

Shortest Impulse Response Measurement Signal That Realizes Constant Normalized Noise Power in All Frequency Bands

YUKI NAKAHARA, YOHEI IYAMA, YUSUKE IKEDA, AND YUTAKA KANEDA, *AES Member*
(ikeda@mail.dendai.ac.jp) (kaneda@c.dendai.ac.jp)

Tokyo Denki University, Adachi-ku, Tokyo 120-8551, Japan

It is desirable that the measured acoustic impulse response has constant normalized noise power (NNP) in all frequency bands. However the conventional measurement signals aimed at achieving this property were derived intuitively, and the theoretical background is insufficient. In this work we first theoretically derived the relational formula that the measurement signals must satisfy for the measured impulse response to have constant NNP over all frequency bands. This formula includes all the measurement signals that achieve constant NNP. We then found the shortest (equivalently, the minimum energy) measurement signal among them. We call this signal the bandwise minimum noise (BMN) signal. Experiments to measure the room impulse responses were carried out. The experimental results confirmed that the impulse responses measured by the BMN signal had almost constant NNP in all frequency bands. Also, it was confirmed that the BMN signal achieved the required NNP for reverberation time measurement with the shortest signal length as compared with the conventional measurement signals.

0 INTRODUCTION

A room impulse response is an important characteristic from which various room acoustic parameters are calculated [1–4]. Measurement signals (or excitation signals) with high energy are used for impulse response measurement to achieve a required noise reduction performance. Conventionally, white spectrum signals, such as maximum length sequence (MLS) [5–7] and time-stretched pulse (TSP, or linearly swept sine) [8–10] signals, have been used as typical measurement signals. An exponentially swept sine (ESS or log-SS) signal [11–14] with high energy in the low-frequency band is currently the predominantly used measurement signal for room acoustic measurements. It improves the signal-to-noise ratio (SNR) in the low-frequency band where room noise has high energy.

The aforementioned measurement signals have predetermined fixed-spectra. In contrast, measurement signals with adaptive spectra reflecting the background noise spectra or measurement system characteristics have been proposed for a more effective noise reduction [15–23]. These signals are called adaptive measurement signals whose characteristics are described in detail in Sec. 1.

Depending on the application, an impulse response may be filtered to obtain frequency band impulse responses that

are then used. In this case, it is generally desirable that the impulse responses in all bands have constant normalized noise power (NNP). However, when a conventional measurement signal is used, the noise power in the measurement result varies depending on the frequency band. This causes various problems in the use of impulse responses. For example, in the measurement of reverberation time, a longer measurement signal is required to suppress NNP to a value below the required level, which is explained in detail in later sections. In this paper we propose a new adaptive measurement signal named the bandwise minimum noise (BMN) signal [24, 25] to solve this problem.

In Sec. 1, the principle of impulse response measurement and conventional adaptive measurement signals are explained. In Sec. 2, the problem of using conventional measurement signals is pointed out. In Sec. 3, the BMN signal, which is the shortest signal that realizes constant NNP over frequency bands, is theoretically derived. In Sec. 4, the synthesis and measurement processes for the BMN swept sine (BMN-SS) signal are explained. In Sec. 5, experimental results demonstrating the effectiveness of using the BMN signal are reported. In Sec. 6, some considerations regarding the BMN signal are discussed. The conclusion is given in Sec. 7.

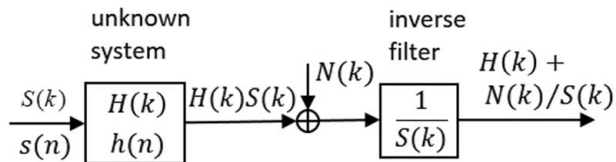


Fig. 1. Principle of impulse response measurement.

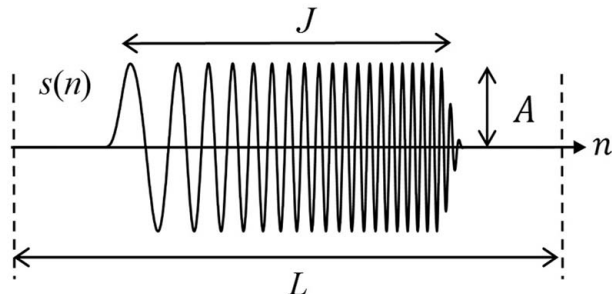


Fig. 2. Diagram of model of swept sine (SS) signal $s(n)$. Effective length J represents the duration in which the SS signal has a constant amplitude A .

1 RESEARCH BACKGROUND

1.1 Principle of Impulse Response Measurement

Fig. 1 shows the principle of impulse response measurement. In Fig. 1, a measurement signal of length L , $s(n)$, is input into an unknown system, where n is the discrete time index. $S(k)$ is the L -point discrete Fourier transform (DFT) spectrum of $s(n)$, where k is the discrete frequency index.

The output of the unknown system is a linear convolution of the impulse response $h(n)$ of the unknown system and $s(n)$. However, by filling the tail of $s(n)$ with zeros or generating $s(n)$ for two cycles, we can obtain the output as the circular convolution of $h(n)$ and $s(n)$. Because the circular convolution corresponds to the product of discrete frequency spectra, the output spectrum of the unknown system is expressed as $H(k)S(k)$, where $H(k)$ is the frequency response of the unknown system derived by the L -point DFT of $h(n)$. In this study the unknown system is an acoustic system, which includes the characteristics of audio equipment such as a loudspeaker.

Then $H(k)$ is obtained by filtering the unknown system output through an inverse filter $1/S(k)$. In an actual measurement environment, noise $N(k)$ (acoustic and electrical) is added to the system output. As a result a noise component $N(k)/S(k)$ is included in the measurement result $\tilde{H}(k)$, i.e.,

$$\tilde{H}(k) = H(k) + N(k)/S(k). \quad (1)$$

In this paper, $N(k)$ added to the system output is assumed to be a stationary random noise, which is called the “background noise.” On the other hand, we call $N(k)/S(k)$ added to the measured $H(k)$ the “noise component.” The measurement result of the desired impulse response $\tilde{h}(n)$ can be derived by the L -point inverse DFT of $\tilde{H}(k)$.

Fig. 2 shows a diagram of the model of a swept sine (SS) signal, which is the most commonly used measurement

signal. In this figure, L is the signal length (samples) for which the signal is defined. The time length at which the signal has a constant amplitude A is called the effective length J (samples).

1.2 Adaptive Measurement Signals

The power spectrum of the background noise $N(k)$ is expressed as

$$P_N(k) = E[|N(k)|^2], \quad (2)$$

($E[\cdot]$: expected value).

Then, the power spectrum of the noise component $N(k)/S(k)$ in Eq. (1) is denoted by $P_{NC}(k)$ and given by

$$P_{NC}(k) = E[|N(k)/S(k)|^2] = P_N(k)/P_S(k), \quad (3)$$

where

$$P_S(k) = |S(k)|^2. \quad (4)$$

Next the signal-to-noise power ratio of the measured frequency response $\tilde{H}(k)$ is given by

$$SNR(k) = \frac{|H(k)|^2}{P_{NC}(k)}. \quad (5)$$

Furthermore the signal-to-noise energy ratio in the octave or fractional-octave frequency band (hereinafter abbreviated as the octave band without distinction), $SNR(p)$, is defined by the following equation:

$$\begin{aligned} SNR(p) &= \frac{\sum_{k \in K_p} |H(k)|^2}{\sum_{k \in K_p} P_{NC}(k)} \\ &= \frac{\sum_{k \in K_p} |H(k)|^2}{\sum_{k \in K_p} \{P_N(k)/P_S(k)\}}. \end{aligned} \quad (6)$$

Here, p is the band index, and K_p represents the discrete frequency range of the p -th band.

Table 1 shows the optimization criteria for three typical adaptive measurement signals, namely the noise whitening (NW) [15, 23]; minimum noise (MN) [16, 17]; and constant SNR (CSN) [18–20] signals. Table 1 also shows the signal power spectrum $P_S(k)$ and the noise component spectrum $P_{NC}(k)$ obtained using each signal.

