PROCEEDINGS of the
International Symposium on Room Acoustics

15 to 17 September 2019 in Amsterdam, Netherlands

The apparent width of spatially extended sources

Clara HOLLOMEY[1]; Bernhard U. SEEBER[1]

[1] Technical University of Munich, Germany

ABSTRACT

The apparent source width (ASW) is an important aspect of sound perception in rooms and has been extensively researched for statistically decorrelated point sources. Spatially extended sound sources differ from point sources in that they do not necessarily yield coherent binaural signals, even when radiating coherently under anechoic conditions. Yet, the relationship between the spatial extent of a sound source and the ASW it elicits is not entirely clear. We investigated the ASW of sound sources extended spatially along a line with stimuli containing different spectro-temporal properties. The results suggest that the spatial extent of the sound source may influence ASW perception.

Keywords: Apparent source width, spatial perception, line source

1. INTRODUCTION

The apparent source width (ASW) is one of the key perceptual attributes of sounds in rooms and thus an important aspect for achieving convincing auralization. Research on its perception focused on spatial impressions elicited by decorrelated point sources (1) and the associated deviations in its main correlate, the interaural coherence (IC, the maximum of the interaural cross-correlation function). However, the relation between the geometric dimensions of sound sources, the ear signals and perception are not fully understood. Grosse et al. (2015) demonstrated the correlation of the ASW with the physical extent of loudspeaker arrangements in an anechoic environment using incoherent noise (2). Yet, there is reason to assume that the spatial extent also affects the perceived source width of coherently radiating sound sources.

Under anechoic conditions, the signals arriving at the listener's ears are the sums of the source signals convolved with the Head Related Transfer Functions (HRTFs) for the left and right ear, respectively. For identical source signals, the IC, and thus, the ASW, will only depend on the decorrelation caused by the HRTFs (3). For coherently radiating sound sources, their effect is usually considered negligible. However, the radiation patterns of spatially extended sound sources are generally less coherent than those of point sources, and depend on their respective physical properties (4). Here, we investigate the perceived source width of spatially extended, coherently radiating sound sources arranged on a line for bandlimited noise, a sequence of harmonic complexes, speech and click trains under anechoic conditions.

2. METHODS

2.1 Stimuli

Spatially extended sound sources were approximated as the superposition of adjacent point sources via their corresponding HRTFs. The HRTFs were measured at a spatial resolution of 0.5 degrees using an artificial head (HEAD Acoustics HMS II.3) with an anatomically formed pinna according to ITU-T P.57.

The maximum extent of the line sources' endpoints was calculated according to

$$extent_{max} = \sqrt{1^2 + \left(\frac{extent}{2}\right)^2}\ [m] \tag{1}$$

[1] Clara.Hollomey@tum.de

Thus, the distance of the sound source center to the center of the listener's head was assumed to be 1 meter, and the spatial extent of the sound source in meters was assumed to correspond to its aperture angle regarding the listener in degrees.

The relative delay of $extent_{max}$ regarding the line source's center point was retrieved with the speed of sound according to

$$delay = \frac{extent_{max}[m]}{343\ [\frac{m}{s}]} \tag{2}$$

and the delays for the HRTFs between its center point and $extent_{max}$ were then calculated via linear interpolation between the line source's center points and its maximum extent. As depicted in Figure 1, this approach resulted in an unequal point source distribution along the line source, emphasizing the center of the sound source. The binaural ear signals were then created by summation of all fractionally delayed left ear and right ear HRTFs, respectively.

