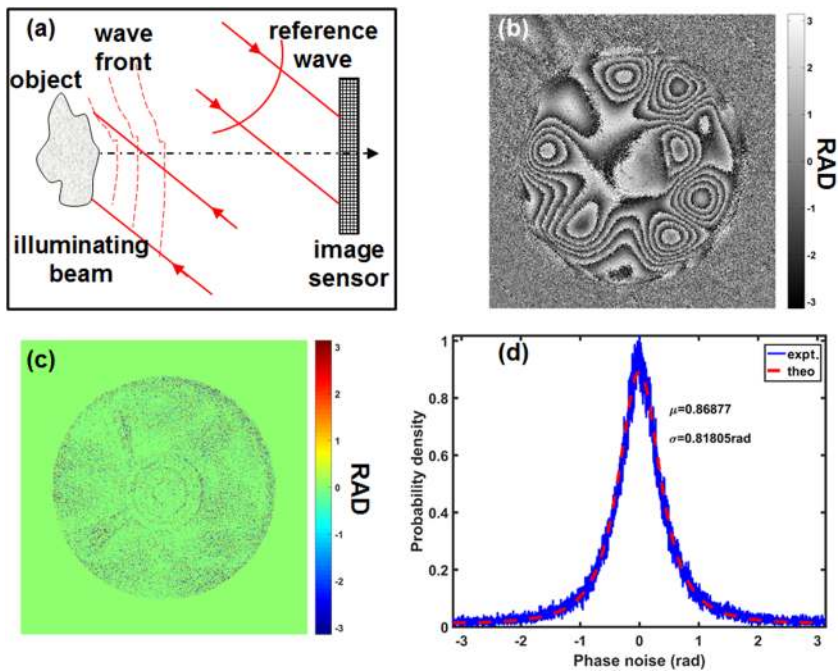
FIG. 1. Illustrations of digital holographic interferometry: (a) basic scheme of the holographic configuration in the Fresnel arrangement, (b) example of the raw phase difference obtained in the case of stroboscopic holography applied to a loud speaker,[24] (c) phase noise from (b), and (d) probability density function of the phase noise (red: theoretical, with $|\mu| = 0.81$ and blue: experimental).

for kernel sizes at 10, 20, and 40 pixels, respectively. Over the past 10 years, deep learning-based artificial intelligence has become a very effective tool for signal and image processing.[26] It follows that alternative approaches to classical filtering schemes, exhibiting almost the same performance, but the reduced computational time could be of very high interest. Section III describes the method we followed in this paper to de-noise phase fringe data, including high speckle noise with a deep learning approach.

III. METHOD

A. Noisy phase data

In order to get phase fringe patterns with a diversity of both signal-to-noise ratio (SNR) and fringe diversity, a realistic numerical simulation was carried out. The goal is to get phase data with fringe pattern diversity and corrupted with speckle noise, with the noise amount being correlated with the fringe density and faithful statistics. The arrangement to produce realistic simulations of speckled phase data is shown in Fig. 2. The simulation principle is described in Ref. 23, but we recall it here briefly. The simulator is based on a double-diffraction[27,28] system in which the inputs are complex valued data with a random phase. The random phase is obtained by considering that the surface related to the target is rough compared to the wavelength of light (visible range). For numerical propagation, the object plane is supposed to be illuminated by a uniform plane wave. A diaphragm with radius R_u is inserted in the back focal plane of the first lens (i.e., the Fourier plane of the system). The value of R_u is adjusted to control the speckle grain size in the phase data. In order to produce N_s pixels per speckle grain in the phase data, the diaphragm is adjusted to $R_u = 1/N_s p_x$ (p_x, pixel pitch in the image plane). The roughness in the input plane is numerically simulated

by considering a surface profile with roughness h having Gaussian statistics and Dirac-type autocorrelation function such that $2\pi h/\lambda$ is higher than 2π.[28] The surface deformation is simulated by using analytical models, such as Gaussian distribution, first, second, and third order polynomials, and Matlab functions, such as "peaks" or "membrane." The surface deformation is then added to the surface roughness. The numerical propagation of the wave front into the double diffraction system provides modulo 2π phase data corrupted with realistic speckle noise.

B. Network training

The network considered in this paper is the one proposed by Zhang et al.[7]. The network includes 59 layers and is based on the concept of the residual network (DnCNN). A noisy image is sent at the input of the network, and the output provides the noise. After
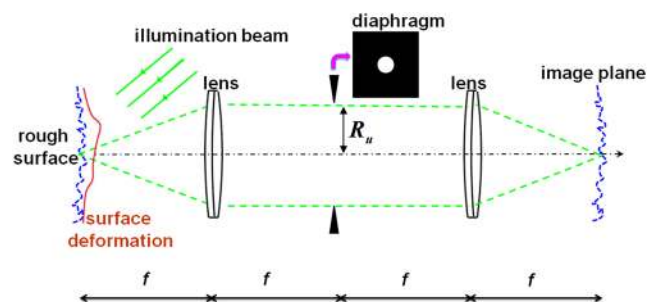

FIG. 2. Scheme to produce speckled phase data with the controlled speckle size (from Ref. 23).

