

---

This is an electronic reprint of the original article.  
This reprint may differ from the original in pagination and typographic detail.

Prawda, Karolina; Schlecht, Sebastian; Välimäki, Vesa

## Robust selection of clean swept-sine measurements in non-stationary noise

*Published in:*  
The Journal of the Acoustical Society of America

*DOI:*  
[/10.1121/10.0009915](https://doi.org/10.1121/10.0009915)

Published: 01/03/2022

*Document Version*  
Publisher's PDF, also known as Version of record

*Published under the following license:*  
CC BY

*Please cite the original version:*  
Prawda, K., Schlecht, S., & Välimäki, V. (2022). Robust selection of clean swept-sine measurements in non-stationary noise. *The Journal of the Acoustical Society of America*, 151(3), 2117-2126.  
<https://doi.org/10.1121/10.0009915>

---

This material is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of the repository collections is not permitted, except that material may be duplicated by you for your research use or educational purposes in electronic or print form. You must obtain permission for any other use. Electronic or print copies may not be offered, whether for sale or otherwise to anyone who is not an authorised user.

## Robust selection of clean swept-sine measurements in non-stationary noise

Karolina Prawda, Sebastian J. Schlecht and Vesa Välimäki

Citation: *The Journal of the Acoustical Society of America* **151**, 2117 (2022); doi: 10.1121/10.0009915

View online: <https://doi.org/10.1121/10.0009915>

View Table of Contents: <https://asa.scitation.org/toc/jas/151/3>

Published by the [Acoustical Society of America](#)

---

### ARTICLES YOU MAY BE INTERESTED IN

[Toward a better understanding of nonoccupational sound exposures and associated health impacts: Methods of the Apple Hearing Study](#)

*The Journal of the Acoustical Society of America* **151**, 1476 (2022); <https://doi.org/10.1121/10.0009620>

[Optimized design of windowed-sinc anti-aliasing filters for phase-preserving decimation of hydrophone data](#)

*The Journal of the Acoustical Society of America* **151**, 2077 (2022); <https://doi.org/10.1121/10.0009823>

[Reconstruction of echoes reaching bats in flight from arbitrary targets by acoustic simulation](#)

*The Journal of the Acoustical Society of America* **151**, 2127 (2022); <https://doi.org/10.1121/10.0009916>

[The sound of one frog calling: The bullfrog's reactions to acoustic stimuli](#)

*The Journal of the Acoustical Society of America* **151**, R5 (2022); <https://doi.org/10.1121/10.0009652>

[Numerical realization of diffuse sound pressure fields using prolate spheroidal wave functions](#)

*The Journal of the Acoustical Society of America* **151**, 1710 (2022); <https://doi.org/10.1121/10.0009764>

[An overview of ambient sound using Ocean Observatories Initiative hydrophones](#)

*The Journal of the Acoustical Society of America* **151**, 2085 (2022); <https://doi.org/10.1121/10.0009836>

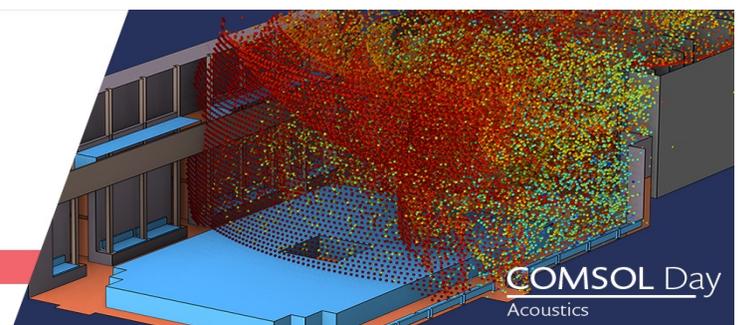
---

See how multiphysics  
simulation is being  
used for acoustics  
and vibrations

at a free, online event

JOIN US »

APRIL 28, 2022



COMSOL Day  
Acoustics

## Robust selection of clean swept-sine measurements in non-stationary noise

Karolina Prawda,<sup>a)</sup>  Sebastian J. Schlecht,<sup>b)</sup>  and Vesa Välimäki 

Acoustics Lab, Department of Signal Processing and Acoustics, Aalto University, FI-02150 Espoo, Finland

### ABSTRACT:

The exponential sine sweep is a commonly used excitation signal in acoustic measurements, which, however, is susceptible to non-stationary noise. This paper shows how to detect contaminated sweep signals and select clean ones based on a procedure called the rule of two, which analyzes repeated sweep measurements. A high correlation between a pair of signals indicates that they are devoid of non-stationary noise. The detection threshold for the correlation is determined based on the energy of background noise and time variance. Not being disturbed by non-stationary events, a median-based method is suggested for reliable background noise energy estimation. The proposed method is shown to detect reliably 95% of impulsive noises and 75% of dropouts in the synthesized sweeps. Tested on a large set of measurements and compared with a previous method, the proposed method is shown to be more robust in detecting various non-stationary disturbances, improving the detection rate by 30 percentage points. The rule-of-two procedure increases the robustness of practical acoustic and audio measurements. © 2022 Author(s). All article content, except where otherwise noted, is licensed under a Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>). <https://doi.org/10.1121/10.0009915>

(Received 15 December 2021; revised 24 February 2022; accepted 9 March 2022; published online 24 March 2022)

[Editor: Ning Xiang]

Pages: 2117–2126

### I. INTRODUCTION

An impulse response (IR) measurement is one of the most common procedures to assess the acoustic qualities of various systems, including physical spaces, such as concert halls and rooms,<sup>1–5</sup> electronic devices,<sup>6</sup> audio software,<sup>7</sup> and more. IR measurements can be conducted using a variety of excitation signals, including impulses, which are produced with sources such as pistols and balloon pops,<sup>8,9</sup> noise-based methods, e.g., maximum-length sequence (MLS)<sup>10–12</sup> and inverse repeated sequence (IRS),<sup>13</sup> and linear and exponential swept-sine signals. Each of these methods has its strong sides and shortcomings.<sup>14</sup>

The exponential swept-sine (ESS) as an excitation signal for measuring IRs was introduced in the form used nowadays by Farina over 20 years ago.<sup>15</sup> Currently, it is widely used as it provides the best consistency and highest robustness of measurements.<sup>14,16</sup> The ESS technique also rejects most of the harmonic distortion,<sup>16</sup> a bane of noise-based methods, such as MLS and IRS.<sup>14,17–19</sup> However, ESS is sensitive to non-stationary noise, which causes artifacts in IRs obtained in the deconvolution process, and may lead to errors in the estimation of acoustic parameters.<sup>14,20–23</sup> The present study investigates the ESS technique and discusses the disturbances that may occur during measurements and negatively affect the resulting IRs. A novel method to discriminate between clean and corrupted sweeps is proposed.

The ESS technique is known for its excellent signal-to-noise ratio (SNR)<sup>14</sup> that results from a long excitation of

low frequencies, which are usually more susceptible to contamination by background noise than high frequencies. This feature of swept-sine signals can be employed to achieve target SNR values for different frequencies by adjusting the time over which specific frequencies are excited.<sup>24–30</sup> The vulnerability of the ESS method to non-stationary noise, however, grows proportionally with the length of the sweep signal. This may force a compromise between lengthening the ESS signal to increase the SNR and shortening it to minimize the risk of the occurrence of non-stationary noise.<sup>19,20</sup> In this light, Stan *et al.*<sup>14</sup> recommend using the swept-sine technique only for measurements in empty, quiet spaces.

Currently, there is no established method to identify non-stationary noise in sweep measurements. Manual detection works only for singular measurements, but in the case of numerous unsupervised measurements, automatic detection is necessary. Guski<sup>31,32</sup> presented an algorithm addressing the problem of automatic classification of contaminated sweeps. Relying, however, on the separation of IR and background noise, this method is prone to errors when estimating decay and noise floor. Therefore, the need for a simpler and more reliable procedure remains.

