Audio Engineering Society

Convention Paper 10463

Presented at the 150th Convention
2021 May 25–28, Online

Characterising non-linear behaviour of coupling capacitors through audio feature analysis and machine learning

Christopher Johann Clarke, Balamurali B T, and Jer-Ming Chen

Science, Mathematics and Technology, Singapore University of Technology and Design

Correspondence should be addressed to Christopher Johann Clarke (christopher_clarke@mymail.sutd.edu.sg)

ABSTRACT

Different electrically-equivalent capacitors are known to impact the sonic signature of the audio circuit. In this study, the non-linear behaviour of five different coupling capacitors of equivalent capacitance (marketed as "audio capacitors"), one at a time, are characterised. A dataset containing the input and output signals of a non-linear amplifier is logged, its audio features are extracted and the non-linear behaviour is analysed. Machine learning is then applied on the dataset to supplement analysis of the Total Harmonic Distortion (THD). The five capacitors' THD performance seem to fall into two categories: below 200 Hz, there is significant standard deviation of 14.1 dBc; above 200 Hz, the capacitors show somewhat similar behaviour, with only 0.01 dBc standard deviation. This separation however, does not hold at regions below 0.2 V. A support vector machine model is trained and classifies the five capacitors well above chance: the best classification at 84% and worst at 36%. The methodology introduced here may also be used to meaningfully assess the complicated behaviour of other audio electronic components.

1 Introduction

Although it is possible to perform sonic manipulation entirely within a Digital Audio Workstation (DAW), many professionals still incorporate the use of analog audio equipment in their set up – studio or otherwise. It has been documented that analog audio circuit will impart a "sonic signature" to the audio [1][2][3]. One non-negligible aspect (amongst other parameters) of the "sonic signature" is non-linear distortion.

The characteristics of non-linear distortion in analog audio circuits can be attributed to many electronic components or design choices in a circuit. One of the components responsible is the coupling capacitor at the output. In 1998, Hood [4] reported that it is probable that capacitors will introduce waveform distortion into any audio line, but direct measurements made at the time did not have sufficient evidence. This neither implies that measurements of the non-linear behaviour cannot be quantified, nor does it comment on the perceptual importance of the non-linearity. To help analog audio circuit designers make decisions for the coupling capacitor construction, this paper analyses the effect that it has on the overall system non-linearity.

1.1 Background and Previous Work

Capacitors are used in many different parts of the audio electronic design, such as High Pass Filtering (HPF) or DC blocking. Gaskell [5] shows that it is possible for university level sound recording students to discern differences in capacitors using the "recordings of various solo instruments as well as several commercial recordings frequently used in the students' ear training

classes". The different capacitors varied in Total Harmonic Distortion plus Noise (THD+N) vs frequency, mostly in the lower frequency ranges. In addition to this, van der Veen and van Maanen [6] provide analytic expressions to the non-linearity of capacitors, for a variety of circuit topologies. Dodds et al. [7] shows that a trained panel of listeners with "critical listening experience" is capable of discernment through detailed test listening methodology (ITU-R BS.1116-1 [8]).

1.2 Objectives and Approach

The objective of this paper is to provide a characterisation and analysis of the non-linear behaviour of coupling capacitors. The (electrically equivalent) coupling capacitor of the circuit will be varied, and the non-linear performance of the different capacitor construction (materials, design, morphology, and other physical specifications) will be compared. The analysis will be further supplemented with machine learning, where a feature map will be constructed by extracting different audio features such as Total Harmonic Distortion (THD), Spurious Free Dynamic Range (SFDR), SIgnal-to-Noise And Distortion ratio (SINAD), and Signal-to-Noise Ratio (SNR). Based on these chosen features, a Support Vector Machine (SVM) model will be trained to identify capacitors used in the non-linear circuit. The analysis and classification of the dataset will indicate difference in the non-linear behaviour across different capacitors construction.