getting the noise, it is necessary to subtract the output from the input image to get the de-noised image. Therefore, the network works as a noise estimator. This network trained on natural images and Gaussian noise is used as a pre-trained network in the present study. In order to train with phase decorrelation speckle noise, the network has been adapted so as to be trained with a set of 40 fringe images sized 1024 × 1024 pixels with a realistic speckle noise. The fringe images are sine and cosine computed from five initial phase patterns from the simulator of Sec. II. We also consider their transposed and phase shifted version with a π/4 phase shift. This leads to a grand total of 40 images. The noise level for the 40 images was set to 2 pixels per speckle grain in the simulator, which corresponds to realistic on-line digital holographic recording conditions (refer also Sec. V). Depending on the local fringe density in patterns, the signal-to-noise ratio of its cosine, in the overall database, varies from 7.31 dB to 11.45 dB. From the database, 384 patches per image were considered, leading to a grand total of 15 360 patches. These 384 patches, extracted from each image, were randomly selected for each of the 1920 epochs of the training. Furthermore, speckle noise realizations of all images were regenerated (see Sec. II) before each new training epoch and no data augmentation was performed. Mini-batch size and initial learning rate were set to 128 and 0.0006, respectively, as summarized in Table I. Therefore, a total of 230 400 iterations were used for the training. Figure 3 shows the illustrations of the selection of the phase fringe patterns and patches from the database. Figure 3(a) shows a selection of 20 cosine images of the noise-free phase data, sized 1024 × 1024 pixels, whereas Fig. 3(b) shows a random selection of 20 × 384 = 7680 patches from the grand total of 15 360 patches.

The cost function, defined as the root-mean-square error (RMSE), was minimized using the stochastic gradient descent. The total duration of the training was about 32 h with 8 core desktop PC including nvidia gtx1080 GPU. Before the training stage, the model was initialized with values corresponding to the original DnCNN model of Zhang,[7] which corresponds to de-noising of natural images with Gaussian noise. In order to validate the model, an evaluation was performed with a second database including 25 phase fringe patterns that have not been seen by the network in the training stage. This database has been used in author's previous works related to phase de-noising in digital holography.[23,29] The model was also
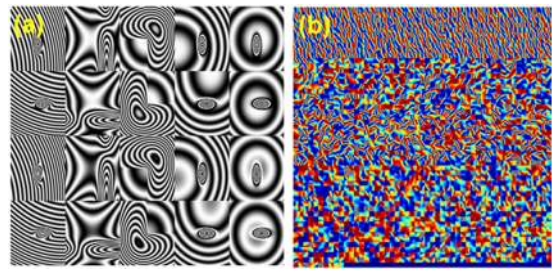
FIG. 3. (a) Set of 20 cosine images of the noise-free phase data, sized 1024 × 1024 pixels and (b) selection of 20 × 384 = 7680 patches from the grand total of 15 360 patches.

compared with two other ones: the original DnCNN model from Zhang,[7] which is a model trained on sets of natural images with Gaussian noise, and the same pre-trained model retrained on the same fringe pattern database that was used for training with realistic speckle noise, but with added Gaussian noise. The behavior of the cost function vs the iteration number is reported in Fig. 4. One can observe that it exhibits a larger variance than in the case of usual network training with Gaussian noise. In order to improve performances, we also introduced iterations in the de-noising process, as shown in the processing scheme reported in Fig. 5. Therefore, 1–5 iterations were taken into account for the three models. The results of the appraisals are provided in Sec. IV. Table I summarizes the parameters that were taken into account for the training of three different network models.

C. Metrics for quantitative appraisal

Quality metrics are required to analyze the performances of the deep learning approach to de-noise speckled phase data. The most adapted criterion for optical metrology is related to the error of the processed phase. The phase error is calculated by subtracting the original simulated noise-free phase fringe pattern from the de-noised one, and then, the standard deviation σ_φ is estimated. The peak-to-valley (PV) of the phase error is also of interest and is simply computed by calculating the difference between the maximum

TABLE I. Parameters of the training of the three network models: natural CNN from Zhang trained with natural images, CNN from Zhang trained with fringes and added Gaussian noise, and CNN from Zhang trained with fringes and realistic speckle noise.

	CNN with natural images[7]	CNN + fringe patterns + Gaussian noise	CNN + fringe patterns + speckle noise
Number of layers	59	59	59
Patch size	50 × 50	50 × 50	50 × 50
Patch number in images	128	384	384
Selection of patches	Random	Random	Random
Learning rate	0.1–0.0001	0.000 6	0.000 6
Epoch number	50	120	1920
Total number of iterations	8000	14 400	230 400
Noise level	>13 dB	11 dB	7.31–11.45 dB
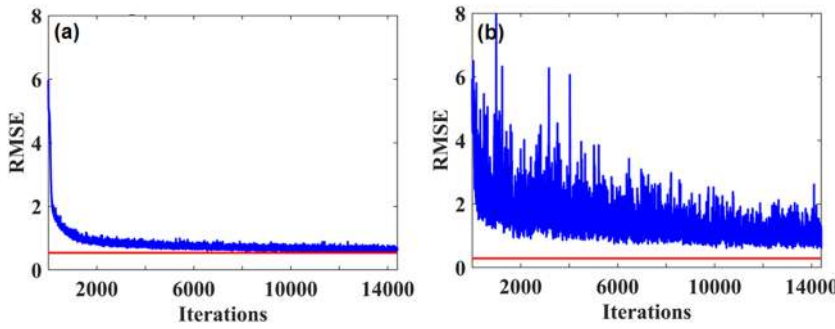
FIG. 4. Evolution of the RMSE cost function of the training process: (a) Gaussian model and (b) speckle model, the horizontal red line indicates the minimum value of the RMSE.

