JASA Express Letters

AUGUST 13 2025

Cropping room impulse responses using unimodal regression of their covariance

Karolina Prawda ; Nils Meyer-Kahlen ; Sebastian J. Schlecht

JASA Express Lett. 5, 081601 (2025)
https://doi.org/10.1121/10.0038960

JASA EXPRESS LETTERS ARTICLE asa.scitation.org/journal/jel

Cropping room impulse responses using unimodal regression of their covariance

Karolina Prawda,[1,a] Nils Meyer-Kahlen,[2] and Sebastian J. Schlecht[3]
[1]AudioLab, School of Physics, Engineering and Technology, University of York, York, United Kingdom
[2]Acoustics Lab, Department of Information and Communications Engineering, Aalto University, FI-02150 Espoo, Finland
[3]Multimedia Communications and Signal Processing, Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), Erlangen, Germany
karolina.prawda@york.ac.uk, nils.meyer-kahlen@aalto.fi, sebastian.schlecht@fau.de

Abstract: The presence of unavoidable background noise limits the signal-to-noise ratio in measured room impulse responses (RIRs). A common solution is to crop the RIR to the time interval where the signal dominates the background noise, but finding the correct onset and truncation points is challenging. It usually requires estimating the sound decay rate and noise floor, which is burdened with uncertainty. In this study, we propose an RIR cropping method based on the covariance between two repeated RIRs and its inherent monotonicity. Evaluation on measured RIRs shows the proposed method is highly robust in different scenarios and outperforms state-of-the-art algorithms.

[Editor: Stefan Bilbao] https://doi.org/10.1121/10.0038960

Received: 24 June 2025 **Accepted:** 29 July 2025 **Published Online:** 13 August 2025

1. Introduction

Measured room impulse responses (RIRs) naturally exhibit a limited signal-to-noise ratio (SNR); at a certain point in the response, the noise energy exceeds that of the decaying RIR. To prevent erroneous estimation of acoustic parameters—such as reverberation time (RT)[1–3]—or energetic measures like clarity and the direct-to-reverberant energy ratio, the noise beyond this point must be removed. Additionally, convolving an RIR containing such noise with an anechoic source signal can result in an audible artifact sometimes referred to as "frozen reverb."[4] When generating training data for machine learning algorithms related to room acoustics, this artifact may also hinder generalization to real-world data. To mitigate issues associated with limited SNR, the most straightforward solution is to crop the RIR at the point where the noise energy begins to dominate over the decaying reverberation,[1,5,6] i.e., the time t when the instantaneous SNR, $\mathrm{SNR}(t)$, is non-positive.

Many solutions to the RIR cropping problem have been proposed; however, they typically require estimating the noise floor level and the decay rate of the RIR, followed by an iterative process to determine the optimal truncation point.[2,4–8] These methods often assume a single-sloped exponential decay[2,9] or require spatial RIRs measured with a microphone array.[4,10] In addition to truncation, determining the onset time of an RIR is a nontrivial task, and errors in onset detection can lead to inaccurate estimation of decay parameters[11] and impair the performance of the aforementioned truncation point detection algorithms.

During acoustic measurements, it is common practice to record RIRs multiple times to mitigate the effects of errors or unpredictable noise events. The literature shows that such repeated measurements are required in various tasks, including non-stationary noise detection[12,13] and removal,[14] assessment of RIR variability,[15–17] speed of sound estimation,[18,19] and anomaly detection.[20,21] Although never strictly identical, consecutively measured RIRs display high similarity to each other where $\mathrm{SNR}(t)$ is high and very low similarity for non-positive $\mathrm{SNR}(t)$.[12,17] In this study, we leverage this relationship to distinguish the time interval containing the RIR from the portion dominated by background noise.

This letter proposes a method for RIR cropping using a pair of consecutively measured responses. By analyzing two signals instead of one, we leverage their covariance to distinguish the useful portion of the RIR from background noise. To stabilize the covariance over time and derive a reliable threshold, we apply smoothing via unimodal regression. The proposed method is evaluated against several baseline approaches for both truncation and onset detection. Compared to state-of-the-art techniques, our method relies on fewer assumptions about the onset or decay behavior of RIRs, making it broadly applicable—even to RIRs that comprise fade-in effects, multi-sloped decay, or very low SNR.

[a]Author to whom correspondence should be addressed.