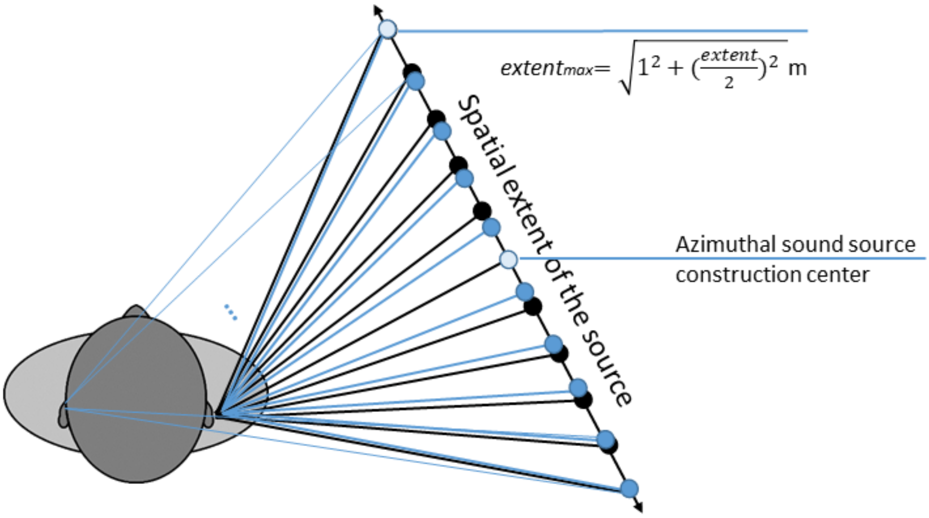

Figure 1: Approximating spatially extended sound sources as a superposition of adjacent point sources via HRTFs. To extend the sound source along a line, the HRTFs were delayed by an angle-dependent delay d computed via interpolation from the center of the sound source (unit-delay) to their maximum extent. The black dots indicate equally spaced point sources, the blue dots indicate the point source positions used in this experiment.

Spatially extended sound sources were constructed around two azimuthal center points, 10 and -70 degrees (counted such that the front of the the listener's head designates 0 degrees, and 10 degrees being slightly to the right), with spatial extents of 0 (point source), 20, 110, and 150 cm. This yielded eight conditions for each of the four stimuli, which were presented in a randomized fashion with ten repetitions at 65 dB SPL via free-field equalized headphones (Sennheiser HD 600).

A click train and white noise, both bandlimited to the range of 200 to 1270 Hz, speech, and a harmonic complex tone were convolved with the delayed and summed HRTFs to generate the stimuli for the experiment. Details on the source signals are given in Table 1.

Table 1 – Overview of the source signals included in the experiment. The onset time refers to the time for a 10 % to 90 % change of the Gaussian envelope

Name	Description	Spectrum	Onset time	Overall Duration
Noise	White Noise	White	50 ms	500 ms
Click train	3 x 30 ms clicks, with 70 ms quiet	White	3 ms	300 ms
Speech	German word “Garten” (taken from a private recording)	Female voice, $f_0 \approx 200$ Hz	30 ms	722 ms
Harmonic Complex	3 successively played harmonic complexes	$f_0 = 150$, 175 and 210 Hz, 8 harmonics each	30 ms	710 ms

2.2 Participants and Procedure

Six participants (median age: 24, ranging from 21 to 28) took part in the experiment. All had normal hearing, with less than 15 dB HL up to 8 kHz as assessed with standard audiometry (Audiometer Madsen Astera2, GN Otometrics A/S, Taastrup, Denmark).

The responses were collected along the lines of Lindemann’s drawing method (5), and as previously employed by Whitmer et al. (6). Participants were provided a graphical user interface (GUI) implemented in MATLAB (Math Works Inc., Nantick, MA) to sketch the spatial percepts elicited by the stimuli from a top view perspective. After listening to a sound, the participants were asked to draw their spatial impression of the sound on the screen by using a computer mouse or the touch screen, and to confirm their selection by pressing the “ENTER” button. In-the-head localizations were accommodated for by providing the participants with a “Near/Far” button to switch between a close-up and a more distant view. The close-up view allowed participants to draw internalized stimuli in an enlarged head. In the case of split images occurring, participants were asked to press the “Split Image” button prior to drawing. A screenshot of the GUI is depicted in Figure 2.

To introduce the participants to the drawing method, the graphical interface and the interpretation of spatial percepts, a familiarization session took place prior to the main experiment. This session covered localization and source width perception, as well as possible artefacts associated with the use of non-individual HRTFs, such as in-the-head localization, and the perception of split images. The stimuli used in this session were the same as in the main experiment whenever possible, and feedback was given when applicable, either directly or in an indicative fashion.

2.3 Analysis

For the analysis, the obtained drawings were approximated either as lines or ellipses in a least squares sense. For the elliptic fit, the algorithm proposed in (7) was used, and the decision on which approximation method would be suitable for a given drawing was based on visual inspection.

The reported source width was calculated for both types of shapes from their respective extreme values (8), as evaluated by calculating their quadrant-dependent angle with the drawing’s x-axis, and is further given in absolute values with regards to the coordinate system of the graphical user interface, which ranged from -1 to +1 for both axes.

