Audio Engineering Society

Convention Paper 9900

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

A New THD+N Algorithm for Measuring Today's High Resolution Audio Systems

Alfred B. Roney[1] and Steve Temme[2]

[1]*The MathWorks, Inc. (formerly Listen, Inc.)*
[2]*Listen, Inc.*

Correspondence should be addressed to Steve Temme (stemme@listeninc.com)

ABSTRACT

We present a mathematical definition of Total Harmonic Distortion + Noise suitable for testing high-resolution digital audio systems. This formal definition of the "distortion analyzer" mentioned in AES17 defines THD+N as the RMS error of fitting a sinusoid to a noisy and distorted sequence of measurements. We present the key theoretical result that under realistic conditions a modern THD+N analyzer is well-described by a Normal probability distribution with a simple relationship between relative error and analysis dwell time.

These findings are illustrated by comparing the output of a commercial distortion analyzer to our proposed method using Monte Carlo simulations of noisy signal channels. We will demonstrate that the bias of a well-designed distortion analyzer is negligible.

1 Introduction

Total Harmonic Distortion + Noise (THD+N) is used throughout the audio industry as a easily-interpreted measure of system fidelity. This metric has not kept up with advances in technology, remaining defined in terms of analog notch filters. Since the filter can have a range of Q and be of arbitrary order, results obtained from different analyzers cannot be compared. Also lacking is a clear and practical definition of the expected repeatability of measurements.

In this paper we will present a modern definition of THD+N that is suitable for implementation as a DSP algorithm. From this definition a model of the repeatability of the measurements will be derived. Finally, we will use Monte Carlo simulations to illustrate and compare the performance of the old and new methods. From the comparison, we will see that the inherent bias of an THD+N estimator can be negligibly small in practice, and far less than the bias introduced by the transition bands of a notch filter.

2 Methods

We call the input signal that we generate the *stimulus*, and use the symbol $u(t)$ to represent it in the continuous-time domain. Starting from the stimulus signal's continuous-time deterministic representation,

$$u_0(t) = a_0 \cos(2\pi f_0 t - \theta_0) \tag{1}$$

we substitute $t \to \frac{k}{m}$ and $f_0 \to \omega_0$ to obtain a deterministic representation of its sampled equivalent.

$$u_k = a_0 \cos\left(2\pi\omega_0 \frac{k}{m} - \theta_0\right) \tag{2}$$

The discrete parameters $k \in \mathbb{Z}$ and $m \in \mathbb{N}$ are the sample index and analysis buffer size, respectively. The amplitude, $a_0 \in (0,1]$, is a unitless proportion of the full-scale amplitude. The frequency, $\omega_0 \in \left[1, \frac{m}{2}-1\right]$ *cycles / window*, corresponds to the bin indices of an m-point discrete Fourier transform. The phase, $\theta_0 \in (-\pi,+\pi]$, is as a unitless radian angle. Together, these parameters comprise a reduced form independent of sampling rate and analog scale that can readily be converted to any desired units if the sampling rate and analog-to-digital scaling factors are known. These units were chosen to prevent the key results from being needlessly complicated by the presence of the conversion factors.

Next, we need a model of the noise and distortion in our *device under test* (DUT). Let X be a random vector constructed from m realizations of a zero-mean Normal distribution.

$$X \sim \mathcal{N}\left(0, \sigma_X^2\right) \tag{3}$$

$$\sigma_X^2 = \mathbf{E}\left[X^2\right] \tag{4}$$

The individual samples are independent and distributed identically. The parameter $\sigma_X \geq 0$ is the standard deviation of the noise model's distribution. X is a realistic model of noise since modern digital audio systems introduce negligible distortion, and a Gaussian process accurately describes an electrical noise floor with an RMS signal level of σ_X.

We combine our deterministic stimulus with our random noise floor model to obtain a random process model of the *resultant*, Y.

$$Y = u_k + X \tag{5}$$

We use capital symbols to denote random processes, and lowercase symbols to represent either deterministic variables or realizations of a random process. Thus, a realization of the resultant is represented as y_k,

$$y_k = u_k + x_k \tag{6}$$

and its RMS level is defined in the usual way.

$$\sigma_Y = \sqrt{\frac{1}{2}a_0^2 + \sigma_X^2} \tag{7}$$

There are two common definitions of THD+N. The first is the square root of the ratio of the non-stimulus energy to that of the total signal.

$$\text{THD+N}_R = \sqrt{\frac{\mathbf{E}\left[X^2\right]}{\mathbf{E}\left[Y^2\right]}} \tag{8}$$

$$= \sqrt{\frac{\sigma_X^2}{\frac{1}{2}a_0^2 + \sigma_X^2}} \tag{9}$$

The subscript R denotes that the resultant signal level serves as the reference. Equation 9 was adopted for THD+N analyzers during the pre-digital audio era. A tuned notch filter output a *residual*, and the RMS levels of the resultant and residual were computed by analog measurement circuits. This convention leads to an intuitive $\text{THD+N}_R \in [0,100)\%$ output range, and remains in common use.