2 Methodology

2.1 Specifications of the Test Circuit

The test circuit (Figure 1, adapted from [9]), was chosen to demonstrate changes in the non-linear characteristics of the system[1] when the coupling capacitor construction is different and, was prepared and housed in a metal chassis to isolate from external electromagnetic interference. This circuit was chosen as it is known to exhibit non-linear distortion [10].

[1] Anecdotally, the author of this paper is able to discern differences between the capacitors. However, perceptual discernment is not within the scope of our investigation.

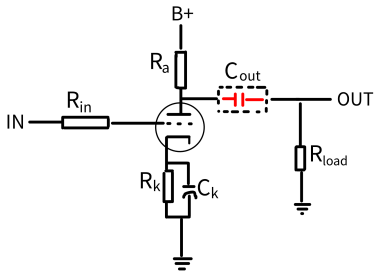

Figure 1: Schematic of the circuit with the capacitor under test in red.

The component specifications of the circuit are $R_{in} = 8.2\,k\Omega$, $R_a = 100\,k\Omega$, $R_k = 1.2\,k\Omega$, $C_k = 33\,\mu F$, $C_{out} = 0.33\,\mu F$, $R_{load} = 100\,k\Omega$, B+ = 250 VDC, Vacuum Tube = ECC83/12AX7. Using a Norman Koren SPICE model [11] for the ECC83/12AX7, the transconductance (g_m) for this single ended [9] amplifier is 1.876 mS with an amplification factor (μ) of 93.5. With a SPICE simulation, the HPF frequency calculated and is below the range of the test conditions (20-20,000 Hz).

2.2 Details of the Capacitors Under Test

The five capacitors chosen for characterisation are all marketed and labelled as "audio capacitors" that have been rated at 0.33 μF with a tolerance of $\pm 1\%$. To ensure that the test variables were electrically identical, any variance in the capacitance, equivalent series resistance (ESR), or equivalent series inductance (ESL) were compensated (until $\pm 0.01\%$ variance achieved) with the use of a low valued capacitor, resistor, or inductor. The manufacturer datasheet (if provided) contains only limited information about the capacitor's construction, as detailed in Table 1.

Table 1: Capacitor labels and plot colours for Figure 2 to 5.

Capacitor		Plot Colour
Capacitor A	Paper in oil, Tin Foil	Red
Capacitor B	Paper in wax, Aluminium Foil	Green
Capacitor C	Plastic, Metalized Foil	Blue
Capacitor D	Oil, Metal Foil	Cyan
Capacitor E	Metalized Polypropylene	Magenta

The five capacitors were placed one at a time as C_{out}, as seen in Figure 1, and the test circuit was measured

using sine wave signals generated from the function generator of an Digilent Analog Discovery 2 from 50 mV to 600 mV (10 mV intervals) at each frequency of 20 to about 20,000 Hz (see eq. (11)). Both the input (oscilloscope channel 1) and output (oscilloscope channel 2) of the Digilent Analog Discovery 2 were logged on a computer. The input signals were checked for distortion and noise to ensure optimum test conditions. Features such as THD, SFDR, SINAD, and SNR were extracted from the output signals to obtain a feature map for the SVM model.

2.3 Feature Extraction and Support Vector Machine Model

THD, SFDR, SINAD, and SNR [12][13][14] are defined in this paper, respectively, as:

$$THD = \frac{\sqrt{(P_{F2})^2 + (P_{F3})^2 + (P_{F4})^2 + (P_{F5})^2 + (P_{F6})^2}}{P_{F1}}, \tag{1}$$

where F_1 is the fundamental, P_{F1} is the power of the fundamental, and n in F_n identifies the n^{th} harmonic,

$$SFDR = \frac{\sqrt{(P_{Fx})^2}}{P_{F1}}, \tag{2}$$

where P_{F1} is the power of the fundamental and P_{Fx} is the power of the greatest non-harmonic component,

$$SINAD = 10log_{10}\frac{P_{F1} + P_N + P_D}{P_{F1}}, \tag{3}$$

where P_{F1} is the power of the fundamental, P_N is the power of noise, and P_D is the power of all harmonics of the signal, and

$$SNR = \frac{P_{signal}}{P_N}, \tag{4}$$

where P_{signal} is the power of the signal and P_N is the power of noise.

