DIGITAL CORRECTION IN SMALL ROOMS

A Perceptual Study Evaluating 3 Small Critical-Listening Room Calibration Systems

Michael P. Hagen

Submitted in partial fulfillment of the requirements for the
Master of Music in Music Technology
in the Department of Music and Performing Arts Professions
Steinhardt School
New York University

Advisor: Paul Geluso

[May 9th, 2021]

ABSTRACT

As DAWs, Microphones, and ADC/DAC Interfaces have become more affordable, accessible, and simple to use, there number of Home Studios and Project Studios have skyrocketed. These are studios that fit into the space available, as opposed to the purpose-built Recording Studios of the past. Because these spaces are often in spare bedrooms, garages, or dorm rooms, it is often not possible to drastically change the acoustics of the space to improve the quality of the listening environment, and instead the engineer turns to the software realm to try to improve the listening experience. In this thesis, a subjective study was performed testing user preference of 2 commercial room correction products (Sonarworks Reference 4 and IK Multimedia's ARC System 3), as well as an Inverse Filtering system of this author's own design, against one another and against the room's uncorrected response. Frequency responses were obtained from the rooms as well, both with processing engaged and without. This data was analyzed to determine the efficacy of each of the systems and to compare the subjective and objective data gathered, and draw conclusions on how well each system operated, and whether it is a viable solution for software-based room correction. It was found that both commercial systems performed well on average but did deteriorate the response of one of the rooms, illustrating that these systems do not always work. On average, Sonarworks was preferred over ARC, but the normal room response was, on average, slightly preferred over ARC across all testing. The Novel Inverse Filtering system was least preferred and was found to perform very poorly in 2 of the 3 rooms. In the other room, it was preferred more than ARC and Sonarworks, but less than the Normal Room Response, indicating that the system is viable but needs significant work before it will be viable as a commercial product. The comparison of the frequency responses for each system indicates that the target that the Inverse filtering response system was correcting towards was drastically different from ARC's or Sonarworks', which is likely the reason for the poor performance.

ACKNOWLEDGEMENTS

I would like to thank all of my instructors at the NYU Steinhardt School of Culture, Education, and Human Development for guiding me along this academic journey, the staff and faculty at the NYU Clive Davis Institute of Recorded Music for giving me a wonderful home and continued learning environment as a staff member, my friends and family for supporting me pushing me along when I may have otherwise dragged my feet, and my wonderful wife for helping me to reach this goal.

TABLE OF CONTENTS

LIST OF FIGURES

1. INTRODUCTION

As the computing power of the average home computer increases and the cost of music technologies decrease, the existence of home studios will continually increase. Home Studios, also called Project Studios, are small recording studios owned by a recording engineer or artist and often housed in their home. These studios live in contrast to the "professional" studios such as Abbey Road, Hit Factory, Sony Music Studios, and other larger operations as a different tier of cost-benefit. While these home studios don't have the same high-quality mixing consoles, equalizers, compressors, and microphones that the larger studios have, they cost the artist far less in terms of studio time (or cost nothing at all when the studio is the artist's). These types of studios use software and cheaper audio processing equipment to make up the gap in audio quality and capability, and the number of products in this subset of music technology industry has exploded. Today an artist can easily purchase high quality recording equipment for under $200 and get started recording at home in hours.

However, while these technological advances have made great strides towards making production more accessible to artists and project studio owners, it cannot change the fact that most of these studios are in the artist's office, garage, or another location in the artist's home. Because of this, the room was not built for critical listening or recording and is fraught with acoustic issues due to the inherent small size, a lack of vibration decoupling and sound reinforcement, and the presence of many parallel surfaces. These issues mar the otherwise high-quality recordings captured by the equipment and imprint a sonic signature that is recognizable by an experienced listener as coming from a poor recording room. Some project studios work around this by building a small booth, or a room within the existing room, in which they can have proper acoustic separation and sound reinforcement to create a high quality, though small, recording environment. Others record with the microphone and source in a small acoustically treated box such as ISOVOX's Home Vocal Booth.

While this can work well for getting higher quality recordings, it does not improve the listening environment of the engineer. Because of this, the engineer may hear errors when mixing that exist not in the recording but their listening environment. Compensating for these errors while mixing can cause serious problems, since the “correction” effectively introduces the inverse error into the recording. The severity of these errors varies from room to room, but it makes it difficult for a Project Studio engineer to trust what they are hearing. Often this leads the engineer to purchase expensive, high quality headphones, which helps but cannot solve the problem. While headphones provide a more controlled listening experience, every time they are put on the slight differences in how they couple to the outer ear change the frequency response slightly. In addition, headphones are not accurate for mixing content that will be heard on speakers, since listening on headphones has no speaker crosstalk and removes the phantom center channel you hear when listening on a stereo pair of speakers. This is particularly troublesome in higher frequencies, where the crosstalk can have an even more significant effect on the recording’s overall balance.

When it is not possible to change the room’s acoustics to improve the quality of the listening environment, software can be used to compensate for the room’s errors before the signal leaves the speakers. This type of software is called “Room Correction”, “Active Room Correction”, or “Digital Room Correction”, and has been around for more than 20 years. Popular products include Sonarworks Reference 4, Dirac Live, and IK Multimedia’s ARC, but many speaker and audio amplifier manufacturers include some type of correction in their products. This thesis will study 3 digital room correction systems and subjectively investigate the preference of, and qualities observed by, experienced listeners listening in project studio and project studio-like environments.

1.1 Motivation, Goals, and Impact

It is the goal of this thesis to objectively and subjectively evaluate Sonarworks Reference 4, IK Multimedia’s ARC System 3, and a novel inverse filtering application written by this author, in terms of their performance in correcting an imperfect listening environment into a “perfected” one (by way of user rating), in terms of 5-point scale ratings of particular audio qualities, such as brightness (again, by way of user rating), and in terms of direct frequency response comparisons (by way of delta frequency response plots generated from impulse responses of the corrected and un-corrected systems). It is this author’s hope that the information learned from this thesis will aid in the development and progression of simple, useful, and affordable room correction for “Bedroom Producers” and Project Studio owners. To that end, the results from this thesis are intended for future development of a novel room correction system that is open source and easily implemented in a DAW or other environment.

2. PRIOR WORK

Room correction is an exceptionally broad topic in audio, so this section serves to orient the reader in this field and to introduce the reader to the main topics that are the subject of this thesis.

2.1 History of Room Correction

2.1.1 Idealized stereo response and acoustics

Traditionally, the frequency response of recording studios and listening rooms is corrected and "tuned" by reducing and equalizing the overall reverberations across all frequencies, eliminating standing waves and flutter echo, and by smoothing out extreme boosts of particular frequencies. These corrections are done in such cases when they would audibly improve the listening experience of the room, and the errors are left alone when they would not be audible to a human. In digital correction, we analyze the room objectively, correcting for errors as defined in the code of that particular correction system, and this often differs from what would be subjectively done by an audio engineer or acoustician to improve a room. Using digital correction, we are able to perfectly correct for every error in a room while listening in one specific location, but often the correction makes the listening experience worse in other locations (Bank, 2018). Many of these corrections are done to correct for errors that a human would not normally hear anyway and are a waste of CPU cycles.

If we take into account this quasi-logarithmic frequency resolution of the human auditory system, we realize that we should only equalize the aspects that lead to audible error, and leave the inaudible imperfections alone (Bank, 2018). In addition to preserving processing power, this prevents such extreme boosts as would overload an amplifier or speaker (Gerzon, 1991). Humans are able to hear peaks in a spectral response more easily than they can hear notches (though broad-band notches are still easily audible), but as the frequencies increase these peaks and notches depend more and more on factors such as position in the room and the material that the surfaces inside are composed of, as well as factors most audio engineers may not consider, such as humidity and temperature.

2.1.2 Acoustic and Analog methods of room correction

Typically, the correction of a room's errors is done in the mid and high frequencies through the use diffuser/absorber panels on the walls, and in the lower frequencies through the use of bass traps/absorbers. Absorber panels are mounted on the walls or ceiling and absorb the acoustic energy that would normally be reflected back into the room, while diffuser panels scatter the sound energy across a wide angle so that very little of the sound energy is reflected back to the listener. Bass traps are typically larger structures filled with some kind of dense material and serve to move in response to low frequency energy in the room, thereby converting the sound energy into kinetic movement and heat through the friction of the material moving against itself (Robjohns & White, 2007). The overall intention of acoustic treatment in a studio is, in general, to reduce the frequency specific reverberation of the room to equalize it over all frequencies (Green and Dunbar, 1947).

Correction can also be done using analog electrical circuits to correct for the measured errors in the room, but this is expensive and time consuming, due to the number of filters that must be implemented to get a great room response. With the advent of digital filtering and correction techniques, correction via analog circuitry is done much less often. (Walker, 1998), (Gerzon, 1991)

2.1.3 Digital correction

Digital correction products have been produced and implemented by B&W Loudspeakers, Marantz, Sigtech, (Gerzon, 1991) Harman International, Anthem, and Audyssey (Olive, Jackson, Devantier, Hunt, and Hess, 2009), and while none of these companies produce a standalone room correction, they all (excluding Sigtech Audio, which no longer exists) still develop products that implement some form of room/loudspeaker response correction. In the last few years, most companies producing loudspeakers, power amplifiers, or AVRs have implemented some form of correction into their systems, including Crown, Denon, Neumann, PMC, JBL and Genelec. Currently, there are two main types of room correction systems available commercially, hardware-based and software-based. The hardware-based systems are the type used in the loudspeakers, power amplifiers, and AVRs mentioned before, while the software-based systems include products by IK Multimedia, Sonarworks, and Dirac. Sonarworks specifically has gained a significant reputation among engineers, due in no small part to endorsements from big names in the audio engineering and production industry.

2.2 Common methods and concerns in implementing digital room correction

2.2.1 Deconvolution

The acoustic response of signals radiated in a room are distorted linearly by the reflections from the walls, which often manifest themselves as reverberations and echoes (Miyoshi, Kaneda, 1988) These reflections from the walls also impart distortions in the frequency response of the room. Since the response of a room has a linear response to changes in volume and does not change depending on the time a signal is input, we can categorize a room as a "Linear Time-Invariant", or LTI, system. Like all LTI systems, it is possible to generate an impulse and record the response of the room, as is typically done when generating an impulse response for convolution reverb. Once we know the frequency response of the room in the form of a transfer function (using an FFT to convert the impulse response into the frequency domain), it is possible to generate a filter that perfectly corrects this response by applying an inverted copy of the room's transfer function to the output signal. This inverted response, or inverse filter, will perfectly correct for every error in the output system, such that if a response were to be recorded in the same spot as the original, it would be completely flat (all other factors identical). This filter can be easily applied to a room's output in the form of convolution with an inverted copy of the room's recorded response, or via deconvolution of the signal with the room's recorded response. (Howison, Lamontagne, Luna, and Newell, 2005). The issue with this type of correction is that there are many small bandwidth errors that humans cannot normally hear, as well as errors and imperfections at frequencies that humans are not as sensitive to. These are errors that even a critical listener would likely not notice, but their correction can cause the room's response to deteriorate in other locations and cause audible notches in the response. To avoid significant notches in response, a positive bias is added to the measured response of the room. (Cecchi, Carini, and Spoors, 2018).