Fig. 3(a) shows a diagram of the model of frequency response $|H(k)|^2$ of an unknown system and the background noise $P_N(k)$ with a large low-frequency component. Figs. 3(b)–3(d) show diagrams of models of power spectra of noise components obtained using NW, MN, and CSN signals, respectively.

The NW signal flattens the noise component spectrum (Fig. 3(b)) and reduces the effect of the low-frequency component, but the total energy of the noise component $\sum_k P_{NC}(k)$ is not small (equal to that obtained using a white spectrum measurement signal [17]). The MN signal has been proven to minimize the total energy of the noise component among all measurement signals having the same energy. While the NW and MN signals use only $P_N(k)$, the CSN signal uses the rough estimate of $|H(k)|^2$ in addition to $P_N(k)$. Therefore the CSN signal can control the relative relationship between the system response and the noise

Table 1. Optimization criteria and power spectra of adaptive measurement signals and obtained noise components.

Signal	Optimization criterion	Power spectrum $P_S(k)$	Noise component spectrum $P_{NC}(k) = P_N(k)/P_S(k)$
NW	Noise whitening	$C_1 \cdot P_N(k)$	$1/C_1$
MN	Minimizing noise power	$C_2 \cdot \sqrt{P_N(k)}$	$\sqrt{P_N(k)}/C_2$
CSN	Constant SNR(k)	$C_3 \cdot P_N(k)/ H(k) ^2$	$ H(k) ^2/C_3$

$C_1, C_2, C_3 = \text{constant}$.

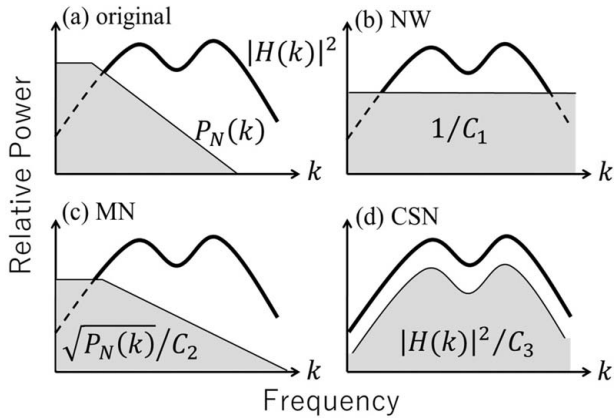


Fig. 3. Diagrams of models of noise reduction performance obtained using conventional adaptive measurement signals. (a) System frequency response $|H(k)|^2$ and original noise spectrum $P_N(k)$. (b), (c), (d) Power spectra of noise components obtained using NW, MN, and CSN signals, respectively. CSN = constant signal-to-noise ratio; MN = minimum noise; NW = noise whitening.

component and achieves constant $SNR(k)$ over the entire frequency range (Fig. 3(d)).

Thus, conventional adaptive measurement signals satisfy various optimization criteria, and the measurement signal suitable for the intended use should be selected.

2 PROBLEM OF CONVENTIONAL MEASUREMENT SIGNALS IN REVERBERATION TIME MEASUREMENT

The reverberation time $T_{30}(p)$, which is a representative characteristic of room acoustics, is calculated on the basis of the p -th octave bandpass-filtered impulse response $\tilde{h}_p(n)$. Here we define the normalized noise power $NNP(p)$ of $\tilde{h}_p(n)$ as the power (mean-square) of the noise component normalized by the peak value of $\tilde{h}_p^2(n)$. In ISO 3382, it is required for obtaining $T_{30}(p)$ that $NNP(p)$ should be -45 dB or lower in all frequency bands. Hereinafter, this condition is referred to as the “ -45 dB condition.”

In the case of using conventional measurement signals, $NNP(p)$ differs for different frequency bands. The examples of the measurement results derived using a conventional measurement signal are shown in Figs. 4(a)–4(c). The figures show $\tilde{h}_p^2(n)$ normalized by the peak value for the 125, 1,000, and 8,000 Hz one-third-octave bands, respectively.

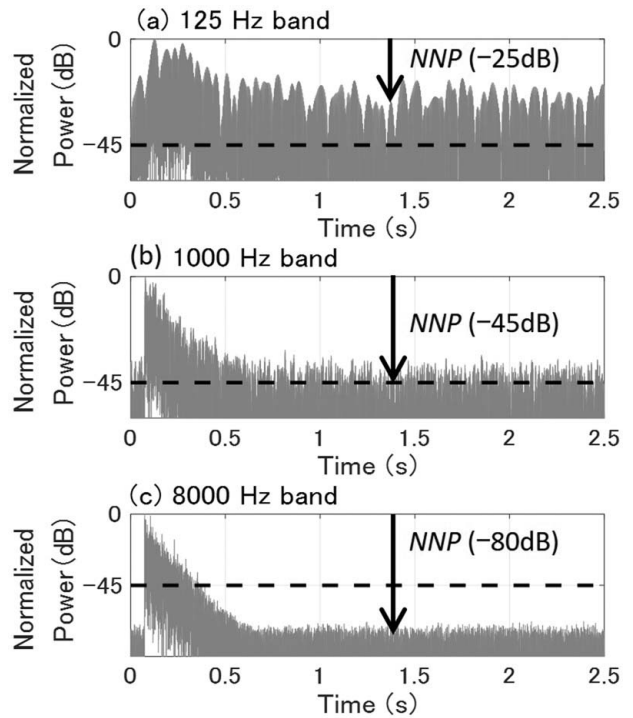


Fig. 4. Examples of instantaneous powers of bandpass-filtered impulse responses $\tilde{h}_p^2(n)$ obtained using conventional time-stretched pulse signal. Normalized noise power $NNP(p)$ significantly varies depending on the frequency band.

From Fig. 4(b), $NNP(p)$ for the 1,000 Hz band is about -45 dB. However $NNP(p)$ is higher than -45 dB for the 125 Hz band, as shown in Fig. 4(a). We denote the maximum value of $NNP(p)$ among all the frequency bands as $(NNP)_{max}$. To reduce $(NNP)_{max}$ to below -45 dB and satisfy the -45 dB condition with the same amplitude and spectrum of measurement signal, the signal length should be increased [17].

For the 8,000 Hz band $NNP(p)$ is much lower than -45 dB, as shown in Fig. 4(c). This seems to be a good result; however a higher signal energy, i.e., a longer signal than necessary, was used to reduce $NNP(p)$.

Thus, when using conventional measurement signals, $NNP(p)$ is high or low depending on the frequency band. (We will discuss in Sec. 6 that constant SNR and constant NNP are different concept, so the CSN signal also cannot realize constant NNP.) As a result, the problem of requiring a long measurement signal arises. To address this problem, constant NNP that is neither too high nor too low in all frequency bands should be realized. In the following section, we will develop a measurement signal that realizes a

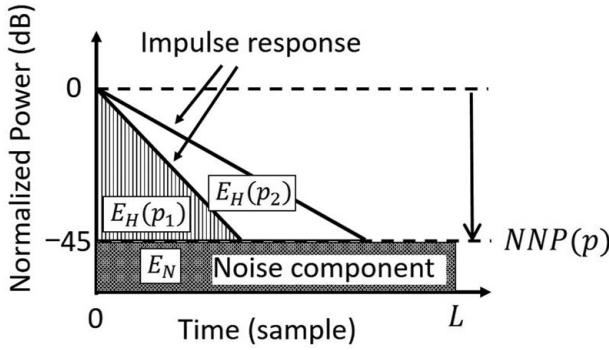


Fig. 5. Diagram of model of instantaneous powers of bandpass-filtered impulse responses $\tilde{h}_p^2(n)$ with $NNP(p)$ of -45 dB. p_1 and p_2 are band indexes. NNP = normalized noise power.

constant NNP of -45 dB in all frequency bands. We refer to this condition as the “constant -45 dB condition.”