This paper proposes to identify clean and contaminated ESS measurements based on their similarity to each other, expressed by the Pearson correlation coefficient (PCC). Used in applications such as pattern recognition<sup>33</sup> and as a criterion for filter optimization,<sup>34</sup> PCC proved to be more advantageous than the mean square error criterion. Similarly, cross correlation was used as a measure to estimate IRs sensitivity to small changes in sound-source position<sup>35</sup> as well as for robust IR measurement against nonlinearities.<sup>36–38</sup> This suggests that parameters related to similarity are good indicators

<sup>a)</sup>Electronic mail: karolina.prawda@aalto.fi

<sup>b)</sup>Also at: Art and Media Lab, Dept. of Media, Aalto University, FI-02150 Espoo, Finland.

of changes in audio signals, even when the environment is not free from noise.

The present work studies the problem of ESS measurements corrupted by non-stationary noise and introduces a procedure called the rule of two (Ro2). Ro2 is a method to identify a pair of clean sweeps, those not contaminated by non-stationary noise, from a series of measurements in a noisy environment. The method is based on the correlation between measured ESS signals. Various factors impacting the correlation are examined. The threshold separating clean sweeps from corrupted ones is determined. The Ro2 procedure is tested on a big dataset of ESS measurements and is compared to another method aimed at detecting impulsive noise in sweep measurements.

The remainder of this paper is organized as follows. Section II discusses the correlation between acoustic signals and describes the proposed method. In Sec. III, the expected contamination, such as stationary noise and time variance, are presented. Section IV elaborates on the types of non-stationary contamination and their effect on the correlation. Section V describes the validation procedure for the proposed method, discusses the experimental results, and compares the proposed method with another technique. Section VI concludes the paper.

## II. METHOD

This section tackles the detection of non-stationary events in an ESS signal and proposes a novel method called Ro2. The correlation of acoustic signals is also discussed.

### A. Problem formulation

Assessing whether the signal obtained during acoustic measurement is free of non-stationary noise or other artifacts is often a difficult task. Therefore, a good practice is to record a few test signals so as to be able to choose the best one, should unexpected acoustic events occur. In this case, recordings of the same conditions of the system under test can be compared to one another.

Given two acoustic measurements  $y_1$  and  $y_2$ , we want to determine whether they are clean or not. Assuming that contamination is a random occurrence, we measure the similarity of  $y_1$  and  $y_2$  as an indicator of contamination: if the similarity is low, then the contamination is indicated (in either one or in both of the signals), whereas a great similarity denotes an uncontaminated pair. We propose PCC as a robust measure of similarity. PCC is defined as

$$\rho_{y_1, y_2} = \frac{\text{cov}(y_1, y_2)}{\sigma_{y_1} \sigma_{y_2}} = \frac{\frac{1}{N-1} \sum_{k=1}^N (y_1(k) - \mu_{y_1})(y_2(k) - \mu_{y_2})}{\sqrt{\frac{1}{N-1} \sum_{k=1}^N (y_1(k) - \mu_{y_1})^2 \frac{1}{N-1} \sum_{k=1}^N (y_2(k) - \mu_{y_2})^2}}, \quad (1)$$

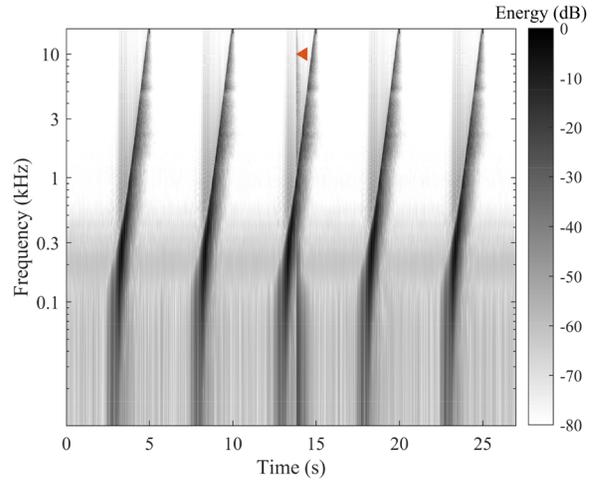


FIG. 1. (Color online) Spectrogram of a measurement consisting of five consecutive swept-sine signals. The arrow points to an impulsive noise event appearing in sweep #3 at about 14 s.

where  $\text{cov}(y_1, y_2)$  is the covariance of signals  $y_1$  and  $y_2$ ,  $\sigma_{y_1}$  and  $\sigma_{y_2}$  are their standard deviations,  $\mu_{y_1}$  and  $\mu_{y_2}$  are the mean values, and  $N$  is the number of samples in the signals. In acoustic measurements, the mean of the measured signals is removed,<sup>39</sup> transforming Eq. (1) to

$$\rho_{y_1, y_2} = \frac{\sum_{k=1}^N y_1(k)y_2(k)}{\sqrt{\sum_{k=1}^N y_1(k)^2 \sum_{k=1}^N y_2(k)^2}}. \quad (2)$$

Assuming that the system under test is free of noise, and neither the system nor the measurement equipment changes between or during the recording of both signals, i.e.,  $y_1 = y_2$ , then  $\rho_{y_1, y_2} = 1$ . Consecutive measurements, however, are never strictly the same, and the PCC of two clean measurements is impacted by two classes of factors: (1) expected disturbances including stationary background noise and time-variances of the measured system and (2) non-stationary occurrences such as impulsive noise or sound dropouts.<sup>31,40</sup>

Note that in the present study the term “clean signal” refers to a measured signal that contains stationary background noise and effects of time variance only, whereas the term “contaminated” is used for the signals containing both expected and unexpected disturbances.

An example of a set of measured ESS signals is shown in Fig. 1, where one of five sweeps is contaminated with impulsive noise. The corresponding PCC matrix is presented in Table I together with the total energy of each sweep in dB. The contaminated signal displays lower similarity with the other sweeps, while also having higher energy than the rest.

### B. Proposed method

The proposed method presents a systematic criterion to distinguish expected disturbances from non-stationary noise

TABLE I. PCC matrix for a series of five sweeps, cf. Fig. 1. Sweep #3 is less similar to the other signals, indicating the presence of non-stationary noise. The largest energy of sweep #3 also suggests the presence of additional noise. The smallest PCC values and the largest energy are highlighted.

Sweep #	1	2	3	4	5	Energy (dB)
1	1.000	0.999	<b>0.995</b>	0.999	0.999	71.76
2	0.999	1.000	<b>0.995</b>	0.999	0.999	71.75
3	<b>0.995</b>	<b>0.995</b>	1.000	<b>0.994</b>	<b>0.994</b>	<b>71.83</b>
4	0.999	0.999	<b>0.994</b>	1.000	0.999	71.74
5	0.999	0.999	<b>0.994</b>	0.999	1.000	71.73

to create a meaningful and robust measure for the level of contamination. The Ro2 method requires a correlation threshold  $\hat{\rho}_{y_1,y_2}$  separating clean signals from contaminated ones. Thus, the Ro2 is

$$\text{if } \rho_{y_1,y_2} > \hat{\rho}_{y_1,y_2} \text{ then } y_1 \text{ and } y_2 \text{ are a clean pair.} \quad (3)$$

When non-stationary noise occurs in the measurement, the PCC value does not point directly towards the contaminated sweep. Thus, when  $\rho_{y_1,y_2} < \hat{\rho}_{y_1,y_2}$ , the measurement should be repeated and the correlation of all captured signals should be estimated. The measurement can end when at least two signals fulfill the requirement in Eq. (3). Section III shows how to determine the threshold  $\hat{\rho}_{y_1,y_2}$ .

### III. EXPECTED CONTAMINATION

In the following, we present the two main sources of “unavoidable” impacts on correlation that are identified here, namely, background noise and time-variance. The following discussion leads to the determination of the expected PCC  $\hat{\rho}_{y_1,y_2}$ , which serves as the detection threshold.