2.2.2 Adaptive Inverse Filters

As stated previously, correction of a room's response based on errors measured at a single point will fail to control the response at points away from the measured point. In fact, the response will often degrade at points other than the measured point when compared to the initial response, depending on the method of correction applied. It is possible to, instead, place multiple microphones into a room and record the error signal at each by subtracting a desired signal from the recorded response, and attempt to minimize the errors in the room's response by creating an adaptive filter that automatically adjusts filter coefficients in order to minimize the sum of the squares of the error signals at each microphone. (Cecchi et al, 2011) This allows the creation of a filter that corrects more for the general response of the room, rather than the exact response at a particular point. While all microphones may pick up a boost at around 300Hz, they may not all hear a boost at 90Hz, and in fact some of the microphones may actually record a notch at that frequency. By attempting to equalize all the microphones as best as possible, the filter will generally improve the response of the room, but without damaging the response in other locations. An adaptive filter of this type can be implemented either as an FIR (including, for example, a Kautz Filter) or IIR filter, though the benefit of using a recursive IIR filter is offset by the complications in the algorithm's adjustment of filter coefficients, partly because for some filter coefficients, the filter may be unstable, which can prevent the algorithm from operating correctly. (Elliott, Nelson, 1989). This type of correction can be easily extended to multichannel systems and can even allow for a degree of crosstalk cancelation, though obviously this only works for a single listening point (Nelson, Hamada, and Elliott, 1992).

2.2.3 Concerns and limitations of room correction products

As previously mentioned, a primary concern for the implementation of room correction systems is the deterioration of the room's corrected response, or worse, the creation of a worse response, in listening positions other than the one corrected for (Norcross, Soulodre, 2004). It is because of this that equalization in one point is not sufficient for correction of a listening system for musical content, even just due to the distance between the listener's ears. Any listening system for audio engineers must correct for the room's response along multiple points, or at the least in an area large enough for the listener to work or listen in (Rubak, Johansen, 2000)

The other primary concern is in the implementation of these filters, though with rapid advancements in computation technology this is much less of a concern than it once was. Previously, it was necessary to use decimation to reduce the sample rate of certain frequency bands to just below the Nyquist frequency order to reduce the computational load on the device (Gerzon 1991) or use an incredibly powerful machine that most audio engineers wouldn't have access to (Nelson, Hamada. Elliott, 1992), but this is no longer necessary for most applications.

2.3 Evaluated Commercial Correction Systems

2.3.1 Sonarworks Reference 4

The 4th iteration of Sonarworks' room correction software brought major DSP improvements allowing for "zero latency" to be introduced when using the software to correct the audio output path. Sonarworks is currently the most popular room correction software available (according to sweetwater.com), used by over 70,000 studios worldwide and endorsed by Grammy nominated engineers and producers mentioned by name on the Sonarworks website (*Sonarworks Reference 4*, 2021). Because the software is closed-source, it is not possible to determine exactly what processing is being done to correct the room's response, but by looking at the patents that they hold, it is apparent that Sonarworks are using a combination of the above methods to accomplish their correction, including adaptive inverse filtering and deconvolution which are clearly portrayed in the patent documentation (Sproģis, 2016; Sproģis, Bēms, and Popeliš, 2018).

2.3.2 IK Multimedia ARC System 3

IK multimedia's latest revision of their room correction software simplifies the room analysis process, allows for customization of the recommended filter settings, and use of 3rd party measurement microphones. One such customization is the ability to adjust the frequency range to be corrected, allowing the user to only correct the frequency ranges that they know the room has issues in. ARC System 3 also allows the user to toggle linear phase mode, allowing the existing phase response of the original signal to be maintained. Unfortunately, there do not appear to be any patents filed by IK Multimedia for room correction, so it is unclear exactly what methods they are using for their room correction, but the parameters and features of the product make is clear that they are using inverse filtering and traditional FIR/IIR filters. (*ARC System 3*, 2021)

2.4 Previous Comparisons and Analysis of Room Correction Methods and Products

2.4.1 Objective Comparison of Products

A comparison of 5 room correction products was conducted in by Olive et al. in 2009 comparing Audyssey Room Equalizer, Anthem Statement D2 Processor, Prototypes Harman International Room Correction 1 and Room Correction 2, and Lyngdorf DPA-1. For this study 20 second mono loops were played through an audio system in which they could switch between the 5 room correction products and a control "No-EQ" setting, without being informed of what each of the 6 options were. The program material was Jennifer Warnes' "Bird on a Wire", Tracy Chapman's "Fast Car", and James Taylor's "That's Why I'm Here", chosen because this material had reliably revealed spectral differences among different loudspeakers and equalizations in previous testing. The sound was reproduced by a pair of B&W 802N 3-way loudspeakers.

Olive et al. objectively analyzed the room response by taking acoustical measurements of each room correction system in each of the 6 listening seats, using a log-sweep with 48 points per octave from 20 Hz up to 20 kHz. For this, they used 6 calibrated omni-directional DBX RTA-M microphones, each positioned in at one of the listening seats so that the positioning would remain constant throughout the measurement process.

Through this measurement, it was found that most of the measured differences occurred below 200 Hz and between 1.5 kHz - 4 kHz, and that the more preferred room corrections produced smoother curves with a higher output below 60 Hz. It was also found that the more preferred room corrections had more downward sloped curves, where the sound pressure level decreased as the frequency increased. The two frequency responses with the flattest curves were the two least preferred correction products. It was also noted that most of the correction products were able to fix a resonance at 48 Hz, all except one were able to correct a wide 2 kHz dip, and 2 added a boost at 80 Hz, which was unnecessary from a correction standpoint based on the initial response of the room (Olive et Al, 2009).

2.4.2 Subjective Comparison of Products

In addition to the objective measurements in the experiment conducted by Olive et al, the main data from their study was subjective data collected from the listeners. From this data, it was found that the listeners preferred the room correction responses with the flattest perceived spectral balances. Olive et al also collected comments on the response of each product, and from this we learn that the listeners felt that the “No EQ” option was the second most colored (Product #6, the least preferred product was considered the most colored. It is clear from this data that the listeners felt they preferred the room which was the least colored, harsh, thin, and the most full and neutral.

As previously mentioned, most of these companies no longer produce a standalone room correction product, and that they have shifted their correction product development to be included in their other products. In addition, notable correction products such as Sonarworks Reference and IK Multimedia’s ARC, the most popular commercial products today, had not been released, and were therefore not part of the study.

3. METHOD

As mentioned in the introduction, this thesis will focused on a subjective and objective analysis and comparison of 3 digital room-correction systems in project studio-like environments. The 3 systems evaluated were Sonarworks Reference 4, IK Multimedia's ARC 3, and a simple inverse filtering script written by this author, the implementation of which be detailed later. Objective measurements in the form of frequency response measurements of the listening environments were compared and analyzed to put hard data behind the perceptions of the listeners. These measurements were compared to measurements taken of the same environments when corrected using each of the 3 most popular standalone products as an attempt to estimate the methods of correction for each product.

The subjective investigation had the listener listen to binaural recordings of 3 different samples of program content played in 3 different rooms with each of the 3 correction methods and a control with no correction (3 samples of material * 3 rooms * 4 correction methods and control = 36 listenings). The listener was, at no point, made aware of what they are listening to (room, correction method, or program material). Following the completion of this section, the listener went through a comparative section where they had the ability to switch between all 4 correction types while listening to a particular piece of program material from a particular room and were asked to put them in order according to how they felt the recordings represent an accurate critical listening space for music production purposes. Afterwards, they were asked to rate each of the systems on a 5-point scale on a few subjective characteristics. They would then proceed to the next material & room combination.

The objective investigation involved an analysis of frequency responses recorded by taking impulse responses in each of the rooms through each of the 3+control correction methods to see what exactly is being done to the program material when it is processed by each of the correction systems, and additionally how closely each of the correction systems match a flat frequency response.

The technical data collected consists of 36 binaural musical recordings and 12 binaural impulse responses. These recordings consist of 3 musical excerpts recorded in 3 small critical listening rooms and corrected with one of 3 room correction systems (and recorded uncorrected as a control. This comes out to 3 excerpts * 3 rooms * (3 correction systems + 1 uncorrected) for a total of 36 musical recordings. The impulse responses come from the same variables, but removing the excerpts, so 3 rooms * (3 correction systems + 1 uncorrected) for a total of 12 impulse responses. The binaural musical recordings were used for qualitative analysis in the form of a questionnaire filled out by expert listeners. The binaural impulse responses were be used as a quantitative analysis looking at frequency response of the uncorrected and corrected versions of each room to see how each of the 3 correction systems operated in each of the 3 tested environments.

3.1 Musical Excerpts

The Musical excerpts chosen are Jennifer Warnes's Bird on a Wire from the album Famous Blue Raincoat, Mr. Brightside by the Killers from the album Hot Fuss, and LOYALTY. by Kendrick Lamar from the album DAMN. Olive et al's 2009 study used Jennifer Warnes' "Bird on a Wire", Tracy Chapman's "Fast Car", and James Taylor's "That's Why I'm Here", and this study keeps "Bird on a Wire" but discards the other two, replacing them with Mr. Brightside and Loyalty which were selected from Izotope's list of reference tracks (*4 Tracks to Reference for Mixing in 2018*, 2021; *4 Popular Mixing Reference Tracks, and Why They Work*, 2021). These tracks encompass a wide tonal and dynamic range, though it is important to mention that it is impossible to represent all genres of music with 3 tracks.

The 3 reference tracks were ripped from the North American CD release of the mentioned albums using Exact Audio Copy, a high-quality CD ripping application with error detection. No errors were detected when ripping the CDs. As files ripped from a CD, the audio files have a sample rate of 44.1Khz and a bit depth of 16 bits. The files were saved from Exact Audio copy as PCM waveform (.wav) files.