2. Methodology

This section presents the proposed methodology for estimating onset and truncation points, discusses RIR and background noise covariance, and explains the smoothing procedure using unimodal regression.

2.1 Covariance and noise energy

Two repeated acoustic measurements x and x' are modeled as follows:[12,17]

$$x(t,f) = h(t,f) + u(t,f) \quad \text{and} \quad x'(t,f) = h'(t,f) + u'(t,f), \tag{1}$$

where h and h' are RIRs and u and u' are stationary background noise terms at time t and frequency f. The covariance between x and x' is

$$\sigma_{x,x'}(t,f) = \mathrm{cov}(x,x'^*) = |\mathbb{E}[(h+u)(h'+u')^*]| = |\mathbb{E}[hh'^*] + \mathbb{E}[hu'^*] + \mathbb{E}[h'^*u] + \mathbb{E}[uu'^*]|, \tag{2}$$

where $\mathbb{E}$ is the expected value and $()^*$ denotes the complex conjugate. Note that in this work, we omit the time-frequency dependency where appropriate for conciseness. For discrete signals, $\mathbb{E}$ is approximated by a short-time average. The short-time covariance can be estimated through the averaged sample covariance

$$\hat{\sigma}_{x,x'}(t,f) = w(t) * |x(t,f)\,x'^*(t,f)| = \sum_k w(t-k)|x(k,f)\,x'^*(k,f)|, \tag{3}$$

where $w(t) = 1/T$ for $-T/2 < t < -T/2$ and 0 otherwise, T is the window length, and $*$ is the convolution operator.

We assume that the RIR and background noise are uncorrelated; thus, $\mathbb{E}[hu'^*] = \mathbb{E}[h'^*u] = 0$. Ideally, the background noise terms are also assumed to be uncorrelated with each other, resulting in $\mathbb{E}[uu'^*] = 0$,[12,17] and thus $\sigma_{x,x'}(t,f)$ would tend to 0 before the onset of the RIR and after the RIR energy decays below the noise floor. Hence, in such an ideal scenario, the onset point t_O and the truncation point t_T would be found at

$$t_\mathrm{O} = \arg\min_t \sigma_{x,x'}(t,f) \geq \epsilon, \quad t_\mathrm{T} = \arg\max_t \sigma_{x,x'}(t,f) \geq \epsilon, \tag{4}$$

i.e., the earliest and the latest point where the covariance exceeds a threshold value ϵ, respectively, such that $t_\mathrm{O} < t_\mathrm{T}$.

However, it is possible for the measured background noise terms u and u' to exhibit some correlation, e.g., arising from electric humming.[12] The two most extreme cases are when the noise terms are fully correlated or fully anticorrelated, $u = \pm u'$, resulting in $\mathbb{E}[uu'^*] = \pm\mathbb{E}[u^2] = \pm\mathbb{E}[u'^2]$. In this study, the case of fully correlated background noise is treated as the worst-case scenario, representing the highest possible noise covariance values. Hence, we set the threshold for onset and truncation point estimation as the first and last time, respectively, when the covariance curve exceeds the covariance of correlated noise terms, $\epsilon = \mathbb{E}[u^2]$:

$$t_\mathrm{O} = \arg\min_t \sigma_{x,x'}(t,f) \geq \mathbb{E}[u^2], \quad t_\mathrm{T} = \arg\max_t \sigma_{x,x'}(t,f) \geq \mathbb{E}[u^2]. \tag{5}$$

To obtain the threshold values in Eq. (5), it is necessary to consider parts of the measured signal where only background noise is present. In acoustic measurements, it is recommended that a certain recording time is allowed before or after the excitation signal is emitted to estimate the background noise energy. Depending on the measurement conditions and the procedure, the length of the noise signal will vary. However, under the assumption of stationary background noise, this should not have a significant impact on $\mathbb{E}[u^2]$, which is evaluated in Sec. 3.2. Additionally, noise energy can be robustly estimated using median of the noise samples.[12,14]

2.2 Unimodal regression

As the short-time covariance curve of a pair of measured RIRs reflects their natural noisiness, choosing the right window size to obtain a good compromise between the level of detail in the covariance curve and its smoothness is crucial. Depending on the signal sparsity in time, the optimal window size may vary between measurements or even throughout the RIR duration; thus, finding the best parameters is challenging. Hence, we propose using unimodal regression for smoothing, as it does not require choosing the window size.

