POLITECNICO
MILANO 1863

SCUOLA DI INGEGNERIA INDUSTRIALE
E DELL'INFORMAZIONE

Executive Summary of the Thesis

# Experimental Validation for Data Driven Near-Field Acoustic Holography

Laurea Magistrale in Music and Acoustic Engineering

**Author:** Alessio Lampis

**Advisor:** Prof. Francesco Ripamonti

**Co-advisors:** Raffaele Malvermi, Juan Sebastian Gonzalez

**Academic year:** 2021-2022

## 1. Introduction

Several techniques for sound localization have been used in acoustics and sound control to resolve noise and vibration problems. One powerful technique is the Near-Field Acoustic Holography (NAH), which provides a meaningful representation of the entire three-dimensional acoustic field radiated by a vibrating source.
NAH is distinguished from other methods by having unlimited spatial resolution and being capable of extracting many physical quantities, such as particle velocity or sound intensity.
The principle of NAH is backpropagating the acoustic sound pressure field captured in a *hologram* into the surrounding air medium using inversion algorithms.
The peculiarity of NAH is that the sound field is measured in a region very close to the source, hence the term *near-field*. The evanescent waves contained in this part of the acoustic field are responsible for the unlimited spatial resolution.
From the first implementation [3] to the most common algorithms used nowadays [1, 6], NAH makes it possible to analyze the vibration field of the sound source. It is usually employed for reconstructing the velocity field on the source surface.

A recent innovation in this field is represented by the adoption of Deep Learning technologies. Specifically, a Convolutional Neural Network (CNN) has been used to retrieve the velocity field of a vibrating source [4, 5]. This data driven approach showed encouraging results but experimental validation is still missing.
This thesis work aimed at validating the CNN with experimental data. To do so, we designed and built an experimental setup to record the sound field of a vibrating source. We focused our study on complex-shaped objects because they are more common in practical applications. We used a violin back plate as test case.
For building the experimental setup, our starting point was the simulation setup employed for training the CNN [4]. Moreover, we analyzed several setups for performing NAH on musical instruments as reference. We identified three main features:

- it is necessary to have free-field conditions, therefore acoustic measurements should be done in an anechoic environment;
- to avoid any scattered field, the dimension of the microphones should be smaller than the wavelength of the sound field;
- to include the evanescent waves, the dis-

tance between microphones and vibrating source should be around 1/4 of the acoustic wavelength or less.

## 2. Experimental Setup

Fig. 1 shows a picture of the experimental setup we built. It is located in the anechoic chamber of the Cremona Campus of Politecnico di Milano, Italy.
The violin plate was suspended horizontally by means of two crisscrossing rubber bands to simulate free boundary conditions.
To excite the plate, we used an LSD V406 electro-dynamic shaker driven by a function generator/power amplifier configuration system. To avoid any additional mass loading due to the shaker, we decoupled the oscillating piston of the shaker to the violin plate by a soft spring, which was preloaded with an elongation of 5 mm, corresponding to imposing a static force of around 0.5 N. The spring was attached below the violin plate, at a point close to the border.
Above the plate, an array of eight Beyerdynamic MM1 microphones was arranged in a row along the width of the plate. The microphones were fixed 3 cm from each other and laid on a plane parallel to the plate and distant 2 cm from the highest arching point. To hold the microphone in position, we designed and built an aluminum frame, which was mounted on a linear actuator. The linear actuator was used to move the aluminum frame, together with the microphones, along the length of the plate. This way, we were able to scan the near-field above the whole source. The movement of the actuator was controlled by an Arduino board and bounded by two limit switches.