For reasons of convenience we define THD+N as the reciprocal of the *signal-to-noise* ratio.

$$\text{THD+N}_F = \sqrt{\frac{\sigma_X^2}{\frac{1}{2}a_0^2}} \tag{10}$$

$$= \frac{\sqrt{2}}{a_0} \times \sigma_X \tag{11}$$

The subscript F denotes that the level of the fundamental is used for reference. A pair of simple formulas convert measured results from one definition to the other.

$$\text{THD+N}_R = \sqrt{\frac{\text{THD+N}_F^2}{1 + \text{THD+N}_F^2}} \tag{12}$$

$$\text{THD+N}_F = \sqrt{\frac{\text{THD+N}_R^2}{1 - \text{THD+N}_R^2}} \tag{13}$$

A digital computer allows a detailed analysis of the signal buffer to estimate the parameters of frequency, phase and amplitude[1]. In the digital domain, we estimate the parameters of u_k and subtract it from the resultant instead of processing y_k with a notch filter. We call this process *synthetic filtering*. It provides a close approximation of an infinitesimally-narrow notch filter without the inconvenience of infinite settling time.

Since Equation 11 shows that THD+N is directly proportional to σ_X, the problem of estimating THD+N becomes the problem of estimating the standard deviation of the Gaussian random variable X that represents the noise floor of the DUT. This means that for sufficiently large m we can model a perfect THD+N estimator as the standard deviation of a standard deviation estimator, which is also a normally-distributed random variable.

$$\text{THD+N} \sim \mathcal{N}\left(\mu_{\text{THD+N}}, \sigma^2_{\text{THD+N}}\right) \tag{14}$$

The mean and standard deviation of the underlying probability distribution for the THD+N estimator are the parameters of Equation 14 [2].

$$\mu_{\text{THD+N}} = \text{THD+N} \tag{15}$$

$$\sigma_{\text{THD+N}} = \frac{\text{THD+N}}{\sqrt{2m}} \tag{16}$$

One important and convenient consequence of Equation 16 is that the uncertainty scales proportionally to the magnitude of the measurement. In order to specify the precision across the full dynamic range of a perfect THD+N analyzer, we can divide the standard deviation by the mean to obtain the relative error.

$$\sigma_{\text{rel}} = \frac{\sigma_{\text{THD+N}}}{\mu_{\text{THD+N}}} \tag{17}$$

$$= \frac{1}{\sqrt{2m}} \tag{18}$$

Equation 18 shows that the relative uncertainty of a perfect THD+N estimator depends solely on the buffer size m when m is sufficiently large.

In order to test this theory, we performed a Monte Carlo simulation that allowed us to precisely specify the amount of noise and distortion in the signal. This was implemented by using the pseudo-random number generation facilities of MATLAB to simulate a Gaussian white noise floor that was added to each sample of a stimulus sinusoid. The signal was then dithered and quantized to 24-bit signed integer code points using a triangular pseudo-random process that conforms to AES17[3]. One analyzer supported buffer-based inputs, so phase was randomized by computing pseudo-random deviates from a uniform distribution. Since the filter-based analyzer did not support buffer-based operation, a sequence of buffers with continuous phase were processed using a virtual strip chart feature, then exported as a spreadsheet for analysis.

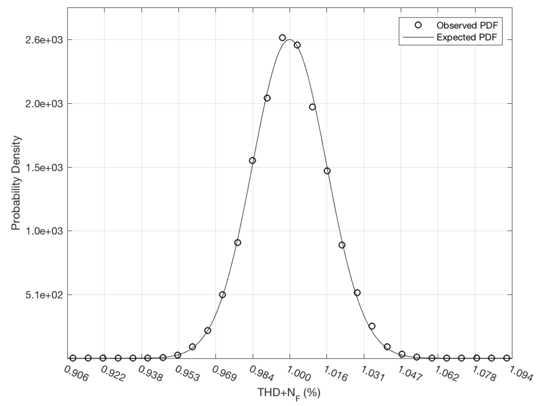

Fig. 1: Comparison of expected PDF to measurements. These data, obtained numerically during qualification of the Monte Carlo signal generator, demonstrate the model's predictive strength.