As the RIR energy decays over time and reaches the noise floor, so does the covariance.[17] At this point, it is certain that the RIR energy will not increase anymore above the noise level; thus, the covariance cannot grow either, and any high $\sigma_{x,x'}$ values from the noisy RIR parts stem from the background noise and do not indicate a useful portion of the signal reappearing. Similarly, the covariance values are low before the onset of the RIR, as in that region, the measurement consists solely of the background noise. Thus, a covariance curve of a pair of RIRs will exhibit only one global maximum.

Exploiting this property, we propose covariance smoothing with unimodal regression, which seeks to minimize a weighted least squares error:

$$\min_{\tilde{\sigma}_{x,x'}} \sum_t w_t(\tilde{\sigma}_{x,x'}(t,f) - \hat{\sigma}_{x,x'}(t,f))^2, \tag{6}$$

subject to the constraint that for each f, there exists one $t_{\max}$ such that[22]

$$\begin{aligned}&\tilde{\sigma}_{x,x'}(t_i,f) \le \tilde{\sigma}_{x,x'}(t_j,f) \le \tilde{\sigma}_{x,x'}(t_{\max},f), \text{ where } t_i \le t_j \le t_{\max}, \text{ and}\\ &\tilde{\sigma}_{x,x'}(t_{\max},f) \ge \tilde{\sigma}_{x,x'}(t_k,f) \ge \tilde{\sigma}_{x,x'}(t_l,f), \text{ where } t_{\max} \ge t_k \ge t_l,\end{aligned} \tag{7}$$

where $\tilde{\sigma}_{x,x'}$ is the covariance curve $\hat{\sigma}_{x,x'}$ smoothed using unimodal regression and weights w_t are strictly positive.[23]

The advantage of this approach is that for each part of Eq. (7), all functions are chosen from a class of isotonic (non-increasing or non-decreasing) functions, without assuming any shape of the target function, as opposed to linear regression,[24] for example. In terms of room acoustics, this means no other model assumption about RIRs need to be made, apart from having a single maximum. Hence, the presence of single- or multi-sloped decays, fade-in effect, or the RIR dynamic range does not violate this assumption. This is in contrast to other truncation methods, which seek to fit a decay model with a single slope to find the truncation point.[5,6] Unimodal regression also does not assume direct sound to be the highest energy component of an RIR, unlike some of the onset finding methods.[11]

Smoothing using unimodal regression is illustrated in the left pane of Fig. 1. The peaks and troughs of the decaying covariance curve are efficiently equalized, forming a steadier trend. Similarly, the covariance values of the noise floor region are evened out.

3. Validation

The proposed method was first validated on a set of simulated RIRs. We used two test cases: in the first one, the RIRs were synthesized by multiplying white Gaussian noise with a decaying exponential function, so that a frequency-independent RT of 1 s was achieved; in the second one, the fade-in effect was simulated by combining two decaying RIRs, one of which was multiplied with a negative gain, imitating a measurement where the source and receiver are in different rooms.[25,26] The fade-in RIRs had a frequency-independent RT of 2 s.

To simulate a typical measurement, we added non-decaying random noise to each of the RIRs to mimic background noise. To consider different scenarios, the background noise terms were identical (fully correlated noise), identical with flipped signs (anticorrelated noise), different random sequences (uncorrelated noise), or different sequences with a correlation of 0.2 (weakly correlated noise). The fully correlated noise condition represents the situation in which both measurements are the same, $x = x'$, i.e., one measurement is used instead of two, and the autocovariance is calculated in Eq. (3). For each condition, the background noise gain was varied to achieve SNRs ($\max_t \mathrm{SNR}(t) = \max_t \mathbb{E}[h(t)]/\mathbb{E}[u(t)]$) between 10 and 30 dB.

3.1 Onset and truncation points

In this case, the energy of noiseless RIRs and the noise energy are known; therefore, the ground truth (GT) onset and truncation points are determined as the times at which the $\mathrm{SNR}(t) = 0$ dB before increasing (onset) and after decaying (truncation). The determination of t_{O} and t_{T} using the proposed method is illustrated in the right pane of Fig. 1 for the example of 30 dB of SNR. The $\tilde{\sigma}_{x,x'}(t,f)$ values are almost identical for all RIR pairs when the $\mathrm{SNR}(t)$ is positive—between around 0.5 and 1 s—but differ for non-positive $\mathrm{SNR}(t)$ depending on the correlation of the noise terms.