and minimum phase errors. However, such criteria have a significant drawback since they are based on the calculation of an average error, which does not account for distortions but may affect the structures in the processed phase data. It follows that the same amount of error can be observed between two restored phase data when one of them may exhibit a very different perceived quality. However, the "quality index," Q_{index}, criterion from Wang and Bovik[30] does account for the perceived quality of restored images. The value of Q_{index} is included in the range [−1, +1], when +1 is reached and when the two images are quite similar. In this paper, the quality index will be used to check the perceived quality. Another metric that can be considered is the peak signal-to-noise ratio (noted P_{SNR}), which is defined as the ratio between the maximum gray level in the image and the standard deviation between the initial and processed image. The P_{SNR} is given in the following equation for images digitized with *nbits*:

$$P_{SNR} = 10 \log 10 \left[\frac{(2^{nbits} - 1)^2}{\frac{1}{MN} \sum_{i,j} (s(i,j) - d(i,j))^2} \right]. \tag{5}$$
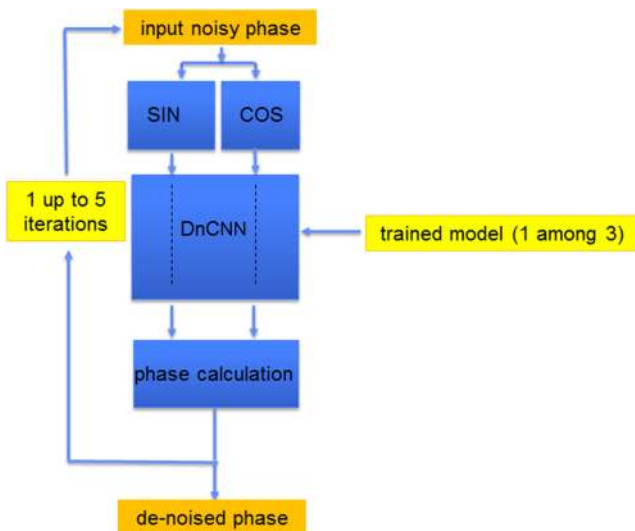
FIG. 5. Scheme of the de-noising process used for the evaluation of the three models. The phase is re-calculated by using an arctangent formula of the sine and cosine images.

In Eq. (5), $s(i, j)$ and $d(i, j)$ refer to the initial noise-free phase data and the enhanced one, respectively. The set of the index values (i, j) relates to the matrix coordinates having M rows and N columns. Finally, a last metric, called the method error,[31] yields the value of the phase error generated by the de-noising algorithm itself when its input is composed of noise-free images. This error can be interpreted as the lower bound of the phase error relative to the database used for the evaluation. It is then computed as the mean value of phase errors when the inputs of the de-noising algorithm are the noise-free images from the database.

D. Challenger algorithms

The deep learning approach is compared to a set of reference de-noising algorithms from the literature. In particular, the BM3D,[32] the WFT2F,[25] and the dual-tree wavelet transform (DTDWT)[33] are selected because of their established performances. The block-matching 3D (BM3D) filter relies both on local and non-local characteristics of images. As a general rule, BM3D is based on the concept of grouping and collaborative filtering. Grouping finds mutually similar image blocks and stacks them together in dedicated arrays. Collaborative filtering produces individual estimates of all grouped blocks by filtering them jointly. The BM3D algorithm is recognized as the state-of-the-art in natural image de-noising. The 2D windowed Fourier transform filter, WFT2F, is based on a local Fourier transform, which may take into account the non-stationary characteristics of the noise. The WFT2F algorithm was designed for phase filtering in speckle and holographic metrology;[25] it can be considered as the dedicated algorithm to process such data.[23] The dual-tree complex wavelet transform is an alternative approach to the classical discrete wavelet transform (DWT) and offers important enhancements in the context of image de-noising. As a main property, such transform is nearly shift invariant and directionally selective in the two and higher dimensions. This property is achieved with a redundancy factor of only 2d for d-dimensional signals, which is substantially lower than the undecimated DWT usually involved in de-noising applications in image processing. In addition, the DTDWT is non-separable but is based on a computationally efficient separable filter bank.

IV. RESULTS

As the first results, we present comparisons of the three selected model networks that correspond to the DnCNN model trained with natural images with Gaussian noise, with noise-free fringe patterns

with added Gaussian noise (an amount around 11 dB), or with realistic speckle noise in phase data from the simulator, respectively. These three models are called models "1," "2," and "3," respectively.