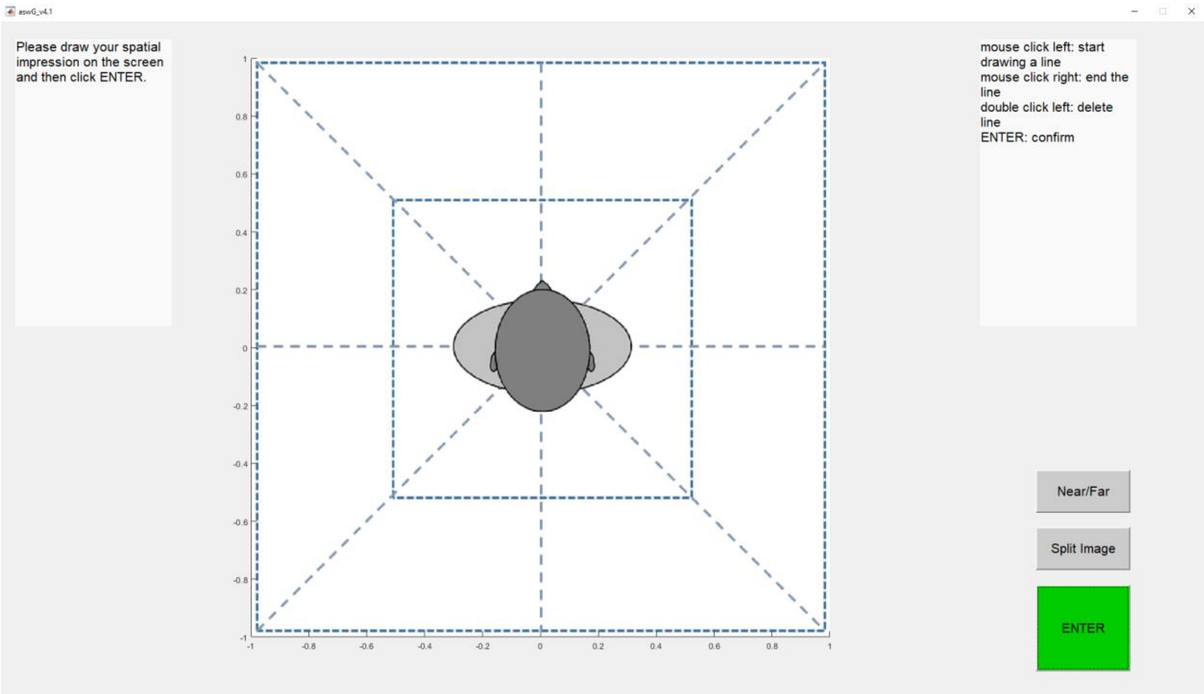

Figure 2 – Screenshot of the graphical user interface as displayed during the experiment. The participants heard a sound and were then asked to sketch their spatial impression either via touchscreen or via computer mouse. The participants confirmed their selection by pressing the “ENTER” button, the “Near/Far” button could be used to obtain a closer view of the head, and the “Split Image” button served as an indication of the according percept.

3. PRELIMINARY RESULTS

The median reported source width across participants is depicted, along with their first and third quartiles, in Figure 3 for a theoretical azimuthal sound source contruction center of 10 degrees, and in Figure 4 for a theoretical azimuthal sound source construction center of -70 degrees.

The overall reported source width was not significantly influenced by neither the type of sound source used for generating the stimuli, their azimuthal position, nor their theoretical spatial extent. However, participants’ responses varied considerably, as illustrated in Figure 5. Participant 1 rated the stimuli as significantly narrower ($p<0.05$) than the other participants, as assessed by a Bonferroni corrected Wilcoxon-Rank sum test between the median responses collapsed over all conditions per participant.

Participants 2, 3, and 4 exhibited similar response patterns to participant 5, in that the point source was usually reported as narrower than the other stimuli that were generated with the same sound sources for both azimuthal positions. A Bonferroni corrected regression analysis over the spatial extent per participant, azimuthal position and type of sound source used for stimuli generation, only yielded significance ($p<0.05$) for the response pattern given by participant 5 for the harmonic complex at an azimuthal position of 10 degrees. The stimuli generated using click trains were in tendency indicated as narrower than the remaining stimuli, but not significantly so.