The relation between noise terms correlation and the onset and truncation point estimation is shown in Table 1. The results are displayed as the difference between the GT onset or truncation point and the estimate given by the proposed method, according to the formula $\Delta t_{\mathrm{X}} = 100\% \times |\hat{t}_{\mathrm{X}} - t_{\mathrm{X}}|/t_{\mathrm{X}}$, where $\mathrm{X} \in \{\mathrm{O}, \mathrm{T}\}$ and $\hat{t}_{\mathrm{X}}$ signifies values obtained with the proposed method, for onset and truncation point determination, respectively.

The validation results show that in cases of uncorrelated, weakly correlated, and anticorrelated noise, the estimation errors do not exceed 8.2% for t_{T} and 4.2% for t_{O}, proving the method is successful for typical and atypical RIRs and in low SNRs. However, the differences grow significantly when the simulated background noise is fully correlated, as the transition from the covariance of the RIR to background noise is less steep, as shown in the right pane of Fig. 1. To remedy this, we reduced the correlation of the noise terms by adding a random noise sequence to one of the RIRs from the

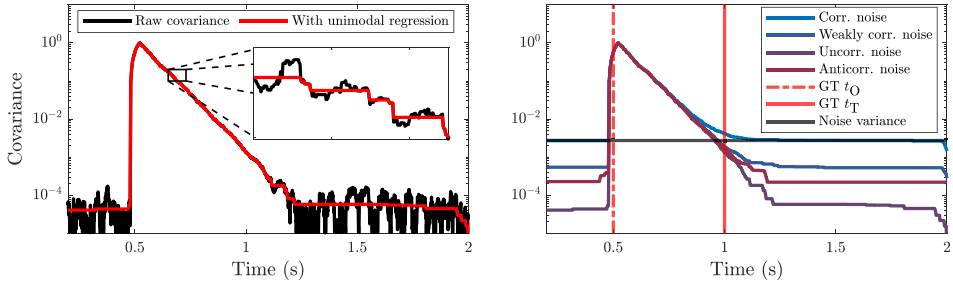

Fig. 1. Left: Covariance of a pair of RIRs and smoothing with unimodal regression. Right: Onset and truncation point estimation using RIR covariance and noise variance.

Table 1. Onset and truncation time differences between GT and proposed method for simulated RIRs with and without fade-in affect, SNRs between 10 and 30 dB, and different correlation of background noise terms. All Δt_O and Δt_T are expressed as percentages of the true times.

Test case	SNR (dB)	10		15		20		25		30	
		Δt_O	Δt_T	Δt_O	Δt_T	Δt_O	Δt_T	Δt_O	Δt_T	Δt_O	Δt_T
No fade-in	Uncorrelated	2.6	4.3	3.8	4.5	4.0	5.1	4.2	5.4	4.2	4.5
	Weakly correlated	3.0	1.1	3.9	2.6	4.2	3.7	4.2	3.6	4.2	3.1
	Anticorrelated	3.2	4.6	3.9	4.3	4.2	4.9	4.2	4.3	4.2	3.9
	Correlated	80.3	176.4	80.3	145.7	80.3	121.1	80.3	101.0	80.3	84.3
	Correlated + uncorrelated	4.0	24.3	4.2	43.6	4.2	73.9	4.2	58.1	4.2	44.9
Fade-in	Uncorrelated	0.9	0.4	1.6	0.9	3.2	0.1	3.9	1.1	4.2	0.2
	Weakly correlated	0.1	4.1	2.2	2.1	3.3	3.6	3.9	2.8	4.2	1.0
	Anticorrelated	0.3	8.2	2.0	1.4	3.2	1.8	3.9	0.0	4.2	0.6
	Correlated	80.3	121.1	80.3	84.3	80.3	58.3	80.3	38.8	80.3	32.1
	Correlated + uncorrelated	3.3	73.8	3.9	84.7	4.2	58.3	4.2	48.5	80.3	32.1

pair. The results in Table 1 for the correlated plus uncorrelated case show that such a trick can offer improvements, especially in the onset point detection, where the errors were reduced to 3–4%. This highlights that the proposed method typically requires two separate measurements instead of a single one.