The features are then normalised using a Kaiser-Bessel window function [15]:

$$\frac{L.sinh(\beta.\sqrt{(1-(Lf/\alpha)^2)})}{I_0.\beta(\beta.\sqrt{(1-(Lf/\alpha)^2)})} \tag{5}$$

where I_0 is the modified Bessel function of the first kind, L is the window duration, and $\beta = \pi\alpha$, where α is a non-negative real number that determines the shape of the window.

An SVM model was then trained using this multi-dimensional dataset to identify the capacitors used in the non-linear circuit. The SVM model was chosen as it is a model capable of handling higher input dimensions, and, no dimensionality reduction was applied as it may risk information loss [16]). The SVM model, trained with a Gaussian Radial Basis Function (RBF) kernel (eq. 10) was chosen to reduce computation time and can be expressed in the following form:

Given a training set of n labeled examples,

$$S_t = (x^{(t)}, y^{(t)}), \text{where } t = 1, ..., n, \tag{6}$$

and parameters of a classifier (θ, θ_0). The distance of each training point to the decision boundary can be measured by:

$$\gamma^{(t)}(\theta, \theta_0) = \frac{y^{(t)}(\theta \cdot x^{(t)} + \theta_0)}{||\theta||} \tag{7}$$

Optimising to maximise the minimum distance to the boundary, we can formulate the primal and dual optimisation as follows:

$$\text{(primal)} \qquad min\frac{1}{2}||\theta||^2 \tag{8}$$

subject to $y^{(t)}(\theta \cdot x^{(t)} + \theta_0) \geq 1$, where $t = 1, ..., n$.

$$\text{(dual)}\ max\sum_{t=1}^{n}\alpha_t - \frac{1}{2}\sum_{t=1}^{n}\sum_{t'=1}^{n}\alpha_t\alpha_{t'}y^{(t)}y^{(t')}\mathbf{K}(x^{(t)} \cdot x^{(t')}) \tag{9}$$

subject to $\alpha_t \geq 0$, where $t = 1, ..., n$.

The Gaussian Radial Basis Function (RBF) kernel [17] used is:

$$\mathbf{K}(\mathbf{x}, \mathbf{x}') = exp\left[-\frac{||\mathbf{x} - \mathbf{x}'||^2}{2\sigma^2}\right] \tag{10}$$

where σ is a free parameter and $||\mathbf{x} - \mathbf{x}'||^2$ is the Euclidean distance.

3 Results and Discussion

Although different audio features (THD, SFDR, SINAD, SNR) are collected, we will mainly focus on THD as it is a familiar (and an intuitive) feature of audio analysis. This allows for the reader to have a better understanding of the inter-capacitor performance. This will be followed by a presentation of the inter-feature scatter plots will be used to demonstrate the complicated behaviour of the relationships among the features and the choice for including machine learning as a supporting strategy for audio analysis.

3.1 Analysis of THD

Figures[2] 2, 3, 4 show the THD (dBc) values across frequency (Hz) and input amplitude (V). The plot was generated from a dataset of 48000 data points, logged in Section 2.2.

The geometry in Figure 2 above is complicated, and can be described somewhat as a section of a saddle surface (cf. catenoid). While the capacitors display similar contours in general, there is noticeable distinctiveness among the surfaces. The top figure in Figure 2 is dominated by Capacitor A (red) on the top surface but, interrupted or dappled indicating surprising complicated behaviour (non-smooth) of the THD values across both frequency and input amplitude: the irregularity of colours is due to the different surfaces interweaving as values cross over each other. The colour on the top surface corresponds to the capacitor with the highest THD at that given frequency and input amplitude. On the underside however, Capacitor E (magenta) occupies the largest area associated with less interruption (in contrast to the top plot), indicating persistently lowest THD values across frequencies and input amplitudes sampled.