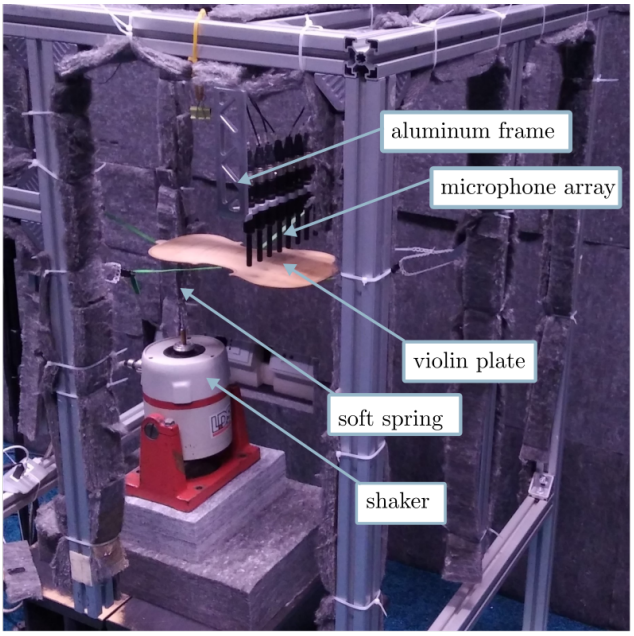


Figure 1: Experimental setup for performing NAH on a violin plate. The aluminum plane is attached to a linear actuator, not visible in this picture.

## 3. Resonance frequencies of the violin plate

To make the violin plate radiate, this was excited in resonance conditions: a stationary sinusoidal force was applied to the plate at a driving frequency equal to one of the plate resonance frequencies.
The resonance frequencies of the plate, also called eigenfrequencies, were identified by computing the *Complex Mode Indicator Function* (CMIF) [2]. The locations of the CMIF peaks corresponded to the eigenfrequencies of the system, as shown in Fig. 2.
To obtain the CMIF, we first measured a set of impulse responses of the violin plate using the roving hammer impact test: an accelerometer was positioned at a fixed point corresponding to the center of mass of the violin plate, and we excited the plate in $M = 88$ points along its surface, with a dynamometric hammer. The signals of the accelerometer and the hammer were recorded simultaneously by a PCB board. The CMIF was computed as

$$\mathrm{CMIF}(f) = \mathbf{\Sigma}(f)^H \mathbf{\Sigma}(f), \tag{1}$$

where $H$ is the Hermitian adjoint operator and $\mathbf{\Sigma}(f) \in \mathbb{R}^M$ derived from the singular value decomposition of the real mobility matrix

$$\Re(\mathbf{H_1}(f)) = \mathbf{U}(f)\mathbf{\Sigma}(f)\mathbf{V}(f)^H, \tag{2}$$

in which

$$\mathbf{H_1}(f) = [H_1^{(1)}(f), H_1^{(2)}(f), \ldots, H_1^{(M)}(f)]. \tag{3}$$

The mobility is the harmonic response of a system expressed as the velocity over a unit of force. The term $H_1^{(m)}(f)$ is the H1 estimator of the mobility [2] of the $m^{th}$ measurement point at frequency $f$.

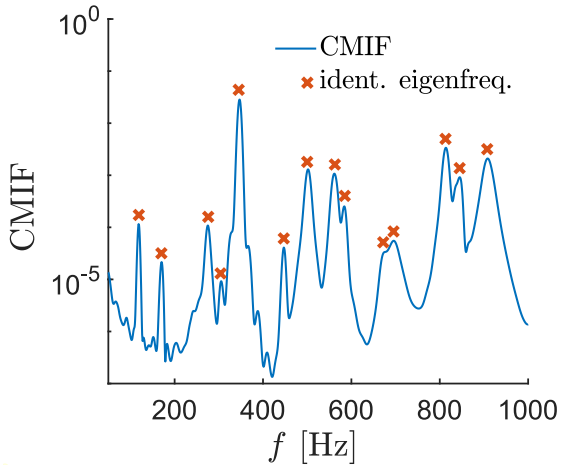


Figure 2: CMIF and the identified eigenfrequencies.

