



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 J.; Välimäki, Vesa

Non-stationary noise removal from repeated sweep measurements

Published in: JASA Express Letters

DOI: 10.1121/10.0028203

Published: 01/08/2024

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. J., & Välimäki, V. (2024). Non-stationary noise removal from repeated sweep measurements. JASA Express Letters, 4(8), Article 081601. https://doi.org/10.1121/10.0028203

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.

AUGUST 16 2024

Non-stationary noise removal from repeated sweep measurements

Karolina Prawda 💿 ; Sebastian J. Schlecht 💿 ; Vesa Välimäki 💿

() Check for updates

JASA Express Lett. 4, 081601 (2024) https://doi.org/10.1121/10.0028203





LEARN MORE



Advance your science and career as a member of the Acoustical Society of America

asa.scitation.org/journal/jel

CrossMark



Karolina Prawda,^{1,a)} D Sebastian J. Schlecht,² D and Vesa Välimäki¹ D ¹Acoustics Lab, Department of Information and Communications Engineering, Aalto University, FI-02150 Espoo, Finland ²Multimedia Communications and Signal Processing, Friedrich-Alexander-Universität Erlangen-Nürnberg, 91054 Erlangen, Germany karolina.prawda@york.ac.uk, sebastian.schlecht@fau.de, vesa.valimaki@aalto.fi

Abstract: Acoustic measurements using sine sweeps are prone to background noise and non-stationary disturbances. Repeated measurements can be averaged to improve the resulting signal-to-noise ratio. However, averaging leads to poor rejection of non-stationary high-energy disturbances and, in the case of a time-variant environment, causes attenuation at high frequencies. This paper proposes a robust method to combine repeated sweep measurements using across-measurement median filtering in the time-frequency domain. The method, called Mosaic, successfully rejects non-stationary noise, suppresses background noise, and is more robust toward time variation than averaging. The proposed method allows high-quality measurement of impulse responses in a noisy environment. © 2024 Author(s). All article content, except where otherwise noted, is licensed under a Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

[Editor: Ning Xiang]

https://doi.org/10.1121/10.0028203

Received: 30 April 2024 Accepted: 17 July 2024 Published Online: 16 August 2024

1. Introduction

The exponentially swept sine (ESS) is an excitation signal used in the measurement of room impulse responses (RIRs) since many decades.^{1,2} It is praised for its ability to achieve a high signal-to-noise ratio (SNR) in the resulting RIRs,³ robustness and consistency of measurements,³ and ease of rejection of harmonic distortion.^{2–5} The ESS is, however, sensitive to non-stationary disturbances, which translate to artifacts in the deconvolved RIRs.^{3,6,7} Additionally, if averaging of repeated ESS measurements is used to enhance the SNR, it often leads to time-variance-induced loss of energy in the resulting signal, especially in high frequencies.^{7,8} This work focuses on non-stationary noise removal from repeated sweep measurements while avoiding energy loss. Such methods are particularly useful when the sound environment cannot be controlled fully, such as in case of public spaces, urban environments, outdoor places, or factories with machinery.⁹

Currently, there is a lack of methods to remove non-stationary noise from acoustic measurements. Only few studies known to the authors developed techniques for this specific problem,^{6,10,11} however, not achieving a complete nonstationary noise removal. The literature regarding the removal of clicks from audio signals, such as speech or music,^{12,13} is more comprehensive, with median filtering being a popular signal-processing technique.¹² Those approaches, though, are unsuitable for acoustic measurements, as the transients are only moderately suppressed and not fully rejected.¹³ Additionally, clicks are just one of several non-stationary noise types, which can contain fairly long or tonal noise events.^{6,7}

In a previous study, we faced the problem of non-stationary noise corrupting background noise energy estimates in a detection algorithm.⁷ We found that using a median instead of a mean as an energy estimator significantly improved the robustness of energy evaluation and, in consequence, the non-stationary noise detection process.⁷ Similarly, the median is used in background subtraction techniques in image processing, where the across-frame temporal median filter allows the separation of moving (non-stationary) objects, such as cars and people, from the stationary background image.¹⁴ Median filtering was also used to remove unwanted reflections in semianechoic RIRs.¹⁵

This letter introduces a method that applies across-measurement median filtering to recorded sweeps. As the method outputs the clean sweeps from fragments of ESS measurements, we call it Mosaic. In analogy with across-frame video filtering, we treat the ESS and sound decay as a "stationary background image," while the non-stationary noise is regarded as a "moving object" and filtered out. We show the advantage of time-frequency filtering over time-domain filtering in achieving a measurement with highly diminished non-stationary disturbances. The background noise energy suppression in the filtered sweeps is estimated. We compare the proposed method with commonly used averaging and with state-of-the-art methods of removing transients from acoustic measurements,^{6,10} showing the superiority of this approach.