#### A. Effect of background noise

The term “background noise” in acoustic measurements refers to any type of unwanted extra sound event. Since this definition includes non-stationary noise, the distinction needs to be made that in this study “background noise” is used to describe only the stationary noise.

The presence of stationary noise in the sweep signals affects their correlation. Therefore, in Eq. (2), we need to consider two noise signals  $n_1$  and  $n_2$  with zero mean. For this subsection, the background noise is the only disturbance such that the measurement signal is a mixture of the signal with the background noise, i.e.,  $y_1 = x + n_1$ , where

$$x = s * h \quad (4)$$

is the convolution of the ESS  $s$  and room impulse response  $h$ , denoted by an asterisk  $*$ . Similarly,  $y_2 = x + n_2$ . The resulting correlation is then

$$\rho_{y_1,y_2} = \frac{\sum_{k=1}^N (x(k) + n_1(k))(x(k) + n_2(k))}{\sqrt{\sum_{k=1}^N (x(k) + n_1(k))^2 \sum_{k=1}^N (x(k) + n_2(k))^2}}. \quad (5)$$

If the noise signals are uncorrelated with the ESS signals as well as with each other, i.e.,  $\sum_{k=1}^N n_1 n_2 = 0$ ,  $\sum_{k=1}^N x n_1 = 0$ , and  $\sum_{k=1}^N x n_2 = 0$ , then, Eq. (5) can be simplified to

$$\rho_{y_1,y_2} = \frac{\sum_{k=1}^N x(k)^2}{\sqrt{\sum_{k=1}^N (x(k)^2 + n_1(k)^2) \sum_{k=1}^N (x(k)^2 + n_2(k)^2)}}. \quad (6)$$

Thus, the signal energies are related by

$$\rho_{y_1,y_2} = \frac{E[x]}{\sqrt{(E[x] + E[n_1])(E[x] + E[n_2])}}, \quad (7)$$

where the energy of a signal is computed as

$$E[x] = \sum_{k=1}^N |x(k)|^2. \quad (8)$$

When the noise signal energies are equal, i.e.,  $E[n_1] = E[n_2] = E[n]$ , the PCC can be estimated using the SNR value,

$$\hat{\rho}_{y_1,y_2} = \frac{E[x]}{E[x] + E[n]} = \frac{\text{SNR}}{1 + \text{SNR}}, \quad (9)$$

where the SNR is expressed in terms of signal energies,

$$\text{SNR} = \frac{E[x]}{E[n]}. \quad (10)$$

In practice,  $E[x]$  is unknown as it is affected by the room impulse response  $h$ . However, it can be inferred from the difference  $E[x] = E[y_1] - E[n_1] = E[y_2] - E[n_2]$ .

#### B. Correlated background noise

Equation (9) provides an expected PCC value based on the assumptions that (1) the sweep responses are identical and (2) the background noise is uncorrelated and stationary. In the following, we discuss these assumptions.

The background noise is likely to contain strong harmonic content caused, for instance, by electric humming. Depending on the phase relation between measurements, harmonic background noise can be strongly correlated.

Let us consider the two extreme cases: when the noise signals  $n_1$  and  $n_2$  are fully correlated positively or negatively (anticorrelated). Thus,  $n_1 = \pm n_2$ , ergo  $\sum_{k=1}^N n_1 n_2 = \pm E[n]$ . This yields the following bounds to Eq. (9):

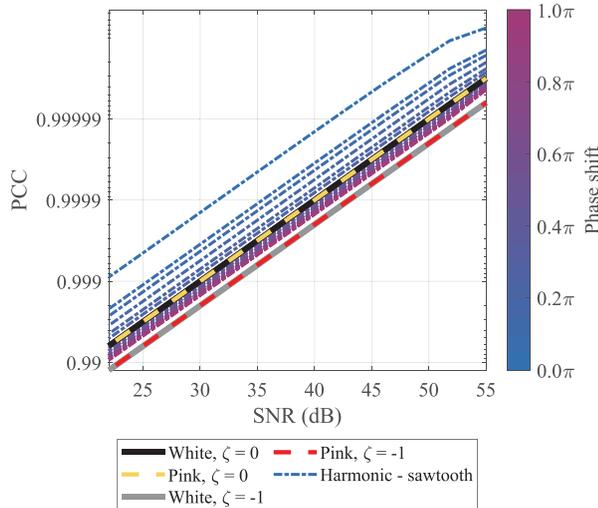


FIG. 2. Comparison of the PCC values for different SNRs of uncorrelated and anticorrelated stationary noise, as well as noise with harmonic content.

$$\frac{E[x] - E[n]}{E[x] + E[n]} \leq \hat{\rho}_{y_1, y_2, \zeta} \leq \frac{E[x] + E[n]}{E[x] + E[n]} = 1, \quad (11)$$

where  $\zeta = \rho_{n_1, n_2}$  is the correlation between two stationary noise terms. Note that perfectly correlated background noise, as part of the measurement signal, is virtually indistinguishable from an ESS.

To this end, an experiment showing the relation between PCC and SNR values was conducted. A 3-s-long ESS was synthesized and convolved with a synthetic IR having reverberation time (RT) of 2 s. This signal was then added to a set of white and pink noise signals having various energies so that different values of SNR could be obtained. The noise signals were either uncorrelated ( $\zeta = 0$ ) or anticorrelated ( $\zeta = -1$ ). The PCC values of these combined signals were calculated using Eqs. (9) and (11).

To simulate background noise with harmonic content, sawtooth waves were added to the aforementioned noise signals. The phase shifts between these signals were randomized between 0 and  $\pi$ . The results of the experiment, shown in Fig. 2, indicate that for clean signals, i.e., without non-stationary noise, PCC calculated as a function of SNR reaches high values close to unity. The results show that the spectral characteristics of the stationary noise have a negligible effect on the correlation as long as the noise does not contain periodic components that may result in sharp peaks or dips in the spectrum, e.g., sine waves.

The results also illustrate that harmonic content in a noise signal can heavily influence the correlation in both positive and negative directions. In Fig. 2, phase shifts for different SNR values create lines parallel to the ones resulting from the assumptions of uncorrelated and anticorrelated stationary noise. Small phase shifts close to 0 produce a highly correlated signal, whereas the increase in phase shift towards  $\pi$  decreases the PCC values, placing the signals

with the biggest shift between the uncorrelated and anticorrelated boundaries.

In principle, estimating the correlation of the background noise between measurements is possible if there is a sufficiently long time interval without any other signals. However, the time intervals between measurements can be several seconds such that the stationarity of the background signals needs to be fulfilled precisely to reliably estimate the correlation. Here, we adopt the worst-case scenario of anti-correlated noise as a lower bound for the expected PCC.

### C. Transfer-function variation

The measured system itself can undergo change. For instance, the position of the microphone and loudspeakers may vary due to vibrations, or the propagation paths can be altered due to variations in the air caused by temperature and humidity fluctuations or air movement (e.g., due to ventilation).<sup>17,31,40-43</sup> Unlike background noise, the measurement variations impact the impulse response  $h$  directly such that the signal model is

$$x = s * (h + v), \quad (12)$$

where  $h$  is some “ideal” room impulse response and  $v$  is the variation of the impulse response between measurements. Thus, the energy relation of two measurements are

$$\begin{aligned} E[y_1] &= E[s * h] + E[s * v_1] + E[n_1], \\ E[y_2] &= E[s * h] + E[s * v_2] + E[n_2]. \end{aligned} \quad (13)$$

The difference between the two measurements is then

$$y_1 - y_2 = s * (v_1 - v_2) + n_1 - n_2. \quad (14)$$

The energy relation is

$$E[y_1 - y_2] = E[s * (v_1 - v_2)] + E[n_1 - n_2]. \quad (15)$$

We choose the variation energy between two measurements such that  $E[v_1] = E[v_2]$  and  $E[s * v_1] = E[s * v_2]$ . Thus,

$$E[s * (v_1 - v_2)] = E[s * v_1] + E[s * v_2] = 2E[s * v_1] \quad (16)$$

since the variations  $v_1$  and  $v_2$  are uncorrelated (as the correlated part belongs to  $h$  by definition). Thus, we define the transfer-function variation factor

$$\tau = \frac{E[s * (v_1 - v_2)]}{E[s * h]} = \frac{2E[s * v_1]}{E[s * h]} = \frac{2E[s * v_2]}{E[s * h]}, \quad (17)$$

where  $E[s * (v_1 - v_2)]$  can be retrieved from the difference  $y_1 - y_2$  using Eq. (15), and  $E[s * h]$  can be retrieved from the measurement as  $E[s * h] = E[y_1] - E[s * v_1] - E[n_1]$  using Eq. (13) and Eq. (16).