## 4. Acquisition procedure

The vibration of the violin plate was driven by the shaker with a sinusoidal force at a frequency equal to an eigenfrequency $f_k$.
The radiating acoustic field was digitally recorded by an eight channels Firestudio Presonus audio interface for a duration of 7 s at a frequency rate of $f_s = 48$ kH. The microphone array was then moved and the recording was repeated. The process continued until the whole surface above the vibrating plate was scanned. A total of $N = 272$ recordings were done, to scan a portion of the near-field with a grid of $8 \times 34$ points. The spatial resolution of this grid was 1 cm along the length and 3 cm along the width of the violin plate.
The acquisition took place in a temperature and hygrometry controlled conditions. Specifically, the temperature was set at around 21°C and the humidity at around 25%. Before the acquisition, the microphone array was calibrated.

## 5. Holographic images

From the recordings, we made a set of holographic images of the acoustic sound field. Each recording represented a “pixel” of the image obtained by the RMS value of the sound pressure. The images were normalized so that their values ranged between 0 and 1.
We observed that above 345 Hz the holographic images were completely corrupted and we discarded them. The measurements relative to these images contained a strong noise component. For this reason, we showed only on the measurements below 345 Hz and within this range we focused on three eigenfrequencies corresponding to the so-called "signature modes" (mode 1, mode 2, and mode 5). The signature modes are of great importance in violin making because luthiers use them for tuning their instruments.
We built a FEM model using *COMSOL Multyphysics®* to obtain a set of simulated holographic images. We also normalized these images. Fig. 3 shows a comparison between experimental and simulated holographic images.
Regarding mode 1, we noted a high resemblance between the images, while mode 2 and mode 5 differed in the top-left part, where a higher level of sound pressure occurred. This could be due to the contribution of an unwanted scattered field. It could be also caused by the preload of the spring, which was attached on the same spot where the higher level of sound pressure was detected.
To investigate this aspect, we measured the signal-to-noise ratio (SNR) as the ratio between the recording of the radiated sound field and the recording of the noise, obtained while the shaker was running without being attached to the violin plate. This analysis showed that the SNR for mode 2 and mode 5 is around 5 dB, confirming that the noise level in these recordings is high.

## 6. Velocity field prediction

We used the CNN to predict the velocity field of the violin plate. Specifically, we exploited two neural networks. One, called *CNN-NAH* [5], has been trained with a dataset of around 200000 hologram images and their relative velocity field images. The vibrating source in this case was an isotropic rectangular plate. The other network, called *SRCNN-NAH* [4], was trained with a violin top plate as a source. Specifically, with a dataset of around 58000 hologram images and their relative velocity field images. Both the datasets have been made with extensive simulation campaigns.
The CNN-NAH is fed with images with a resolution of $16 \times 64$ points, while the SRCNN-NAH with images with a resolution of $8 \times 8$ points. To obtain such resolutions from the experimental holographic images, we spatially sampled them with a biharmonic spline interpolation method.