^{a)}Author to whom correspondence should be addressed.

2. Methodology

This section presents the proposed method for non-stationary noise removal. We discuss both time-domain (Mosaic-T) and time-frequency-domain (Mosaic-TF) variants of median filtering across repeated measurements.

2.1 Median filtering across repeated measurements

Given the possibility of recording an erroneous or corrupted signal, it is common practice to capture more than one sweep during an acoustic measurement session, especially when the environment is noisy. In the previous method, called the Rule of Two (Ro2), we recommended repeating the measurement until at least a pair of clean (i.e., free from non-stationary noise) sweeps is captured.⁷ Performing sufficiently many repetitions, however, is not always feasible, and all captured sweeps may be contaminated. Here, we take advantage of having multiple repetitions of the same measurement to remove non-stationary noise and recover a clean sweep from several contaminated ones.

Given a set of K repeated sweep measurements, $x_1(n), ..., x_K(n)$, the per-sample median filtering in Mosaic-T is conducted:

$$x_{\mathrm{M}}(n) = \mathrm{median}(x_1(n), \dots, x_K(n)). \tag{1}$$

Here, *n* is the discrete time, $x(n) = s(n) * h(n) + u(n) + u_{ns}(n)$, where s(n) is the excitation ESS, h(n) is the RIR, * is the convolution operator, u(n) is the stationary background noise, i.e., noise that does not change its properties during the course of a measurement,⁷ and $u_{ns}(n)$ is the non-stationary noise term. For repeated measurements, the excitation signal is highly similar as it is played from a loudspeaker source, and the measured RIRs only change slightly between measurements.⁷ In contrast, the stationary background noise is independent between measurements.

Median filtering acts as a robust estimator, especially stable against out-of-distribution values, as its breakdown point is 0.5. This means that if less than 50% of data points are contaminated, the estimator will result in the correct value.¹⁶ Therefore, to recover a clean sweep from a series of K measurements based on Eq. (1), at each time n, at least $\lceil K/2 \rceil$ samples need to be free from non-stationary noise, where $\lceil \cdot \rceil$ is the rounding-up operator. In other words, there is no requirement for any of the K sweeps to be completely free of non-stationary noise as long as the disturbances do not occur at the same time. It is crucial, however, that all sweeps are time aligned before Mosaic is applied. Therefore, in this work, we pre-process all ESSs by upsampling by the factor of ten,¹⁷ time-aligning based on the lag of maximum values of cross correlation between two sweeps, and downsampling to the original sampling rate.

An example of non-stationary noise removal from a series of K=3 ESS signals is presented in Figs. 1(a) and 1(b). Every signal was contaminated with a synthesized disturbance: a broadband click of 0.1 s duration (sweeps 1 and 3) or a sine tone of 1 kHz and lasting for 0.6 s (sweep 2). The noise events are marked in red in Fig. 1(a). Mosaic-T following Eq. (1) removes all the disturbances, recovering a clean sweep, as shown in Fig. 1(b).

Figure 1(c) illustrates a histogram of sample values distribution for ESSs from Fig. 1(a) between 2.3 and 3 s from the start of each sweep—at the time of the tonal disturbance in sweep 2. The sample-value distributions are different



Fig. 1. Series of sweep signals contaminated with clicks and tones in (a) time and (d) time-frequency domains, with contaminations marked in red, and sweeps recovered using (b) Mosaic-T and (e) Mosaic-TF. (c) Distribution of sample values between 2.3 s and 3 s from sweeps in (a). (f) Corresponding distributions for sweeps in (d).

when only sweeps 1 and 3 are considered, as they are clean in the given time, to when all sweeps are analyzed. Including the contaminated sweep visibly broadens the distribution. However, the median values are unaffected, showing that, regardless of the non-stationary noise presence, less than 50% of the samples are contaminated.