3.2 Noise covariance

Under the assumption of stationary background noise, the length of the noise signal should not have an impact on the threshold in Eq. (5). Here, we evaluate this hypothesis for the simulated RIRs.

The $\mathbb{E}[u^2]$ of the RIRs was calculated for the noise lengths of 100 ms, 200 ms, 500 ms, 700 ms, and 1 s and used in the case of uncorrelated background noise. The resulting onset and truncation times were then compared to the GT in the same way as in Sec. 3.1. The Δt_O and Δt_T values for each SNR condition were averaged over all analyzed noise lengths and are shown in Table 2, with standard deviations indicated only when they were greater than zero. The results show excellent agreement with the numbers for the uncorrelated noise cases in Table 1, proving that the length of noise signal for threshold estimation has minimal impact on the accuracy of the proposed method, as long as the assumption of stationary background noise is met.

4. Evaluation

We evaluated the proposed method on two datasets: *Arni*, which contains RIRs measured repeatedly in the same room over a long period of time[27] and the multi-room transition dataset (*MRTD*),[28] which contains RIRs from multi-room scenarios, and three repetitions of each measurement. The details of the measurement setups are described elsewhere.[27,29] Naturally, no GT data are available for real, measured RIRs. Thus, the evaluation focuses on whether any plausible estimate is returned and on the spread of these estimated values.

As shown in Sec. 3, the proposed method works best with a pair of RIRs. Therefore, in the case of *Arni*, we used two consecutive measurements to determine the onset and truncation points. Repetitions in *MRTD* often contained non-stationary noise, such as transients or speech,[14] which could hinder the evaluation. Therefore, in this study, we used the output of the Mosaic-TF method,[14] as it removes such noise, and the RIR which had the highest correlation to the Mosaic RIR, as this indicated that it contained the least noise of the three repetitions.[12] For both datasets, the threshold $\mathbb{E}[u^2]$ was determined from the first RIR from each pair by calculating the variance of 0.5 s of the measurement, where only the background noise was present.

4.1 RIR truncation

To the authors' best knowledge, there exists no method prior to the proposed one that estimates the onset and the truncation point of an RIR at the same time. Therefore, we split the evaluation into two parts.

First, we assess the estimation of the truncation point. For this purpose, the most established approach found in the literature is the Lundeby method,[5] which determines the decay by finding a single slope and the background noise level by fitting a constant to a short-time averaged RIR. The fit is performed iteratively. The truncation point is then established as an intersection point between the energy decay curve and the noise floor level, but it is recommended to leave 5–10 dB of safety margin above the noise floor level before truncation,[1,5] which is consistent with the suggestion from standards that describe RT estimation.[30,31] In this work, we use the Lundeby method implementation provided in the pyfar library.[32]

Figure 2(a) presents the results of the truncation point estimation on 595 RIRs from the *Arni* dataset. Since all of the RIRs were measured in the same room under the same conditions, we expect very similar truncation times, subject to

Table 2. Effect of noise length in onset and truncation point determination. All $\overline{\Delta t_O}$ and $\overline{\Delta t_T}$ are expressed as percentages. Standard deviations are only indicated when they are nonzero.

SNR (dB)	10		15		20		25		30	
	$\overline{\Delta t_O}$	$\overline{\Delta t_T}$	$\overline{\Delta t_O}$	$\overline{\Delta t_T}$	$\overline{\Delta t_O}$	$\overline{\Delta t_T}$	$\overline{\Delta t_O}$	$\overline{\Delta t_T}$	$\overline{\Delta t_O}$	$\overline{\Delta t_T}$
No fade-in	2.7	4.3 ± 0.1	3.8	4.5 ± 0.1	4.0	5.1	4.2	5.4	4.2	4.5
Fade-in	0.9	0.4	1.6	0.9	3.2	0.1	3.9	1.1	4.2	0.2

variations due to time variance and fluctuations in the background noise level. The results show that the truncation point values are very similar for both methods, with the proposed approach consistently resulting in a later truncation. The observed differences between frequency bands are expected because of different decay times and background noise levels. However, since we do not know the GT t_T, we cannot assess which method was more accurate in estimating it.

Figure 2(a) shows that in lower frequencies, 250–500 Hz, the proposed method shows more variability than Lundeby, but starting from 1 kHz, the standard deviations of both distributions are almost identical, showing that the performance of the proposed method is comparable to the baseline in such typical conditions. An example of an *Arni* RIR envelope with both truncation points marked is depicted in Fig. 2(c), confirming that both methods result in similarly plausible values.