3 SHORTEST MEASUREMENT SIGNAL REALIZING CONSTANT NNP :BMN SIGNAL

3.1 Desired SNR for Realizing Constant NNP

Fig. 5 shows a diagram of the model of instantaneous powers of the impulse responses $\tilde{h}_p^2(n)$ normalized by their peak values in the frequency bands p_1 and p_2 when $NNP(p)$ is -45 dB. $E_H(p_1)$ and $E_H(p_2)$ denote the impulse response energies (the sum of instantaneous powers) that depend on the instantaneous power decay characteristic of the frequency band p . The slower the decay, the higher the energy.

In contrast the noise component energy E_N is $10^{-4.5} \cdot L$ in all frequency bands, where L is the impulse response length (samples) that is equal to the measurement signal length. Hence the SNR of the p -th frequency band in the time domain when $NNP(p)$ is -45 dB, $SNR_D(p)$, is given by the energy ratio as

$$SNR_D(p) = \frac{E_H(p)}{E_N} = \frac{E_H(p)}{10^{-4.5} \cdot L}. \quad (7)$$

This equation means that $NNP(p)$ of all bands is controlled to be -45 dB by controlling the SNR of each band p to $SNR_D(p)$ of Eq. (7). Hereafter $SNR_D(p)$ refers to the “desired SNR.”

3.2 Shortest Measurement Signal That Realizes the Desired SNR

The SNR of the p -th frequency band can also be expressed using the frequency spectra of the system response and noise component as shown in Eq. (6). By letting the spectral representation (Eq. (6)) be equal to $SNR_D(p)$ we obtain

$$\frac{\sum_{k \in K_p} |H(k)|^2}{\sum_{k \in K_p} \{P_N(k)/P_S(k)\}} = SNR_D(p). \quad (8)$$

This equation shows the relationship that the power spectrum $P_S(k)$ of the measurement signal must satisfy to realize the constant -45 dB condition. However, the $P_S(k)$ that sat-

isfies Eq. (8) is not unique. There are many choices. In the following, the shortest (i.e., having the minimum energy) measurement signal among the signals satisfying Eq. (8) is derived.

On the basis of the principle of the MN signal [17], the measurement signal with the power spectrum proportional to $\sqrt{P_N(k)}$ minimizes the noise component energy $\sum_k \{P_N(k)/P_S(k)\}$ among the measurement signals having the same energy $\sum_k P_S(k)$. At the same time this measurement signal is the shortest among the signals that realize the specified noise component energy (the proof is in APPENDIX A).

Here, we call the shortest measurement signal that realizes $SNR_D(p)$ (i.e., satisfies Eq. (8)) as the BMN signal and denote its power spectrum as $P_{BMN}(k)$. Then, $P_{BMN}(k)$ should be

$$P_{BMN}(k) = C(p) \cdot \sqrt{P_N(k)}, \quad (9)$$

where $C(p)$ is a constant for $k \in K_p$.

By substituting Eq. (9) for $P_S(k)$ in Eq. (8) we obtain

$$C(p) \frac{\sum_{k \in K_p} |H(k)|^2}{\sum_{k \in K_p} \sqrt{P_N(k)}} = SNR_D(p). \quad (10)$$

By solving Eq. (10) for $C(p)$ we obtain

$$C(p) = SNR_D(p) \frac{\sum_{k \in K_p} \sqrt{P_N(k)}}{\sum_{k \in K_p} |H(k)|^2}. \quad (11)$$

Finally by substituting Eq. (11) into Eq. (9) we obtain

$$P_{BMN}(k) = SNR_D(p) \frac{\sum_{k \in K_p} \sqrt{P_N(k)}}{\sum_{k \in K_p} |H(k)|^2} \sqrt{P_N(k)}. \quad (12)$$

In practice $|H(k)|^2$, $E_H(p)$, and $P_N(k)$ in Eqs. (7) and (12) are not known in advance, so their estimates obtained by short-time preliminary measurements are used.

4 SYNTHESIS OF BMN-SS SIGNAL

Fig. 6 shows the BMN-SS signal synthesis procedure.

(1) The power spectrum of background noise is estimated from a 1–2-s recording, and the result is denoted by $\hat{P}_N(k)$.

(2) A preliminary measurement of the impulse response and frequency response is carried out, and the measurement results are denoted by $\hat{h}(n)$ and $\hat{H}(k)$, respectively. In this preliminary measurement a high SNR is not required. For example even if $\hat{H}(k)$ contains a noise component of -10 dB and the denominator of Eq. (12) increases 1.1-fold, the resulting error will be about 0.5 dB ($\approx 10 \cdot \log_{10}(1.1)$). Therefore, the measurement signal length L_1 used in the preliminary measurement can be very short compared with the signal length L_2 used in the main measurement.

(3) The estimated impulse response $\hat{h}(n)$ is bandpass-filtered into octave (or one-third-octave) bands, and its energy $E_H(p)$ is calculated. Then the desired SNR $SNR_D(p)$ is obtained using Eq. (7) with $L = L_1$.

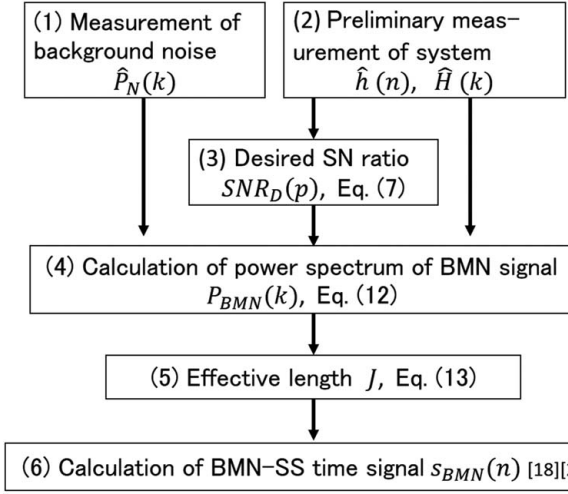


Fig. 6. BMN-SS signal synthesis procedure. BMN = bandwise minimum noise; BMN-SS = bandwise minimum noise swept sine; SNR = signal-to-noise ratio.

(4) The power spectrum of the BMN signal, $P_{BMN}(k)$, is calculated using $\hat{P}_N(k)$ and $\hat{H}(k)$, instead of $P_N(k)$ and $H(k)$, respectively, in Eq. (12).

(5) The effective length J of the measurement signal is calculated using Eq. (13) based on Parseval’s equation [26],

$$J = \frac{1}{A_{eff}^2} \cdot \frac{1}{L_1} \cdot \sum_{k=0}^{L_1-1} P_{BMN}(k), \quad (13)$$

where A_{eff} represents the effective value (RMS) of the measurement time signal ($= A / \sqrt{2}$ for the SS signal in Fig. 2). A_{eff} is determined by a user to produce a proper sound level.

(6) The BMN time signal $s_{BMN}(n)$ is synthesized on the basis of the BMN power spectrum $P_{BMN}(k)$. In this study, the measurement signal is an SS signal (we can choose pseudonoise instead), and the signal length L_2 is determined to be $2J$. The number of discrete frequencies L_1 of $P_{BMN}(k)$ is changed to L_2 by resampling in the frequency domain. Refer to [18] or [20] for details of the method of synthesizing an SS time signal from its power spectrum.

5 EXPERIMENT

To verify the implementability of the BMN signal and compare the performance of the BMN signal with those of the conventional measurement signals, experiments to measure the room impulse response were carried out.

5.1 Experimental Conditions

Table 2 shows experimental conditions. The experiment was carried out in rooms with volumes of 460 m^3 (room 1) and 70 m^3 (room 2). A full-range loudspeaker (BOSE MM-101) and a subwoofer (YAMAHA NS-SW300) were used. Fig. 7 shows the frequency responses of the loudspeaker and subwoofer. The loudspeaker and subwoofer were placed 5 m from the microphone.

Table 2. Experimental conditions.