Validation is achieved with the database used in Ref. 23, which is constituted with 25 phase fringe patterns with realistic speckle decorrelation noise and cosine signal-to-noise ratio varying from 3 dB to 12 dB. As such fringe patterns have also been used for the training, we have inserted a rotation of 90° of the images of the validation database to avoid any similar patterns when de-noising. Therefore, it follows that there was no data augmentation in images before training.

Figure 6 shows the comparisons and demonstrates that the original DnCNN network without any specialized retraining (model "1") yields bad results for de-noising phase fringe patterns with realistic speckle noise, even when iterations are performed. The results for model "2" show an important improvement with a score of about 0.07 rad at the second iteration. The best performance is obtained with model "3" with a phase error at 0.031 rad at the 5th iteration, whereas the state-of-the-art WFT2F algorithm provides a phase error at 0.025 rad (dashed line in Fig. 6). The method error for models "1," "2," and "3" vs the iteration number is shown in Fig. 7. Similarly as in Fig. 6, the dashed line corresponds to the WFT2F algorithm.

Figure 7 shows that the method error (a minimum bound of the rms phase error) for model "1" has weak values for the two first iterations. The reason is that, practically, the network does not see any presence of noise in the images. As a consequence, the input image remains unchanged at the output of the network. For model "3," the method error is less than that of WFT2F until the second iteration and is at 0.008 rad. The method error for model "2" is worse than that for model "3," whatever the number of iterations. Therefore, for metrology purpose, model "2" should not be retained for any phase de-speckling processing. Note that increasing the number of iterations leads to an increase in the method error at an amount that can reach two times the method error from WFT2F, for five iterations. This means that too much iterations do contribute to degrade
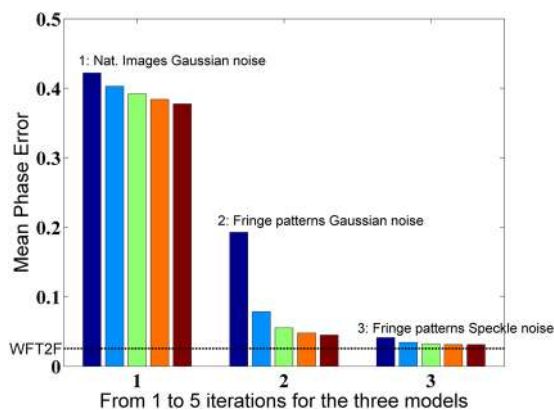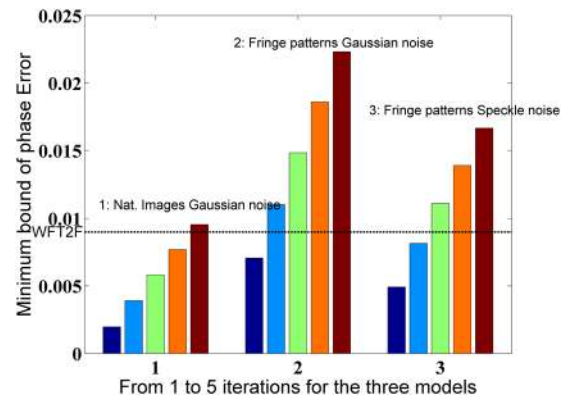
FIG. 7. Method error (a minimum bound of phase error vs the iteration number) for the three compared models. For each model, the color code is dark blue, blue, green, orange, and brown to 1, 2, 3, 4, and 5 iterations, respectively. The horizontal dotted line is the method error of the WFT2F algorithm at 0.009 rad.

the phase data. In the following, we present the results obtained for three iterations, which is a good compromise between the metrological performance (low method error) and the de-noising performance.

In Figs. 8 and 9, "DL" (deep learning approach) relates to the network training with realistic speckle noise (model "3"). The DL for 3 iterations is compared with the three best methods for speckled phase de-noising: WFT2F, BM3D, and DTDWT. Figure 8(a) shows the ranking of the rms phase error, Fig. 8(b) the ranking of the Q_{index}, and Fig. 8(c) the ranking of the PSNR for the four algorithms. Figure 8(d) shows the rms phase error for the 4 algorithms, Fig. 8(e) the Q_{index}, and Fig. 8(f) the PSNR vs the input SNR in the cosine image.

Figure 8 shows that DL is very close to the WFT2F for each of the three appraisal metrics. The DTDWT algorithm yields relatively good results although below that from DL. Note that the second place of DL is obtained with a phase error at 0.032 rad. In Fig. 8(d), DL and WFT2F have very similar behaviors for the SNR input varying from 6 dB to 12 dB. For a SNR from 3 dB to 6 dB, the difference between DL and WFT2F slightly increases. This can be explained by the fact that the DL algorithm was initially trained with high values of the input SNR.

In Fig. 9, the peak-to-valley error (PV) is provided for the four algorithms: DL, WFT2F, BM3D, and DTDWT. Figure 9 shows that the PV for DL is better than those obtained with WFT2F, BM3D, and DTDWT.

Note that in this paper, we considered 2 pixels per speckle grain. When increasing the number of pixels per speckle grain, that is, when increasing the speckle size, the SNR in the phase data decreases. Hence, when increasing the speckle grain size, the SNR decreases.[34] It follows that the DnCNN requires to be re-trained with the adapted set of data. However, the ranking observed in Figs. 8 and 9 is not changed. From Figs. 8 and 9, for such SNR values, the DL de-noiser is state-of-the-art.

