Audio Engineering Society

Convention Paper 9827

Presented at the 143rd Convention
2017 October 18–21, New York, NY, USA

Use of Repetitive Multitone Sequences to Estimate Nonlinear Response of a Loudspeaker to Music

Pascal Brunet[1], William Decanio[1], Ritesh Banka[1], and Shenli Yuan[2]

[1] *Samsung Research America, DMS Audio-Lab, Valencia CA 91355*
[2] *CCRMA, Stanford University, Stanford CA 94305*

Correspondence should be addressed to Pascal Brunet (p.brunet1@samsung.com)

ABSTRACT

Aside from frequency response, loudspeaker distortion measurements are perhaps the most commonly used metrics to appraise loudspeaker performance. The stimuli utilized for many types of distortion measurements are not complex waveforms such as music or speech, thus the measured distortion characteristics of the Device Under Test may not reflect the performance of the device when reproducing typical program material. The topic of this paper is the exploration of a new multitone sequence stimulus to measure loudspeaker system distortion. This method gives a reliable estimation of the average nonlinear distortion produced with music on a loudspeaker system, and delivers a global objective assessment of the distortion for a DUT in a normal use case.

1 INTRODUCTION

Aside from frequency response measurements, distortion measurements are one of the most widely used metrics to appraise the performance of audio systems. With respect to loudspeaker systems, distortion measurements are routinely used as a basis for making performance comparisons, diagnostic tools and production quality monitoring. Nonlinear distortion implies that the output of a device cannot be entirely described by the result of a linear operation (transfer function) applied on the input signal. The most frequently encountered type of distortion measurement is harmonic distortion. Harmonic distortion is a type of nonlinear distortion whereby a system exhibits an output response that is not linearly related to the sine wave applied to its input. When a loudspeaker system generates harmonic distortion, it adds frequency components to its output that are harmonically related (i.e. integral multiple) to the frequency of the sine present at its input.

Historically, the popularity of harmonic distortion measurements likely stems from the simplicity of the equipment required, as well as the ease with which the measurement may be performed. Typically harmonic distortion is measured by applying a swept sine wave stimulus to a Device Under Test (DUT) while observing the output spectrum. The amplitude spectrum of the fundamental response as well as any resultant distortion products can be graphed and used as a basis of comparison or figure of merit for the performance of the DUT. The results of harmonic distortion measurements can also be a useful diagnostic tool. For exam-

ple, inspecting a graph of second- and third-harmonic distortion components can reveal insight into the mechanisms of distortion in transducers [1].
There are some common pitfalls that can arise with respect to the measurement and evaluation of harmonic distortion from loudspeaker systems. Generally speaking, a swept sine doesn't account for distortion terms that fall outside of system bandwidth nor fall on the fundamental. Also, a simple spectral analysis cannot distinguish between noise and weak sine components. This necessitates the distortion measurement be conducted in an environment with a low noise floor, such as an anechoic chamber, to avoid contamination issues that can affect accuracy of the measurement.
Many harmonic distortion measurements use a sine wave stimulus that sweeps from low to high frequencies. This may cause misleading results when attempting to characterize the harmonic distortion and acoustic output of powered systems. The leading low-frequency portion of the sweep can activate a system's compressor/limiter protection yielding different SPL/distortion characteristics than would be observed during operation with a high to low sweep or a actual program material. This is particularly true for active systems which have significant equalization boost applied to low frequencies.
Similarly, for resonant systems, sweep-up can give different results than sweep-down and both are biased [2]. Although not exclusive to harmonic distortion measurements, the correlation between measured harmonic distortion and the perception of audibly objectionable performance is problematic [3, 4, 5, 6]. This is further compounded by the differences of signal characteristics (crest factor, test spectrum) between typical test stimuli and actual program material. As previously noted, for the case of harmonic distortion measurements, the stimuli utilized for many types of distortion measurements are not complex waveforms such as music or speech, thus the measured distortion characteristics of the DUT may not typically reflect the performance of the device when reproducing music. Distortion can be estimated for music (and program material in general) using input-output coherence analysis [7], but this technique cannot discriminate between noise and distortion and is totally dependent on the music used.
While harmonic distortion measurements essentially apply a single frequency tone as a stimulus, multitone distortion measurements simultaneously apply many frequency components, which facilitates the production of intermodulation distortion (sum and difference components), as well as harmonic distortion. Multitone stimuli possess signal characteristics that are closer to real music than single-tone stimuli. Further, as will be shown, the spectrum and crest factor properties of a multitone sequence can be tailored to closely mimic music, providing a more realistic means to assess the distortion characteristics of the system under test, while staying general enough to predict average results.