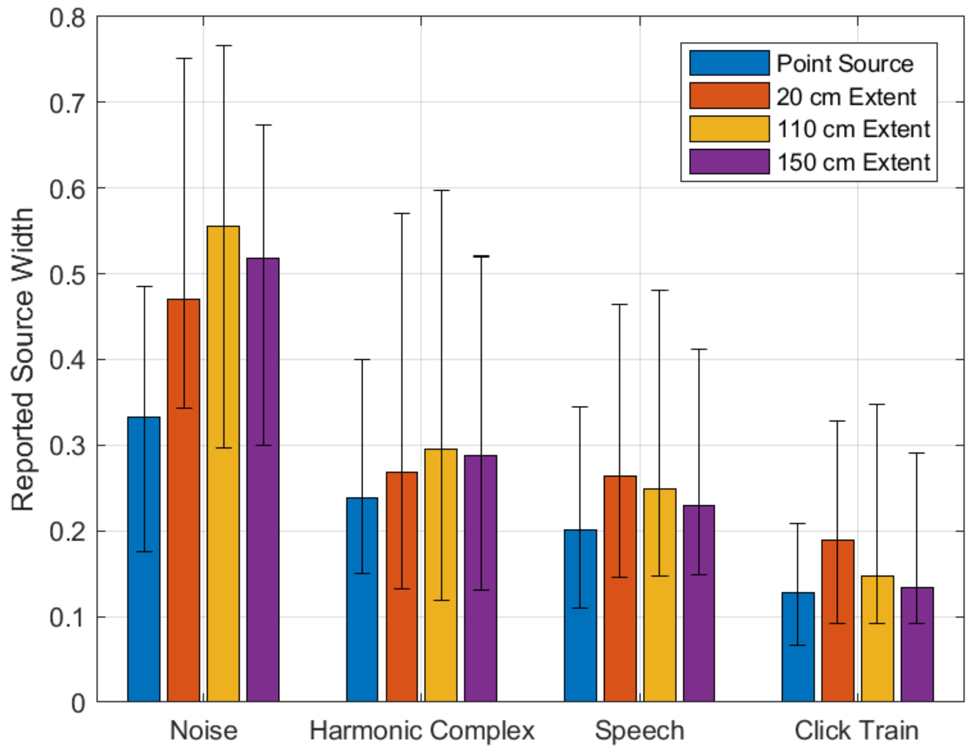

Figure 3 - The median absolute reported source width for all responses per condition at the azimuthal stimulus construction center of 10 degrees regarding the listener. The error bars indicate the associated first and third quartiles.