Table II summarizes the computation times obtained with DL, WFT2F, BM3D, and DTDWT. Note that GPU computation for WFT2F, BM3D, and DTDWT was not available. Therefore, the
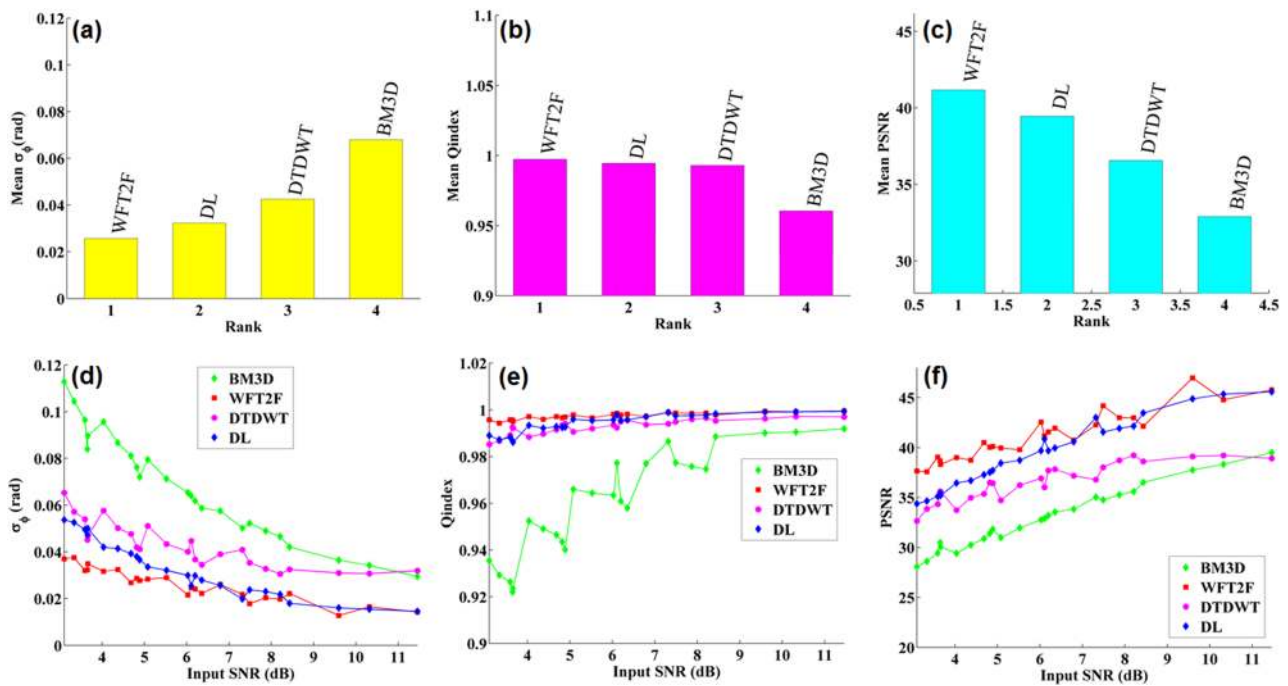

FIG. 6. Mean phase error (rad) vs the iteration number for the three compared models. For each model, the color code is dark blue, blue, green, orange, and brown for 1, 2, 3, 4, and 5 iterations, respectively. The horizontal dotted line is the mean phase error method from the WTF2F algorithm.

FIG. 8. Rankings and trends of the four algorithms vs metrics: (a) ranking of the rms phase error, (b) ranking of the Q_{index}, (c) ranking of the PSNR, (d) rms phase error for the five algorithms vs the input SNR, (e) Q_{index} vs the input SNR, and (f) PSNR vs the input SNR.

four algorithms were run on a laptop PC Intel core i5 at 2.30 GHz equipped with MATLAB, and the computation times were estimated with this equipment. Of course, one may assume that their transposition into GPU would give better computation time; however, the hierarchy would have to be checked.

It appears that the computation time is lower for DL compared to WFT2F.
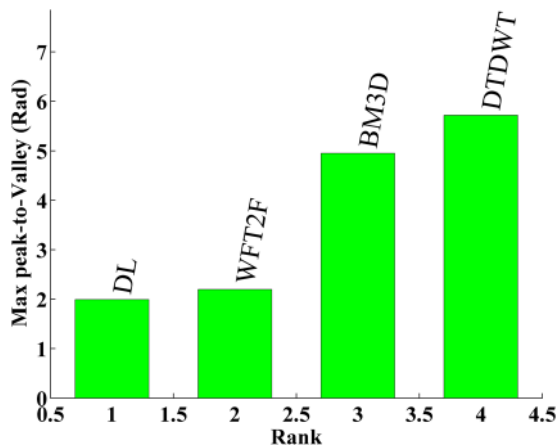

FIG. 9. Rankings of the four algorithms vs the peak-to-valley metric.