3 Results

Figure 1 shows that the model describes the signal process with high accuracy when the underlying noise floor is a Gaussian process. The data used to obtain this figure were extracted directly from the noisy channel model, prior to the addition of dither and subsequent quantization. Thus, they represent the underlying signal model precisely and illustrate the statistical properties of an ideal THD+N estimator.

When a notch filter is used to remove the stimulus from the resultant, the transition bands of the filter remove considerable energy from the residual signal. Figure 2 uses data collected from a widely-deployed commercial instrument to illustrate the effect. It is important to note that although the bias remains essentially constant with respect to dwell time since it is dominated by the width of the filter transition bands, the bias is obscured by the spread of the measurement for short dwell times. The notch filter also increases the spread of the PDF when compared to our proposed method, to which we now turn our attention.

Figure 3 presents a repeat of the experiment using a research prototype of the synthetic filter method. Of key importance is the close match of the shapes of the theoretical and observed probability distributions. The parameter estimation process introduces a small positive bias, but the data suggests that this error can

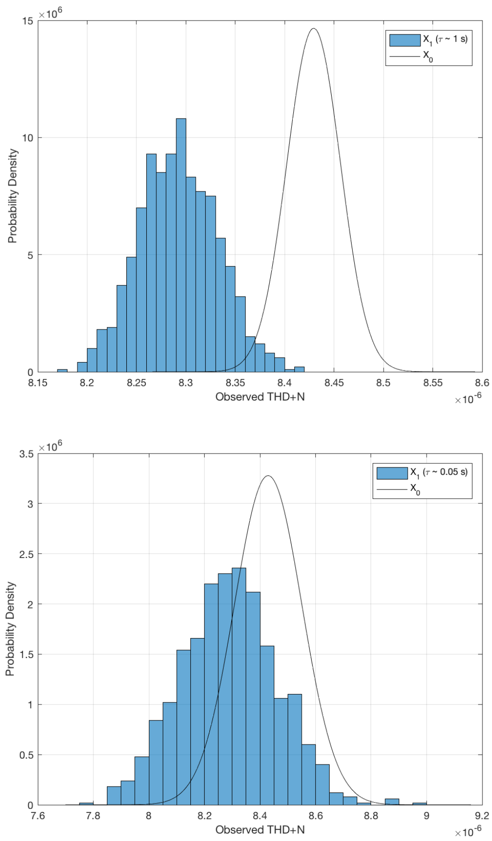

Fig. 2: Measurement distribution for notch filter method. Here X_0 represents the ideal process, and X_1 represents the observed process, a sequence of measurements output by a widely-deployed commercial distortion analyzer system. The analysis dwell time is represented by τ *seconds*.

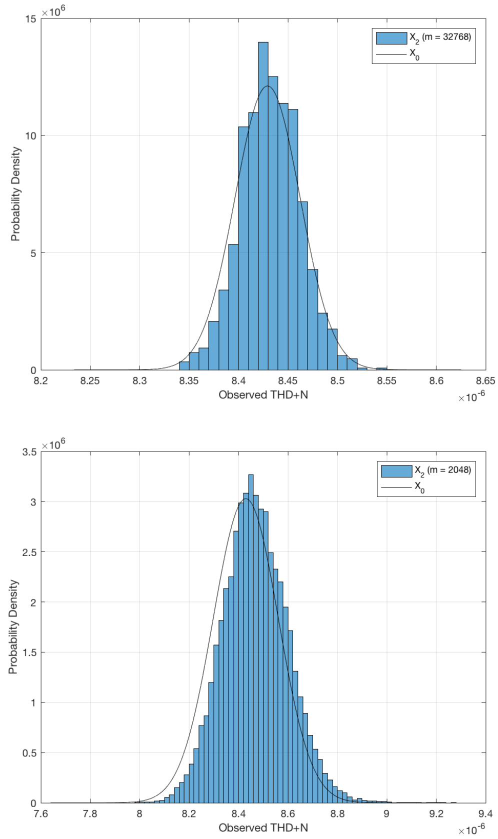

Fig. 3: Measurement distribution for synthetic filter method. Here X_0 represents the ideal process, and X_2 represents the observed process. The analysis dwell time is represented by m *samples*.

be made negligibly small by increasing the analysis dwell time. Thus, for a given level of error tolerance, less dwell time is required when using the synthetic filter method. Surprisingly, we will see that even better algorithms than this prototype exist.