Room volumes	460 m ³ (room 1), 70 m ³ (room 2)
Loudspeaker	BOSE MM-101
Subwoofer	YAMAHA NS-SW300
Measurement signals	BMN-SS, TSP, ESS, CSN-SS
Sound pressure level of measurement signal	80 dB at distance of 5 m
Background noise level	about 60 dB (in both rooms)
Sampling frequency	48 kHz
Frequency band	63–8,000 Hz one-third-octave band

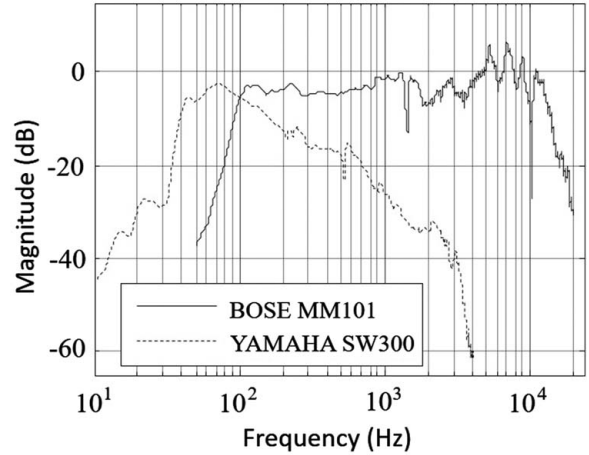


Fig. 7. Frequency responses of loudspeaker and subwoofer used in experiment.

To compare the performance characteristics of the BMN-SS signal with those of conventional measurement signals, measurements using the TSP (linearly SS) and ESS (exponentially SS) signals having fixed-spectra as the most popular measurement signals were performed. Also, the CSN-SS signal having an adaptive spectrum was adopted for comparison. This is because only the CSN signal among the conventional adaptive signals, the same as the BMN-SS signal, uses a rough estimate of the frequency response of an unknown system in addition to the background noise spectrum. (In APPENDIX B, it is explained that the signal recommended in ISO 18233 is CSN.)

The amplitude and effective length of all measurement signals were set to be identical so that the signals had the same energy. The measurement frequency bands were one-third-octave bands with central frequencies of 63–8,000 Hz.

5.2 Experimental Results

The room impulse response was measured in room 1 using the BMN-SS signal in accordance with the process shown in Fig. 6. First background noise was recorded for 2 s to estimate its power spectrum. Estimated noise spectrum $\hat{P}_N(k)$ was smoothed using a Hanning window with one-third-octave width. Then in a preliminary measurement, the room impulse response $\hat{h}(n)$ and frequency response $\hat{H}(k)$

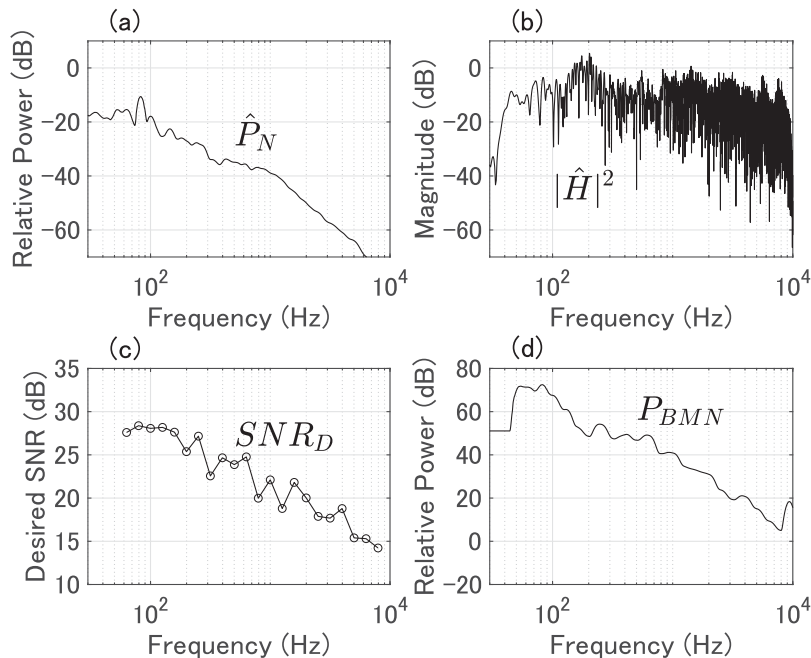


Fig. 8. Experimental results obtained in room 1. (a) Power spectrum of background noise, $\hat{P}_N(k)$. (b) Rough estimation of frequency response of room acoustic system $|\hat{H}(k)|^2$. (c) Desired SNR, $SNR_D(p)$ in each frequency band. (d) Calculated power spectrum of BMN signal, $P_{BMN}(k)$. BMN = bandwise minimum noise; SNR = signal-to-noise ratio.

were measured using the ESS signal with a signal length of $L_1 = 2^{16}$ samples (1.4 s). (For the calculation of the CSN-SS, $|\hat{H}(k)|^2$ was smoothed with a one-third-octave window.) Figs. 8(a) and 8(b) show $\hat{P}_N(k)$ and $|\hat{H}(k)|^2$, respectively, obtained in the preliminary measurement. Next the desired SNR $SNR_D(p)$ was calculated from the bandpass-filtered impulse response $\hat{h}_p(n)$. Fig. 8(c) shows the result. In the low-frequency band, the energy $E_H(p)$ was high because the impulse response was long, and thus the desired SNR became high.

By using $\hat{P}_N(k)$, $|\hat{H}(k)|^2$, and $SNR_D(p)$, we determine the spectrum of the BMN signal, $P_{BMN}(k)$, shown in Fig. 8(d). The effective length J of the measurement signal was calculated to be 261,690 samples (≈ 5.5 s) using Eq. (13). Then, the BMN-SS time signal $s_{BMN}(n)$ was synthesized with the signal length $L_2 = 2J$, i.e., 523,380 samples (≈ 11 s).

Fig. 9(a) shows the waveform $s_{BMN}(n)$ of the BMN-SS signal. Because this BMN-SS signal was designed in the frequency domain [18] (we can also design in the time domain), it has small ripples on the waveform (where the maximum peaks at both ends are out-of-band components).

Fig. 9(b) shows the spectrogram of the BMN-SS signal. For comparison Fig. 9(c) shows the spectrogram of the ESS signal. The BMN-SS signal had a longer sweep time in the low-frequency band and thus a higher energy than the ESS signal. Fig. 9(d) shows the time waveform of the impulse response obtained using the BMN-SS signal.

Fig. 10 shows the instantaneous power of the bandpass-filtered impulse response in the 125, 1,000, and 8,000-Hz bands obtained using the BMN-SS signal. NNP became about -45 dB in each frequency band.

5.3 Performance Comparison of BMN-SS Signal With Conventional Signals

Fig. 11 shows $NNP(p)$ of the impulse response in each one-third-octave band measured using the BMN-SS and conventional signals. As shown in Fig. 11, the $NNP(p)$ obtained using the BMN-SS signal (thick line) is controlled to be almost -45 dB in the target frequency bands (63–8,000 Hz). The maximum $NNP(p)$ among all the frequency bands, $(NNP)_{max}$, derived using the BMN-SS signal is -44 dB in the 3,981-Hz band. If a variation below 1 dB is acceptable, this result meets the constant -45 dB condition.

In contrast, $NNP(p)$ of the impulse responses measured using the TSP (thin gray line) and ESS (thick gray line) signals are much lower than -45 dB in the high-frequency range. However, in the low-frequency range, $NNP(p)$ are higher than -45 dB, and the condition required to measure the reverberation time T_{30} is not satisfied. This is because, in the low-frequency range, the spectrum of background noise is significantly large, but the fixed-spectral signals, TSP and ESS, do not have sufficient energy.