The PCC of the transfer-function variation is

$$\begin{aligned} \rho_{y_1, y_2} &= \frac{\sum_{k=1}^N y_1(k)y_2(k)}{\sqrt{\sum_{k=1}^N y_1(k)^2 \sum_{k=1}^N y_2(k)^2}} \\ &= \frac{E[s * h]}{\sqrt{(E[s * h] + E[s * v_1])(E[s * h] + E[s * v_2])}} \\ &= \frac{E[s * h]}{\sqrt{(1 + \tau/2)^2 (E[s * h])^2}} = \frac{1}{1 + \tau/2}. \end{aligned} \quad (18)$$

Therefore, the transfer-function variation factor  $\tau$  serves as a tolerance parameter for the expected PCC from Eq. (11),

$$\hat{\rho}_{y_1, y_2, \zeta, \tau} = \frac{\hat{\rho}_{y_1, y_2, \zeta}}{1 + \tau/2}. \quad (19)$$

The effect of time variances on the impulse response measurements can be modeled with time-stretching<sup>40</sup> or by introducing sinusoidal jitter to the signal.<sup>17</sup> The complexity and unpredictability of such variations, however, might render these experiments insufficient to predict  $\tau$  values correctly. Therefore, in this study, transfer-function variation is estimated from the measured signals in Sec. V.

#### IV. NON-STATIONARY NOISE EVENTS

During an acoustic measurement, various non-stationary disturbances can occur. Such artifacts are, e.g., impulses, low-frequency noises, and sound dropouts, which originate from door slams, heavy vehicles moving outside of the measured space, and errors in measurement software.

This section examines how different types of non-stationary noise impact the PCC values, depending on their energy, frequency content, and time of occurrence. The effect of contamination on the correlation threshold estimation is also discussed.

##### A. Impulsive noise

The relation between the energy added to the sweep and the drop in PCC values can be concluded from Eq. (5) when we consider that one of the signals is contaminated with additional non-stationary noise  $n_{ns}$ , which is also assumed to be zero-mean and uncorrelated with both sweeps and stationary noise signals. Following the same reasoning leading from Eq. (5) to Eq. (7), we arrive at the following formula:

$$\rho_{y_1, y_2} = \frac{E[x]}{\sqrt{(E[x] + E[n_{ns}] + E[n])(E[x] + E[n])}}, \quad (20)$$

where  $E[n_{ns}]$  is the energy of non-stationary noise.

The theoretical values of correlation estimated with the aforementioned formula are presented in Fig. 3. Equation (20) predicts the general trend of decrease in the PCC with the growing energy difference between the clean and

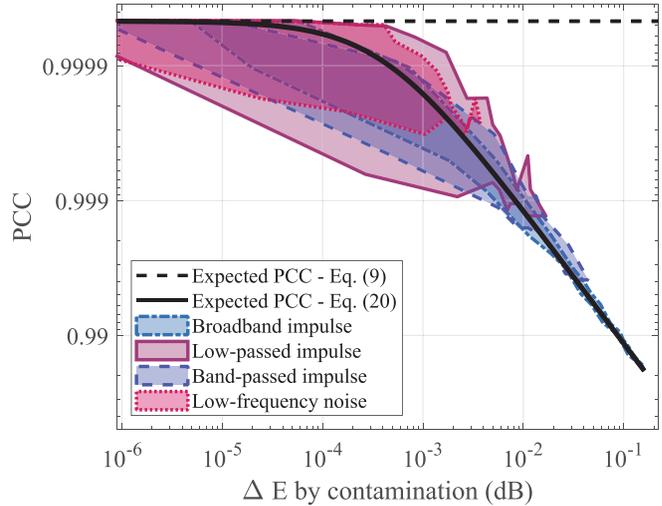


FIG. 3. Comparison of the PCC values for different non-stationary noise types for synthetic sweep with pink stationary noise and SNR of 84 dB. The areas cover the minimal and maximal values of measured correlation for the respective type of disturbance and  $\Delta E$ .

contaminated ESS. The energy difference  $\Delta E$  is a quotient of the energy of the two signals,

$$\Delta E = \frac{E[y_1]}{E[y_2]} = \frac{E[x] + E[n] + E[n_{ns}]}{E[x] + E[n]}. \quad (21)$$

In the experiment, a synthetic sweep signal containing stationary noise, as described in Sec. III B, with an SNR of 84 dB was further contaminated with impulsive noise and low-frequency noise. Broadband, lowpassed, and band-passed impulses served as impulsive disturbances. Lowpass-filtered white Gaussian noise lasting 1 s was used to elevate the noise floor of the measurement. One hundred signals of each type of non-stationary noise were used. The disturbances appeared at different times within the sweep, and their energy varied as well to obtain various changes to the contaminated signal's energy.

All signals used in the experiment had a different frequency content: the broadband impulse spanned across all frequencies, the lowpassed one had its cutoff frequency at 100 Hz, the bandpassed extended between 500 and 5000 Hz, whereas the white noise was lowpass filtered at 300 Hz.

The results of this experiment, presented in Fig. 3, show that increasing the signal's energy causes the PCC values to drop in accordance with Eq. (20). They also reveal that the correlation between the sweeps may vary for disturbances that add the same amount of energy. This phenomenon is especially prominent for the narrowest-band disturbances, such as low-passed impulse and low-frequency noise, which might be because non-stationary disturbance is not completely uncorrelated, but displays similarity to either the ESS or stationary noise. This is especially possible in the low-frequency region, where the sound is usually less diffuse.<sup>44</sup>

Note that the energy differences represented by  $\Delta E$  and mentioned in Table I are often small ( $<0.1$  dB) and may fall below the uncertainty of the measurement equipment

uncertainty. Therefore, the Ro2 should be applied to measurements performed within a short time, using the same measurement equipment, and the same settings.

### B. Median-based background noise energy estimation

Another problem related to the presence of non-stationary noise in the measurement is the possibility of contaminating the background noise used to estimate the SNR and thus, PCC threshold. A wrongly estimated noise energy  $E[n_1]$  leads to an underestimated PCC threshold and thus to the incorrect classification of clean and contaminated sweeps.

When a non-stationary random event contaminates the background noise, the affected samples carry more energy than the clean ones. Therefore, the contamination skews the amplitude distribution of noise in the positive direction. The nature of the stationary noise does not allow for the use of the PCC values as a discriminant for finding the non-stationary noise, as with sweep signals. However, if the amplitude distribution of the noise signal is Gaussian, i.e.,  $n_1(k) \sim \mathcal{N}(0, \sigma^2)$ , a robust estimator can be used.

The energy of Gaussian noise is essentially a scaled mean value of the squared signal. The drawback of such an estimator is its high sensitivity to outliers, resulting in false estimations for contaminated signals. The median, however, is less influenced by the outliers than the mean, since its breakdown point, i.e., the maximum proportion of contaminated observations that do not force the estimator to result in an aberrant value, is higher than that of mean: The breakdown point for the median is 0.5 whereas it is 0 for the mean.<sup>45-47</sup> This means that if at least 50% of samples are not contaminated, the median values are not skewed.