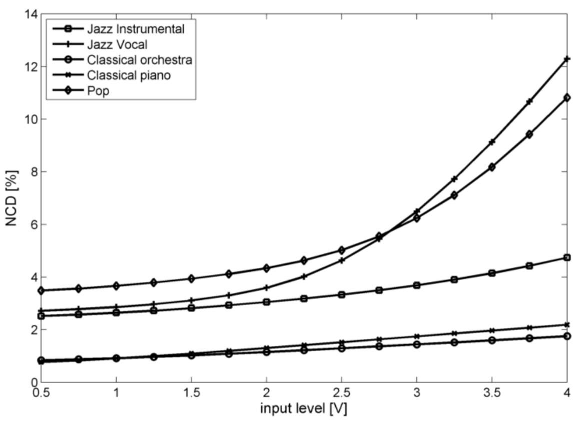

Fig. 1: Distortion vs input level for different pieces of music. Distortion is estimated using input-output coherence analysis.

2 METHOD

2.1 Introduction and motivation

As seen in Fig.1, distortion varies greatly with the music piece played. Strictly speaking, NonLinear Distortion (NLD) depends on the power spectrum and amplitude distribution of the stimulus signal. Therefore, in order to address the practically infinite variety of program material (music, speech, movie soundtracks...) that can be played on an audio system, a statistical approach is needed. The goal of our random multitone technique is to predict the NLD that will occur on average with program material.

2.2 Theory and justification of new approach

In this section we explain the method used to estimate the Stochastic Nonlinear Distortion (SNLD) of a system using a sequence of random-phase multitones. Our DUT model is described in Fig. 2. We consider

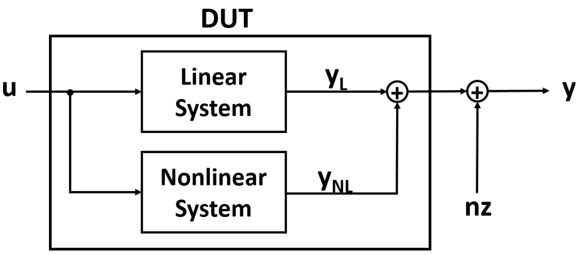

Fig. 2: General model for nonlinear DUT - u: stimulus, y_L: linear response, y_{NL}: nonlinear response, nz: output noise.

the DUT as a deterministic system with smooth nonlinearities (Volterra system). This model does not include random processes like rub and buzz, loose particles, air leak, port noise, etc. A signal through this kind of nonlinear system yields the sum of a linear response plus nonlinear contributions and measurement noise [7]. This additive model allows for extraction of the nonlinear contributions from a periodic stimulus with statistical processing.

To illustrate this concept, we have fed a full multitone with constant amplitude spectrum and random-phase into a simulated driver and calculated its sound pressure Frequency Response Function (FRF) (Fig. 3). With a linear model of the driver, the FRF is smooth, independent of the random-phase. With a nonlinear model, the FRF shows random contributions added to the linear FRF that are uncorrelated along the frequencies and look like noise. Two different random realizations of the multitone result in two different FRF illustrating the dependence on the phase pattern. For each FRF measurement, the stimulus consists of two successive periods: the first one to reach steady state, long enough to let the transient die and the second for the FRF calculation. The method described below is inspired from [2].