Above 200 Hz, the overall THD response of the system sampled across frequency and input amplitude is complicated, and there is no clear distinction among the five capacitors. Below 200 Hz however, the standard deviation of THD among the five capacitors is 14.1 dB while the standard deviation of THD above 200 Hz is 0.01 dB; this strongly contrasting behaviour motivates us to further analyse frequency regions below 200 Hz.

[2]All coloured plots in this paper will be available at *https://christopherclarkesutd.github.io/*. The online plots will also allow for features such as rotation and zoom to provide more clarity, and facilitate more contribution from better insight.

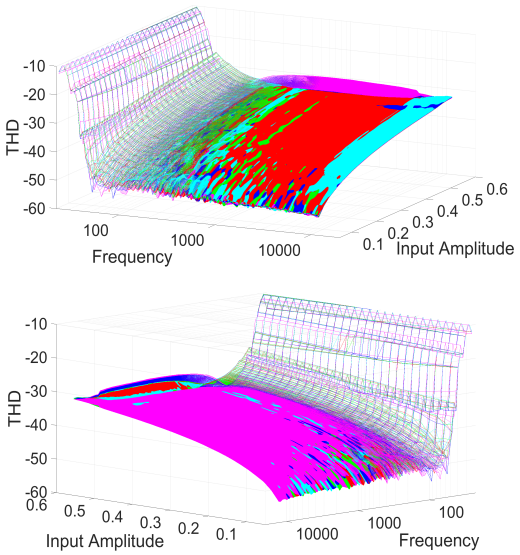

Figure 2: Surface plot[2] of THD (dBc) over input amplitude (V) and frequency (Hz) (legend indicated in Table 1). The non-homogeneous colours indicate intermingling surfaces which are capacitors with the highest THD values at the given frequency and input amplitude. The bottom plot is the same dataset as the top plot but rotated to show the underside. Here, Capacitor E (magenta) mostly dominates, indicating rather less intermingling of the surfaces.

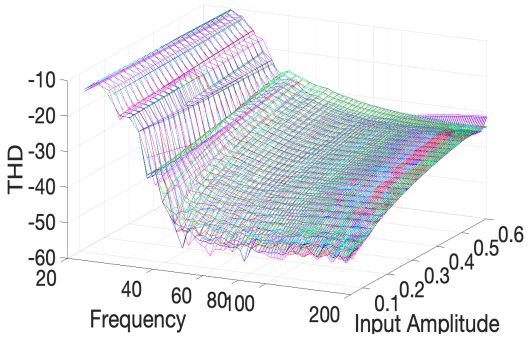

Figure 3: Surface plot[2] of the THD within 20-200 Hz (legend indicated in Table 1). Showing separation of surfaces which are then depicted in detail in Figure 4.

Figure 3 replots the 20-200 Hz frequency region. Below 200Hz, separate surfaces are observed that converge at about 200 Hz (depending on the input amplitude). There is an observable structure with sharp upward slope at low frequencies. A similar behaviour was reported earlier in [18], where capacitors with different electrode materials were sampled across the frequency range with 1 V_{RMS}. Capacitor B (green plot) shows the highest THD for this frequency region.

The distinct surfaces also show a difference across input amplitude. In the upper group of surfaces of Figure 3, the group's trend maintains the slope of increasing THD with increasing amplitude, while the lower group of Figure 3 shows a decrease in gradient of THD at the region of 0.5 V.