V. APPLICATION TO WIDE-FIELD HOLOGRAPHIC VIBROMETRY

Digital holographic vibrometry is a powerful approach to study the vibrations of structures when subject to stationary, transient, or non-stationary excitations.[35] The experimental setup is based on high-speed recording of digital holograms in the Fresnel configuration, as depicted in Fig. 10. The light is emitted from a continuous Diode Pumper Solid State (DPSS) laser at λ = 532 nm with a maximum power at 6 W. The sensor is a high-speed camera from Photron, with pixel pitch at p_x = 20 μm and maximum spatial resolution including 1024 × 1024 pixels at a low frame rate. At 100 kHz, the spatial resolution is reduced to 264 × 384 pixels. The exposure time is set to 1 μs. The negative zoom is adjusted to capture holograms from a rectangular area sized 30 × 15 cm^2 (~450 cm^2) when fulfilling the Shannon requirements for the hologram sampling. The structure is an aluminum plate excited with
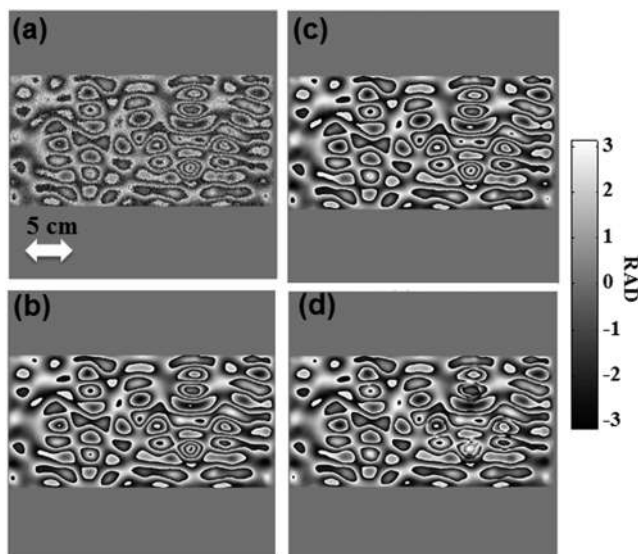
TABLE II. Iteration number and computation times obtained with DL, WFT2F, BM3D, and DTDWT.

Algorithm	Iterations	Computation time (s)
DL	3	210
WFT2F	1	278
BM3D	3	130
DTDWT	3	11
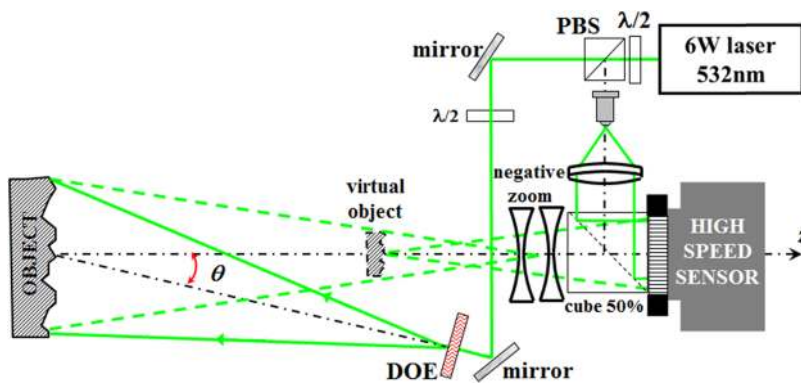
FIG. 10. (a) Experimental setup for wide-field holographic vibrometry (PBS: polarizing beam splitter and DOE: diffractive optical element, $\lambda/2$ half-wave plate).


FIG. 11. (a) Experimental noisy Doppler phase for vibration at 17 512 Hz, (b) Doppler phase processed with WFT2F, (c) Doppler phase processed with DL, and (d) Doppler phase processed with DTDWT. The color bar indicates the range of the color map in the modulo 2π phase images.

a mechanical shaker at a frequency of 17 512 Hz. The beam to illuminate the structure is spatially expanded by using a dedicated diffractive optical element (DOE) to yield a rectangular light shape.

The digital holograms are numerically reconstructed with 2048 × 2048 data points, and the phases are extracted. The phase differences between two consecutive instants are calculated to yield the Doppler phase.[35] Since the Doppler phases are speckled due to speckle decorrelation between the two instants, de-noising is required. Considering Sec. IV, Doppler phases were processed by both DL, WFT2F, and DTDWT algorithms. Figure 11(a) shows the modulo 2π noisy Doppler phase from the digital holographic process. The useful part of the phase map includes 1006 × 501 data points. Figure 11(b) shows the phase map processed from the WFT2F algorithms, and Fig. 11(c) exhibits the phase map processed from the DL approach. In Fig. 11(d), the phase map processed from the DTDWT algorithm is shown. As can be seen, the DTDWT algorithm generates phase dislocations in Fig. 11(d), whereas this is not the case for the other algorithms. It appears that the DTDWT is not adapted to such phase maps. The computation time for the DL approach is around 230 s, whereas it is around 300 s for the WFT2F.