$(NNP)_{max}$ for the CSN-SS signal (thin line) is lower than those derived using the TSP and ESS signals (it is -41.8 dB in the 63-Hz band). Because the CSN-SS signal has an adaptive spectrum, i.e., it has a high power in a band in which the background noise power is high, the increase in $(NNP)_{max}$ is reduced. However, it still does not meet the -45 dB condition.

To make $(NNP)_{max}$ obtained using the CSN-SS signal equal to that obtained using the BMN-SS signal that satisfies the -45 dB condition, it is necessary to reduce the $NNP(p)$ by 2.2 dB. This requires the signal length to be increased by a factor of $10^{(2.2/10)}$ (about 1.7). Similarly, for

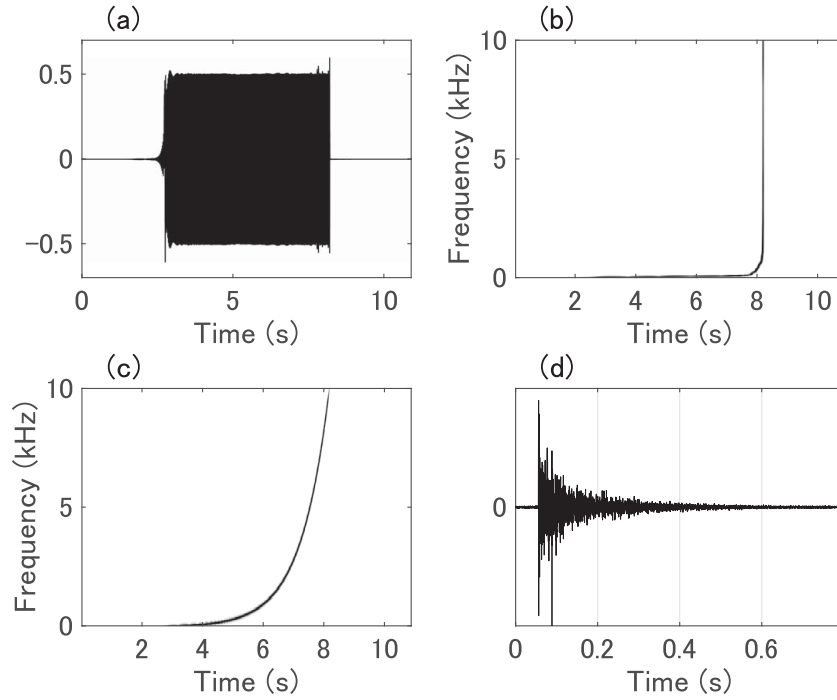


Fig. 9. (a) Waveform of BMN-SS signal, $s_{BMN}(n)$. (b) Spectrogram of BMN-SS signal. (c) Spectrogram of conventional ESS signal. (d) Impulse response obtained using BMN-SS signal. BMN-SS = bandwise minimum noise swept sine; ESS = exponentially swept sine.

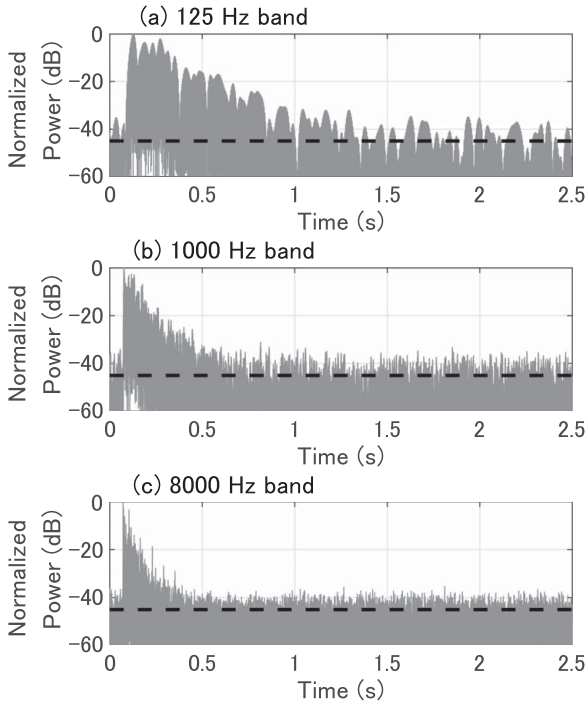


Fig. 10. Instantaneous power of bandpass-filtered impulse response in 125, 1,000, and 8,000-Hz bands obtained using BMN-SS signal. BMN-SS = bandwise minimum noise swept sine.

the $(NNP)_{max}$ obtained using the ESS and TSP signals to be the same as that obtained using the BMN signal, it is necessary to increase the signal length by 7.2 and 355 times, respectively, as shown in Table 3. Thus the BMN

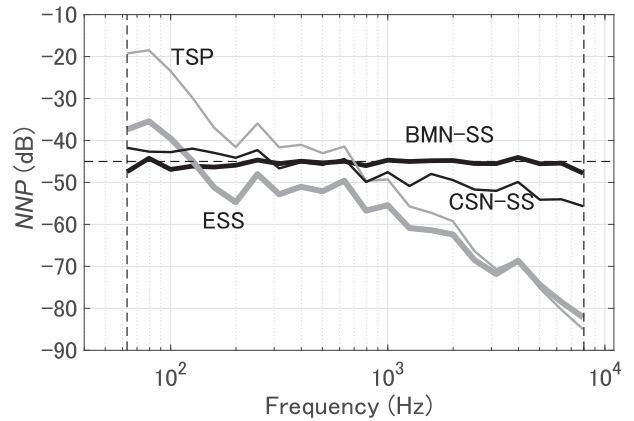


Fig. 11. Normalized noise power $NNP(p)$ of impulse response in each frequency band measured using BMN-SS, TSP, ESS, and CSN-SS signals in room 1. The $NNP(p)$ derived using the BMN-SS signal is almost -45 dB in all the frequency bands. BMN-SS = bandwise minimum noise swept sine; CSN-SS = constant signal-to-noise ratio swept sine; ESS = exponentially swept sine; NNP = normalized noise power; TSP = time-stretched pulse.

signal, compared with conventional measurement signals, can significantly reduce the signal length required to satisfy the -45 dB condition.

Figs. 12(a) and 12(b) show the noise spectrum $\hat{P}_N(k)$ and frequency response $|\hat{H}(k)|^2$ in room 2, and Figs. 12(c) and 12(d) show the desired SNR, $SNR_D(p)$, and BMN spectrum $P_{BMN}(k)$ calculated from them, respectively. The BMN-SS signal length was $L_2 = 946,328$ samples (≈ 20 s). Fig. 13 shows $NNP(p)$ of each band obtained using each measurement signal. We can see that $NNP(p)$ obtained

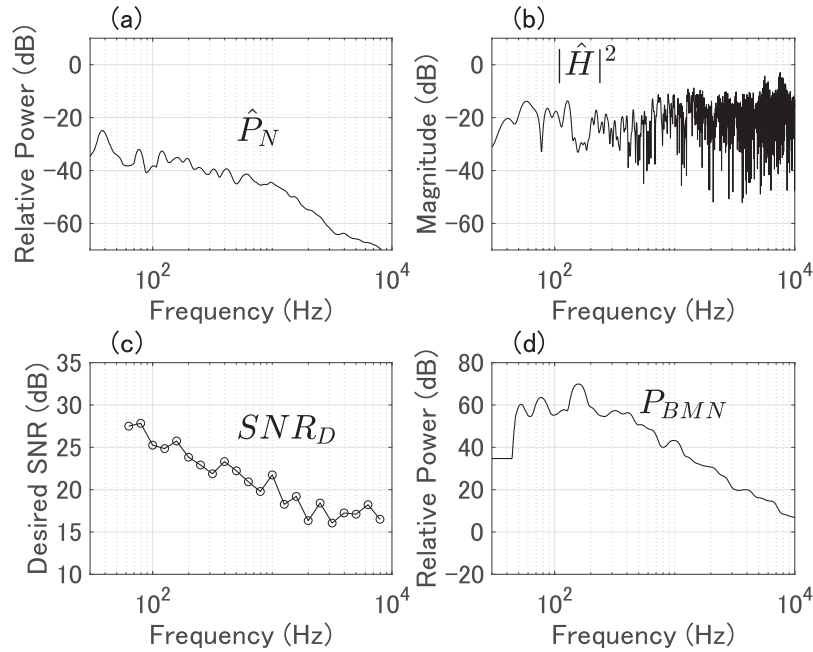


Fig. 12. Experimental results obtained in room 2. (a) Power spectrum of background noise, $\hat{P}_N(k)$. (b) Rough estimation of frequency response of room acoustic system $|\hat{H}(k)|^2$. (c) Desired SNR $SNR_D(p)$ for each frequency band. (d) Calculated power spectrum of BMN signal, $P_{BMN}(k)$. BMN = bandwise minimum noise; SNR = signal-to-noise ratio.