3.2 Correction Systems

This study makes use of 3 different systems for correcting the response of a room. 2 of these systems are commercial products, and one is a MATLAB utility that implements smoothed inverse filtering by compensating a given audio file.

3.2.1 Commercial Correction Systems

The 2 commercial correction systems used are IK Multimedia ARC System 3 and Sonarworks Reference 4 Studio Edition. These products were chosen because they are the two most popular software-based room correction applications. Both applications were run on an early 2013 MacBook Pro with 8GB of RAM and a 2.4 Ghz Intel Core I7 Processor. Both applications were purchased directly from the manufacturer, and a student discount was used for Sonarworks. At the time of writing, IK Multimedia do not provide student discounts.

3.2.2 Offline Correction System based on Inverse-Filtering

The offline correction system is written in MATLAB and corrects for a rooms response by generating a filter that compensates for the frequency response of a room (measured using impulse responses). This filter then applied to program content prior to playback (offline) through deconvolution. Because this utility functions by compensating for the coloration caused by the playback system and does not directly correct the playback system itself, it is inaccurate to call this a correction system. It would be more correct to call it a compensation system, however for this thesis it will be referred to as a correction system in its use as a method of correcting for the room's response as a comparison to the commercial products.

Below, the normal coloration caused by an audio playback system is depicted in Figure 1.

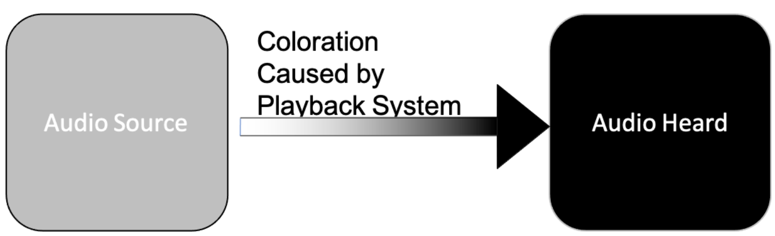


Figure 1. Illustration of the normal coloration imposed by a playback system

In any system the playback system consists of all the software and hardware between the audio source and its destination (typically a listener, or in this case, a binaural microphone).

In Figure 2 below is a flowchart depicting the use of the MATLAB utility.

Figure 2. Illustration of the processing done by the MATLAB Inverse Filtering System and the subsequent coloration imposed by the playback system

When using the inverse filtering system, of the playback system "darkening" the signal that is played from the computer, the playback system "darkens" a "lightened" version of the audio, and as a result the audio that is heard through the playback system is spectrally identical to the audio file on the computer. It should be noted that these descriptions of "lightening" and "darkening" are abstractions designed to simplify the function of this utility, and do not reflect "light" or "dark" audio qualities.

This process is accomplished through deconvolution, the process illustrated in Section 2.2.1. An IR of the playback system is obtained, and the source audio is deconvolved with the IR of the playback system. Whereas convolution emphasizes common frequencies and minimizes un-common ones, deconvolution is the opposite. Deconvolution removes common frequencies and emphasizes the un-common ones, so in this way it can be seen as a "spectral subtraction", removing the traits of a system from an audio signal when given the characterization of that system.

Convolution and Deconvolution can be accomplished in a few ways, the two primary distinctions being direct convolution/deconvolution and fast convolution/deconvolution. For this utility it was decided that fast deconvolution would be preferable, as it is much faster when dealing with longer audio files, and most printed/bounced mixes are not short. Fast convolution is done by converting the audio files from the time domain into the frequency domain using a Fast Fourier Transform, and then doing point-wise multiplication instead of convolution, which is much faster than convolution as a mathematical operation. When doing deconvolution instead of convolution, the operator is point-wise division instead of multiplication, but the process is otherwise identical. The convolved audio then exists in the frequency domain, so an IFFT (Inverse Fast Fourier Transform) is used to convert from the frequency domain back into the time domain. The output of this process is identical to the output of direct convolution, and the process is much faster. It is also during this process that the smoothing factor is applied to the Impulse response thereby “resisting” the narrow peaks and troughs of the playback system IR and decreasing the strength of the deconvolution on the input audio. This factor is applied by scaling the mean of the impulse response and adding that to the impulse response itself before deconvolving. This method comes from Dr. Erol Kalkan’s paper and function in “Deconvolution of Two Discrete Time Signals in the Frequency Domain”. Dr. Kalkan is a research structural engineer using deconvolution to analyze and predict the seismic response of buildings in terms of spectral acceleration, but the smoothing that he uses in his deconvolution is not limited to architectural models. This method works well in testing for deconvolution of playback system frequency responses. The full formula for the correction algorithm inspired by Dr. Kalkan is:

$$= real\left(ifft\left(\frac{fft(x)\circ conj(fft(y))}{\left(fft(y)\circ conj(fft(y))\right)+\left(z\circ mean\left(fft(y)\circ conj(fft(y))\right)\right)}\right)\right)$$

Figure 3. Inverse Filtering System formula inspired by Dr. Erol Kalkan

where x is the signal vector, y is the impulse response vector, and z is the smoothing strength (0.0-1.0). The MATLAB code used for processing is Included in Appendix C.

The processing is as follows:

- Compute FFTs of signal (x) and impulse response (y), (X) and (Y) respectively.
- Generate variants of X and Y compensated for phase manipulations (X_1 and Y_1 respectively)
- Generate the compensation factor S, which is the smoothing strength value multiplied by the mean of the compensated IR Y_1): z*mean(Y_1)
- Smooth the signal in the frequency domain by doing element wise division of the signal by the sum of the compensated IR and the compensation factor: X_1./(Y_1+S)
- Convert the signal back into the time domain via IFFT and take the real output.

3.3 Collection of Impulse Responses

The Corrected and Uncorrected Impulse responses were recorded using Pro Tools 2018 installed on and run from the aforementioned MacBook Pro. To collect the impulse responses, a Neumann KU 100 Dummy Head Binaural microphone was used to record the response of the room as a Binaural Room Impulse Response (BRIR) using a Zoom H6 as the ADC/DAC. BRIRs were recorded for the uncorrected room, as well as for each correction system. The head was be placed in the normal listening position for the room, and this location was noted on a floorplan of the room. This position was used as the location for all recordings in the room. A 1 second long 15Hz-20KHz sine sweep was played from Pro Tools out of the left and right studio monitors simultaneously and recorded back into Pro Tools via the KU-100. The Recorded sweep was deconvolved with the original sweep to create the impulse responses for each room as BRIRs.

3.4 Collection of Binaural Recordings

To collect the binaural recordings, the musical content was played from Pro Tools 2018 from the main speakers in each room and recorded back into Pro Tools. For Sonarworks and ARC the plugin version of the software was used on a track, and for the offline correction system the compensated audio was used instead of the original source audio. As with the impulse responses, the ADC/DAC used was a Zoom H6, and the microphone will be a Neumann KU 100 Dummy Head Binaural microphone.

3.5 Web-Based Survey

The survey was filled out anonymously by participants via Qualtrics. The Audio Files were housed on Google Drive as .wav files and were inserted into each page of the survey using iframes pointed at each of the files.

4. ANALYSIS

4.1 Survey Protocol

The survey was filled out entirely online, listening to the binaural audio examples over headphones. The survey presented the participant with a musical excerpt and the ability to switch between 4 different versions of the excerpt. These 4 versions were the binaural recordings for a particular room for that musical excerpt, so for each listening the only variable will be the correction system (including the recording with no correction system). In total, there were 9 listenings that each participant went through, one for each musical excerpt/room pairing (3 rooms * 3 musical excerpts). An estimated completion time of 30-45 minutes was provided to participants, and the participants were directed to not listen to every recording in its entirety and to only listen for as long as was necessary to answer the questions on the page.

The participant was first asked to order the 4 systems (3 correction systems and uncorrected recording) according to which they believe represents an accurate critical listening space for music production purposes. (First being the most accurate critical listening space, last being the least accurate critical listening space). After answering this question, they will be asked to rate each of the 4 systems on a ten-point scale on a variety of qualitative characteristics. The characteristics were: Brightness, Bass Balance, Left-Right Balance, Stereo Width, and Clarity. After answering all of the questions for that room and musical excerpt pairing, they will move on to the next room and musical excerpt pairing, until they have answered all the questions for all 12 room and musical excerpt pairings. At the end of the data collection period, the survey results were exported from Qualtrics XM and imported into a Google Sheets spreadsheet for processing, averaging, and ranking.

An example page from the survey can be found in Appendix A.

4.2 System Preference

The participant's system preference was converted into a preference percentage by assigning a point value to each position in the rating. A rating of 1 (most preferred system) was given a value of 4, a rating of 2 was given a value of 3, a rating of 3 was given a value of 2, and a rating of 4 was given a value of 1. These points were added up across all rating groups and individually for each room and divided by the total of points for that set, giving an easier to read percentage of how many points of the total each system held. The results are shown in Figure 4 below, and the highest result in each column is highlighted:

	Total - Preference Rating	Research Lab - Preference Rating	Edit Suite 4 - Preference Rating	Production Suite B - Preference Rating
Sonarworks	28.0%	21.3%	30.0%	32.7%
Inverse	18.4%	26.0%	14.0%	15.3%
ARC	26.7%	23.3%	32.7%	24.0%
Normal	26.9%	29.3%	23.3%	28.0%

Figure 4. Preference Rating of Each Processing System

Because each of the three correction systems is intended to correct a rooms response for any genre or type of music, the preference rating across all listenings for each song was not collected.

The "winner" across all tests was Sonarworks Reference 4, despite its last place result in the Research Lab. Due to its high preference rating in Edit Suite 4 and Production B, it still comes out on top, followed by the Normal Room's response, then IK Multimedia's ARC System 3, and finally the Novel Inverse Filtering system.

It notable that that the inverse filtering system was rated terribly in both Edit Suite 4 and Production Suite B, but did very well in the Research lab, and is the highest rated correction system in that room (since that room's most preferred "system" was the uncorrected room). In this room it was rated 4.6% higher than Sonarworks and 2.6% higher than ARC, which is a higher preference than ARC had over Sonarworks in its highest rated room (Edit Suite 4). It was expected that the inverse filtering system would consistently be ranked lower than all correction systems, so this is a very interesting result.

ARC received a similar preference rating in the Research lab and Production Suite B, but performed extremely well in Edit Suite 4, where it was the highest ranked system. The data also shows that this was the room most improved by room correction software, with a preference of 9.3% over the uncorrected room.