In order to appraise the difference between the WFT2F and the DL processing, the phase difference between the results from Figs. 11(b) and 11(c) was calculated and is shown in Fig. 12(a). In Fig. 12(a), the difference is mainly around 0 rad and may reach 1.5 rad for few localized data points. The reason for that is not clearly explained. Figure 12(b) shows the histogram of the phase difference. The histogram looks like the Gaussian curve with zero mean, and the standard deviation is found at 0.107 rad. We can conclude that the phase difference is acceptable. Of course, since the ground truth is not available with real experimental data, one cannot get the error map for both algorithms. However, considering Fig. 9, the PV of the DL is better than that of the WFT2F.
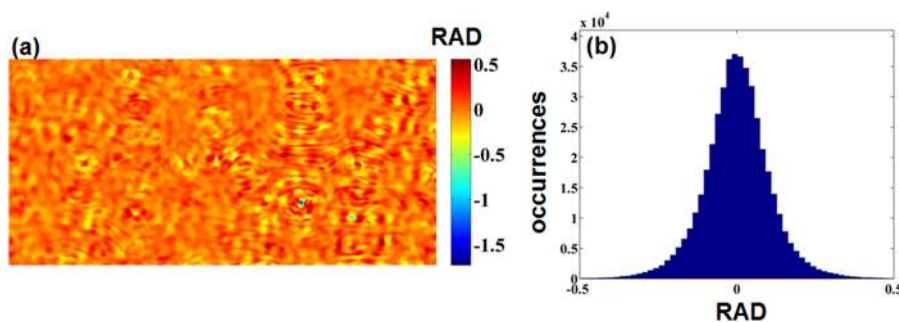
FIG. 12. (a) Phase difference between the results from Figs. 11(b) and 11(c) and (b) histogram of the phase difference in (a).

Therefore, probably, this is a consequence for the particular fringe pattern in Fig. 11(a) to yield the 1.5 rad phase difference between the two methods. Note also that the WFT2F is constrained by the width of the analyzing window (20 pixels) and probably this does not allow the WFT2F to correctly process high spatially transient phase variations, such as in Fig. 11(a). This could be ameliorated by decreasing the window size, but at the cost of increasing the standard deviation of the phase error. Finally, the compromise for the correct parameters and/or suitable algorithm is not so trivial and could be considered in the perspective of the expected application. As an example, in the field of vibroacoustics, for the problem of force identification,[36] the inverse problem is required to process data from full-field measurements, as provided in this section. The discussion related to the required overall quality and accuracy of the measured input data remains opened and needs to be deeply investigated in the future.

From Figs. 11 and 12, we can conclude that the DL, trained on the database with realistic speckle noise, exhibits high efficiency and is state-of-the-art, when de-noising phase data that were not in the learning database.

VI. CONCLUSION

This paper demonstrates that a state-of-the-art deep-learning-based algorithm trained to de-noise phase maps from digital holographic interferometry can be obtained when training the network with realistic noisy phase data. To do this, faithful speckle noise conditions, having non-Gaussian statistics and non-stationary property, being amplitude-dependent, and exhibiting spatial correlation length, must be simulated. As a result, training a network with data including simply added stationary Gaussian noise does not provide a robust de-noiser. In order to achieve the state-of-the-art de-noiser, the deep neural network was trained with a database constituted with 40 noisy phase maps, including various orientations and fringe densities. With three iterations, the performance obtained with the proposed approach yields comparable standard deviation to the 2D windowed Fourier transform algorithm and better peak-to-valley. The computation time of the proposed approach is smaller than that of the 2D windowed Fourier transform, especially for experimental noisy phase data. Processing of experimental data from wide-field digital holographic vibrometry at a 100 kHz frame rate yielded promising results. Future works concern reducing the computation time, for example, by moving to the Tensorflow software environment, and deep investigation of the architecture could be carried out by testing other network configurations, for example, when reducing the number of layers. Finally, the overall quality and accuracy of the processed data must be investigated to determine which phase error could be tolerated in the related metrology problems.

REFERENCES

[1] J. W. Goodman and R. W. Lawrence, "Digital image formation from electronically detected holograms," Appl. Phys. Lett. **11**, 77 (1967).

[2] T. Kreis, *Handbook of Holographic Interferometry: Optical and Digital Methods* (Wiley, 2004).

[3] U. Schnars and W. Jueptner, *Digital Holography* (Springer, 2005).

[4] T. C. Poon, *Digital Holography and Three-Dimensional Display: Principles and Applications* (Springer-Verlag, 2010).

[5] P. Picart, *New Techniques in Digital Holography* (ISTE-Wiley, 2015).

[6] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," Nature **521**, 436–444 (2015).

[7] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, "Beyond a Gaussian denoiser: Residual learning of deep CNN for image de-noising," IEEE Trans. Image Process. **26**(7), 3142–3155 (2017).

[8] T. Zeng, H. K.-H. So, and E. Y. Lam, "Computational image speckle suppression using block matching and machine learning," Appl. Opt. **58**, B39–B45 (2019).

[9] J. P. Krishnan, J. M. Boucas-Dias, and V. Katkovnik, "Dictionary learning phase retrieval from noisy diffraction patterns," Sensors **18**(11), 4006 (2018).

[10] G. Choi, D. H. Ryu, Y. J. Jo, Y. S. Kim, W. Park, H.-S. Min, and Y. P. Park, "Cycle-consistent deep learning approach to coherent noise reduction in optical diffraction tomography," Opt. Express **27**, 4927–4943 (2019).