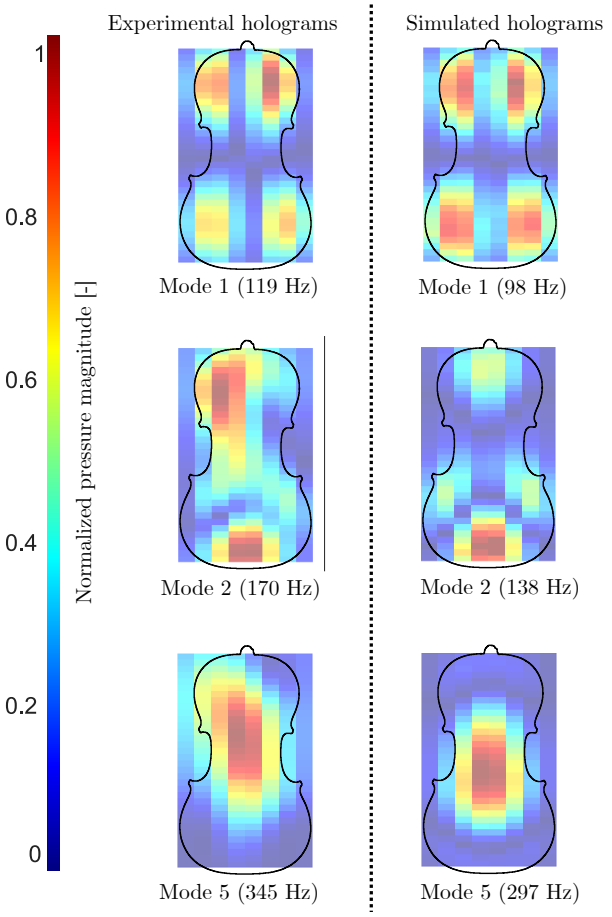


Figure 3: Normalized acoustic pressure magnitude of the near-field of a violin back plate. On the left, are experimental holographic images. On the right, are simulated holographic images. From top to bottom, signature modes (mode 1, mode 2, and mode 5).

The predictions were computed in terms of mobility and were defined as images with a resolution of $16 \times 64$ points.
The experimental validation of the networks was assessed by comparing the predictions to the experimental velocity field, which was retrieved by the real part of the mobility matrix $\mathbf{H_1}(f_k)$ at the resonance frequency $f_k$. To obtain a ground truth (GT) image, we spatially sampled the velocity field on a grid of $16 \times 64$ points.
Fig. 4 shows the comparison between the predictions and the GT. Normalized magnitude values were considered. The CNN-NAH network was able to correctly predict the modal pattern of modes 1 and 5. Regarding the latter, the typical ring nodal line was slightly altered on the top left. Mode 2 failed to be predicted with such network, even if the lower half of the image showed a nodal line similar to the GT. Regarding the SRCNN-NAH, the modal pattern of mode 1 was clearly recognizable. However, the predictions of mode 2 and 5 were less accurate and some of the nodal lines were detectable.
A qualitative analysis of these predictions pointed out that the CNN-NAH provided a more clear estimate of the velocity fields. However, in the regions of the predictions where the holographic images had an higher sound pressure level, the accuracy was lower. The modal patterns of the SRCNN-NAH predictions were recognizable if compared to the GT, but the mode 5 was the most difficult to identify.
We computed the *Normalized Cross-Correlation* (NCC) to quantify the similarity between predictions and GT. This metric is defined as

$$\mathrm{NCC}(\hat{Y}(f_k), Y(f_k)) = \frac{\hat{Y}(f_k)^T Y(f_k)}{||\hat{Y}(f_k)||_2 ||Y(f_k)||_2}, \tag{4}$$

where $\hat{Y}(f_k) \in \mathbb{R}^{16\times64}$ denotes the prediction and $Y(f_k) \in \mathbb{R}^{16\times64}$ the GT. The NCC is a scalar bounded between 0 and 1. When it is equal to 1, the two quantities are equal.

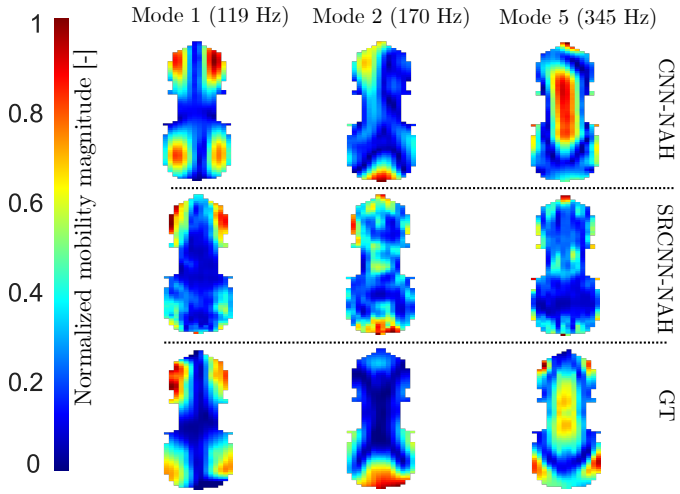


Figure 4: Prediction of the velocity field in terms of normalized mobility magnitude. From top to bottom, the two Deep Learning networks, and the ground truth (GT).