2.2 Median filtering across time and frequency

Next, we consider median filtering in time and frequency, as the non-stationary disturbances are often sparse in time (clicks) or frequency (tonal disturbances, speech), appearing as either vertical or horizontal stripes in the signal's spectrogram.¹⁸ The spectral representation of an audio signal consists of complex numbers, however, for which the median does not exist. Thus, we propose to treat the real and imaginary parts independently, performing the Mosaic-TF median filtering as follows:

$$X_{\rm M}(n,f) = {\rm median}({\rm Re}(X_1(n,f),...,X_K(n,f))) + \iota \,{\rm median}({\rm Im}(X_1(n,f),...,X_K(n,f))),\tag{2}$$

where X_M is the short-time Fourier transform (STFT) of the median-filtered signal x_M [cf. Eq. (1)] over time *n* and frequency *f*, $X_1, ..., X_K$ are STFTs of single measurements $x_1, ..., x_K$, and *i* is the imaginary unit. In this case, it is possible to reject the non-stationary noise that overlaps in time but occupies different frequency bins.

Figure 1(d) shows the STFTs of a series of sweeps contaminated with synthetic tonal disturbances: a sine of 2 kHz lasting 3 s in sweep 1, a 0.6-s-long 1-kHz tone in sweep 2, and a 2-s-long 5.6-kHz tone in sweep 3. Again, the occurrences of non-stationary noise events are marked in red. The disturbance removal was performed according to Eq. (2). The result of Mosaic-TF is depicted in Fig. 1(e), showing the STFT of a perfectly cleaned ESS.

Figure 1(f) illustrates the distributions of magnitude values in 1-kHz frequency bin, where the tonal disturbance appears in sweep 2. Similarly, as in Fig. 1(c), when only the non-contaminated bins from sweeps 1 and 3 are considered, the distribution is different from when bins of all sweeps are analyzed. Notably, the number of high-magnitude bins increased with the presence of contamination. The medians of both distributions, however, remain unchanged, showing that less than half of all bins contain non-stationary noise. For this reason, the noise removal is perfect in Fig. 1(e).

3. Validation on measurements

In the following, we evaluate the proposed non-stationary noise removal method on two sets of measured ESS signals. We compare Mosaic with commonly used averaging in terms of contamination rejection, noise energy suppression, and robustness toward time variance.

3.1 Evaluation data

The validation of the proposed method was performed on two datasets of sweep measurements: *Arni* and *MRTD. Arni* contains sweeps recorded in the variable acoustic laboratory of the Acoustics Lab of Aalto University in Espoo, Finland. The ESSs were 3-s-long and spanned frequencies from 20 Hz to 20 kHz. Each measurement was repeated five times (K = 5). A more detailed description of the measurement procedure and equipment is found in the literature.^{7,19} We used the Ro2 on *Arni* to identify the sweeps corrupted by non-stationary noise.⁷

MRTD comprises sweeps measured in several spaces of the Otakaari 5 building in Espoo, Finland. The details of measurement setup and locations are presented elsewhere.²⁰ The excitation signals were 2-s-long ESSs repeated three times for each configuration (K= 3). The noisy environment caused non-stationary noise to appear in many of the measured sweeps, with impulsive noise and speech being the most common disturbances.

3.2 Non-stationary noise removal using Mosaic

The performance of Mosaic-T, Mosaic-TF, and averaging is compared in Fig. 2 on the example of three sweeps from the *MRTD* database. Each ESS is contaminated with either speech [cf. Fig. 2(a)] or impulsive noise [cf. Figs. 2(b) and 2(c)].

The results of averaging process are depicted in Fig. 2(d), showing remains of the speech signal (between 0.4 and 1 s) and transients (around 1.1 s) still present. If Mosiac-T is used, the non-stationary noise events are removed, but harmonic distortion is enhanced, as illustrated in Fig. 2(e). As all the disturbances happen at different times and different frequencies in the measured sweeps, they are successfully removed by Mosaic-TF, as shown in Fig. 2(d).