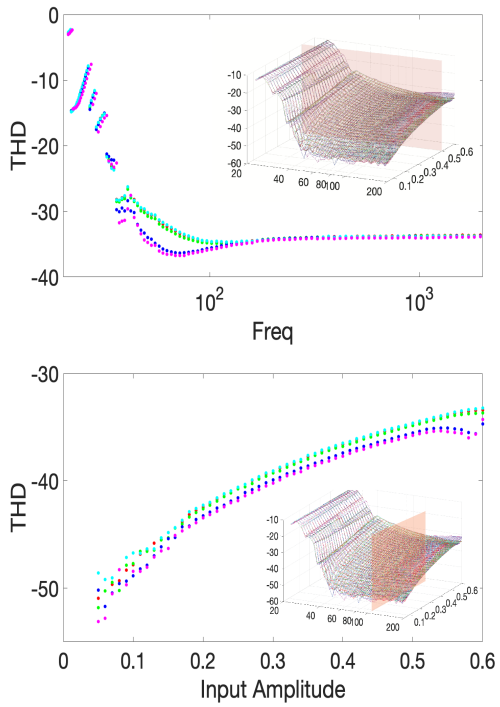

Figure 4: A slice of the surface plot[2] in Figure 3 sampled at 0.5 V (top) and about 100 Hz (bottom) (legend indicated in Table 1). Clear separation of Capacitors A, B, D in one group (red, green, cyan), and Capacitor C, E (blue, magenta) in another group.

Figure 4 shows a slice of Figure 3 at 0.5 V in the upper plot, and a slice of Figure 2 at about 100 Hz. This plot further demonstrates the separation and convergence of the surfaces, highlighting the clear separation of Capacitors A, B and D (red, green, cyan) in one group, and Capacitors C and E (blue, magenta) in another group. The discontinuities in the lowest frequencies indicate further complicated behaviour in the THD. The lower plot shows the similarity in the THD values, which indicates a rough equivalence at lower input amplitudes. The lower figure in Figure 4 shows the separation of surfaces along the input amplitude, and is sampled at about 100Hz. In the region below 0.2 V input amplitude, the THD values are not as separated as before and have an inconsistent distribution. This shows substitutability across capacitors if design specifications are within this amplitude range and THD.

3.2 Inter-feature complexity and the Motivation for Machine Learning

While THD alone shows many interesting and distinguishing behaviours between capacitors, this is merely one aspect of characterizing capacitors. In the context of characterising inter-capacitor non-linear behaviour, more audio features are required to provide a complete picture of the system's characteristics at a given frequency and input amplitude. However the inclusion of more audio features, such as SFDR, SINAD, and SNR (cf. Section 2.3), greatly increases the complexity of visual assessment. A representative sample subset of characteristic plots are shown in Figure 5. These plots illustrate the difficulty and limitations of visual inspection to assess and separate the performance of the system with the change in capacitors. To meaningfully address the complexity of such rich multidimensional data (and to get around the limitation of visual inspection), a machine learning approach is incorporated; the SVM model proposed will provide greater understanding of the multidimensional data.

3.3 Support Vector Machine Classification

The results in Figure 6 were obtained from a slightly reduced dataset of 40,040 data points of the original dataset, with 8,008 for each capacitor. This reduction was to exclude frequencies where the THD calculation would return the results of aliased signals due to the nature of sampling rates. A random set of 70% of the data was selected to be used for training the SVM, while the remaining 30% was used as the test set.

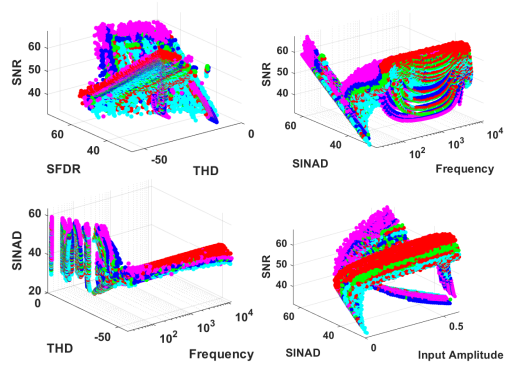

Figure 5: An indicative subset of inter-feature plots present the infeasibility of visual analysis in multidimensional data prompting the use of SVM (legend indicated in Table 1).

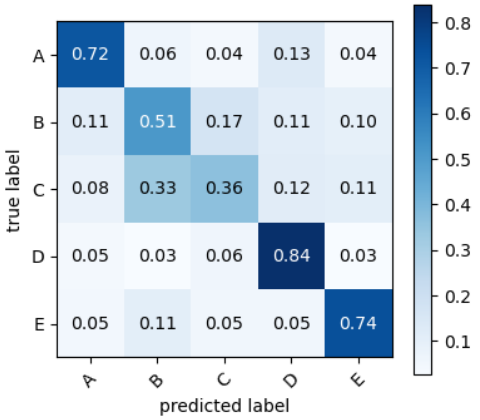

Figure 6: The confusion matrix for the separation of the test dataset used for SVM. The poorest performing is the classification of Capacitor C, which can sometimes be misclassified as Capacitor B.