For squared samples from the Gaussian distribution, the mean and median are related by a constant scaling factor  $b_\chi = 1.4826$ .<sup>47</sup> Thus, the robust noise energy estimate is

$$E[n_1] = \sum_{k=1}^N n_1(k)^2 = Nb_\chi^2 \text{median}(n_1(1)^2, \dots, n_1(N)^2). \quad (22)$$

To demonstrate the effectiveness of this method, a 2-s-long noise signal with Gaussian distribution and different values of noise power was contaminated at random times with impulsive noise (as described in Sec. IV A) and 200-ms-long lowpassed Gaussian noise. The non-stationary disturbances were scaled by random factors to achieve various effects on the noise energy. Then, the mean and scaled median values of noise energy were compared to a target value—the energy of noise without contamination.

The results presented in Fig. 4 show that while the mean energy value can change by tens of decibels in the presence of non-stationary noise, the median remains essentially unchanged and very close to the target (within 1 dB). Additionally, using the median instead of the mean does not add any processing to the detection process, making it the recommended procedure for calculating the background noise energy.

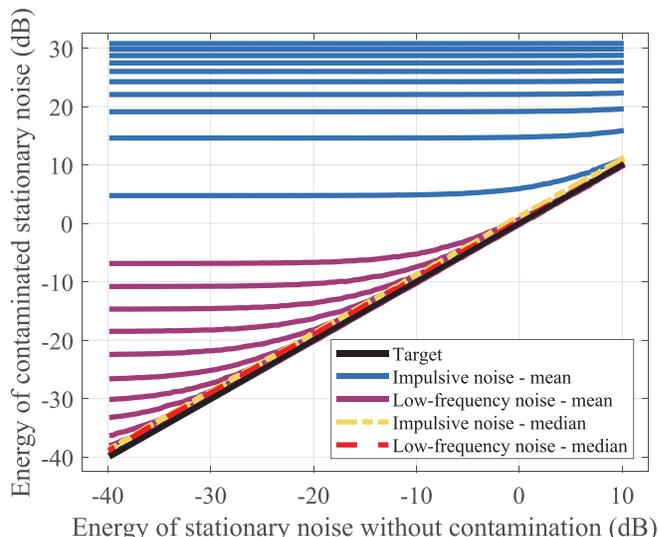


FIG. 4. Comparison between the mean and median energy of a contaminated noise signal. The colored solid lines show non-stationary disturbances of different energies corrupting background noise signals, with the loudest impulse at the top and the quietest at the bottom. The median energy remains closest to the target in all cases.

### C. Sound dropouts

The software-related dropouts do not add energy to the contaminated signal, but reduce it instead. Additionally, skipping the samples creates two cropped sweeps, one of which is shifted with respect to the clean sweep. Therefore, the energy difference introduced by the dropout is of less importance to the correlation between two signals than the time at which the skipping occurs.

To estimate the effect of the time of the dropout on the correlation, the synthesized sweeps, as used in Sec. III B, were contaminated with sound dropouts. The dropouts were simulated by deleting small portions of the signal, ranging from one to ten samples, at different times throughout the ESS and shifting the remaining portion of the sweep by the respective number of deleted samples. The dropouts were broadband disturbances, since a discontinuity is an impulsive event affecting all frequencies.

The relation between the drop in the PCC values at the time at which the samples were skipped is depicted in Fig. 5. The results show that if the dropout happens in the beginning of the sweep, the correlation between the clean and corrupted sweeps is very low. However, if the contamination occurs later in the signal, the PCC drop is less prominent. Additionally, the dropouts appearing after the ESS has finished playing (in the present case, after 3.0 s) affect the correlation only marginally.

### V. VALIDATION

In this section, the database used for testing the proposed method is presented. The results of using Ro2 on this dataset are presented, and the transfer-function variation is determined. The proposed method is also compared with

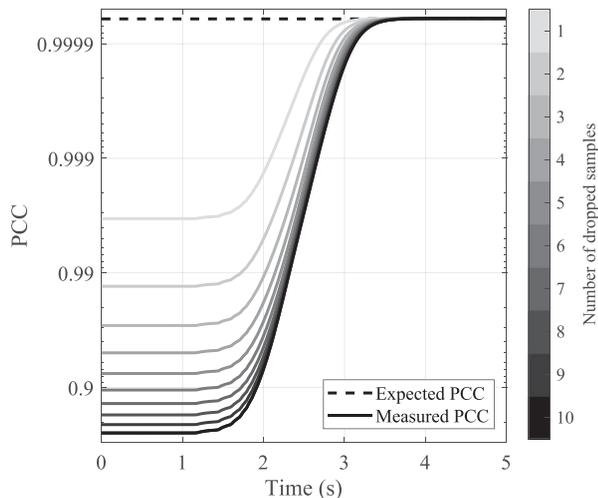


FIG. 5. Effect of sound dropouts on the PCC between sweep signals. The solid lines show the effect for the number of dropped samples from one (topmost plot) to ten (bottommost plot), whereas the dashed line indicates the PCC without dropouts.

another procedure for coping with non-stationary noise in acoustic measurements.

### A. Validation database

Ro2 was validated on a database of swept-sine measurements collected in the *Arni* room at the Acoustics Lab of Aalto University, Espoo, Finland.<sup>5,48</sup> *Arni* is a rectangular room, with dimensions 8.9 m × 6.3 m × 3.6 m (length, width, and height, respectively). The room’s walls and ceiling are equipped with acoustic panels that can switch their state between open and closed, changing the amount of absorption and thus varying the acoustics within the space. A view of the space and measurement equipment is shown in Fig. 6.

The equipment used during the measurements included a 01 dB LS01 omnidirectional loudspeaker (sound source), two G.R.A.S. 1/2-in. diffuse-field microphones of type 40AG, two G.R.A.S. 1/2-in. free-field microphones of type 46AF, one Brüel & Kjær 1/2-in. diffuse-field microphone of type 4192, a G.R.A.S. power module of type 12AG, a measurement laptop, and a MOTU UltraLite mk3 Audio Interface. The measurement signal was a 3-s-long ESS<sup>19</sup> that was played five times for each panel configuration with 2 s of silence in between to allow the sound to fully decay. The total number of measurements was 5342, amounting to 26 710 sweeps recorded with each microphone.

Due to the size of the database and the time required for its collection, the measurements were conducted automatically, without human supervision. Therefore, when an unwanted acoustic event occurred, no action was taken to discard the corrupted recording and repeat the measurement. This approach led to many sweeps being contaminated with non-stationary noise of unknown origin, type, and energy. Examples of sounds recorded during the measurements are available online.<sup>49</sup>

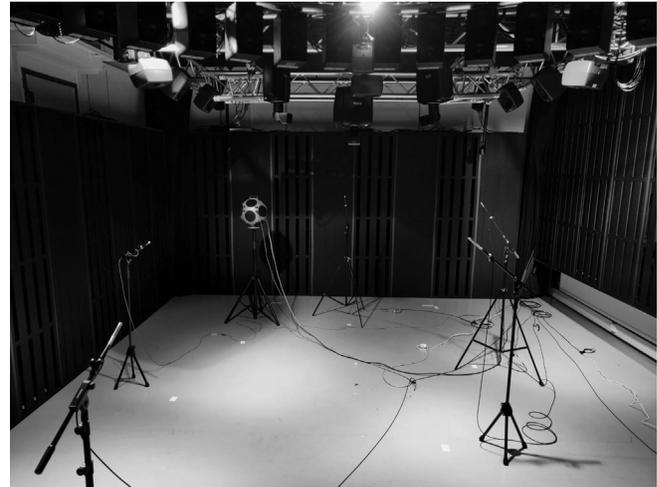


FIG. 6. Variable acoustics laboratory *Arni* and the equipment used in the measurements.