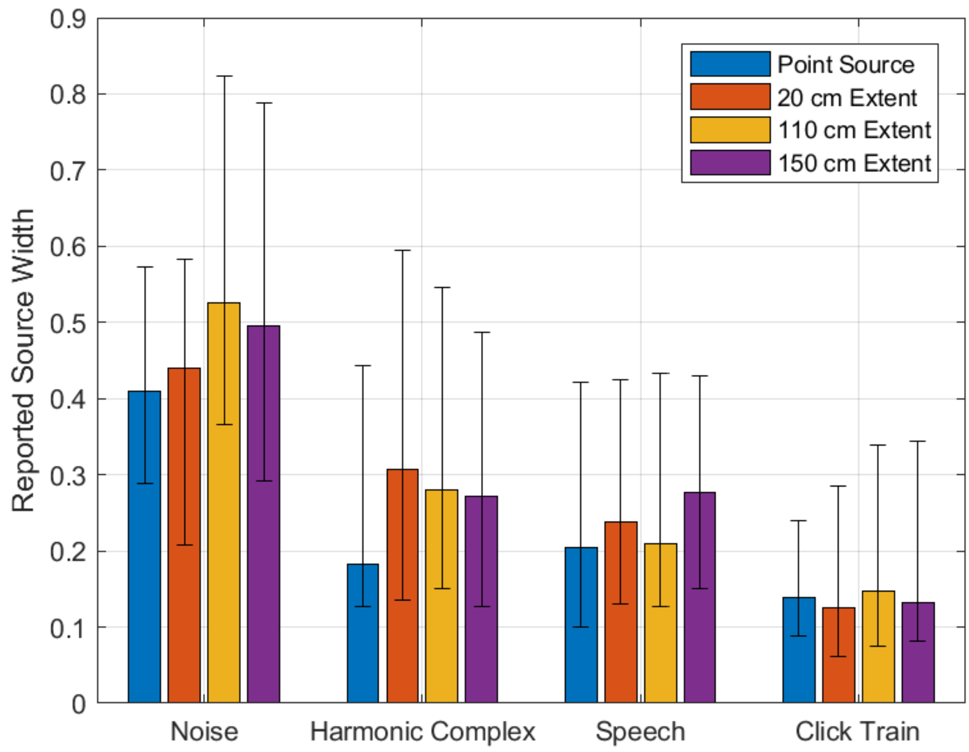

Figure 4 – same as Figure 3, for an azimuthal stimulus construction center of -70 degrees

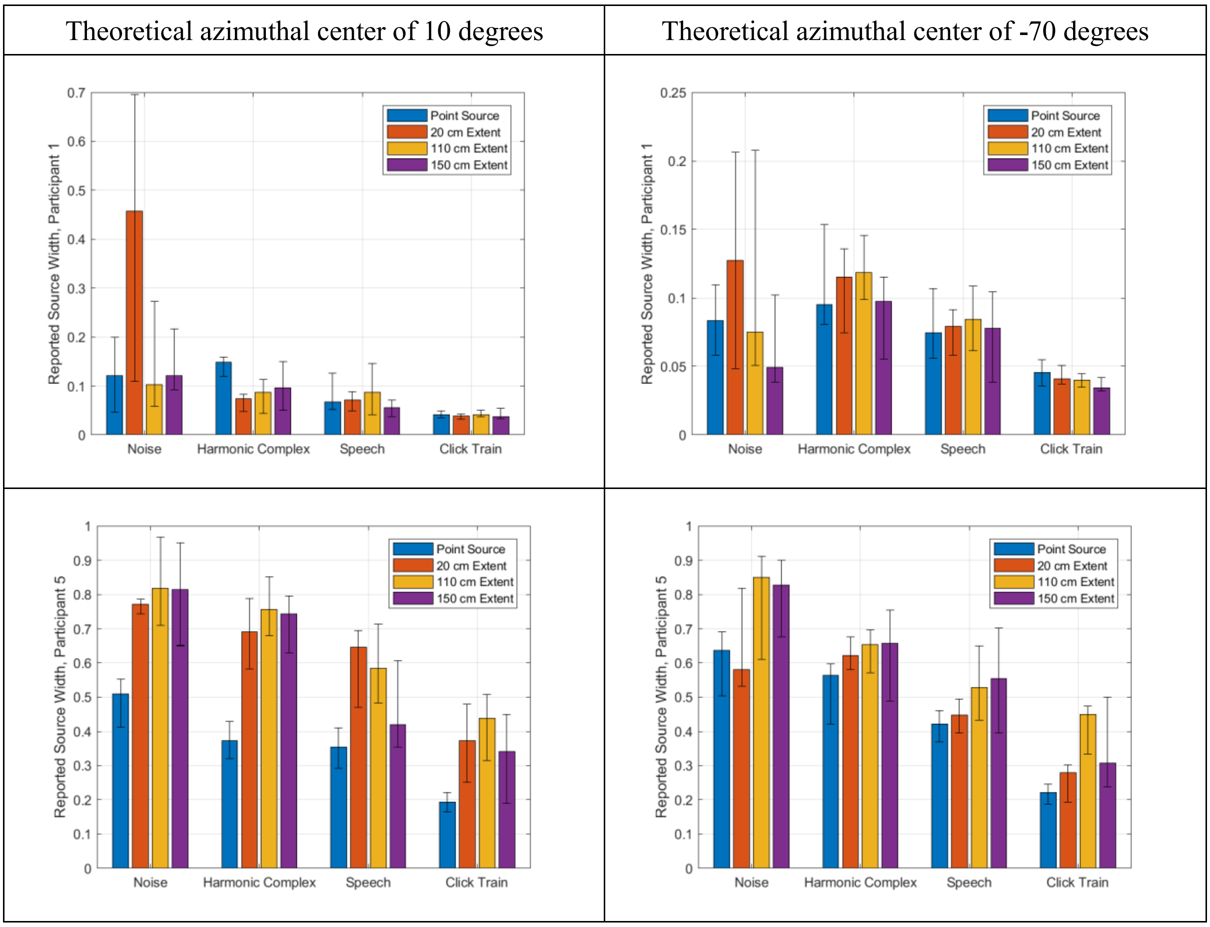

Figure 5 - The response patterns for participant 1 and 5 for an azimuthal stimulus construction center of 10 degrees (left) and -70 degrees (right).