The confusion matrix in Table 6 indicates that we have successfully separated capacitors A, B, C, D, and E; Capacitor D and Capacitor E were best classified at 84% and 74% respectively, while the poorest classification was observed for Capacitor C at 36%, however this result is still above chance (20%).

In the classification of Capacitor C, the SVM model misclassified 33% of the test data as Capacitor B. Overall results could have been improved if there was less misclassification in Capacitor C. This also suggests that there could be similarity between Capacitors B and C, indicating substitutability between these two capacitors, if so required. This claim, however, requires further investigation.

The results are also in agreement with [5], [7], and [18], demonstrating that there is an empirical difference in system performance with varying capacitor construction. In [5] and [18], the varying capacitors provided differences in the system responses. The measurements made in this paper provide an extension, in both frequency and input amplitude, to the reported measurements in the aforementioned works. These differences support the listening test results in [7], as the participants demonstrated discernment.

3.4 Further Use of the SVM Model

The SVM model proposed enables the analog audio circuit designer to gain insight on the behaviour of different capacitors in the system. For example, if a designer wanted to check for similarities between two coupling capacitors (of different construction), capacitors could be tested using this methodology to check for misclassification in the confusion matrix. If misclassification is present, that would imply that the capacitors are similar in non-linear behaviour.

4 Conclusion

A methodology to analyse audio electronic circuit performance under varying capacitor construction is proposed. A circuit was chosen to meet the requirements of producing non-linear distortion. A dataset is obtained by sampling the system with sine waves at discrete points along the frequency spectrum and across a range of input amplitudes. Audio features, such as THD, SFDR, SINAD, and SNR, were extracted from this dataset. The THD results are visually analysed to understand THD behaviour at different regions of frequency and input amplitude. To cope with the complexity and richness of data, however, an SVM model was trained with the multidimensional dataset. The separability of capacitors in a system from their empirical features is shown. Further work will be put into improving the success ratio through optimisation of the machine learning model.

Here, we demonstrate a method to analyse non-linear behaviour in coupling capacitors. The separability in the machine learning model informs the similarities and differences in the overall performance of the system when the capacitors of different construction are used. Above 200 Hz, the capacitors show somewhat similar behaviour, with only a 0.01 dBc of standard deviation in Total Harmonic Distortion. Figure 2 shows that Capacitor A (red) occupies the largest area on the top surface of the plot and Capacitor E (magenta) occupies the largest area on the bottom surface of the plot, indicating the highest THD and lowest THD respectively. Below 200 Hz, there is a greater standard deviation of 14.1 dBc, indicating a greater difference in non-linear behaviour. In addition to that, at the region of lowest frequencies, there are points of discontinuities. Visual examination of the surfaces show two separated groups, with three capacitors (Capacitors A (red), B (green), D (cyan)) in one group and two capacitors (Capacitors C (blue), E (magenta)) in another group. With increasing amplitude, there is separation in the rate of the increase of the THD. This separation however, does not hold at the region below 0.2 V, where THD values are inconsistently distributed. A support vector machine model is trained (70% to 30% train to test split), and classifies the five capacitors above chance; the best classification at 84% and worst at 36%.

The method described allows the analog audio circuit designer to gain a meaningful understanding of the changes in the complicated non-linear multidimensional characteristics of the system. This impacts the sonic signature of the equipment. With characterisations of more capacitors with different construction, perhaps a database could be developed for analog design to supplement existing SPICE simulations and Monte Carlo analyses – giving us a better understanding of what contributes to sonic signatures (perceptual and/or electronic) of different audio electronic devices. This database could be extended to include other factors such as mechanical resonance [7] and other components such as resistors, inductors, transformers, transistors, and vacuum tubes.