### B. Ro2 measurement and selection procedure

The Ro2 detection proceeds as follows: before every measurement, a short period of silence (background noise) is captured, and its energy is calculated from Eq. (22). Next, an ESS is captured, and its energy is calculated as well. In the event that the noise and sweep signal lengths are different, their energies cannot be compared, and thus, the signal power can be used instead. The procedure is repeated so that two sweeps are measured. The expected PCC value is then computed from Eq. (9), and the lower bound for the expected PCC is calculated from Eq. (11). Next, the tolerance resulting from transfer-function variation [obtained from Eq. (18)] is applied according to Eq. (19). Finally, the detection threshold is compared to the sweeps’ PCC estimated from Eq. (1).

If the measured PCC is on or above the threshold, both sweeps are classified as clean, and the measurement can end. If, on the other hand, the correlation is below the threshold, the presence of non-stationary noise is indicated. The measurement must continue until two sufficiently highly correlated sweeps are obtained. The ESSs which display low correlation with the clean sweeps are marked as contaminated and are discarded.

### C. Transfer-function variation estimation

The transfer-function variation factor was estimated for all measurements based on the difference between signals, as in Eqs. (15)–(17). Since this was done for both clean and contaminated sweeps, many values are skewed in the positive direction due to non-stationary disturbances. To eliminate outliers, values of  $\tau$  that were higher than three median absolute deviations (MADs) from the median were discarded.

The distribution of transfer-function variation factors is displayed in Fig. 7. The figure shows that abnormal values of  $\tau$  start just above the adopted threshold, with a prominent rise in the number of outliers between  $10^{-3}$  and  $10^{-2}$ .

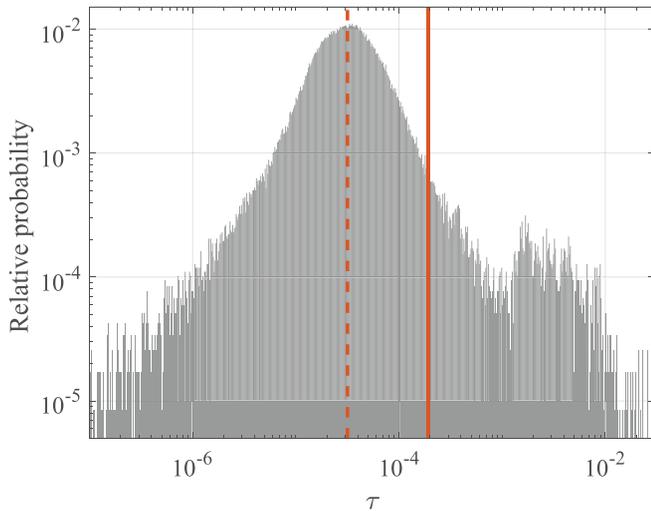


FIG. 7. (Color online) Histogram of measured values of transfer-function variation. The dashed line marks the median, whereas the solid line shows the threshold of three MADs separating the outliers.

In the present study,  $\tau = 0.00019$  (indicated in Fig. 7 with a solid line) was used to set the tolerance to the PCC threshold. Thus, the detection criterion is not as strict as when using the SNR-based threshold. Note, however, that the value of  $\tau$  depends on, among others, the length of the sweep, the time between consecutive measurements, and the characteristics of air movement within the measured space. Therefore, although the  $\tau$  presented here may be used as a guideline for similar conditions, ideally it should be estimated for each measurement scenario separately.

#### D. Correlation-based detection

In the present study, a period of silence before the emission of each sweep was used for background noise energy estimation. It was certain that there would be no late part of the decay or long-ringing modes present in that part of the measurement.

The validation results are presented in Fig. 8, where a two-dimensional histogram shows the relative probability of the expected and measured PCC values. The distribution reveals two clusters: the first one is located along the diagonal, where both the expected and measured PCC values are similar. It contains clean signals and includes the largest number of occurrences (the color bar limits the probabilities to increase readability). The second cluster consists of clearly contaminated sweeps. It is located below the diagonal in Fig. 8, meaning that the measured PCC is considerably lower than the values expected based on Eq. (9).

The SNR-motivated prediction of PCC values is indicated in Fig. 8 with a dashed blue line. Such a threshold is visibly too strict since the majority of measured signals fall below it. This proves that the tolerances due to transfer-function variation and background noise correlation need to be considered in Ro2. The lower bound for the expected PCC,  $\hat{\rho}$  with  $\zeta = -1$ , is indicated in Fig. 8 with a dash-dotted blue line. Although most of the signals are now

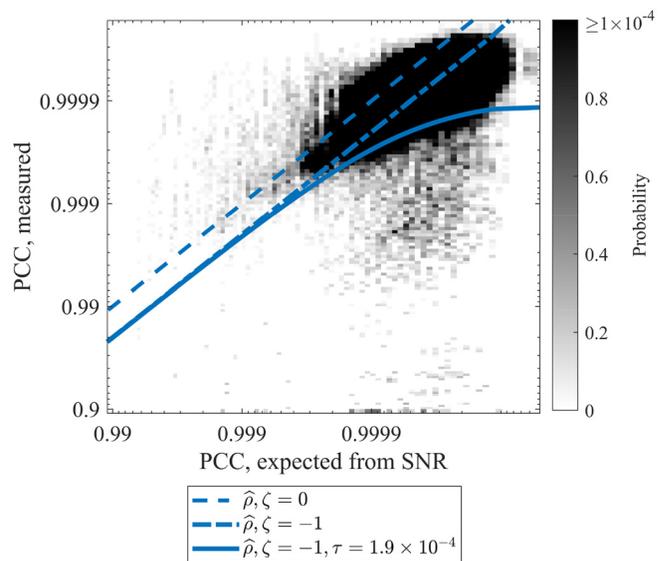


FIG. 8. (Color online) Two-dimensional histogram of expected and measured values of PCC. The dashed line represents the expected values based on the SNR, the dash-dotted line shows the lower bound for  $\hat{\rho}$  with  $\zeta = -1$ , whereas the solid line represents a threshold including transfer-function variation,  $\hat{\rho}$  with  $\zeta = -1$ , and  $\tau = 1.9 \times 10^{-4}$ , which is the proposed threshold.

classified as clean, a large number of measurements still lies below this threshold.

The final detection threshold  $\hat{\rho}$  with tolerance  $\tau$  accounting for time variance is marked with a solid blue line in Fig. 8.  $\hat{\rho}$  with  $\zeta = -1$ , and  $\tau = 1.9 \times 10^{-4}$  identifies most of the sweeps from the top cluster as clean, whereas the signals from the bottom cluster are considered contaminated. This threshold also considers excessive time variation as contamination, due to which sweeps that do not otherwise contain non-stationary noise are discarded. The threshold represented by  $\hat{\rho}$  when  $\zeta = -1$  and  $\tau \neq 0$  is recommended when using the Ro2 procedure.

#### E. Comparison with a previous method

The proposed correlation-based detection is compared with the procedure developed by Guski,<sup>31,32</sup> since it is the only other method created specifically for the purpose of identifying impulsive noise in sweep measurements. In this approach, the detection is conducted by first separating the sweep with the IR from the background noise using the iterative approach by Lundeby *et al.*<sup>50</sup> Then, the logarithmic ratio between the maximum value and the root mean square value of the stationary noise is calculated. If the ratio is in the range of values typical for Gaussian noise, i.e., 12–14 dB, the measurement is classified as clean. However, if the ratio is higher, namely, 20 dB or more, contamination is indicated. Therefore, in the present study, the value of 20 dB is used as the threshold discriminating between clean and corrupted sweeps for the Guski method. The implementation provided in the ITA Toolbox<sup>51</sup> was employed when testing this procedure.