2.3 Stimulus creation, structure and parameters

As an input signal, we use a full multitone with random-phase:

$$u(t) = \sum_{k=-N/2+1}^{N/2-1} U_k e^{j\omega_k t} \quad (1)$$

The use of harmonic frequencies ensures a periodic signal: $\omega_k = 2\pi k \frac{f_s}{N} t$, where N/f_s is the multitone period. The phases $\angle U_k$ are random, uniformly distributed on

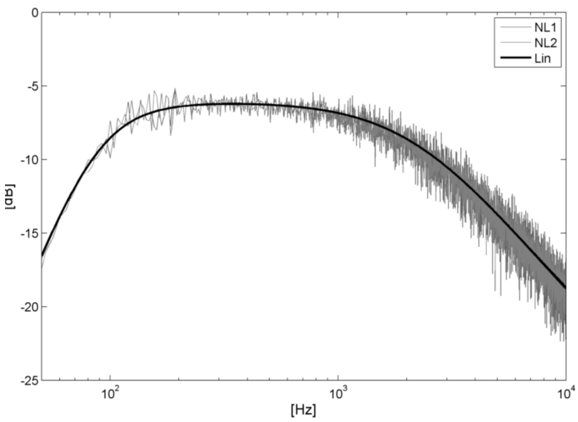

Fig. 3: FRF linear system (Lin) vs. nonlinear system with two random multitone realizations (NL1, NL2). Nonlinear contributions look like added 'noise'.

$[0, 2\pi[$. That phase distribution ensures a signal with random values and an amplitude distribution that tends asymptotically to a Gaussian law when $N \to \infty$. Each random-phase realization yields a specific time pattern. The amplitude spectrum $|U_k|$ is constant and user-defined. The signal has a zero mean ($|U_0| = 0$). Several multitone periods are then concatenated to form a sequence. A general sequence is composed of $M \times (P+1)$ periods. Each group of $P+1$ periods uses a particular realization of the random-phase pattern, and M successive realizations leads to $M \times (P+1)$ periods. For each realization, the first period is disregarded as it contains settling transients for the system. Periodic signals allow the output noise to be averaged out and different realizations of the random-phase pattern allow the nonlinear distortion to be estimated.

2.4 Spectral shaping (music profile)

To predict the NLD that will occur with program material, the stimulus needs to have a similar power spectrum. Statistical studies have been performed on music and speech, and different recommendations for test signals exist. We tried the following profiles (Fig. 5): 'IEC' for IEC 60268-1 1985, [8], 'AESw' for AES preprint 4277 [9], and 'Pop' from an average of different popular music. Simulation with a nonlinear driver model shows that the NLD varies significantly from one spectral profile to the next (Fig. 6). The data suggest that the bass content is determinant for distortion.

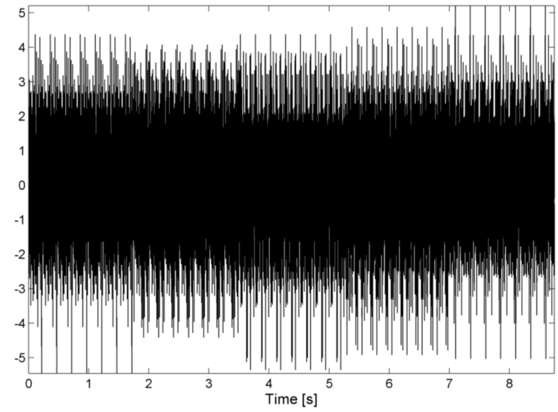

Fig. 4: Example of random-phase multitone. Each of the 5 blocks of 7 periods is created with a different random phase realization, resulting in a particular time pattern.