[11] H. Chen, Y. Zhang, W. Zhang, P. Liao, K. Li, J. Zhou, and G. Wang, "Low-dose CT via convolutional neural network," Biomed. Opt. Express **8**, 679–694 (2017).

[12] Q. Zhang, Q. Yuan, J. Li, Z. Yang, and X. Ma, "Learning a dilated residual network for SAR image despeckling," Remote Sens. **10**(2), 196 (2019).

[13] W. Jeon, W. Jeong, K. Son, and H. Yang, "Speckle noise reduction for digital holographic images using multi-scale convolutional neural networks," Opt. Lett. **43**(17), 4240–4243 (2018).

[14] C. Chen and Z. Xu, "Aerial-image denoising based on convolutional neural network with multi-scale residual learning approach," Information **9**, 169 (2018).

[15] G. Chierchia, D. Cozzolino, G. Poggi, and L. Verdoliva, "SAR image despeckling through convolutional neural networks," in *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS)* (IEEE, 2017), pp. 5438–5441.

[16] P. Wang, H. Zhang, and V. M. Patel, "SAR image despeckling using a convolutional neural network," IEEE Signal Process. Lett. **24**(12), 1763–1767 (2017).

[17] Y. Ma, X. Chen, W. Zhu, X. Cheng, D. Xiang, and F. Shi, "Speckle noise reduction in optical coherence tomography images based on edge-sensitive cGAN," Biomed. Opt. Express **9**, 5129–5146 (2018).

[18] T. L. Bobrow, F. Mahmood, M. Inserni, and N. J. Durr, "DeepLSR: A deep learning approach for laser speckle reduction," Biomed. Opt. Express **10**, 2869–2882 (2019).

[19] K. Yan, Y. Yu, C. Huang, L. Sui, Q. Kemao, and A. Asundi, "Fringe pattern denoising based on deep learning," Opt. Commun. **437**, 148–152 (2019).

[20] F. Hao, C. Tang, M. Xu, and Z. Lei, "Batch denoising of ESPI fringe patterns based on convolutional neural network," Appl. Opt. **58**(13), 3338–3346 (2019).

[21] U. Schnars and W. Jüptner, "Direct recording of holograms by a CCD target and numerical reconstruction," Appl. Opt. **33**(2), 179–181 (1994).

[22] D. C. Ghiglia and M. D. Pritt, *Two-Dimensional Phase Unwrapping: Theory, Algorithms and Software* (Wiley, 1998).

[23] S. Montresor and P. Picart, "Quantitative appraisal for noise reduction in digital holographic phase imaging," Opt. Express **24**(13), 14322–14343 (2016).

[24] J. Leval, P. Picart, J.-P. Boileau, and J.-C. Pascal, "Full-field vibrometry with digital Fresnel holography," Appl. Opt. **44**(27), 5763–5772 (2005).

[25] Q. Kemao, "Windowed Fourier transform for fringe pattern analysis," Appl. Opt. **43**(17), 2695–2702 (2004).

[26] G. Barbastathis, A. Ozcan, and G. Situ, "On the use of deep learning for computational imaging," Optica **6**, 921–943 (2019).

[27] J. C. Dainty, *Laser Speckle and Related Phenomena* (Springer-Verlag, 1984).

[28] J. W. Goodman, *Speckle Phenomena in Optics* (Roberts and Company Publishers, 2006).

[29] V. Bianco, P. Memmolo, M. Leo, S. Montresor, C. Distante, M. Paturzo, P. Picart, B. Javidi, and P. Ferraro, "Strategies for reducing speckle noise in digital holography," Light: Sci. Appl. **7**, 48–58 (2018).

[30] Z. Wang, A. C. Bovik, and L. Lu, "Why is image quality assessment so difficult?," in *Proceedings of the IEEE ICASSP* (IEEE, 2002), Vol. 4, pp. 3313–3316.

[31] S. Montresor and P. Picart, "Error analysis for noise reduction in 3D deformation measurement with digital color holography," J. Opt. Soc. Am. B **34**(5), B9–B34 (2017).

[32] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising with block-matching and 3D filtering," Proc. SPIE **6064**, 606414 (2006).

[33] I. Selesnick, R. G. Baraniuk, and N. G. Kingsbury, "The dual-tree complex wavelet transform," IEEE Signal Process. Mag. **22**(6), 123–151 (2005).

[34] P. Picart and S. Montresor, "Contrast transfer function of de-noising algorithms," Opt. Express **27**, 23336–23356 (2019).

[35] L. Lagny, M. Secail-Geraud, J. Le Meur, S. Montresor, K. Heggarty, C. Pezerat, and P. Picart, "Visualization of travelling waves propagating in a plate equipped with 2D ABH using wide-field holographic vibrometry," J. Sound Vib. **461**, 114925 (2019).

[36] C. Faure, F. Ablitzer, C. Pézerat, and J. Antoni, "Vibration source characterization using force analysis technique and a Bayesian regularization," in Proceedings of the 23rd International Congress on Sound and Vibration, Athen, Greece, July 10–14, 2016.