4.3 Characteristic Rating

The subjective characteristic data was also collected and converted into box plots which are shown in the following section. Subjective observations on 5 characteristics were reported by the participants. These characteristics were:

- Brightness, where a 1 is dull, 3 is desirable, and a 5 is too bright
- Bass Balance, where a 1 is lacking bass, 3 is well balance, and a 5 is bass-heavy
- Left-Right Balance, where a 3 is centered, 1 is left heavy, and a 5 is right heavy
- Stereo Width, where a 1 is very narrow, 3 is an even stereo field, and a 5 is a hole in the middle
- Clarity, where a 1 is extremely poor, and a 5 means instruments are well defined

After collection, the responses to the Left-Right Balance question were discarded, as they were all skewed to the left (towards a value of 1), which was due to an imbalance during recording using the KU-100 dummy head. However, because the error affected all recordings equally, it was still possible to make valid comparisons between systems in the other criteria.

4.3.1 Characteristic Ratings across all Rooms

The characteristic ratings for each system across all rooms are reported as box plots in figure 5 below.

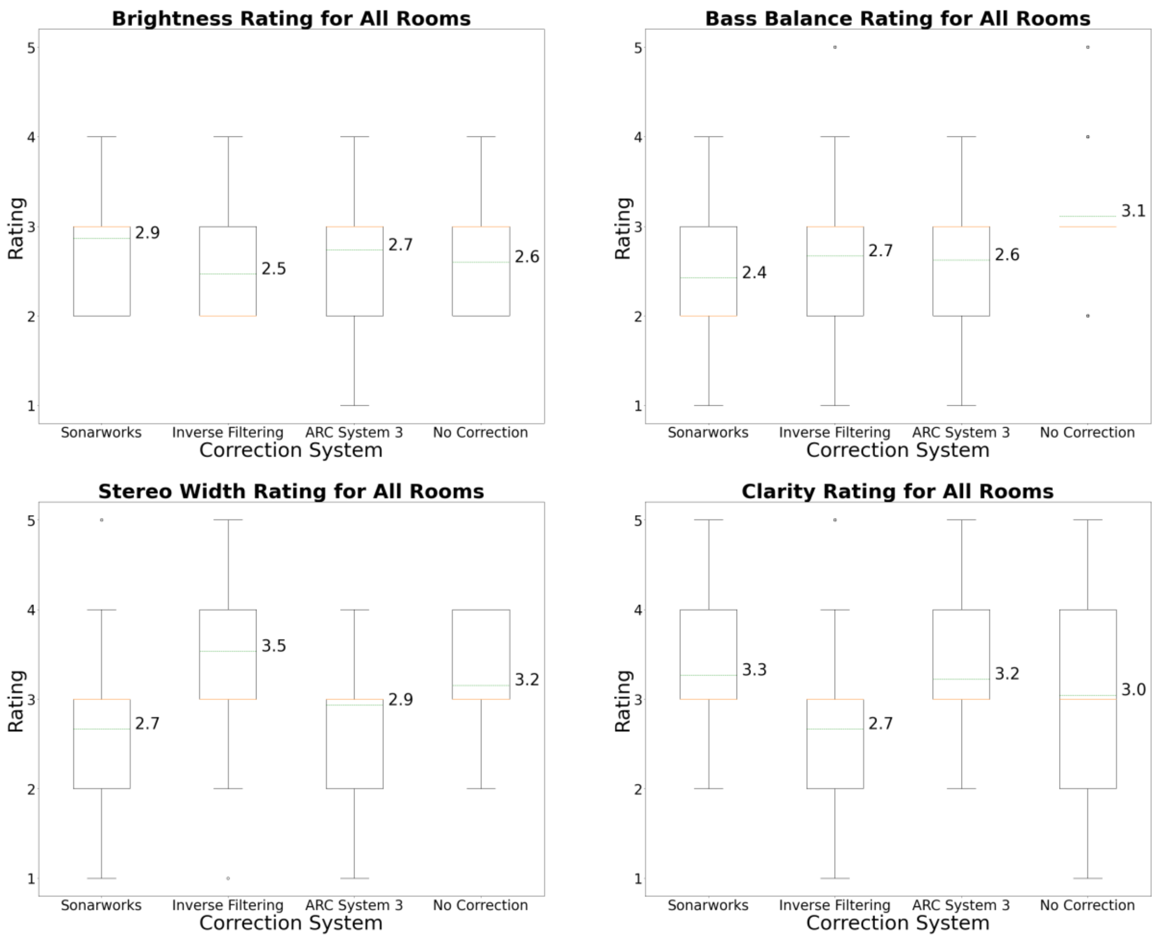


Figure 5. Characteristic Ratings for Each System Across All Rooms

In the boxplots, the orange line represents the median value, and the dotted green line represents the mean value.

It is of interest that the only system that had the best average score in two qualities (Brightness and Clarity), Sonarworks, was the system that was most preferred in terms of ranking. The Inverse filtering system had the lowest average in 3 qualities, which could explain why it was the least preferred system.

The average perceived brightness and clarity of the room was raised by both Sonarworks and ARC but reduced by the inverse filtering system. This suggests that there is a correlation between the Brightness and Clarity of the room. Additionally, all 3 correction systems reduced the average perceived level of bass in the rooms, away from what was considered balanced. Small rooms often have a stronger bass response than desired so it seems logical that the bass should be reduced, but the results show that the most balanced bass response was, on average, the uncorrected system, where it was slightly above the desired level. All of these spectral changes will be analyzed in Section 4.6 via the recorded frequency responses of the correction systems.

Additionally, the spread of results shown by the boxplots is relatively consistent, with a difference between the first and third quartile almost always being 1 (No Correction being the exception, with a difference of 2 between the first and third quartiles for Clarity Rating and a difference of 0 between the first and third quartiles for Bass Balance. Because this data represents the responses for all 3 rooms, and the preference table has shown that the preference differed greatly between rooms, it is more important to look at the per-room data in the following sections.

4.3.2 Characteristic Ratings for Edit Suite 4

The characteristic ratings for each system in Edit Suite 4 are reported as box plots in figure 6 below.

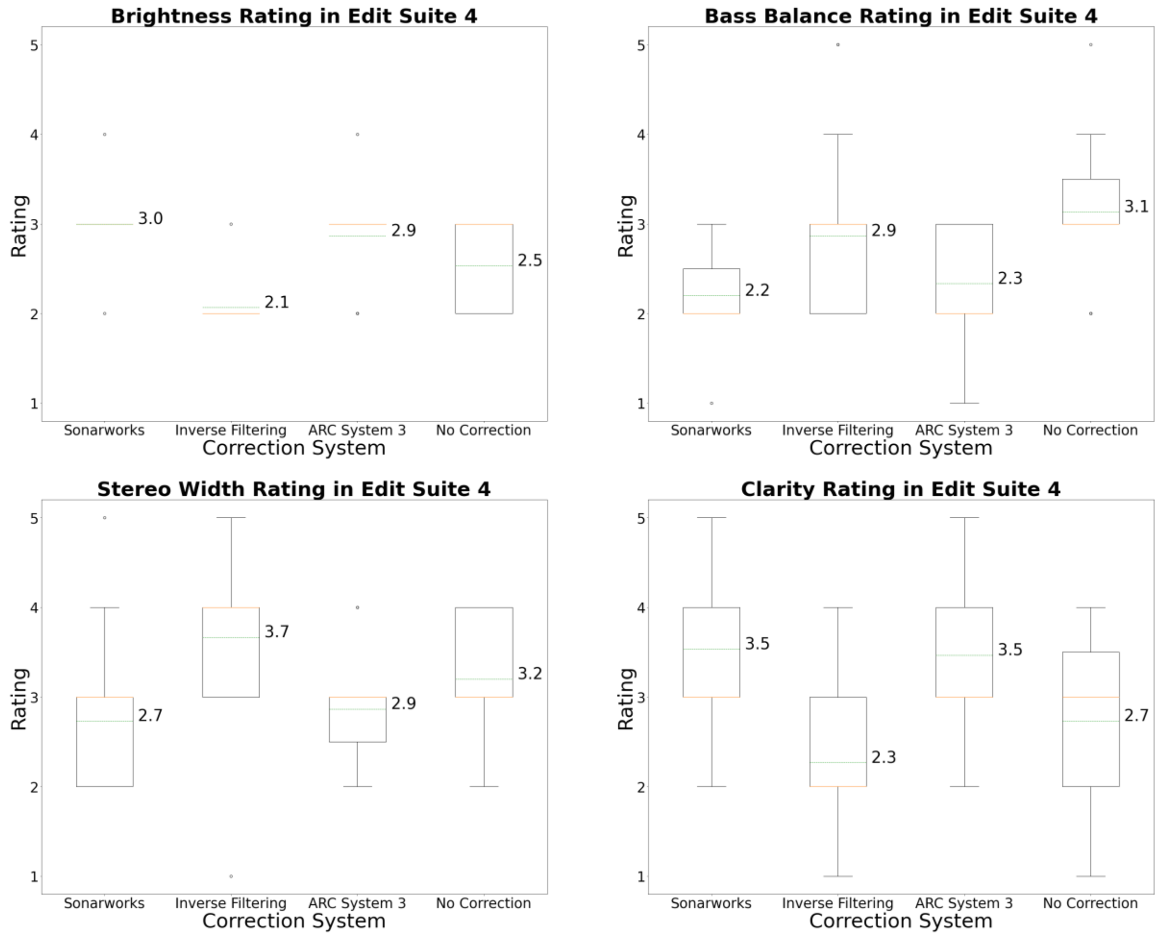


Figure 6. Characteristic Ratings for Each System in Edit Suite 4

Here we can see that Sonarworks, which was the second most preferred system, had the best average rating in 2 qualities, while the most preferred ARC only had the best average in 1. Looking at the spread of results for these two systems, it they appear to have a similar variation, though in the Bass Balance there is a clear difference. While Sonarworks had a smaller Q1-Q3 spread, it that spread is between 2 and 2.5, while the Q3 value for ARC is sitting at the ideal value of 3. It is possible that this is part of the reason why ARC was preferred in this room, and the closeness of these ratings is reflected in the closeness of the preference scores as well.

Compared to the “For All Rooms” boxplots, these plots are much narrower (smaller distance between the first and third quartiles), though the “No Correction” plot for Clarity is very similar to the plot “For All Rooms”. It is also interesting that the results for Brightness in the 3 correction systems show very little variation, possibly indicating that the correction systems are all correcting toward (and to some extent achieving) an idealized high frequency response.

In general, the inverse filtering system performed poorly in terms of rating but was similar in terms of consistency to the commercial systems, indicating that while the inverse filtering system may have corrected the audio toward a different ideal than the commercial systems.

4.3.2 Characteristic Ratings for Production Suite B

The characteristic ratings for each system in Production Suite B are reported as box plots in figure 7 below.