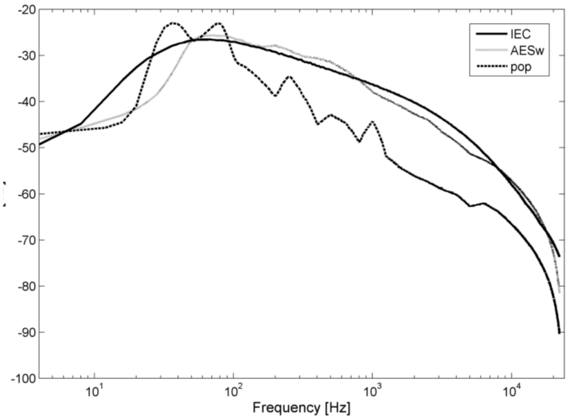

Fig. 5: Amplitude spectra for different flavors of multitone signals

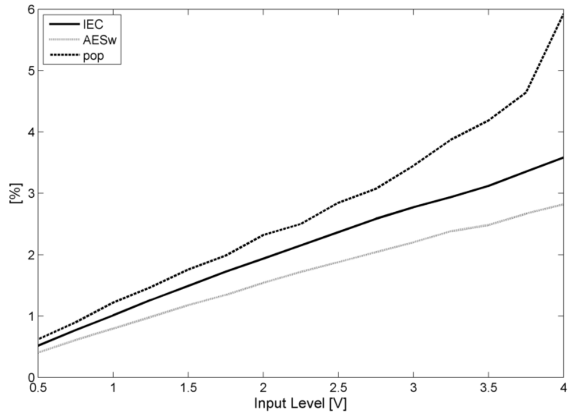

Fig. 6: NLD for different flavor of multitone signals

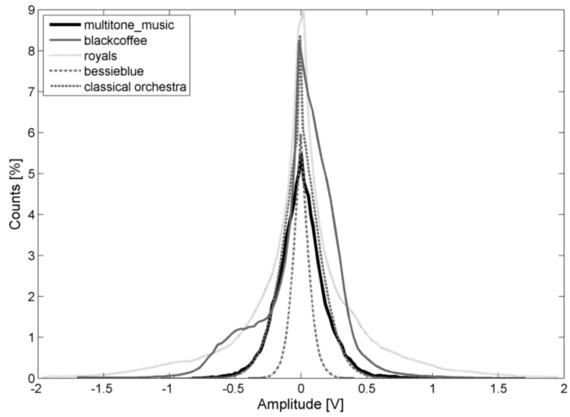

Fig. 7: Amplitude distribution of different pieces of music compared to multitone with Laplace distribution.

2.5 Amplitude shaping: crest factor, Laplace vs Gaussian distribution

The random-phase multitone has a Gaussian amplitude distribution. It is known that music and speech Probability Density Functions (PDF) are closer to Laplace distribution [10]. That can be seen in Fig. 7. To better imitate typical audio content we shape the amplitudes of the signal to obtain a Laplace PDF. The process is the following:

- x= random multitone signal
- From Gaussian PDF ($x \sim \mathcal{N}(0,1)$) to Uniform PDF ($u \sim \mathcal{U}(-\frac{1}{2},\frac{1}{2})$):

$$u = \mathrm{erf}(\frac{x}{\sqrt{2}}) \quad (2)$$

where $\mathrm{erf}(x)$ is the Gauss error function,

- From Uniform PDF to Laplace PDF ($y \sim \mathscr{L}(0,1)$):

$$y = -\mathrm{sign}(u)\ln(1-2|u|) \quad (3)$$

- Re-shape the amplitude spectrum, while keeping the phases:

$$Y = \mathrm{DFT}(y) \quad (4)$$

$$\tilde{y} = \mathrm{DFT}^{-1}(\frac{AY}{|Y|}) \quad (5)$$

where A is the target amplitude spectrum and DFT stands for Discrete Fourier Transform.

The last step is necessary because the nonlinear process distorts the amplitude spectrum. The crest-factor goes from $4\,(12\,\mathrm{dB})$ up to $7\,(17\,\mathrm{dB})$.

2.6 Analysis method: FFT and statistical analysis

The analysis method is based on two assumptions:

1. The DUT is deterministic: same signal in, same signal out. Therefore the output variance on a sequence of periodic input signal for a given random realization is only due to the output noise.
2. The added NLD depends on multitone phase pattern. the variance over different random realizations is the sum of the variance due to nonlinearities and variance due to noise.

The statistical analysis of a sequential random multitone uses these two assumptions to estimate:

1. The best linear approximation (BLA) of the input-output FRF of the DUT,
2. The measurement noise power spectrum,
3. The stochastic NLD power spectrum,
4. A global figure of merit that quantifies the total amount of stochastic nonlinearities (SNLD).