Further work can also be done to assess other models. This work is an *ab initio* investigation towards the understanding of the methodology applied, and thus, only an SVM was used. Moving forward, investigations of different machine learning topologies, varying the amount of audio features and understanding which features hold more significance will be conducted, and reported.

References

[1] Gottinger, B., *Rethinking distortion: Towards a theory of 'sonic signatures'*, Ph.D. thesis, 2007.

[2] Moore, A., Till, R., and Wakefield, J., "An Investigation into the Sonic Signature of Three Classic Dynamic Range Compressors," in *Audio Engineering Society Convention 140*, 2016.

[3] Lindell, B., "The Sound amp; Color of the Fairchild 670 Compressor Plug-In," 2013, retrieved from http://www.uaudio.com/blog/puremix-fairchild. Last accessed 14 March 2021.

[4] Hood, J., *Valve and Transistor Audio Amplifiers*, Elsevier Science, 1998, ISBN 9780750633567.

[5] Gaskell, R.-E., "Capacitor "Sound" in Microphone Preamplifier DC Blocking and HPF Applications: Comparing Measurements to Listening Tests," in *Audio Engineering Society Convention 130*, 2011.

[6] van der Veen, M. and van Maanen, H., "Non-Linear Distortions in Capacitors," in *Audio Engineering Society Convention 124*, 2008.

[7] Dodds, P., Duncan, P., and Williams, N., "Audio Capacitors. Myth or Reality?" in *Audio Engineering Society Convention 124*, 2008.

[8] ITU-R, "ITU-R Recommendation BS.1116-3. Methods for the Subjective Assessment of small Impairments in Audio Systems including Multi channel Sound Systems," 2015.

[9] Rozenblit, B., *Tubes and circuits*, Transcendent Sound, Inc., Kansas City, MO, 2012, ISBN 9781477532867, pp. 159, 2012.

[10] Bussey, W. and Haigler, R., "Tubes versus transistors in electric guitar amplifiers," in *ICASSP '81. IEEE International Conference on Acoustics, Speech, and Signal Processing*, volume 6, pp. 800–803, 1981, doi:10.1109/ICASSP.1981.1171205.

[11] Koren, N., "Improved vacuum tube models for SPICE, Part 1," 2003, retrieved from http://www.normankoren.com/Audio/ Tubemod-spice_article.html. Last accessed 14 March 2021.

[12] Analog Devices Inc., E., *Data Conversion Handbook*, Elsevier Science, 2004, ISBN 9780080477015.

[13] Sheingold, D., *Analog-digital Conversion Handbook*, Analog Devices technical handbooks, Prentice-Hall, 1986, asin B000JL2FOO.

[14] Analog Devices Inc., E. and Zumbahlen, H., *Linear Circuit Design Handbook*, Elsevier Science, 2011, ISBN 9780080559155.

[15] Harris, F. J., "On the use of windows for harmonic analysis with the discrete Fourier transform," *Proceedings of the IEEE*, 66(1), pp. 51–83, 1978, doi:10.1109/PROC.1978.10837.

[16] Zenil, H., Kiani, N. A., and Tegnér, J., "Quantifying loss of information in network-based dimensionality reduction techniques," *Journal of Complex Networks*, 4(3), pp. 342–362, 2015, ISSN 2051-1310, doi:10.1093/comnet/cnv025.

[17] Vert, J., Tsuda, K., and Schölkopf, B., *A Primer on Kernel Methods*, pp. 35–70, MIT Press, Cambridge, MA, USA, 2004.

[18] Kaye, Z., "Selecting capacitors to minimize distortion in audio applications," *Analog Design Journal, Signal Chain*, pp. 1–4, 2020.

Appendix

Logarithmic spacing of sampled frequencies

The spacing of the frequencies was optimised to provide granularity at lower frequencies, as pilot studies showed a region of interest between 20-200 Hz. 78 of the 150 frequencies are below 200 Hz.

Frequencies are described with the expression:

For the n^{th} frequency F_n,

$$F_n = exp\left[0.0658n\right] + 20 \qquad (11)$$