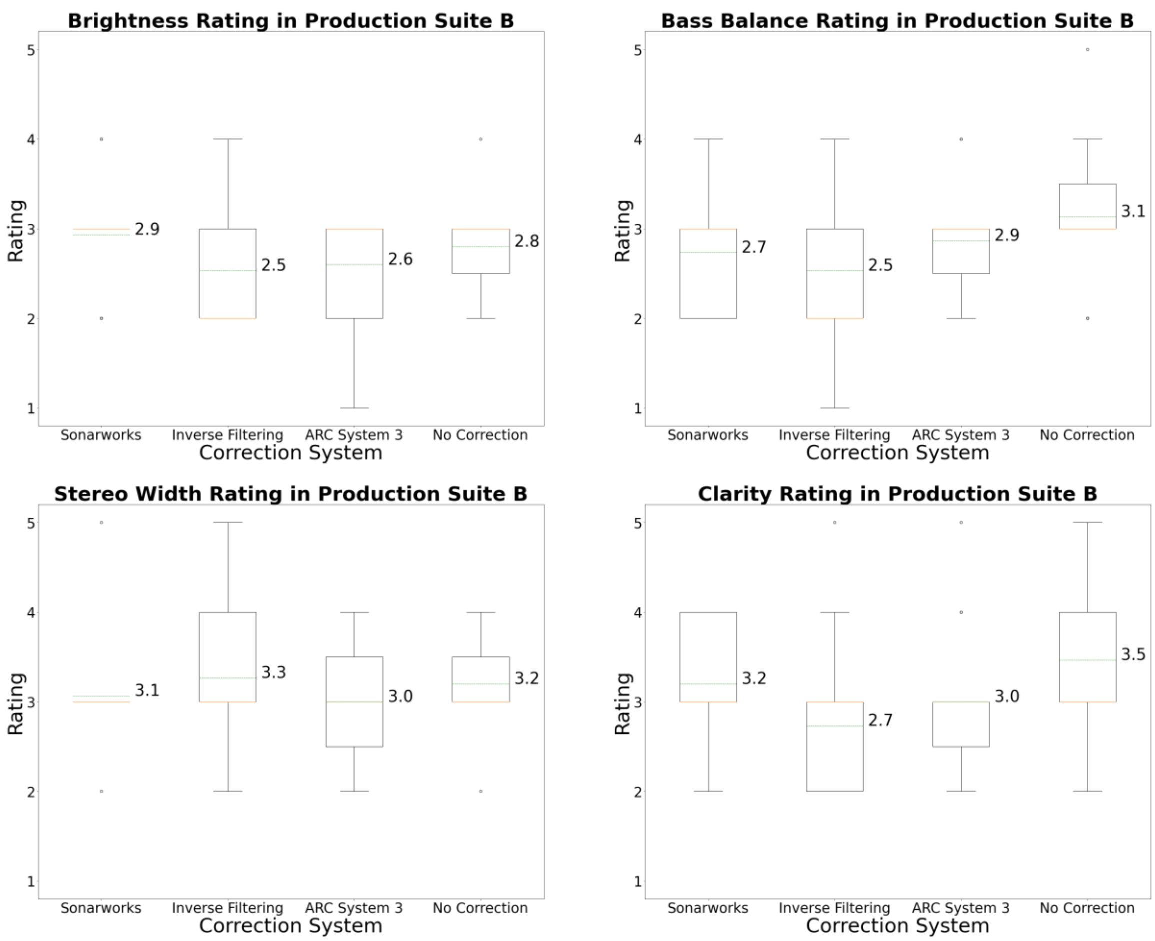


Figure 7. Characteristic Ratings for Each System in Production Suite B

There is much more variation in these plots compared to the Edit Suite 4 plots. Due to this, looking at the average values is misleading. ARC had the best average score in 2 characteristics, while Sonarworks only had 1. However, Sonarworks also had extremely consistent results in both Brightness and in Stereo Width, which may be part of why it was preferred in general over ARC in this room.

The results for the uncorrected room have a very small spread around the idealized value, excepting the Clarity rating, and this could by why the uncorrected response is rated higher than ARC (though lower than Sonarworks) in this room.

The Inverse filtering system again performs poorly, with the worst average result in every category and the largest spread. This is reflected in the huge gap in preference between ARC (3^{rd}) and Inverse Filtering (4^{th}) in this room with preference ratings of 24.0 and 15.3 respectively.

4.3.2 Characteristic Ratings for Research Lab

The characteristic ratings for each system in the Research Lab are reported as box plots in figure 8 below.

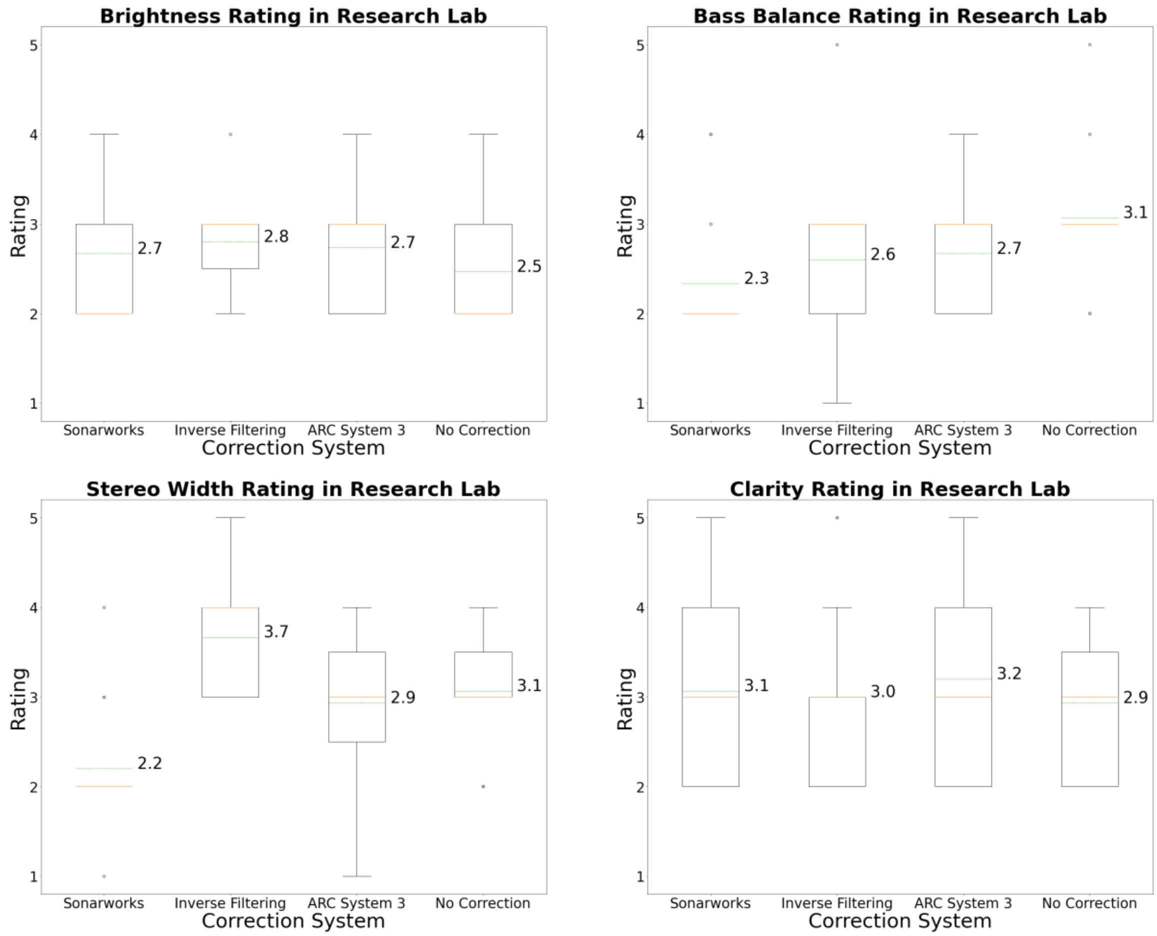


Figure 8. Characteristic Ratings for Each System in Research Lab

In the Research Lab the results for Sonarworks are on average worse than in the other two rooms for Brightness and Stereo Width, and the spread is much larger on the clarity and Brightness compared to the other rooms as well. Conversely, The Inverse Filtering system and ARC had relatively good results. This was the only room in which the normal response of the room was most preferred, and this is backed up by the relatively close to ideal average and narrow spread of the ratings for the characteristics above. The Inverse filtering system performed reasonably well in this room as well, certainly compared to its performance in the other two rooms.

Looking at the rating data above, it seems that the failures of each correction system to Improve the response of the room were likely room-specific, as each system had significantly different results, however, it is possible to draw some conclusions. It appears that all 3 correction systems deteriorated the Bass balance of the room, reducing it past the ideal and down to the “Lack of Bass” region of the scale. While the Brightness results were generally improved over the uncorrected room and the Clarity rankings were higher, the uncorrected room response was preferred. The spectral changes imposed by each correction system will be evaluated in Section 4.6.3 below.

4.4 Processing of Frequency Responses

The collected impulse responses were uploaded to google drive and processed using a python script running via Google Colaboratory to generate the 3 types of frequency response plots.

4.5 Frequency Responses of each room + correction system (and uncorrected)

For the standard frequency response plots for each room, the impulse response was loaded from the Google Drive folder it was stored in, converted into the frequency domain via an FFT, then each channel was plotted such that the x axis was frequency, and the y axis was the amplitude of each frequency in dB. The lines on the plot were also smoothed using local averaging to create a plot would better represent what a human would actually hear.

The standard frequency plots can be found in Appendix B and are not referenced here because the information they contain is better summarized in the Delta Frequency Response plots in the following section.

4.6 Δ Frequency Responses (Correction System – Uncorrected Response)

To allow for easier interpretation of the frequency responses, difference plots were used to show the change in Frequency response induced by each correction system. These Plots were generated by converting the corrected and uncorrected impulse responses of each room into the frequency domain using an FFT, and then subtracting the uncorrected response from the corrected response so that you are left with the changes that the correction did to the room. These were summed to mono to allow them to be plotted on the same graph.

4.6.1 Change in Frequency Response in Edit Suite 4

The Change in Frequency Response in Edit Suite 4 induced by each correction system is shown in figure 9 below.