Table 3. Comparison of BMN-SS signal length with conventional signal lengths (room 1).

Measurement signal	Maximum normalized noise power $(NNP)_{max}$	Required signal length*
BMN-SS	-44.0 dB (in 3,981-Hz band)	1.0
CSN-SS	-41.8 dB (in 63-Hz band)	1.7
ESS	-35.4 dB (in 80-Hz band)	7.2
TSP	-18.5 dB (in 80-Hz band)	355

* Signal length normalized by BMN signal length required to make $(NNP)_{max}$ equal to that derived using BMN signal. BMN-SS = bandwise minimum noise swept sine; CSN-SS = constant signal-to-noise ratio swept sine; ESS = exponentially swept sine; TSP = time-stretched pulse.

Table 4. Comparison of BMN-SS signal length with conventional signal lengths (room 2).

Measurement signal	Maximum normalized noise power $(NNP)_{max}$	Required signal length
BMN-SS	-44.5 dB (in 3,162-Hz band)	1.0
CSN-SS	-40.7 dB (in 63-Hz band)	2.4
ESS	-34.3 dB (in 158-Hz band)	10.5
TSP	-20.4 dB (in 158-Hz band)	257

* BMN-SS = bandwise minimum noise swept sine; CSN-SS = constant signal-to-noise ratio swept sine; ESS = exponentially swept sine; TSP = time-stretched pulse.

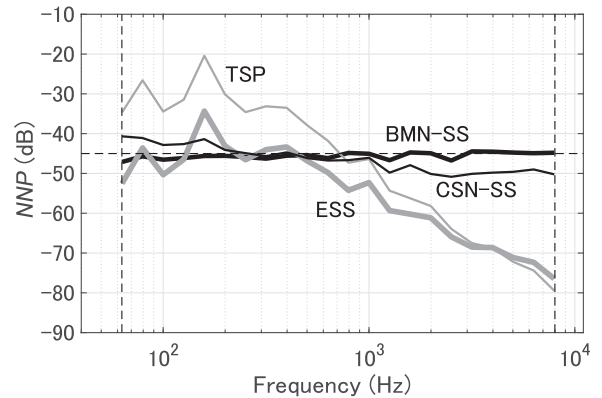


Fig. 13. Normalized noise power $NNP(p)$ of impulse response in each frequency band measured using BMN-SS, TSP, ESS, and CSN-SS signals in room 2. The $NNP(p)$ derived using the BMN-SS signal is almost -45 dB in all the frequency bands. BMN-SS = bandwise minimum noise swept sine; CSN-SS = constant signal-to-noise ratio swept sine; ESS = exponentially swept sine; TSP = time-stretched pulse.

using the BMN-SS signal (thick line) is almost constant at -45 dB. On the other hand, the $(NNP)_{max}$ obtained using the other measurement signals are higher than -45 dB. Table 4 shows $(NNP)_{max}$ obtained using each measurement

signal and the signal length required to reduce $(NNP)_{max}$ to the same value as that obtained using the BMN-SS signal. The CSN-SS, ESS, and TSP signals require lengths 2.4, 10.5, and 257-fold that of the BMN-SS signal, respectively.

As can be seen in Fig. 12(a), the background noise power in room 2 is not very high in the low-frequency range. However, the level of the frequency response $|\hat{H}(k)|^2$ shown in Fig. 12(b) decreases at around 160 Hz, which causes a significant increase in the $NNP(p)$ derived using the conventional fixed-spectrum measurement signals as shown in Fig. 13. On the other hand, since the CSN-SS and BMN-SS

signals utilize the characteristics of $|\hat{H}(k)|^2$, it is possible to compensate for the decrease in frequency response at around 160 Hz. Therefore, the increase in $NNP(p)$ at this band is small. However, $NNP(p)$ derived using the CSN-SS signal increases in the low-frequency range.

Unlike fixed-spectrum measurement signals, i.e., TSP and ESS, adaptive measurement signals, i.e., CSN-SS and BMN-SS, require preparation involving background noise recording, preliminary measurement, and signal synthesis. The preparation time was about 4.4 s in this experiment (2 s for noise recording, 1.4 s for system premeasurement, and about 1 s for signal synthesis calculation). To satisfy the -45 dB condition, the BMN-SS signal length was 11 s in the first experiment, so the total measurement time was 15.4 s. On the other hand, the ESS signal should be 7.2 times longer, that is, 79 s. Therefore even considering the preparation time, the BMN signal can significantly reduce the measurement time of the impulse response required for calculating T_{30} as compared with conventional measurement signals.

6 DISCUSSION

6.1 NNP

The definition of NNP is not unique depending on the normalization factor. In this paper, the squared impulse response $\tilde{h}_p^2(n)$ was normalized by its peak value in accordance with ISO 3382. However it has been pointed out that noise power cannot be normalized properly with this method owing to the effect of the initial peak shape of $\tilde{h}_p^2(n)$ [27]. Therefore the peak value after smoothing by moving average or the maximum value after applying the model of exponential decay has been proposed as the alternative normalization factor [27]. Regardless of which normalization factor is used to define the NNP, the desired SNR corresponding to each can be calculated so that the BMN signal can realize constant NNP that corresponds to each normalization factor with the shortest signal.

6.2 SNR

There are various definitions for the SNR, but in general SNR is not proportional to NNP. This is because there is no proportional relationship between the peak value of $\tilde{h}_p^2(n)$ (normalized factor) for the NNP calculation and the energy of $\tilde{h}_p^2(n)$ for the SNR calculation. Among various definitions of SNR we use the SNR of the octave band impulse response defined by Eq. (6) in the frequency domain. Its time domain model representation is shown in Fig. 5. From this figure, we can see that SNR is not proportional to NNP. Therefore even if SNR is controlled to be constant, constant NNP cannot be realized. In this paper desired SNR is used as a parameter for controlling NNP.

6.3 Environmental Conditions

6.3.1 Time Variance

The time variance of the background noise spectrum degrades the performance of the BMN signal. It is required that after the preliminary measurement, the background

noise should maintain stationarity until the main measurement is completed. In our experiments, a stationarity of about 11 or 20 s was required. Performance degradation due to time variance can be detected programmatically and resolved by remeasurement unless time variance frequently occurs. Such performance deterioration due to environmental variance is a problem common to all adaptive measurement signals.

6.3.2 Nonlinearity

A typical nonlinear factor of the acoustic measurement system is the nonlinearity of an amplifier or a loudspeaker. When an SS signal is used, the nonlinear distortion causes a harmonic distortion and generates an artifact in the impulse response. However, this artifact occurs in the noncausal part of the impulse response and is separable [18, 28]. Even if it cannot be separated and overlaps with the linear response, the energy is sufficiently smaller than the linear component, so the effect of the artifact on the impulse response energy decay curve (Schroeder curve) is small.

6.3.3 Other Conditions

The environmental conditions such as the room response $H(k)$, background noise spectrum $P_N(k)$, and reverberation time $T_{30}(p)$ affect the performance difference between the BMN signal and conventional signals. Thus the results in Tables 3 and 4 depend on the environmental conditions.