Figure 2(b) presents the outcome of the evaluation on 50 RIR pairs from the *MRTD* dataset. As each RIR pair was captured at a different location, consistent t_T values are not expected. The proposed method produced a truncation point for every evaluated condition, with similar values over different measurements, suggesting the absence of large estimation errors. The Lundeby method, on the other hand, failed in 84% of the test cases, resulting in t_T of 0 s. This is likely due to the low SNR in many RIRs, and the presence of multi-sloped decays. Both of those issues violate the model assumptions of the Lundeby method and make the noise floor level and decay slope estimation difficult. The examples of RIR envelopes from *MRTD* are illustrated in Fig. 2(d) and 2(e), showing that the proposed method works well even in very low SNR conditions, as well as in a scenario where finding a single decay slope is challenging.

4.2 Onset point determination

Next, we compare the proposed method to four onset time detection approaches presented in Ref. 11: using the time of the maximum absolute value of the RIR (M), reducing it by 5 ms (M_5), mean over time (D_E), and threshold (E). We use RIRs from the *MRTD* dataset—the same 50 pairs as in Sec. 4.1—which include both line-of-sight scenarios and those with occlusions that introduce a fade-in effect.[25,26]

Figure 3 presents the results for onset point detection using the five compared methods. As with truncation point detection in measured RIRs, GT t_O is unavailable, so no method's result can be definitively deemed correct. However, Fig. 3a shows that the proposed method and D_E yield the most consistent onset estimates across six frequency bands. Since the speed of sound in air is assumed frequency-independent, we expect t_O to remain constant or nearly constant across frequencies—an outcome that indicates robustness. In contrast, both M and M_5 display significant variability across

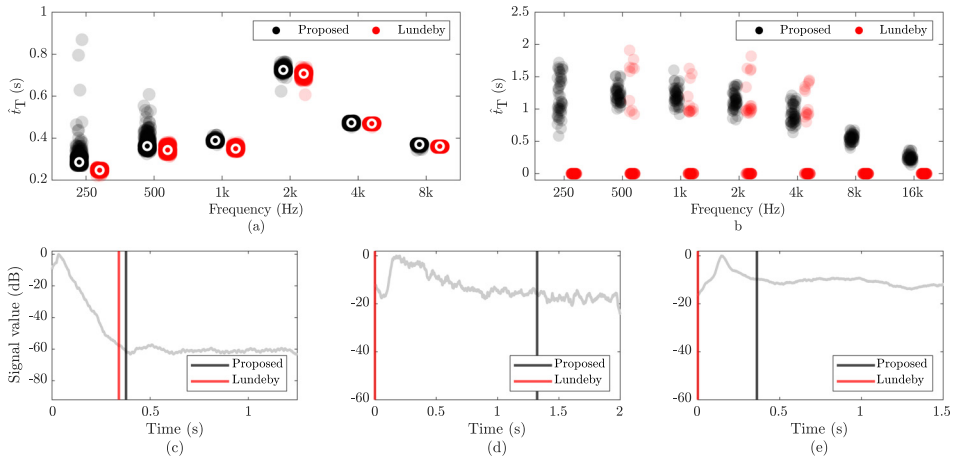

Fig. 2. Top: Truncation points over frequency for (a) *Arni* dataset and (b) *MRTD* dataset. (a) The white circles mark the medians of t_T distributions. Bottom: Examples of RIR envelopes with indicated truncation points for (c) *Arni* and (d, e) *MRTD*.

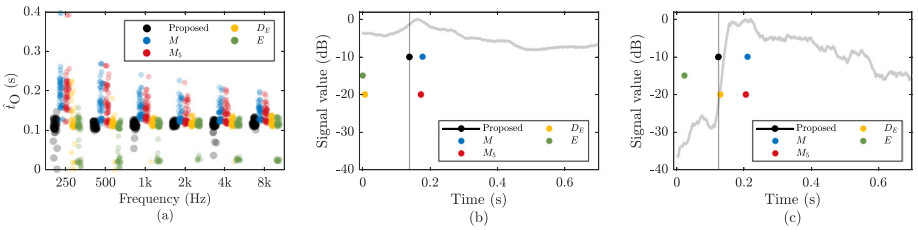

Fig. 3. (a) Onset points for over frequency for *MRTD* dataset. (b, c) Examples of *MRTD* RIR envelopes with onset points marked.

frequencies, suggesting less reliable estimates. Additionally, the distributions of t_O values within each frequency band are much wider for M and M_5, further indicating a lack of robustness. While D_E and E show improved stability, both methods—particularly E—yield erroneously low onset times for a subset of RIRs.