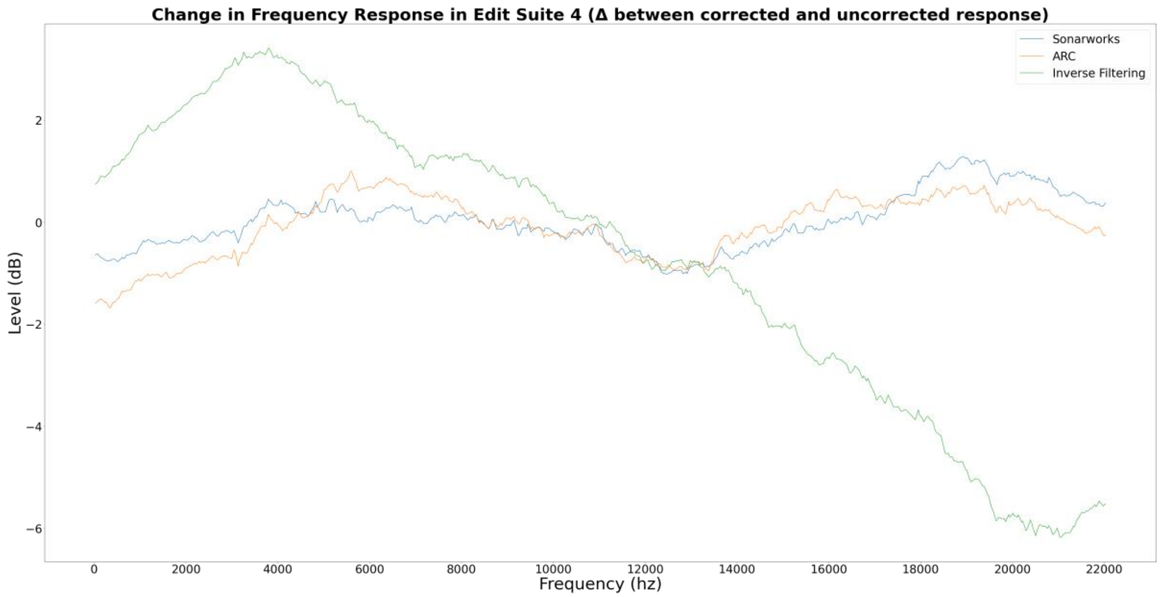


Figure 9. Change in Frequency Response in Edit Suite 4 Induced by Each Correction System

It is immediately clear that the Inverse Filtering correction system varied greatly from the other two in scale, boosting and cutting frequencies across a range of 9dB, while the other two methods varied by less than 3dB. However, it can also be seen that all 3 systems made some similar correction moves, like the spikes around 4800 Hz and 8800 Hz, as well as the dip around 7250 Hz. Sonarworks and ARC are extremely similar, especially in the 8 KHz – 14 KHz range, where the two lines are almost on top of one another. The differences will be more easily seen in the comparison plots in the following section. The closeness of the changes made by the two correction systems also aligns with their close preference scores.

Comparing this plot to the subjective data analyzed earlier, we do see that bass frequencies were reduced by Sonarworks and ARC, matching the participant's perception, but that the same frequencies were boosted by the Inverse Filtering system, so it would seem that the perception that the bass frequencies were reduced came from the fact that the mid and high frequencies up to 4000 Hz were boosted significantly more. Additionally, the perception that the brightness of the room was increased by Sonarworks and ARC and reduced by the Inverse Filtering system appears valid, as the former two system boosted frequencies starting at 13 KHz while the Inverse Filtering system reduced the same frequencies.

4.6.2 Change in Frequency Response in Production Suite B

The Change in Frequency Response in Production Suite B induced by each correction system is shown in figure 10 below.

Figure 10. Change in Frequency Response in Production Suite B Induced by Each Correction System

Just like Edit Suite 4, we see similar boost to the high end, but the bass response here is quite different. Additionally, the high-end boosting done by ARC is not reflected in the room's characteristic rating, where ARC was rated only slightly brighter than the Inverse Filtering system, and much lower than the room's normal response.

In this case the Inverse Filtering system cut the bass frequencies by about 2 dB and boosted the upper end between 3 KHz and 6 KHz. From there it flattens out and then cuts frequencies above 14 KHz severely. It is possible that this cut is due to the frequency response of the KU-100 Binaural head, but a comparison would need to be made using a different binaural microphone (or by deconvolving the impulse responses used to train the Inverse filtering system with an impulse response capturing the KU-100's response alone) in order to test this.

4.6.3 Change in Frequency Response in Research Lab

The Change in Frequency Response in the Research Lab induced by each correction system is shown in figure 11 below.

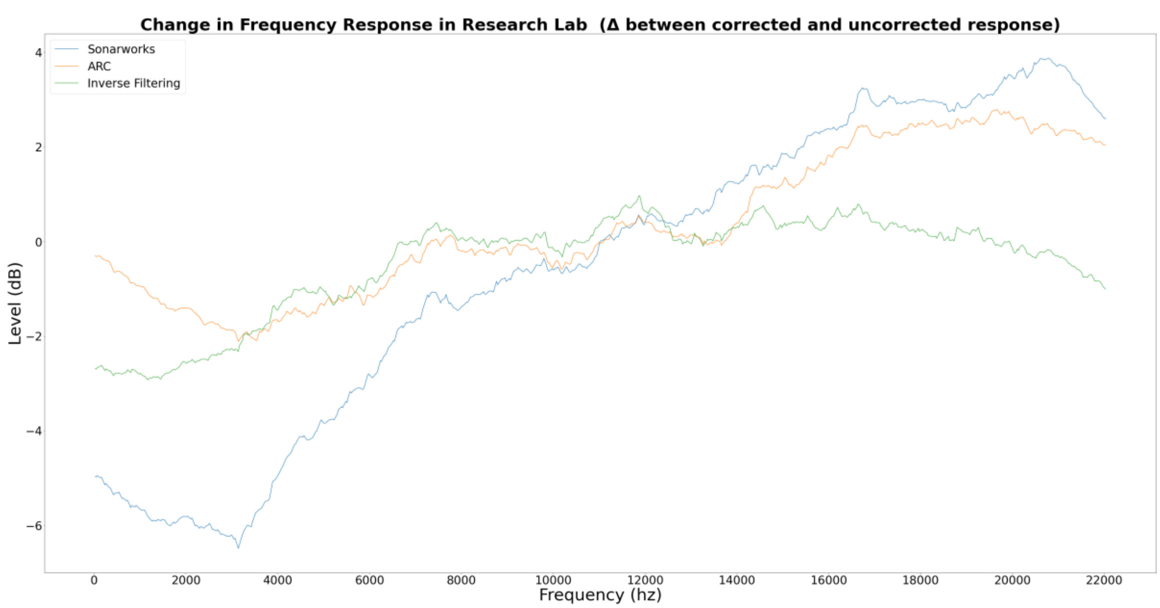


Figure 11. Change in Frequency Response in Research Lab Induced by Each Correction System

Looking at both this plot and the plot for Edit Suite 4, there is a striking similarity between the shape of the frequency response change made by Sonarworks here and that made by the Inverse Filtering system in Edit Suite 4. The two lines appear to be simply flipped horizontally. It is also interesting to see that for the first time the Inverse filtering system is not making huge changes. It more or less follows the changes made by ARC, with a larger lower frequency cut and a slight cut in high frequencies instead of ARC's significant boost in those same frequencies. Sonarworks was least preferred in this room, while in all others the Inverse Filtering System was least preferred. It is significant that the only time Sonarworks was less preferred than the normal room, its frequency response looks so similar to that of the Inverse filtering system, which in the other rooms is less preferred than the normal room's response.

More than that, the poor performance of Sonarworks in terms of its characteristics ratings appears to be corroborated here as well, where the bass response was drastically reduced. However, despite it having a much higher boost in the high end than the other two systems, it is behind both of them in perceived brightness.

4.7 Δ Frequency Responses (Differences between Correction Systems)

To allow for easier evaluation in the differences between the frequency changes induced by each correction system, delta frequency response plots were also generated comparing the correction systems to one another. To create these, the same process was followed as creating the Correction System – Uncorrected Response Delta Frequency Responses, after the uncorrected response was subtracted from each corrected response, the Correction System – Uncorrected Response Delta Frequency Responses were subtracted from one another. These graphs more easily illustrate the differences between the spectral changes induced by one correction system vs another correction system. This was done for each room, and again graphed on the same plot.

4.7.1 Comparison of the Change in Frequency Response in Edit Suite 4

The Comparison of the change in frequency response between each system in Edit Suite 4 is shown in Figure 12 below.

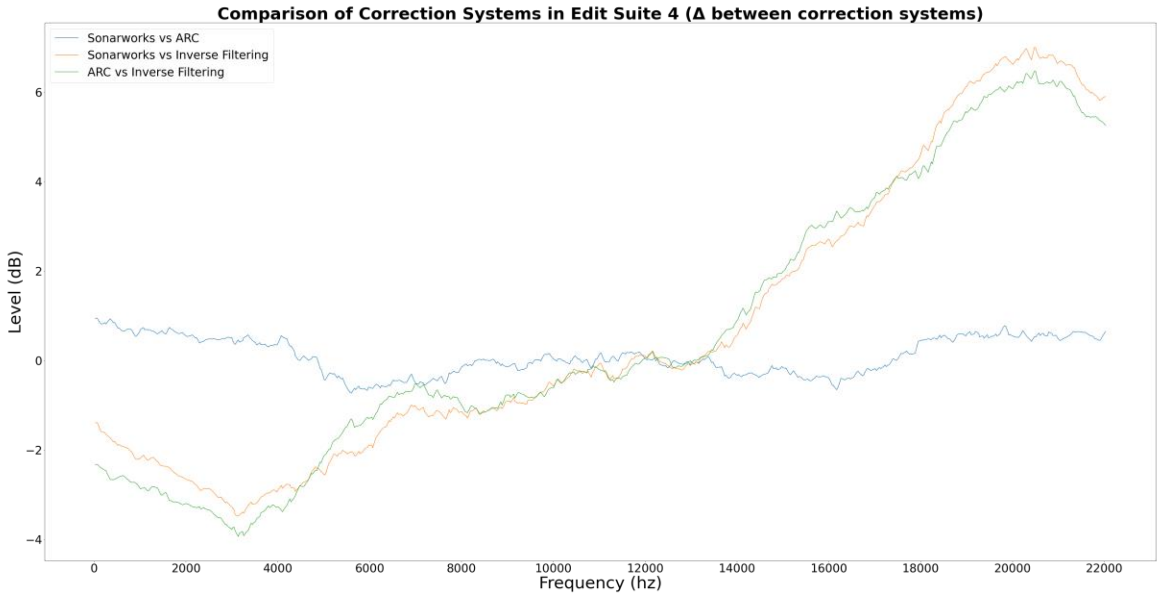


Figure 12. Comparison of the Change in Frequency Response in Edit Suite 4 Induced by Each Correction System

In this plot is clear that the observation earlier that Sonarworks and ARC were very similar in terms of the changes that they made in Edit Suite 4 is true. The Sonarworks vs ARC line is very close to 0, showing that they are nearly identical. Sonarworks cut less of the bass frequencies compared to ARC but cut more of the frequencies between 5 KHz and 6 KHz and between 14 KHz and 17 KHz.