After compensating for any input-output delay and disregarding the first period for each realization, we have $P \times M$ frames to analyze. The DFT for each frame is calculated without time weighting. The windowing and associated spectral leakage is avoided thanks to the periodicity of the multitone within each frame. For each multitone frame, the output spectrum for a given period p of a given realization m is expressed by:

$$Y^{[m,p]} = G_{BLA}U^{[m]} + Y_S^{[m]} + NZ^{[m,p]} \quad (6)$$

where:

- The frequency index has been dropped for convenience,
- $Y^{[m,p]}$ is the DFT of the output period,
- G_{BLA} is the linear FRF (BLA) , independent of the input random phases
- $U^{[m]}$ is the DFT of the input period, constant for the realization m
- $Y_S^{[m]}$ is the stochastic spectrum, dependent of the input random phases
- $NZ^{[m,p]}$ is the noise spectrum for the frame

Averaging is done first along the P periods of each realization m:

$$G^{[m,p]} = \frac{Y^{[m,p]}}{U^{[m]}} \quad (7)$$

$$G^{[m]} = \frac{1}{P}\sum_{p=1}^{P} \hat{G}^{[m,p]} \quad (8)$$

$$\sigma_{nz}^{2[m]} = \frac{1}{P(P-1)}\sum_{p=1}^{P} \left|\hat{G}^{[m,p]} - \hat{G}^{[m]}\right|^2 \quad (9)$$

where $G^{[m,p]}$ is the FRF estimate for the period, $G^{[m]}$ is the average FRF for the realization, and $\sigma_{nz}^{2[m]}$ is the sample noise variance.

Then, the BLA of the system FRF is obtained by averaging over M random realizations:

$$G_{BLA} = \frac{1}{M}\sum_{m=1}^{M} G^{[m]} \quad (10)$$

$$\sigma_{BLA}^2 = \frac{1}{M(M-1)}\sum_{m=1}^{M} \left|G^{[m]} - G_{BLA}\right|^2 \quad (11)$$

$$\sigma_{BLA,nz}^2 = \frac{1}{M^2}\sum_{m=1}^{M} \sigma_{nz}^{2[m]} \quad (12)$$

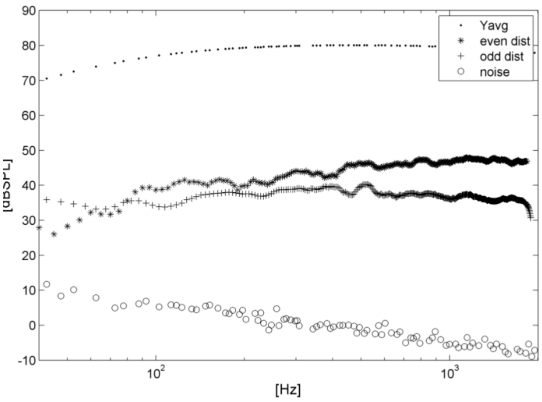

Fig. 8: Separate estimation of even- and odd-order NLD

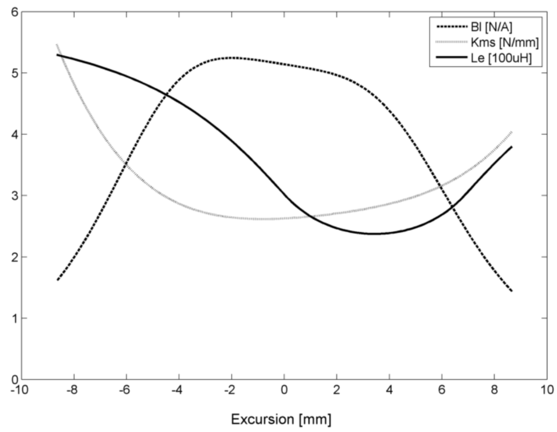

Fig. 9: Large signal parameters of simulated driver.

where G_{BLA} is the BLA of the system FRF, σ^2_{BLA} is the total variance of the FRF, and $\sigma^2_{BLA,nz}$ is the noise variance.