These concerns are further illustrated in Fig. 3(b) and 3(c), which show RIR envelopes with the fade-in effect. For these signals, the maximum values occur after the true onset, leading to large errors when using M and M_5. In Fig. 3(b) both D_E and E predict onset times that are clearly too low, likely due to very low SNR. Figure 3(c) shows that the proposed method and D_E perform similarly, estimating t_O near the correct point. However, the slightly rising noise floor prior to the onset causes E to underestimate the t_O. The proposed method accurately identifies the onset time, demonstrating superior robustness.

5. Conclusion

This study presents a method for RIR cropping based on the covariance between a pair of RIRs, smoothed using unimodal regression. The threshold for estimating onset and truncation points is set as the variance of the background noise term—covariance values above this indicate a positive instantaneous SNR and thus the useful portion of the signal. The method assumes only that the covariance has a global maximum and that the background noise is stationary, eliminating the need to estimate the decay slope.

Validation on simulated data shows the proposed method achieves errors below 8% even at very low SNRs, provided the background noise terms in the RIR pair are not fully correlated. When noise is fully correlated, i.e., when covariance is calculated from a single RIR, large errors occur, demonstrating that using a pair of consecutively measured RIRs yields more robust estimates.

Evaluation on measured RIRs shows the proposed method performs comparably to baseline truncation methods on high-SNR, single-sloped RIRs. In low-SNR, multiple-slope conditions, it proves significantly more robust. For onset detection, the proposed method outperforms the baselines, demonstrating strong robustness and resistance to fade-in effects as well as slight increases in noise floor energy immediately before the RIR onset.

In conclusion, the proposed approach is the most robust and reliable among all considered methods for both onset and truncation detection. This letter demonstrates that using the unimodal covariance of a pair of RIRs is an effective and reliable strategy for RIR cropping and highlights the value of repeated RIR measurements.

Author Declarations

Conflict of Interest

The authors have no conflict of interest to disclose.

Data Availability

The RIRs from the *Arni* dataset are available from the authors upon reasonable request, while the RIRs from the *MRTD* dataset are openly available online[28] and the related code is available at https://github.com/KPrawda/RIR-cropping.

References

[1]M. Guski and M. Vorländer, "Comparison of noise compensation methods for room acoustic impulse response evaluations," Acta Acust. **100**(2), 320–327 (2014).
[2]M. Janković, D. G. Ćirić, and A. Pantić, "Automated estimation of the truncation of room impulse response by applying a nonlinear decay model," J. Acoust. Soc. Am **139**(3), 1047–1057 (2016).
[3]M. Guski and M. Vorländer, "Measurement uncertainties of reverberation time caused by noise," in *Proceedings of the Joint 40th Italian Annual Conference on Acoustics and the 39th German Annual Conference on Acoustics*, Merano, Italy (March 18–21, 2013).
[4]P. 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).
[5]A. Lundeby, T. E. Vigran, H. Bietz, and M. Vorländer, "Uncertainties of measurements in room acoustics," Acta Acust. united Ac. **81**(4), 344–355 (1995).
[6]D. G. Ćirić and M. A. Milošević, "Optimal determination of the truncation point of room impulse responses," Build. Acoust **12**(1), 15–29 (2005).