As shown in Tab. 1, the values of the NCC reflected the previous comments: the similarity between the predictions and the GT was the highest for mode 1, while decreased for modes

2 and 5, even if was relatively high for mode 5 of CNN-NAH.

| Mode | CNN-NAH | SRCNN-NAH |
|---|---|---|
| **1** | **0.9** | **0.91** |
| **2** | **0.7** | **0.74** |
| **5** | **0.84** | **0.74** |

Table 1: Values of the NCC for evaluating the similarity between the predictions and the GT.

## 7. Conclusions

In this thesis work we designed and developed a complete experimental setup for performing NAH with a violin plate as a sound source. We collected experimental data and obtained holographic images, used for retrieving the velocity field on the violin plate surface. To do so, we exploited a novel methodology based on Convolutional Neural Networks. Specifically, we used two neural networks called CNN-NAH and SRCNN-NAH. The first one was able to predict the modal pattern of the violin plate vibration with more accuracy. The second one predicted the velocity field of mode 1 correctly and the ones of the other modes with less precision.
This result showed that the CNN-NAH is more robust to experimental noise than the SRCNN-NAH. Even if the CNN-NAH was trained with rectangular plates sources rather than violin plates, its predictions are more clear. In fact, it is worth noting that the major difference between the two networks was the dimension of the input images: the SRCNN-NAH had less input points and therefore might be more sensible to noise, suggesting that the accuracy of the predictions to be directly linked to this characteristic. Moreover, we noted a high pressure level on the holographic images of mode 2 and mode 5, presumably caused by the spring preload and the presence of the shaker. We believe that also this aspect is responsible for the accuracy of the predictions. With that said, this data driven approach is based on simulated dataset. Therefore, we suggest that experimental data should be taken into account in future developments. For example, by modeling the experimental noise in a novel training set.

## 8. Acknowledgments

The author is indebted to the professors, researchers and academic staff of the Polimi Sound And Vibration Laboratory (PSVL) and for making available the anechoic room of the Cremona Campus of Politecnico di Milano, Italy.

## References

[1] Mingsian R Bai. Application of bem (boundary element method)-based acoustic holography to radiation analysis of sound sources with arbitrarily shaped geometries. *The Journal of the Acoustical Society of America*, 92(1):533–549, 1992.

[2] David J Ewins. *Modal testing: theory, practice and application*. John Wiley & Sons, 2009.

[3] Julian D Maynard, Earl G Williams, and Y Lee. Nearfield acoustic holography: I. theory of generalized holography and the development of nah. *The Journal of the Acoustical Society of America*, 78(4):1395–1413, 1985.

[4] Marco Olivieri, Mirco Pezzoli, Fabio Antonacci, and Augusto Sarti. Near field acoustic holography on arbitrary shapes using convolutional neural network. In *2021 29th European Signal Processing Conference (EUSIPCO)*, pages 121–125. IEEE, 2021.

[5] Marco Olivieri, Mirco Pezzoli, Raffaele Malvermi, Fabio Antonacci, and Augusto Sarti. Near-field acoustic holography analysis with convolutional neural networks. In *INTER-NOISE and NOISE-CON Congress and Conference Proceedings*, volume 261, pages 5607–5618. Institute of Noise Control Engineering, 2020.

[6] Angie Sarkissian. Method of superposition applied to patch near-field acoustic holography. *The Journal of the Acoustical Society of America*, 118(2):671–678, 2005.