First, the detection rate for non-stationary disturbances from Sec. IV is compared. Since the signals were

TABLE II. Comparison of non-stationary-noise-detection methods for synthesized ESS signals. The better result is highlighted in each row.

Non-stat. noise type	Detected Guski	Detected Ro2
Broadband impulse	78%	<b>95%</b>
Low-passed impulse	48%	<b>95%</b>
Band-passed impulse	63%	<b>95%</b>
Low-frequency noise	30%	<b>95%</b>
Dropouts	0%	<b>75%</b>

synthesized, it was known that  $\zeta = 0$  and  $\tau = 0$ . Thus, the strictest threshold could be used for the Ro2. The results for each type of non-stationary noise are presented in Table II. They show that the Ro2 is superior in terms of separating clean and contaminated sweeps, with detection rates being higher for each type of disturbance.

The last row in Table II reveals that both methods perform worst when sound dropouts occur in the ESS. In the case of the Ro2, however, only the dropouts occurring after about 3.5 s are undetected (cf. Fig. 5). Guski’s method, however, was unable to detect this kind of disturbance altogether. This was an expected result since Guski’s method was not intended for identifying such a type of non-stationary noise.

For the remaining types of non-stationary noise, the Ro2 did not correctly identify the sweeps containing disturbances of low energy (cf. Fig. 3). The Guski method, on the other hand, proved inconsistent in this regard, wrongly classifying the ESSs corrupted with both low- and high-energy non-stationary noise.

The comparison was also performed on the dataset of measured sweeps. The signals that were marked as contaminated by the Ro2 were further analyzed by a human annotator. The measurements were checked in terms of audibility of non-stationary disturbances as well as their visibility in spectrograms, since often the signal itself may mask the contamination, rendering it inaudible. The sweeps falling below the detection threshold due to the excessive transfer function variation were not incorporated in further experiments, as the Guski method was insensitive to time variance.

In the annotation process, 283 contaminated signals were selected to be analyzed with the Guski method. The results of the comparison are shown in Table III. The total number of measurements marked as contaminated by the Ro2 served as a reference, constituting 100% of detected non-stationary disturbances. Seventy percent of these signals were also correctly identified by Guski’s method, while 30% were false positives. The human annotation revealed that the majority of unidentified disturbances were short low-frequency noise bursts. The Guski procedure also overlooked a small number of ESSs including impulsive noise.

Both experiments show that the Ro2 outperforms Guski’s method regardless of the type of contamination. Its efficiency and robustness prove that it is the best available method for separating clean sweeps from those containing

TABLE III. Comparison of non-stationary-noise-detection methods for measured ESS signals.

Non-stat. noise type	Guski	Ro2
Detected	177 (70%)	<b>283 (100%)</b>
Low-freq. noise undetected	77 (27%)	<b>0 (0%)</b>
Impulsive noise undetected	9 (3%)	<b>0 (0%)</b>

non-stationary noise. The Ro2 method can thus be recommended for acoustic measurements in situations where non-stationary noise may occur, which include most practical scenarios.

## VI. CONCLUSION

The paper introduces a novel method, called the rule of two or Ro2, to identify a pair of clean exponential swept sines in a series of repeated sweep measurements. The classification is based on the similarity between the ESS signals, expressed by means of Pearson’s correlation coefficient. A detection threshold separates signals containing expected contamination, such as background noise and time variance, from those contaminated by non-stationary noise. This study also shows that using the median to estimate the background noise energy helps avoid the bias caused by non-stationary events.

If the resulting PCC value between two measured sweeps is above the threshold, the measurements are marked as clean, and both signals can be used in further analysis. If, on the other hand, the correlation is lower than the threshold, the presence of non-stationary noise is indicated and the signals must be discarded. Therefore, the measurement should be repeated until a pair of highly correlated ESSs is found.

In the large set of thousands of experiments reported in this study, the Ro2 procedure proved to be reliable and easily applicable in acoustic measurements. It also performed better than the previous established procedure for non-stationary noise detection, proving its robustness and efficiency. The Ro2 procedure increases the reliability of practical acoustic and audio measurements using sine sweeps.

## ACKNOWLEDGMENTS

This work was supported by the “Nordic Sound and Music Computing Network—NordicSMC,” NordForsk Project No. 86892.

<sup>1</sup>M. R. Schroeder, “New method of measuring reverberation time,” *J. Acoust. Soc. Am.* **37**(3), 409–412 (1965).

<sup>2</sup>A. J. Berkhout, D. de Vries, and M. M. Boone, “A new method to acquire impulse responses in concert halls,” *J. Acoust. Soc. Am.* **68**(1), 179–183 (1980).

<sup>3</sup>J. Pätynen, S. Tervo, and T. Lokki, “Analysis of concert hall acoustics via visualizations of time-frequency and spatiotemporal responses,” *J. Acoust. Soc. Am.* **133**(2), 842–857 (2013).

<sup>4</sup>R. H. C. Wenmaekers, C. C. J. M. Hak, and M. C. J. Hornikx, “How orchestra members influence stage acoustic parameters on five different concert hall stages and orchestra pits,” *J. Acoust. Soc. Am.* **140**(6), 4437–4448 (2016).