It is also easier to tell the differences between the Inverse filtering system's processing and the other two compared to the previous graphs. It is clear that there is a boost (since "vs Inverse Filtering" plots are subtracting the Inverse Filtering changes from the other two, a dip in those lines equates to the Inverse Filtering system boosting those frequencies compared to the other two systems) from ~8 KHz to 10 KHz and another smaller one around 11 KHz, and we can see the massive boost of the system from 20 Hz up to ~3 KHz.

4.7.2 Comparison of the Change in Frequency Response in Production Suite B

The Comparison of the change in frequency response between each system in Production Suite B is shown in Figure 13 below.

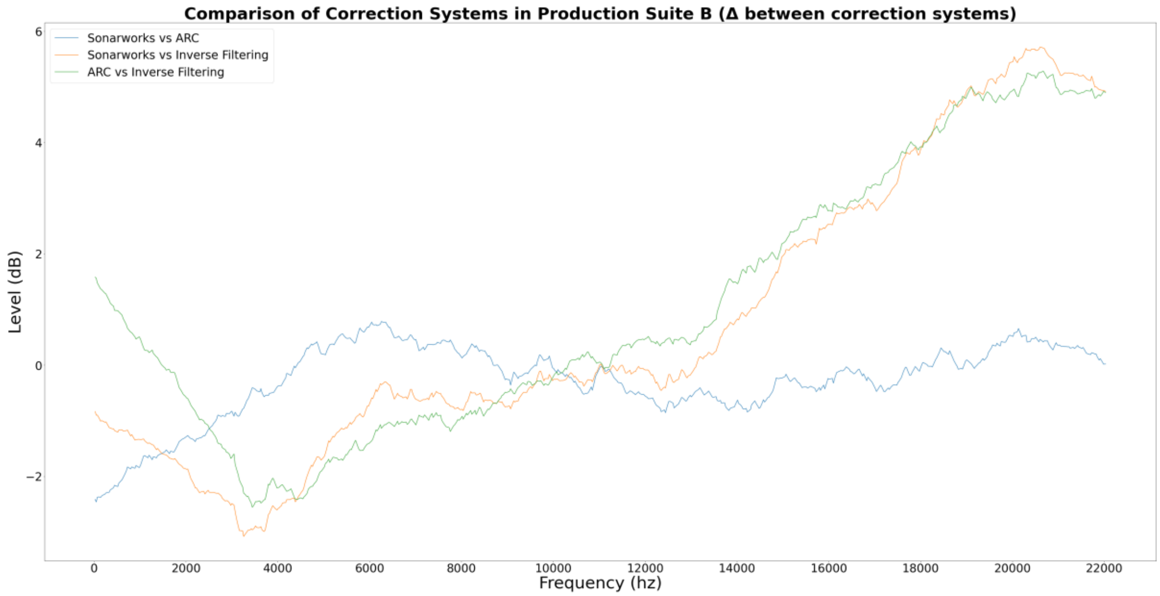


Figure 13. Comparison of the Change in Frequency Response in Production Suite B Induced by Each Correction System

Again, the inverse filtering graphs are drastically different from the Sonarworks vs ARC one, but in Production B the lines have a flatter section in the middle, showing that between 6 KHz and 12 KHz the Inverse filtering system was more similar to the other two. It is especially similar to Sonarworks, though the Characteristic rating scores for the two are drastically different.

The Sonarworks vs ARC plot here is less flat than in Edit Suite 4, showing the much larger cut that Sonarworks made in the lower frequency compared to ARC and the 1dB smaller cut Sonarworks made between 5 KHz and 8 KHz compared to ARC. From about 10 KHz to 18 KHz there is a difference of ~ 1 dB between the two, showing the larger boost made by ARC, and which then becomes a slight difference the other way (Sonarworks boosting more than ARC) at 19 KHz.

4.7.3 Comparison of the Change in Frequency Response in the Research Lab

The Comparison of the change in frequency response between each system in the Research Lab is shown in Figure 14 below.

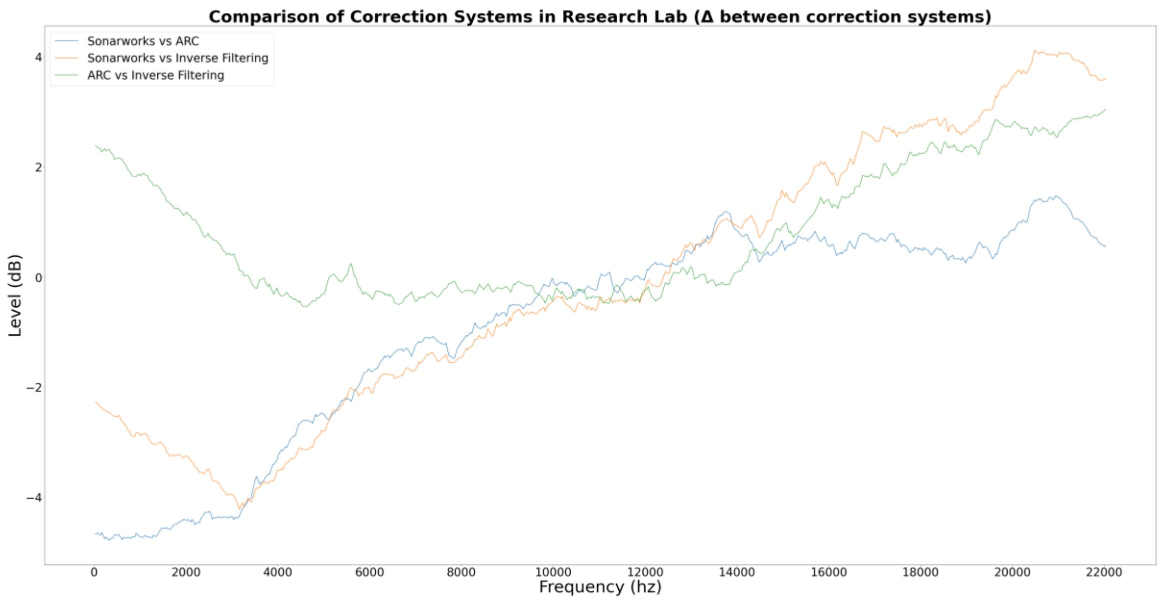


Figure 14. Comparison of the Change in Frequency Response in Research Lab Induced by Each Correction System

In this graph we see ARC and the Inverse filtering system hovering around 0 between ~3.75 KHz and 14 KHz, showing the only time that the Inverse filtering system performed successfully. While not as small as the difference between Sonarworks and ARC in the other two rooms, it is still a significant similarity. Above and below this range, ARC is boosting the up to the highest and lowest frequencies reproduced by up to 2dB. The characteristic ratings reflect a perceived higher bass response in the ARC corrected room over the Inverse Filtering system corrected room, but the larger high frequency boost by ARC is not reflected in the characteristic data. In fact, Sonarworks, which is boosting the high frequencies the most, is rated lower than both ARC and the Inverse Filtering system in terms of brightness.

Although the frequency response changes implemented by Sonarworks differed so greatly from the other two systems, there were still some similarities. Starting at around 14KHz Sonarworks and ARC made very similar changes, up until Sonarworks made a significant boost around 20 KHz, a range that a correction system should logically never manipulate due to it being above the range of hearing for almost all humans.

5. CONCLUSIONS

As was evident starting with Figure 4, Software based room correction can be hit-or-miss. In 2 out of the 3 rooms, some form of software-based correction generally preferred over the uncorrected room, indicating that software-based room correction is certainly capable of improving the listening experience in a small critical listening room, however, there was no one-size fits all product that improved the room's response. In each room, a different correction system was preferred over the others (In the Research Lab, the Inverse Filtering system was the most preferred *correction system*, but the uncorrected response was most preferred), further illustrating the complexity of creating software that performs well in all spaces and hinting at the different processing that each piece of software is "doing under-the-hood". In only one room (Edit Suite 4) did both commercial products improve the listening experience of the room, with listeners generally preferring Sonarworks Reference 4 over IK Multimedia's ARC.

On average, Sonarworks Reference 4 was preferred over IK Multimedia's ARC System 3, which was found to, on average, produce a brighter and clearer result than ARC. This was backed up by the frequency response plots, which showed similar frequency response changes induced by ARC and Sonarworks in the mid-frequency range, but that Sonarworks boosted higher frequencies more than ARC. In all three systems, the average perceived level of bass was reduced according to the listeners, which is a logical correction for small critical listening rooms where low frequency buildup tends to be a problem.

However, while software-based room correction *may* be able to improve the listening experience in a small critical listening room, it may also deteriorate that listening experience. Though very powerful, and, in this author's option, easy to use, it is not a stand-in for traditional acoustic treatments. If, however, such measures are impractical or unavailable for the listener, such as in an apartment or a university dorm, software-based correction may be a viable solution to improve the listening experience of the room.

5.1 Evaluation of Novel Inverse Filtering System

While the Inverse Filtering system was preferred over Sonarworks Reference 4 and IK Multimedia's ARC in one room (Research Lab), it was generally the least preferred of all 4 testing groups, far below the uncorrected response of the room. There are many possible reasons for this, but looking at Figures 9-11 indicate that the Inverse Filtering system was making much larger frequency response changes than ARC or Sonarworks were, and this author believes that to be a large part of why the Inverse Filtering system worsened the listening experience of Edit Suite 4 and Production Suite B. Additionally, the Inverse Filtering system did not use multiple impulse responses to capture the frequency response of the room to be corrected, which both ARC and Sonarworks did. It is possible that recording multiple impulse responses and averaging them could have improved the processing done by the inverse filtering system.

5.2 Future Work

Future studies with a larger pool of test spaces could narrow down further on what makes each correction system improve or deteriorate the listening experience in a particular room. In this study, Edit Suite 4 was the smallest room (107 Sq Ft), followed by Production Suite B (164 Sq Ft), and finally the Research Lab (180 Sq Ft), was the largest, which (coupled with the preference rating results) could indicate that the software is more effective in smaller sized listening environments, but more rooms would need to be included in the testing to draw such a conclusion.