Based on the results in Fig. 2, we recommend that when the measurement environment is noisy, contamination removal is carried out on signal STFTs, according to Eq. (2). Time-domain median filtering might also be used, as harmonic distortion is easily rejected in the deconvolution process. Averaging, however, proved unsuitable for post-processing ESSs containing non-stationary noise. The optimal time-frequency resolution highly depends on the nature of the disturbance. The filtering domain should be chosen such that the sparsity of the disturbances is increased so that repeated occurrence becomes unlikely.

3.3 Background-noise energy suppression

The aim of averaging repeated sweep measurements is to enhance the SNR of the resulting signal by suppressing the background-noise energy. A similar effect can be obtained using across-measurement median filtering. However, the





Fig. 2. (a)-(c) Spectrograms of single contaminated sweeps. (d) Averaged sweep. (e) Mosaic-T. (f) Mosaic-TF. (a) Red rectangle marks the occurrence of speech. (b), (c) Blue rectangles indicate clicks.

median filtering is not as efficient as averaging for stationary noise with Gaussian distribution, resulting in backgroundnoise energy suppression lower by up to a factor of $\pi/2 \approx 3.92 \text{ dB.}^{21}$

To evaluate the success of the proposed method in reducing the background-noise energy compared to averaging, we analyzed the energy of 900 1-s-long measurement snippets from *Arni* dataset, choosing the portion of the signal where neither the sweep nor sound decay was present. To account for the possibility of non-stationary noise contaminating the analyzed snippets, we estimate the background-noise energy using median:⁷

$$\mathbb{E}[u^2] = \sum_{n=1}^{N} u(n)^2 = N b_{\chi}^2 \operatorname{median}(u(1)^2, \dots, u(N)^2),$$
(3)

where $b_{\chi} = 1.4826^{7,16}$ and N is the total number of samples in the background noise signal. We then analyzed the difference in background-noise energy of a single sweep compared to median-filtered or averaged signals when either three or five ESSs were used. The statistical data of this comparison is presented in Table 1.

The results show that all algorithms perform well in line with the theory, both when K=3 (expected background-noise energy suppression is 5.62 dB for median and 9.54 dB for mean) and K=5 (expected result is 10.06 dB and 13.98 dB for median and mean, respectively). Mosaic-TF consistently outperforms Mosaic-T. However, the standard deviations, also given in Table 1, show considerable variability in the performance of all three methods.

3.4 Energy loss due to time variance

In room acoustics, the measurement environment is never strictly time-invariant due to small changes in temperature, humidity, and air movement.⁸ Therefore, a common drawback of cross-measurement averaging is the loss of signal energy (either in the sweep or in the deconvolved RIR) and is most prominent in the high frequencies.⁸

The time-variance-induced energy loss is illustrated in Fig. 3 on an example of ESSs from *MRTD* dataset. The energy of median-filtered and averaged signals is visibly lower than the single-sweep energy toward the end of the ESS, i.e., while the high frequencies are played. This difference is not attributed to the effect of background-noise energy suppression, which should influence the entire sweep in the same way, and not only the high frequencies.

Figure 3 suggests that the time-variance-induced energy loss is more prominent when averaging is used compared to the median-filtering. To test this hypothesis, we compared the energy loss between single sweeps and either averaged or median-filtered sweeps from Arni (900 conditions, K=5) and MRTD datasets (324 conditions, K=3).

Table 1. Background noise energy suppression for different methods and different numbers of used sweeps.

Method K	Mosaic-T		Mosaic-TF		Time-domain averaging	
	3	5	3	5	3	5
Mean (dB) SD (dB)	6.9 6.6	11.1 7.5	7.5 7.0	11.7 7.9	9.8 8.9	14.6 9.8



Fig. 3. (Left) Comparison of signal energy between a single upward sweep, a sweep obtained using Mosaic-TF, and a time-domain averaged sweep. (Right) The energy difference between a single sweep and a median-filtered or an averaged sweep. Dashed vertical line marks the time when the sweep stopped playing after reaching 20 kHz.

The relatively stable environment during data collection for *Arni* results in energy losses around 0.4 dB, with no significant difference between the averaged and median-filtered signals. Much more time-variant measurements for *MRTD* show that using the median-filtering contributes to high frequencies being attenuated by 1.5 dB on average, while using averaging contributes to energy losses of 2.3 dB on average. This shows that Mosaic-TF is more robust toward time-variance-induced energy losses in sweeps than averaging.