- <sup>5</sup>G. Götz, S. J. Schlecht, A. Martinez Ornelas, and V. Pulkki, "Autonomous robot twin system for room acoustic measurements," *J. Audio Eng. Soc.* **69**(4), 261–272 (2021).
- <sup>6</sup>T. Schmitz and J.-J. Embrechts, "Hammerstein kernels identification by means of a sine sweep technique applied to nonlinear audio devices emulation," *J. Audio Eng. Soc.* **65**(9), 696–710 (2017).
- <sup>7</sup>P. Malecki, K. Sochaczewska, and J. Wiciak, "Settings of reverb processors from the perspective of room acoustics," *J. Audio Eng. Soc.* **68**(4), 291–301 (2020).
- <sup>8</sup>J. S. Abel, N. J. Bryan, P. P. Huang, M. Kolar, and B. V. Pentcheva, "Estimating room impulse responses from recorded balloon pops," in *Proceedings of the Audio Engineering Society 129th Convention*, San Francisco, CA (November 4–7, 2010).
- <sup>9</sup>J. Pätynen, B. F. Katz, and T. Lokki, "Investigations on the balloon as an impulse source," *J. Acoust. Soc. Am.* **129**(1), EL27–EL33 (2011).
- <sup>10</sup>M. R. Schroeder, "Integrated-impulse method measuring sound decay without using impulses," *J. Acoust. Soc. Am.* **66**(2), 497–500 (1979).
- <sup>11</sup>J. Borish and J. B. Angell, "An efficient algorithm for measuring the impulse response using pseudorandom noise," *J. Audio Eng. Soc.* **31**(7), 478–488 (1983).
- <sup>12</sup>D. D. Rife and J. Vanderkooy, "Transfer-function measurement with maximum-length sequences," *J. Audio Eng. Soc.* **37**(6), 419–444 (1989).
- <sup>13</sup>C. Dunn and M. J. Hawksford, "Distortion immunity of MLS-derived impulse response measurements," *J. Audio Eng. Soc.* **41**(5), 314–335 (1993).
- <sup>14</sup>G.-B. Stan, J.-J. Embrechts, and D. Archambeau, "Comparison of different impulse response measurement techniques," *J. Audio Eng. Soc.* **50**(4), 249–262 (2002).
- <sup>15</sup>A. Farina, "Simultaneous measurement of impulse response and distortion with a swept-sine technique," in *Proceedings of the Audio Engineering Society 108th Convention*, Paris, France (February 19–22, 2000).
- <sup>16</sup>A. Torras-Rosell and F. Jacobsen, "A new interpretation of distortion artifacts in sweep measurements," *J. Audio Eng. Soc.* **59**(5), 283–289 (2011).
- <sup>17</sup>S. Müller and P. Massarani, "Transfer-function measurement with sweeps," *J. Audio Eng. Soc.* **49**(6), 443–471 (2001).
- <sup>18</sup>P. Guidorzi and M. Garai, "Impulse responses measured with MLS or swept-sine signals: A comparison between the two methods applied to noise barrier measurements," in *Proceedings of the Audio Engineering Society 134th Convention*, Rome, Italy (May 4–7, 2013).
- <sup>19</sup>M. Müller-Trapet, "On the practical application of the impulse response measurement method with swept-sine signals in building acoustics," *J. Acoust. Soc. Am.* **148**(4), 1864–1878 (2020).
- <sup>20</sup>A. Farina, "Advancements in impulse response measurements by sine sweeps," in *Proceedings of the Audio Engineering Society 122nd Convention*, Vienna, Austria (May 5–8, 2007).
- <sup>21</sup>D. C. A. Pantić and D. Radulović, "Transient noise effects in measurement of room impulse response by swept sine technique," in *Proceedings of the 10th International Conference on Telecommunication in Modern Satellite Cable and Broadcasting Services (TELSIKS)*, Nis, Serbia (October 5–8, 2011), pp. 269–272.
- <sup>22</sup>P. Guidorzi, L. Barbarelli, D. D'Orazio, and M. Garai, "Impulse responses measured with MLS or swept-sine signals applied to architectural acoustics: An in-depth analysis of the two methods and some case studies of measurements inside theaters," in *Proceedings of the 6th International Building Physics Conference (IBPC)*, Torino, Italy (August 25–27, 2015), pp. 1611–1616.
- <sup>23</sup>E. Segerstrom, M.-L. Lee, and S. Philbert, "Evaluating four variants of sine sweep techniques for their resilience to noise in room acoustic measurements," in *Proceedings of the Audio Engineering Society 147th Convention*, New York, NY (October 21–24, 2019).
- <sup>24</sup>H. Ochiai and Y. Kaneda, "Impulse response measurement with constant signal-to-noise ratio over a wide frequency range," *Acoust. Sci. Technol.* **32**(2), 76–78 (2011).
- <sup>25</sup>H. Ochiai and Y. Kaneda, "A recursive adaptive method of impulse response measurement with constant SNR over target frequency band," *J. Audio Eng. Soc.* **61**(9), 647–655 (2013).
- <sup>26</sup>Y. Nakahara and Y. Kaneda, "Effective measurement method for reverberation time using a constant signal-to-noise ratio swept sine signal," *Acoust. Sci. Technol.* **36**(4), 344–346 (2015).
- <sup>27</sup>Y. Kaneda, "Noise reduction performance of various signals for impulse response measurement," *J. Audio Eng. Soc.* **63**(5), 348–357 (2015).
- <sup>28</sup>Y. Nakahara and Y. Kaneda, "Improvement of efficiency in reverberation time measurement method using constant signal-to-noise ratio swept sine signal," *Acoust. Sci. Technol.* **37**(3), 133–135 (2016).
- <sup>29</sup>A. Richard, C. L. Christensen, and G. Koutsouris, "Sine sweep optimization for room impulse response measurements," in *Proceedings of Forum Acusticum*, Lyon, France (December 7–11, 2020), pp. 147–154.
- <sup>30</sup>Y. Nakahara, Y. Iiyama, Y. Ikeda, and Y. Kaneda, "Shortest impulse response measurement signal that realizes constant normalized noise power in all frequency bands," *J. Audio Eng. Soc.* **70**, 24–35 (2021).
- <sup>31</sup>M. Guski, "Influences of external error sources on measurements of room acoustic parameters," Ph.D. thesis, RWTH Aachen University, Aachen, Germany, 2015.
- <sup>32</sup>M. Guski and M. Vorländer, "Impulsive noise detection in sweep measurements," *Acta Acust. united Ac.* **101**(4), 723–730 (2015).
- <sup>33</sup>R. O. Duda, *Pattern Classification and Scene Analysis* (Wiley, New York, 1973).
- <sup>34</sup>J. Benesty, J. Chen, and Y. Huang, "On the importance of the Pearson correlation coefficient in noise reduction," *IEEE Trans. Audio Speech Lang. Process.* **16**(4), 757–765 (2008).
- <sup>35</sup>R. Prislán, J. Brunskog, F. Jacobsen, and C.-H. Jeong, "An objective measure for the sensitivity of room impulse response and its link to a diffuse sound field," *J. Acoust. Soc. Am.* **136**(4), 1654–1665 (2014).
- <sup>36</sup>A. C. S. Orcioni and S. Cecchi, "On room impulse response measurement using orthogonal periodic sequences," in *Proceedings of the 27th European Signal Processing Conference (EUSIPCO)*, Coruna, Spain (September 2–6, 2019), pp. 1–5.
- <sup>37</sup>A. Carini, S. Cecchi, and S. Orcioni, "Robust room impulse response measurement using perfect periodic sequences for Wiener nonlinear filters," *Electronics* **9**(11), 1793 (2020).
- <sup>38</sup>A. Carini, S. Cecchi, A. Terenzi, and S. Orcioni, "A room impulse response measurement method robust towards nonlinearities based on orthogonal periodic sequences," *IEEE/ACM Trans. Audio Speech Lang. Process.* **29**, 3104–3117 (2021).
- <sup>39</sup>Note that including mean does not influence PCC values.
- <sup>40</sup>P. Svensson and J. L. Nielsen, "Errors in MLS measurements caused by time variance in acoustic systems," *J. Audio Eng. Soc.* **47**(11), 907–927 (1999).
- <sup>41</sup>T. Niederdränk, "Maximum length sequences in non-destructive material testing: Application of piezoelectric transducers and effects of time variances," *Ultrasonics* **35**(3), 195–203 (1997).
- <sup>42</sup>F. Georgiou, M. Hornikx, and A. Kohlrausch, "Auralization of a car pass-by inside an urban canyon using measured impulse responses," *Appl. Acoust.* **183**, 108291 (2021).
- <sup>43</sup>M. Vorländer and M. Kob, "Practical aspects of mls measurements in building acoustics," *Appl. Acoust.* **52**(3), 239–258 (1997).
- <sup>44</sup>I. Chun, B. Rafaely, and P. Joseph, "Experimental investigation of spatial correlation in broadband reverberant sound fields," *J. Acoust. Soc. Am.* **113**(4), 1995–1998 (2003).
- <sup>45</sup>D. L. Donoho and P. J. Huber, "The notion of breakdown point," in *A Festschrift Erich L. Lehmann* (Wadsworth, Belmont, CA, 1983), pp. 157–184.
- <sup>46</sup>C. Leys, C. Ley, O. Klein, P. Bernard, and L. Licata, "Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median," *J. Exp. Soc. Psychol.* **49**(4), 764–766 (2013).
- <sup>47</sup>P. J. Rousseeuw and C. Croux, "Alternatives to the median absolute deviation," *J. Am. Stat. Assoc.* **88**(424), 1273–1283 (1993).
- <sup>48</sup>K. Prawda, S. J. Schlecht, and V. Välimäki, "Evaluation of reverberation time models with variable acoustics," in *Proceedings of the 17th Sound and Music Computing Conference*, Torino, Italy (June 24–26, 2020), pp. 145–152.
- <sup>49</sup>For more information, see <http://research.spa.aalto.fi/publications/papers/jasa-el-ro2/> (Last viewed March 21, 2022).
- <sup>50</sup>A. Lundebj, T. E. Vigran, H. Bietz, and M. Vorländer, "Uncertainties of measurements in room acoustics," *Acta Acust. united Ac.* **81**(4), 344–355 (1995).
- <sup>51</sup>M. Berzborn, R. Bomhardt, J. Klein, J.-G. Richter, and M. Vorländer, "The ITA-Toolbox: An open source MATLAB toolbox for acoustic measurements and signal processing," in *Proceedings of the 43th Annual German Congress on Acoustics (DAGA)*, Kiel, Germany (March 6–9, 2017), pp. 222–225.