As stated in SEC. 6.3.1, the disadvantage of the BMN signal is that its performance is affected by the time variance of the background noise spectrum. With a signal having a fixed spectrum such as ESS signal, the measurement signal to be used should be determined by weighing this disadvantage with the advantages of shortening the measurement time.

7 CONCLUSIONS

It is desirable that the measured impulse response has constant NNP (normalized noise power) over all frequency bands in various applications, such as reverberation time measurement, auralization, and sound field control. However the conventional adaptive measurement signals aimed at achieving this property were derived intuitively, and the theoretical background was insufficient [18, 21, 28].

In this work we first theoretically derived the relational formula that the measurement signals must satisfy for the measured impulse response to have constant NNP over all frequency bands. This formula includes all the measurement signals that achieve constant NNP. Then the shortest (equivalently, the minimum energy) measurement signal was derived among signals that satisfy the relational formula. We call this signal the BMN signal.

The BMN signal is the shortest measurement signal that measures an impulse response having constant NNP over all frequency bands. In addition, because signal length reduction and noise suppression are alternative concepts, the BMN signal minimizes the constant NNP among all measurement signals having the same signal length.

To verify the implementability of the BMN signal and compare its performance with those of the conventional measurement signals, experiments to measure the room impulse responses were carried out. To obtain reverberation time T_{30} , ISO 3382 requires an NNP of -45 dB or less for the impulse response in all frequency bands.

The most efficient way to achieve this requirement is to make NNP in all frequency bands constant. Experimental results confirmed that when using the BMN signal, almost constant NNP could be achieved. In addition, the length of the BMN signal required to satisfy the ISO requirement was about 1/10 to 1/7 that of the ESS signal, which was a typical measurement signal.

The BMN signal is an adaptive measurement signal. It requires a preliminary measurement of the background noise spectrum and a rough estimation of the acoustic system to be measured. However even including the preliminary measurement time of 4.4 s, the required measurement time was shorter than that of the conventional signals under our experimental conditions.

Because the effectiveness of the BMN signal depends on environmental conditions such as the acoustical system response and the background noise spectrum and its stationarity, it may be necessary to perform preliminary measurements in the actual measurement environment, evaluate candidate measurement signals, and determine the signal to be used.

8 ACKNOWLEDGMENT

This study was partially supported by a Grant-in-Aid for Scientific Research from the Japan Society for the Promotion of Science (JSPS) (20K11872).

9 REFERENCES

- [1] M. R. Schroeder, "New Method of Measuring Reverberation Time," *J. Acoust. Soc. Am.*, vol. 37, pp. 409–412 (1965 Mar.). <https://doi.org/10.1121/1.1909343>.
- [2] H. Kuttruff, "Measuring Techniques in Room Acoustics," in *Room Acoustics*, pp. 238–243 (Elsevier Applied Science, Essex, UK, 1991).
- [3] ISO, "Acoustics - Measurement of Room Acoustics Parameters – Part 1: Performance Spaces," ISO 3382-1:2009(E) (2009).
- [4] M. R. Schroeder, "Integrated-Impulse Method Measuring Sound Decay Without Using Impulses," *J. Acoust. Soc. Am.*, vol. 66, no. 2, pp. 497–500 (1979 Aug.). <https://doi.org/10.1121/1.383103>.
- [5] J. Borish and J. B. Angel, "An Efficient Algorithm for Measuring the Impulse Response Using Pseudorandom Noise," *J. Audio Eng. Soc.*, vol. 31, nos. 7/8, pp. 478–488 (1983 Aug.).
- [6] D. D. Rife and J. Vanderkooy, "Transfer-Function Measurement With Maximum-Length Sequences," *J. Audio Eng. Soc.*, vol. 37, no. 6, pp. 419–444 (1989 Jun.).
- [7] C. Dunn and M. O. Hawksford, "Distortion Immunity of MLS-Derived Impulse Response Measurements," *J. Audio Eng. Soc.*, vol. 41, no. 5, pp. 314–335 (1993 May).
- [8] A. J. Berkhout, D. de Vries, and M. M. Boon, "A New Method to Acquire Impulse Response in Concert Halls," *J. Acoust. Soc. Am.*, vol. 68, no. 1, pp. 179–183 (1980 Jul.). <https://doi.org/10.1121/1.384618>.
- [9] N. Aoshima, "Computer-Generated Pulse Signal Applied for Sound Measurement," *J. Acoust. Soc. Am.*, vol. 69, no. 5, pp. 1484–1488 (1981 May). <https://doi.org/10.1121/1.385782>.
- [10] Y. Suzuki, F. Asano, H. Kim, and T. Sone, "An Optimum Computer-Generated Pulse Signal Suitable for the Measurement of Very Long Impulse Response," *J. Acoust. Soc. Am.*, vol. 97, no. 2, pp. 1119–1123 (1995 Feb.). <https://doi.org/10.1121/1.412224>.
- [11] D. Griesinger, "Beyond MLS-Occupied Hall Measurement With FFT Techniques," presented at the *101st Convention of the Audio Engineering Society* (1996 Nov.), paper 4403.
- [12] A. Farina, "Simultaneous Measurement of Impulse Response and Distortion With a Swept-Sine Technique," presented at the *108th Convention of the Audio Engineering Society* (2000 Feb.), paper 5093.
- [13] A. Farina, "Advancements in Impulse Response Measurements by Sine Sweeps," presented at the *122nd Convention of the Audio Engineering Society* (2007 May), paper 7121.
- [14] A. Novak, P. Lotton, and L. Simon, "Synchronized Swept-Sine: Theory, Application and Implementation," *J. Audio Eng. Soc.*, vol. 63, no. 10, pp. 786–797 (2015 Oct.). <https://doi.org/10.17743/jaes.2015.0071>.
- [15] S. Weinzierl, A. Giese, and A. Lindau, "Generalized Multiple Sweep Measurement," presented at the *126th Convention of the Audio Engineering Society* (2009 May), paper 7767.
- [16] N. Moriya and Y. Kaneda, "Impulse Response Measurement That Maximizes Signal-to-Noise Ratio Against Ambient Noise," *Acoust. Sci. Tech.*, vol. 28, no. 1, pp. 43–45 (2007 Jan.). <https://doi.org/10.1250/ast.28.43>.
- [17] Y. Kaneda, "Noise Reduction Performance of Various Signals for Impulse Response Measurement," *J. Audio Eng. Soc.*, vol. 63, no. 5, pp. 348–357 (2015 May). <https://doi.org/10.17743/jaes.2015.0024>.
- [18] S. Müller and P. Massarani, "Transfer-Function Measurement With Sweeps," *J. Audio Eng. Soc.*, vol. 49, no. 6, pp. 443–471 (2001 Jun.).
- [19] ISO, "Acoustics — Application of New Measurement Methods in Building and Room Acoustics," ISO 18233:2006(E), 6.2.2.1, p. 8 (2006).
- [20] H. Ochiai and Y. Kaneda, "A Recursive Adaptive Method of Impulse Response Measurement With Constant SNR Over Target Frequency Band," *J. Audio Eng. Soc.*, vol. 61, no. 9, pp. 647–655 (2013 Sep.).
- [21] Y. Nakahara and Y. Kaneda, "Effective Reverberation Time Measurement Method by Controlling Background-Noise Levels in Measured Impulse Response," *J. Acoust. Soc. Jpn.*, vol. 72, no. 7, pp. 358–366 (2016 Jul.) (in Japanese).

[22] E. Mommertz and S. Müller, “Measuring Impulse Responses With Digitally Pre-Emphasized Pseudorandom Noise Derived From Maximum-Length Sequences,” *Appl. Acoust.*, vol. 44, no. 3, pp. 195–214 (1995). [https://doi.org/10.1016/0003-682X\(94\)00019-R](https://doi.org/10.1016/0003-682X(94)00019-R).

[23] C. Huszty and S. Sakamoto, “Time-Domain Sweeplets for Acoustic Measurements,” *Appl. Acoust.* vol. 71, no. 10, pp. 979–989 (2010 Oct.) <https://doi.org/10.1016/j.apacoust.2010.06.012>.