4. Comparison with non-stationary noise removal method

In the previous section, we showed that the proposed method is superior to frequently used averaging when it comes to recovering a clean ESS signal in the presence of non-stationary noise. In the following, we compare the Mosaic with the non-stationary noise removal techniques proposed by $\acute{\rm Ciric}$ *et al.*⁶ and Telle and Vary.¹⁰

Ćirić *et al.*⁶ refer to their procedures as A, B, and C. In procedure A, the contaminated part of the measured sweep is replaced with the corresponding normalized fragment of the excitation ESS. In procedure B, the magnitude of the contaminated part is multiplied by the magnitude of the excitation sweep, while the phase stays the same. Procedure C requires prior knowledge of the background noise and the transient itself, as they are used in inverse filtering when deconvolving sweeps into RIRs. Telle and Vary¹⁰ use weighted averaging, where weights for each of K signals are determined based on signal power. The signal with the highest power is assumed contaminated, and thus receives the lowest weight.

In this work, we used a series of five sweeps from *Arni*, where one of the signals is contaminated with transient noise. We carried out procedures A, B, and C on a single sweep, and weighted averaging and proposed Mosaic-TF using K=5 measurements. For procedure C, we estimated the background noise from the parts of measurement without sweep playing, and an RIR obtained from one of the clean sweeps was used as an equivalent of the transient noise.

The comparison results are depicted in Fig. 4. The sweeps in Fig. 4(a) represent the measured signals contaminated with a transient (marked by a blue rectangle). The proposed method, shown in Fig. 4(b) exhibits complete suppression of the impulsive artifact and an added advantage of enhanced SNR. Weighted averaging is illustrated in Fig. 4(c), showing remains of the transient still present. procedure A, depicted in Fig. 4(d), removes the transient, but the



Fig. 4. (a) Measured contaminated sweep, sweeps obtained using (b) Mosaic-TF, (c) weighted average, 10 and (d)–(f) procedures A, B, and C, respectively (Ref. 6). (a) Blue rectangle marks the occurrence of the transient noise.

21 August 2024 07:58:28

information about the sound decay in the contaminated region is lost. Moreover, no fade-in and fade-out in the replaced part of the sweep leads to the appearance of clicks. procedure B, shown in Fig. 4(e), rejects most of the contaminating impulse but also produces clicks. Procedure C completely failed the test, as the non-stationary noise in Fig. 4(f) is identical to the original transient from Fig. 4(a).

To quantify the noise removal performance of the compared methods, we calculated the Pearson correlation coefficients (PCCs) between the sweeps shown in Fig. 4 and one of the clean sweeps from the measurement. PCC is an indicator of contamination in a measurement,⁷ and thus evaluates the success of non-stationary noise removal methods. The PCC values for the compared techniques are displayed in the top-left part of each pane in Fig. 4. Mosaic-TF shows the highest PCC value and the biggest improvement from the initial PCC of 0.9983 in Fig. 4(a). Additionally, Mosaic is applied to the sweeps without prior knowledge of any signal parameters (power,¹⁰ contamination location⁶). The proposed method appears to be superior to all four procedures and thus should be used when the removal of non-stationary noise is necessary.

5. Conclusion

This letter proposes a method to remove non-stationary noise from repeated acoustic swept-sine measurements. The proposed across-measurement median filtering, called Mosaic, rejects disturbances and recovers a clean sweep suitable for further processing, when less than half of the measurements are contaminated at each sample time or frequency bin. We show that performing the median filtering in the time-frequency domains gives the best results.

Comparison with commonly used across-measurement averaging shows the advantages of the proposed method in terms of non-stationary noise rejection and robustness against the time-variance-induced energy loss. In the test against other techniques for non-stationary noise removal, the proposed method offered significant improvements without further corrupting the resulting sweep and with no prior knowledge of the signal or contamination.

This letter teaches that, when measuring in an uncontrolled environment, it is recommended to use multiple shorter ESSs over one very long ESS. The ESSs should still be long enough for obtaining a sufficient SNR and harmonic distortion rejection. The proposed technique can virtually remove any time-frequency sparse non-stationary noises given sufficiently many repetitions.

Author Declarations

Conflict of Interest

The authors have no conflict of interest to disclose.