Finally, the variance due to nonlinearities can then be deduced by subtracting the noise variance from the total variance and the NLD for one realization can be estimated:

$$\sigma_S^2 = \max(\sigma^2_{BLA} - \sigma^2_{BLA,nz}, 0) \qquad (13)$$

$$\hat{Y}_S = M\sigma_S U_0 \qquad (14)$$

$$SNLD = \frac{\|\sigma_S U_0\|}{\|G_{BLA} U_0\|} \qquad (15)$$

where σ_S^2 is the BLA variance due to NLD, $\hat{Y}_S$ is the estimated NLD spectrum (for one realization), U_0 is the multitone amplitude spectrum (constant) and $\|.\|$ stands for the 2-norm. See [2] for detailed explanations and justifications.

2.7 Alternatives

In this paper we use full multitones where all harmonics are excited. Other variants can be used where only some selected harmonics are active. The use of multitone with odd-only harmonics gives the means to distinguish between odd and even order nonlinearities (see Fig. 8). When odd-only harmonics are excited, even-order nonlinearities produce only terms at even frequencies, which are easily detected in the output spectrum. Similarly, odd nonlinearities produce odd-order terms only. These can be detected as well if some odd harmonics are not excited (e.g. one randomly chosen in every 3 successive harmonics).
As an other alternative, an harmonic grid with randomly placed non-excited frequencies makes it possible to estimate the NLD by measuring the output level at the non-excited frequencies and interpolating over neighboring frequencies. This method enables a fast evaluation of the NLD with only one realization, with the drawback of increased variability of the results.

3 RESULTS

3.1 Simulations

First, we tested the method on a simulated loudspeaker. The model used is nonlinear and uses the large signal parameters depicted in Fig. 9 and measured on a KLIPPEL system. We use a nonlinear state-space model that calculates recursively the output samples from the input samples and the excursion dependent parameters $K_{ms}(x), Bl(x), L_e(x)$ of the loudspeaker [11].

We played music and multitone signals through the simulated loudspeaker at increasing rms levels. A pink noise of prescribed rms level has been added at the output to simulate measurement noise. The peak excursion was kept under 8 mm to stay within the operating range of the parameters. The music NLD was measured using input-output coherence analysis [7] and the resulting distortion metric is denoted Non-Coherent Distortion

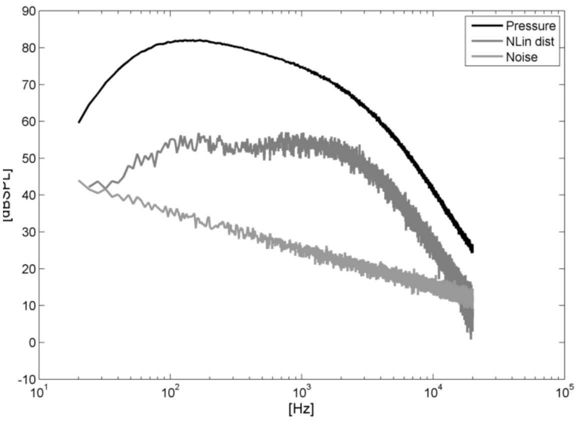

Fig. 10: Simulated driver with output noise. Output level = 104 dBSPL, Noise = 60 dBSPL, SNLD = 8%.

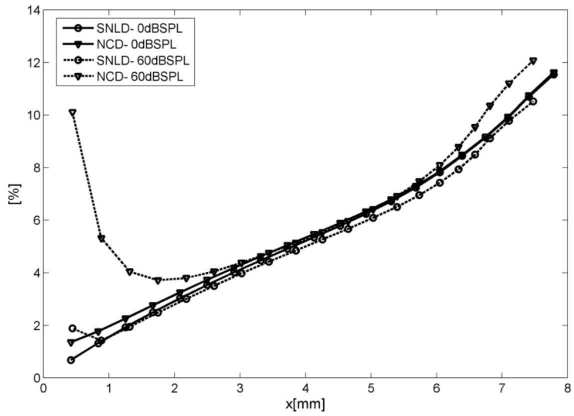

Fig. 11: Comparison SNLD vs. NCD with two different output noise levels)

(NCD). NCD is the ratio of non-coherent power to the output power. Fig. 10 shows how the SNLD algorithm is able to separate the BLA response, the output noise and the SNLD spectrum. A comparison with NCD on a pseudo-noise stimulus (with same power spectrum and amplitude distribution as the multitone) shows a good match with SNLD when the output noise level is low (Fig. 11). When the noise level increases, a bias appears on NCD for low Signal to Noise Ratio (SNR). It is worth noting that SNLD is much less sensitive to noise.