4. DISCUSSION

Increases in the spatial extent of line sources seem to weakly affect the perceived source width. While the relations between spatial extent and reported ASW were mostly not statistically significant, the reported source width did in tendency increase with spatial extent for most participants and conditions. This could point to a weak, non-monotonic influence of spatial extent on ASW. An explanation for the obtained results is that only specific, possibly frequency- and participant-dependent interactions between the linearly delayed point sources and the HRTFs influence the resulting line source's perceived width.

Artefacts associated with the use of non-individual HRTFs, such as the unreliable externalization of sound sources, may have affected the results. For some participants and conditions, the drawings alternated between indicating internalized and externalized percepts. For example, the comparably high inter-quartile range of the reported source width by participant 1 for the noise stimulus depicted in Figure 5 is a consequence of that participant indicating either relatively narrow, internalized or wider, externalized spatial impressions for the associated conditions. If, and how in-the head-localization affects the perceived source width is, however, not yet entirely understood.

The click train being generally perceived as the narrowest sound source is in line with previous research on source width perception (e.g. 9, 10). The relatively short duration of 30 ms per click may have affected the build up of a widened spatial percept.

5. CONCLUSION

This study investigated the effects of the spatial extent of coherently radiating linear sound sources on their perceived source width. Overall, the results suggest that the spatial extent of a line source may influence the perception of its width. However, further research is needed to clarify that issue.

Such research could make use of individually measured HRTFs to obtain more nuanced results on the effects of spatial extent on ASW. Moreover, the stimuli used in this experiment took solely the azimuthal extent of sound sources into account. Yet, most physical oscillators exhibit a certain height, too. This physical property of sound sources could be accounted for by incorporating the associated HRTFs to the set of stimuli. Finally, the spatial extent of sound sources may influence other attributes of spatial hearing, besides the ASW, which could be explored in further research.

ACKNOWLEDGEMENTS

This work was funded by BMBF grant 01 GQ 1004B (Bernstein Center for Computational Neuroscience, Munich) and the Bavarian Elite master program “Neuroengineering”.

REFERENCES

1. Käsbach J. Characterizing apparent source width perception. PhD thesis, Technical University of Denmark; 2016.
2. Grosse J, Hungar F, Klockgether S, van de Par S. Wahrgenommene Quellbreite einer Lautsprecheranordnung in Abhängigkeit der physikalischen Quellbreite. DAGA 2015, 41. Jahrestagung für Akustik, Nürnberg.
3. Potard G, Burnett I. Control and measurement of apparent sound source width and its applications to sonification and virtual auditory displays. Proceedings of ICAD 04-Tenth Meeting of the International Conference on Auditory Display, Sydney, Australia, July 6-9, 2004
4. Goriely A. Vibrations of a Rectangular Membrane. Wolfram Demonstrations Project, 2011 http://demonstrations.wolfram.com/VibrationsOfARectangularMembrane/ (online ressource, ass. 13/07/2019 15:45)
5. Blauert J, Lindemann W. Auditory spaciousness: some further psychoacoustic analyses, The Journal of the Acoustical Society of America 80(533), 1986
6. Whitmer W M, Seeber B U, Akeroyd M A, The perception of apparent auditory source width in hearing-impaired adults, The Journal of the Acoustical Society of America 135(6), 2014
7. Fitzgibbon A.W, Pilu M, Fisher R.B. "Direct Least Squares Fitting of Ellipses" IEEE Trans. PAMI, Vol. 21, pages 476-480; 1999
8. Seeber, B. A new method for localization studies. Acustica. 2002;88:446–450.
9. Blauert J. Spatial Hearing – The Psychophysics of Human Sound Localization. 2nd enlarged edition, The MIT Press, Harvard, MA; 1997.
10. Hirvonen, T.,Pulkki, V. Perceived spatial distribution and width of horizontal ensemble of independent noise signals as function of waveform and samplelength. Audio Engineering Society Convention 124.Audio Engineering Society, 2008.