Data Availability

The raw sweep data are available from the authors upon request. The sound examples that support claims made in this work and related code are available online (https://github.com/KPrawda/mosaic_noise_removal).

References

- ¹S. Norcross and J. Vanderkooy, "A survey of the effects of nonlinearity on various types of transfer-function measurements," in *Proceedings* of the Audio Engineering Society 101st Convention, New York, NY (1995).
- ²A. Farina, "Simultaneous measurement of impulse response and distortion with a swept-sine technique," in *Proceedings of the Audio Engineering Society 108th Convention*, Paris, France (2000).
- ³G.-B. Stan, J.-J. Embrechts, and D. Archambeau, "Comparison of different impulse response measurement techniques," J. Audio Eng. Soc. **50**(4), 249–262 (2002).

⁴S. Müller and P. Massarani, "Transfer-function measurement with sweeps," J. Audio Eng. Soc. **49**(6), 443--471 (2001).

- ⁵M. 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).
- ⁶D. Ćirić, 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 (2011), pp. 269–272.
- ⁷K. Prawda, S. J. Schlecht, and V. Välimäki, "Robust selection of clean swept-sine measurements in non-stationary noise," J. Acoust. Soc. Am. **151**(3), 2117–2126 (2022).

⁸B. N. J. Postma and B. F. G. Katz, "Correction method for averaging slowly time-variant room impulse response measurements," J. Acoust. Soc. Am. 140(1), EL38–EL43 (2016).

⁹F. 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).

- ¹⁰A. Telle and P. Vary, "A novel approach for impulse response measurements in environments with time-varying noise," in *Proceedings of the 20th International Congress on Acoustics (ICA)*, Sydney, Australia (2010), pp. 1–5.
- ¹¹P. Massé, T. Carpentier, O. Warusfel, and M. Noisternig, "A robust denoising process for spatial room impulse responses with diffuse reverberation tails," J. Acoust. Soc. Am. 147(4), 2250–2260 (2020).
- ¹²C. Chandra, M. Moore, and S. Mitra, "An efficient method for the removal of impulse noise from speech and audio signals," in *Proceedings* of 1998 IEEE International Symposium on Circuits and Systems (ISCAS), Monterey, CA (1998). Vol. 4, pp. 206–208.

21 August 2024 07:58:28



- ¹³P. C. Yong and S. Nordholm, "Real time noise suppression in social settings comprising a mixture of non-stationary and transient noise," in Proceedings of the 25th European Signal Processing Conference (EUSIPCO), Kos, Greece (2017), pp. 588–592.
- ¹⁴Y. Zhang, W. Zheng, K. Leng, and H. Li, "Background subtraction using an adaptive local median texture feature in illumination changes urban traffic scenes," IEEE Access 8, 130367–130378 (2020).

¹⁵F. Denk, B. Kollmeier, and S. D. Ewert, "Removing reflections in semianechoic impulse responses by frequency-dependent truncation," J. Audio Eng. Soc. **66**(3), 146–153 (2018).

¹⁶P. J. Rousseeuw and C. Croux, "Alternatives to the median absolute deviation," J. Am. Stat. Assoc. 88(424), 1273–1283 (1993).

¹⁷V. Välimäki and S. Bilbao, "Giant FFTs for sample-rate conversion," J. Audio Eng. Soc. 71(3), 88–99 (2023).

¹⁸L. Fierro and V. Välimäki, "Enhanced fuzzy decomposition of sound into sines, transients, and noise," J. Audio Eng. Soc. 71(7), 468–480 (2023).

- ¹⁹K. Prawda, S. J. Schlecht, and V. Välimäki, "Calibrating the Sabine and Eyring formulas," J. Acoust. Soc. Am 152(2), 1158–1169 (2022).
- ²⁰P. Götz, G. Götz, N. Meyer-Kahlen, K. Y. Lee, K. Prawda, E. A. P. Habets, and S. J. Schelcht, "A multi-room transition dataset for blind estimation of energy decay," in *Proceedings of the 18th International Workshop on Acoustic Signal Enhancement (IWAENC)*, Aalborg, Denmark (2024).
- ²¹L. Yin, R. Yang, M. Gabbouj, and Y. Neuvo, "Weighted median filters: A tutorial," IEEE Trans. Circuits Syst. II: Analog Digital Signal Process. **43**(3), 157–192 (1996).