A comparison between multitones and different music pieces shows that the distortion figures are in the same range (Fig. 12). It is interesting to note the great range of distortions obtained, due to different spectral and temporal shapes. The differences between NCD and SNLD are due to variance in the spectral domain and amplitude domain. Fine tuning the characteristics of the multitone can mimic different kinds of music like jazz, classical, metal, etc. SNLD has the nice property of having a steady rise with increasing input levels. That is not the case with NCD.

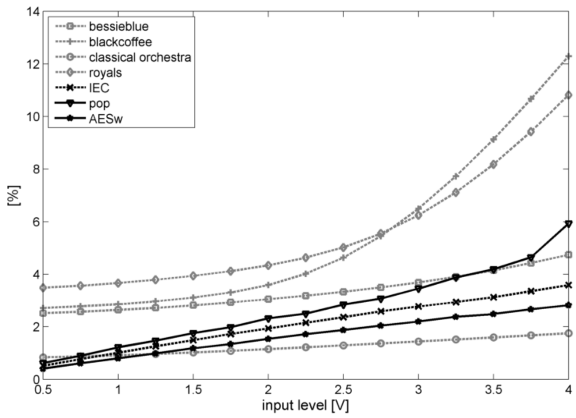

Fig. 12: Music NCD vs Multitone SNLD

It is worth noting that multitone SNLD is a statistical analysis that uses a random input signal and is based on a sample variance. Therefore the SNLD values are also random and have a confidence interval that increases with the mean value. In other words, the variance of the distortion estimates increase with the expected value (Fig. 13). That is consistent with what we have seen with music signals (Fig. 1). Increasing the number

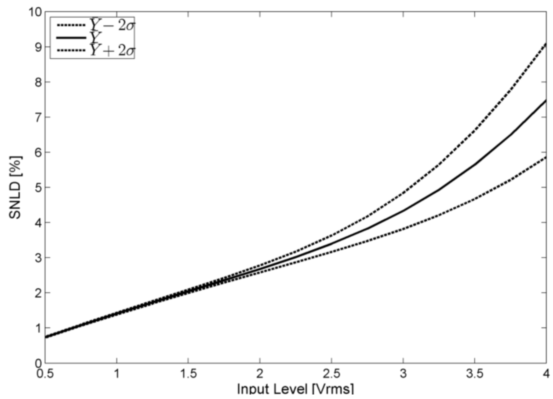

Fig. 13: Multitone SNLD 95% confidence interval. $\bar{Y}$ = mean value, σ = std-deviation

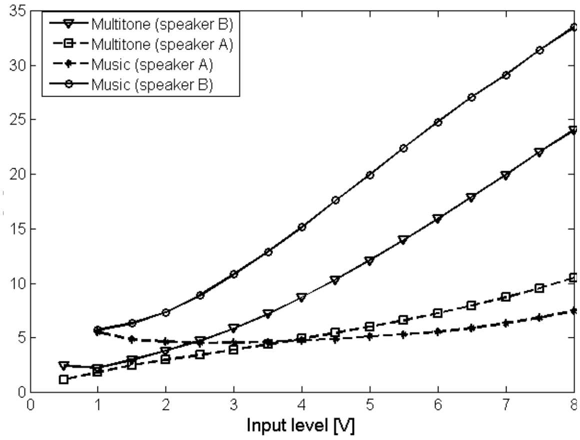

Fig. 14: Comparison of multitone SNLD vs. music NCD on speaker A and B.

of multitone realizations will narrow the confidence interval of the SNLD estimates.

3.2 Measurements

To verify our simulation results, we performed distortion measurements on a loudspeaker using a multitone signal. National Instrument's USB-4431 acquisition card was used to play and record the signal. A GRAS ½″microphone was placed 1 m away from the acoustic center of the loudspeaker. We developed our own custom software in LabVIEW to perform the measurements. The software can play different kinds of stimuli such as multitone and wav files. The user can also select a multitone with different kinds of spectral shapes. The measurements were done in a standard listening room. Because the multitone method relies on averaging, we did not need to perform it in an anechoic chamber and the measurements can be done in a standard listening room. We compared a high-end two-way bookshelf (speaker A) with a low-end system similar in configuration, shape and size (speaker B).