[24] Y. Nakahara and Y. Kaneda, “Improvement of Reverberation Time Measurement Signal Controlling Background Noise Level,” in *Proceedings of the Spring Meeting Acoustical Society of Japan*, pp. 1113–1114 (Yokohama, Japan) (2016 Mar.) (in Japanese).

[25] Y. Nakahara and Y. Kaneda, “Faster Method of Reverberation Time Measurement Using Signal Realizing a Constant Noise Level for Each Frequency Band,” *Acoust. Sci. Tech.*, vol. 37, no. 5, pp. 247–249 (2016 Sep.). <https://doi.org/10.1250/ast.37.247>.

[26] A. V. Oppenheim and R. W. Schaffer, “The Discrete Fourier Transform,” in *Digital Signal Processing*, p. 125 (Prentice Hall, New Jersey, 1975).

[27] C. C. J. M. Hak, R. H. C. Wenmaekers, and L. C. J. van Luxemburg, “Measuring Room Impulse Responses: Impact of the Decay Range on Derived Room Acoustic Parameters,” *Acta Acust. United Acust.*, vol. 98, no. 6, pp. 907–915 (2012 Nov./Dec.). <https://doi.org/10.3813/AAA.918574>.

[28] A. Richard, C. L. Christensen, and G. Koutsouris, “Sine Sweep Optimization for Room Impulse Response Measurements,” in *Proceedings of the e-Forum Acusticum* (Lyon, France) (2020 Apr.). <https://doi.org/10.48465/fa.2020.0163>.

APPENDIX A

EQUIVALENCE BETWEEN MEASUREMENT SIGNAL THAT MINIMIZES NOISE COMPONENT ENERGY AND SHORTEST MEASUREMENT SIGNAL

A measurement signal spectrum with the energy E_1 in a certain frequency range K is denoted as $S_1(k)$. Namely,

$$\sum_{k \in K} |S_1(k)|^2 = E_1. \quad (14)$$

The energy of the noise component, E_{NC} , in the frequency range K measured using $S_1(k)$ is given by

$$E_{NC} = \sum_{k \in K} \{P_N(k) / |S_1(k)|^2\}. \quad (15)$$

The measurement signal that gives the minimum energy of the noise component, E_{NCmin} , among the signals $S_1(k)$ with the energy E_1 is denoted as $S_{min}(k)$. Namely,

$$\sum_{k \in K} |S_{min}(k)|^2 = E_1, \quad (16)$$

$$\begin{aligned} E_{NCmin} &= \sum_{k \in K} \{P_N(k) / |S_{min}(k)|^2\} \\ &\leq \sum_{k \in K} \{P_N(k) / |S_1(k)|^2\}, \end{aligned} \quad (17)$$

for all $S_1(k)$.

Here a signal that gives the same noise component energy E_{NCmin} as $S_{min}(k)$ is denoted as $S_2(k)$. Namely,

$$\sum_{k \in K} \{P_N(k) / |S_2(k)|^2\} = E_{NCmin}. \quad (18)$$

The energy of $S_2(k)$ is denoted as E_2 .

$$\sum_{k \in K} |S_2(k)|^2 = E_2. \quad (19)$$

Here another measurement signal $S_2'(k)$ is defined by

$$S_2'(k) = \sqrt{\frac{E_1}{E_2}} S_2(k). \quad (20)$$

Then the energy of $S_2'(k)$ is given by

$$\sum_{k \in K} |S_2'(k)|^2 = E_1. \quad (21)$$

Therefore $S_2'(k)$ satisfies the definition of $S_1(k)$; thus the energy of the noise component derived using $S_2'(k)$ is higher than E_{NCmin} .

$$E_{NCmin} \leq \sum_{k \in K} \{P_N(k) / |S_2'(k)|^2\} \quad (22)$$

By substituting Eqs. (20) and (18) into Eq. (22), we obtain

$$\begin{aligned} E_{NCmin} &\leq \sum_{k \in K} \left\{ P_N(k) / \left| \sqrt{\frac{E_1}{E_2}} S_2(k) \right|^2 \right\} \\ &= \frac{E_2}{E_1} \sum_{k \in K} \{P_N(k) / |S_2(k)|^2\} \\ &= \frac{E_2}{E_1} E_{NCmin}. \end{aligned} \quad (23)$$

By dividing both sides of Eq. (23) by E_{NCmin} , we obtain

$$1 \leq E_2 / E_1 \quad (24)$$

and

$$E_1 \leq E_2. \quad (25)$$

Thus the signal energy E_1 of $S_{min}(k)$ is lower than or equal to the energy E_2 of $S_2(k)$ that gives the same noise component energy E_{NCmin} . In other words, $S_{min}(k)$ is the measurement signal with the lowest energy among the measurement signals giving the same noise component energy as $S_{min}(k)$. Among the time signals with the same effective value, the signal $S_{min}(k)$ with the lowest energy is the shortest measurement signal.

APPENDIX B

EQUIVALENCE OF SIGNAL DESCRIBED IN ISO 18233 AND CSN SIGNAL

In ISO 18233 [19], the method that “approximates the shape of the spectrum of the excitation signal, as captured at the receiver position, to that of the ambient noise prevailing there” is described without a math-

ematical formula. “The (power) spectrum of the excitation signal, as captured at the receiver position” can be represented by $|H(k)|^2|S(k)|^2$. The ISO method requires that this equals to the ambient noise (power) spectrum. That is $|H(k)|^2|S(k)|^2 = P_N(k)$. Thus the required signal spectrum is $|S(k)|^2 = P_N(k)/|H(k)|^2$. This is the power spectrum of the CSN signal as shown in Table 1.

THE AUTHORS



Yuki Nakahara



Yohei Iiyama



Yusuke Ikeda



Yutaka Kaneda

Yuki Nakahara was born in 1989. He studied information and communication engineering at Tokyo Denki University, Tokyo, Japan. He received an M.Eng. degree for the work of acoustic signal processing from the graduate school of engineering, Tokyo Denki University in 2016. He is currently working at ALPS ALPINE Co., Ltd.

Yohei Iiyama was born in 1995. He studied information and communication engineering at Tokyo Denki University, Tokyo, Japan. He received an M.Eng. degree for the work of acoustic signal processing from the graduate school of engineering, Tokyo Denki University in 2020. He is currently working at FUJIFILM Medical IT Solutions Co., Ltd.

Yusuke Ikeda was born in Tokyo, Japan, in 1978. He received a Ph.D. (International Information and Communication Studies) from Waseda University in 2007. He worked as a researcher at ATR Corporation, NICT, and Kyoto University and as an assistant professor at Tokyo Denki University and Waseda University. In 2017, he joined Tokyo Denki University, Tokyo, Japan. He is currently an associate professor of acoustics and communication laboratory

with the department of Information Media, faculty of Future Sciences, Tokyo Denki University. His research interests include spatial acoustics and sound field visualization and acoustical measurement and modeling. He is a member of the Acoustical Society of Japan, Audio Engineering Society, ACM SIGGRAPH, and IEEE.

Yutaka Kaneda was born in Osaka, Japan, in 1951. He received B.E., M.E., and Doctor of Engineering degrees from Nagoya University, Nagoya, Japan, in 1975, 1977, and 1990. In 1977 he joined Nippon Telegraph and Telephone Corporation (NTT), Musashino, Tokyo, Japan, where he engaged in research on acoustic signal processing. In 2000 he joined Tokyo Denki University, Tokyo, Japan. He is currently a professor emeritus at the Department of Information and Communication Engineering, Tokyo Denki University. His research interests include acoustical measurements and microphone array signal processing. Dr. Kaneda is a member of the Acoustical Society of Japan, the Acoustical Society of America, Audio Engineering Society, Institute of Electrical and Electronics Engineers, and the Institute of Electronics, Information and Communication Engineers of Japan.