[7]J. S. Abel and N. J. Bryan, "Methods for extending room impulse responses beyond their noise floor," in *Proceedings of the Audio Engineering Society 129th Convention*, San Francisco, CA (November 4–7, 2010).
[8]M. Chen and C.-M. Lee, "The optimal determination of the truncation time of non-exponential sound decays," Buildings **12**(5), 697 (2022).
[9]M. Karjalainen, P. Antsalo, A. Mäkivirta, T. Peltonen, and V. Välimäki, "Estimation of modal decay parameters from noisy response measurements," J. Audio Eng. Soc. **50**(11), 867–878 (2002).
[10]C. Hold, T. McKenzie, G. Götz, S. J. Schlecht, and V. Pulkki, "Resynthesis of spatial room impulse response tails with anisotropic multislope decays," J. Audio Eng. Soc. **70**(6), 526–538 (2022).
[11]G. Defrance, L. Daudet, and J.-D. Polack, "Finding the onset of a room impulse response: Straightforward?," J Acoust. Soc. Am. **124**(4), EL248–EL254 (2008).
[12]K. 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).
[13]K. Prawda, S. J. Schlecht, and V. Välimäki, "Short-term rule of two: Localizing non-stationary noise events in swept-sine measurements," in *Proceedings of the Audio Engineering Society 154th Convention*, Espoo, Finland (May 13–15, 2023).
[14]K. Prawda, S. J. Schlecht, and V. Välimäki, "Non-stationary noise removal from repeated sweep measurements," JASA Express Lett. **4**(2), 081601 (2024).
[15]R. Prislan, 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).
[16]B. 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).
[17]K. Prawda, S. J. Schlecht, and V. Välimäki, "Short-time coherence between repeated room impulse response measurements," J. Acoust. Soc. Am. **156**(2), 1017–1028 (2024).
[18]F. Satoh, Y. Hidaka, and H. Tachibana, "Influence of time variance on room impulse response measurement," J. Acoust. Soc. Am. **105**(2), 1367 (1999).
[19]K. Prawda, N. Meyer-Kahlen, and S. Schlecht, "Exploring sauna impulse responses," in *Proceedings of the INTER-NOISE and NOISE-CON Congress and Conference* (2024), Vol. 270, pp. 8145–8154.
[20]M. Wang, S. Clarke, J.-H. Wang, R. Gao, and J. Wu, "Soundcam: A dataset for finding humans using room acoustics," in *Proceedings of the Thirty-Seventh Annual Conference on Neural Information Processing Systems*, New Orleans, LA (December 10–16, 2023).
[21]K. Prawda, "Sensitivity of room impulse responses in changing acoustic environment," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Hyderabad, India (April 6–11, 2025), pp. 1–5.
[22]M. Frisén, "Unimodal regression," J. R. Stat. Soc. Ser. D (Statistician) **35**(4), 479–485 (1986).
[23]M. J. Best and N. Chakravarti, "Active set algorithms for isotonic regression: A unifying framework," Math. Program. **47**(3), 425–439 (1990).
[24]B. Zadrozny and C. Elkan, "Transforming classifier scores into accurate multiclass probability estimates," in *Proceedings of the 8th ACM SIGKDD International Conference Knowledge Discovery and Data Mining, KDD '02*, Edmonton, Canada (July 23–26, 2002), pp. 694–699.
[25]H. Pu, X. Qiu, and J. Wang, "Different sound decay patterns and energy feedback in coupled volumes," J. Acoust. Soc. Am. **129**(4), 1972–1980 (2011).
[26]K. Y. Lee, N. Meyer-Kahlen, G. Götz, U. P. Svensson, S. J. Schlecht, and V. Välimäki, "Fade-in reverberation in multi-room environments using the common-slope model," in *Proceedings of the 2024 AES 5th International Conference on Audio for Virtual and Augmented Reality*, Redmond, WA (August 19–21, 2024).
[27]K. Prawda, S. J. Schlecht, and V. Välimäki, "Time variance in measured room impulse responses," in *Proceedings of the 10th Convention of the European Acoustics Association*, Turin, Italy (September 11–15, 2023), pp. 1639–1646.
[28]P. Götz, G. Götz, N. Meyer-Kahlen, K. Y. Lee, K. Prawda, E. A. P. Habets, and S. J. Schlecht, "Multi room transition dataset" (2024), https://doi.org/10.5281/zenodo.11388246.
[29]P. 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 (September 9–12, 2024).
[30]L. Faiget, C. Legros, and R. Ruiz, "Optimization of the impulse response length: Application to noisy and highly reverberant rooms," J. Audio Eng. Soc. **46**, 741–750 (1998).
[31]ISO 3382-1, "Acoustics—Measurement of room acoustic parameters— Part 1: Performance spaces" (International Organization for Standardization, Geneva, Switzerland, 2001).
[32]"pyfar - Python packages for acoustics research," https://pyfar-gallery.readthedocs.io/en/latest/ (Last viewed January 14, 2024).