As expected from our simulations, the measurement of NLD vs input level with multitone are consistent with music (Fig. 14). As pointed out before, the distortion on music is limited for low input levels, due to the sensitivity to noise of the NCD method.

4 DISCUSSION

For many years, harmonic distortion has ruled the roost as a mean for quantifying the distortion performance of loudspeaker systems. The multitone concept presented allows:

- Quantifying harmonic and intermodulation distortion using a stimulus with spectral and crest factor properties very similar to actual music.
- Tailoring spectrum and crest factor properties of the sequence to nearly match any genre of music. This opens the door for further study and possibly an industry consensus for adaptation of a standardized multitone sequence for loudspeaker distortion measurements.
- Averaging during the measurement to improve SNR and allow measurements in situ.
- Separation of noise, odd- and even-order distortion.

Areas of further investigation:

- The method investigated presents a time-averaged distortion. Is there a way to similarly measure instantaneous distortion? What are the perceptual differences of instantaneous distortion versus average distortion?
- Application of perceptual distortion models to the results of this method [12].

5 CONCLUSION

This paper discussed a new multitone sequence stimulus to measure loudspeaker system distortion. This new sequence exhibits spectral and amplitude distribution properties that can be the same as real music and thus provides a more balanced global assessment of loudspeaker system nonlinear distortion when compared to more traditional distortion methods. The new sequence supports measurement averaging and the ability to distinguish noise from distortion components. Further, odd- and even-order distortion contributions are separable. In addition to relative "real music" distortion comparisons between DUT's, this method is advantageous in that it provides significant insight into trade-offs between distortion performance and system parameters such as compressor/limiter settings when subjected to real music.

6 Acknowledgments

Samsung Electronics and Samsung Research America supported this work. The authors would like to thank the entire staff of Samsung's US Audio Lab who helped with the measurement setup, offered insightful suggestions, reviewed the manuscript and contributed to its content.

References

[1] Voishvillo, A., Terekhov, A., Czerwinski, E., and Alexandrov, S., "Graphing, Interpretation, and Comparison of Results of Loudspeaker Nonlinear Distortion Measurements," *J. Audio Eng. Soc*, 52(4), pp. 332–357, 2004.

[2] Pintelon, R. and Schoukens, J., *System identification: a frequency domain approach*, John Wiley & Sons, 2012, ISBN 9780470640371.

[3] Fielder, L. D. and Benjamin, E. M., "Subwoofer Performance for Accurate Reproduction of Music," *J. Audio Eng. Soc*, 36(6), pp. 443–456, 1988.

[4] Keele Jr., D. B., "Development of Test Signals for the EIA-426-B Loudspeaker Power Rating Compact Disk," in *Audio Engineering Society Convention 111*, 2001.

[5] Voishvillo, A., "Measurements and Perception of Nonlinear Distortion—Comparing Numbers and Sound Quality," in *Audio Engineering Society Convention 123*, 2007.

[6] Starobin, B., "CEA-2010 : Standard Method of Measurement for Powered Subwoofers," 2010.

[7] Temme, S. and Brunet, P., "A New Method for Measuring Distortion Using a Multitone Stimulus and Noncoherence," *J. Audio Eng. Soc*, 56(3), pp. 176–188, 2008.

[8] "IEC 60268-1. Sound system equipment," 1985.

[9] Chapman, P. J., "Programme Material Analysis," in *Audio Engineering Society Convention 100*, Audio Engineering Society, 1996.

[10] Kotz, S., Kozubowski, T., and Podgorski, K., *The Laplace distribution and generalizations: a revisit with applications to communications, economics, engineering, and finance*, Springer Science & Business Media, 2012.

[11] Brunet, P., *Nonlinear System Modeling and Identification of Loudspeakers*, Ph.D. thesis, Northeastern University, 2014.

[12] Pahomov, M. and Na, Y. H., Sr., "Method for Objective Evaluation of Nonlinear Distortion," in *Audio Engineering Society Convention 139*, 2015.