Figure 4 presents an analysis of the error using a pre-production prototype of a new distortion analyzer slated for inclusion in a future version of Listen's SoundCheck software. This analyzer uses a proprietary stimulus estimator that out-performs the research prototype presented in Figure 3. These results clearly show that when the residual is dominated by the noise floor, the estimator behaves in a quasi-ideal fashion. The positive bias is completely swamped by the noise-induced uncertainty. As the noise RMS decreases the dither and quantization noise establish a minimimum measured value. It is important to note that the uncertainty decreases as the noise distribution narrows from Normal. This is evidenced by the reduction in relative error when the triangular dither noise and uniform quantization errors become dominant.

4 Discussion

The data presented in Section 3 clearly demonstrate that using notch filters to remove the stimulus corrupts the measurement by subtracting a significant portion of the residual energy. Furthermore, any modulation products of the stimulus that fall within the stopband of the notch filter will be removed. Thus, any traditional notch filter instrument is blind in a region critical to modern digital systems.

The sole disadvantage of using a synthetic filter-based instrument is that comparisons to legacy data must take into account the different types of bias and uncertainty. Filter-based instruments are also susceptible to this inconvenience, however, due to the range of Q allowed by AES17[3]. To address this problem, the new algorithm analyzed in Figure 4 can be configured to use a frequency domain weighting curve that closely approximates the notch filter from the instrument used for comparison. Unfortunately, this notch emulation introduces the same bias mechanism and increases the uncertainty of the measurement. Furthermore, any low-frequency modulation products of the stimulus will be filtered out when this emulation is enabled. It is for this reason that we recommend that new testing protocols should adopt the THD+N definition presented in this paper, specifying filter-derived measurements only

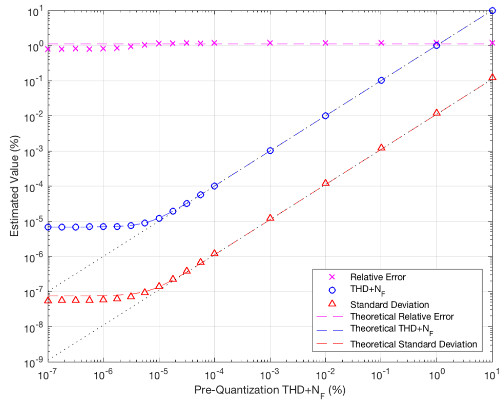

Fig. 4: Estimated THD+N vs model THD+N, with error analysis. This figure was generated by analyzing 4096 realizations of Y per expected value of THD+N with a beta version of Listen's new THD+N analyzer. Each realization consisted of 4096 samples, the frequency was 20 *cycles / buffer*, the amplitude was 1, and the phase was drawn from a Uniform distribution over the half-open interval $(-\pi, +\pi]$. At a sampling rate of 44.1 *kHz*, this is approximately 93 *ms* of a 215 *Hz* stimulus. AES 17-compliant triangular dither was added prior to 24-bit signed integer quantization. Notch filter emulation was disabled.

when a direct comparison to a historical datum must be made.

5 Summary

By using synthetic filters instead of a traditional notch filter, we can obtain a near-ideal estimate of the residual error. It can easily be shown that uncertainty in the parameter estimates introduces a small positive bias into the measurement. The data we have presented shows that this systematic error is negligible for all but the most demanding cases. Therefore, we have left the proof of this non-ideal behavior as an exercise for the reader.

In conclusion, we have presented the definition and error analysis of a distortion analyzer suitable for use in the development of modern digital audio systems. Our comparisons clearly demonstrate the superiority of synthetic filters to traditional notch filters when compared against an objective theoretical benchmark. While it is understandable that a notch filter may be simulated to enable comparison to historical data, or for establishing a measurement baseline when converting from notch filters to synthetic filters, it is the professional opinion of the authors that the industry should transition to synthetic filters when new testing protocols are developed.

References

[1] Kay, S. M., *Fundamentals of Statistical Signal Processing: Estimation Theory*, volume 1, Prentice Hall PTR, Upper Saddle River, New Jersey, 1993.

[2] J. F. Kenney, E. S. K., *Mathematics of Statistics*, volume 1, D. Van Nostrand Company, Inc., Princeton, NJ, 3 edition, 1954.

[3] "AES17-2015: AES standard method for digital audio engineering - Measurement of digital audio equipment," 2015.