It is this author's hope that the inverse filtering system can be improved into a viable, open-source solution for simple room correction. As suggested, limiting the processing both in terms of amplitude and across frequency ranges would greatly improve the efficacy of the Inverse Filtering system. By reducing the system's boosting and cutting to a smaller range of ~3dB would produce a much less drastic, and likely more useful result. Constraining the filtering to frequencies that the listener is likely to notice would likely improve the processing as well. Neither of these changes would be particularly difficult to implement but were outside the scope of this study.

The other glaring issue with the Inverse Filtering system is that it is offline-only, while the commercial systems can be used in a DAW with minimal latency. This greatly reduces the usefulness of the Inverse Filtering system even if it were to perform better than Sonarworks or ARC. In order to make the Inverse Filtering system a viable solution for room correction, code would need to be rewritten in C so that it could be turned into a plugin, which was again outside the scope of this study.

REFERENCES

4 Tracks to Reference for Mixing in 2018. (January 2021). https://www.izotope.com/en/learn/4-tracks-to-reference-for-mixing-in-2018

4 Popular Mixing Reference Tracks, and Why They Work. (January 2021).

ARC System 3. ARC System 3 - Acoustic Correction System. (January 2021). https://www.ikmultimedia.com/products/arc3/.

Bank, B. (2018). Multichannel Equalization and Crosstalk Cancellation Using Fixed-Pole IIR Filters. *Journal of the Audio Engineering Society, 66*(11), 901-909. doi:10.17743/jaes.2018.0061

Cecchi, S., Carini, A., & Spors, S. (2017). Room Response Equalization—A Review. *Applied Sciences, 8*(1), 16. doi:10.3390/app8010016

Cecchi, S., Palestini, L., Peretti, P., Romoli, L., Piazza, F., & Carini, A. (2011, March). Evaluation of a Multipoint Equalization System based on Impulse Responses Prototype Extraction. *Journal of the Audio Engineering Society*, *59*(3), 110-123.

Elliott, S. J. & Nelson, P. A. (1989, November) “Multiple-point equalization in a room using adaptive digital filters,” *Journal of the Audio Engineering Society*. 37, 899-907.

Gerzon, M. (1991). Digital room equalization. *Studio Sound*.

Green, L., & Dunbar, J. Y. (1947). Recording Studio Acoustics. *The Journal of the Acoustical Society of America*, *19*(3), 412–414.

Howison, W., Lamontagne, C., Luna, B., & Newell, D. (2005). Naive Room Response Deconvolution. *Openstax CNX*. Retrieved from http://cnx.org/contents/bc02964a-9f9c-4a21-8123-354b1b127647@5.1/Naive_Room_Response_Deconvolut

Kalkan, E. (2017, February 02). Deconvolution of two discrete time signals in frequency domain - File Exchange - MATLAB Central. Retrieved from https://www.mathworks.com/matlabcentral/fileexchange/60644-deconvolution-of-two-discrete-time-signals-in-frequency-domain

Karjalainen, M., & Paatero, T. (2006, July). Equalization of Loudspeaker and Room Responses Using Kautz Filters: Direct Least Squares Design. *EURASIP Journal on Advances in Signal Processing*, *2007*(1). doi: 10.1155/2007/60949

Miyoshi, M., & Kaneda, Y. (1988, February). Inverse filtering of room acoustics. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, *36*(2), 145–152. doi: 10.1109/29.1509

Nelson, P., Hamada, H., & Elliott, S. (1992, July). Adaptive inverse filters for stereophonic sound reproduction. *IEEE Transactions on Signal Processing*, *40*(7), 1621–1632. doi: 10.1109/78.143434

Norcross, S. G., Soulodre, G. A., & Lavoie, M. C. (2004, October). Subjective Investigations of Inverse Filtering. *Journal of the Audio Engineering Society, 52*(10), 1003-1028. Retrieved from http://www.aes.org/e-lib/browse.cfm?elib=13022

Olive, S., Jackson, J., Devantier, A., & Hunt, D. (2009, October). The Subjective and Objective Evaluation of Room Correction Products. *127th Convention of the Audio Engineering Society*.

Robjohns, H., & White, P. (2007, December). Room for Improvement - A Guide To DIY Studio Acoustics. Retrieved from https://www.soundonsound.com/techniques/room-improvement

Rubak, P. & Johansen, L. G. (2000, February). Design and Evaluation of Digital Filters Applied to Loudspeaker/Room Equalization. *108th Convention of the Audio Engineering Society.*

Sonarworks Reference 4. (January 2021). http://sonarworks.com/reference/.

Sprogis, K. (2016). *U.S. Patent No. US9380400B2*. Washington, DC: U.S. Patent and Trademark Office.

Sproģis, K., Bēms, H., & Popeliš, M. (2018). *U.S. Patent No. US10021484B2*. Washington, DC: U.S. Patent and Trademark Office.

Walker, R. (October 1998). Equalization of Room Acoustics and Adaptive Systems in the Equalization of Small Room Acoustics. *15th International Conference of the Audio Engineering Society.*

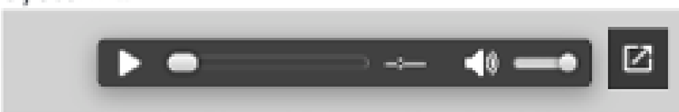
APPENDIX A – Example of Survey Questions

Please listen to the following 4 audio recordings. You may listen to them in any order, and as many or few times as you wish.

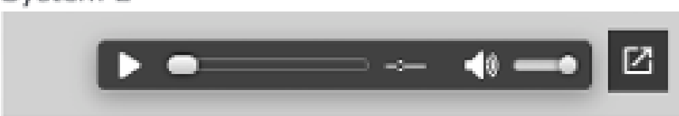
You do NOT need to listen to any of the recordings in their entirety, you only need to listen as much as is needed to answer the questions below.

The audio players may take a few seconds to load, please be patient

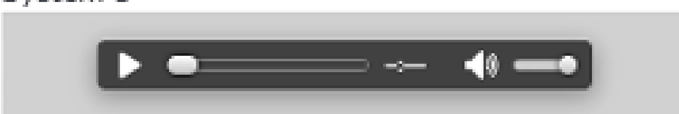
System 1



System 2



System 3



System 4

Please put the 4 systems in order according to how you feel they represent an accurate critical listening space for music production purposes

First system = most accurate critical listening space

System 1

System 2

System 3

System 4

Please rate System 1 in the following perceived qualities

1 2 3 3 4 5

Brightness (1 is dull, 3 is desirable, 5 is too bright)

Bass Balance (1 is lacking bass, 3 is well balance, 5 is bass-heavy)

Left-Right Balance (3 is centered, 1 is left heavy, 5 is right heavy)

Stereo Width (1 is very narrow, 3 is an even stereo field, 5 is a hole in the middle)

Clarity (1 is extremely poor, 5 means instruments are well defined)

Please rate System 2 in the following perceived qualities

1 2 3 3 4 5

Brightness (1 is dull, 3 is desirable, 5 is too bright)

Bass Balance (1 is lacking bass, 3 is well balance, 5 is bass-heavy)

Left-Right Balance (3 is centered, 1 is left heavy, 5 is right heavy)

Stereo Width (1 is very narrow, 3 is an even stereo field, 5 is a hole in the middle)

Clarity (1 is extremely poor, 5 means instruments are well defined)

Please rate System 3 in the following perceived qualities

1 2 3 3 4 5

Brightness (1 is dull, 3 is desirable, 5 is too bright)

Bass Balance (1 is lacking bass, 3 is well balance, 5 is bass-heavy)

Left-Right Balance (3 is centered, 1 is left heavy, 5 is right heavy)

Stereo Width (1 is very narrow, 3 is an even stereo field, 5 is a hole in the middle)

Clarity (1 is extremely poor, 5 means instruments are well defined)

Please rate System 4 in the following perceived qualities

1 2 3 3 4 5

Brightness (1 is dull, 3 is desirable, 5 is too bright)

Bass Balance (1 is lacking bass, 3 is well balance, 5 is bass-heavy)

Left-Right Balance (3 is centered, 1 is left heavy, 5 is right heavy)

Stereo Width (1 is very narrow, 3 is an even stereo field, 5 is a hole in the middle)

Clarity (1 is extremely poor, 5 means instruments are well defined)

← →

APPENDIX B – Individual Frequency Response Plots

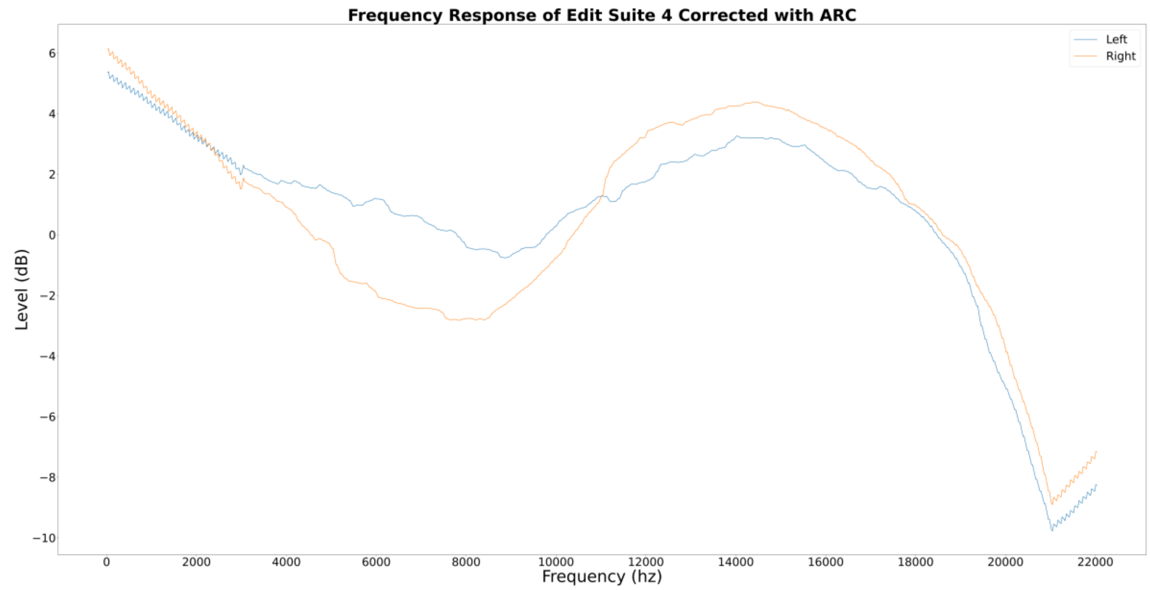


Figure 15. Frequency Response of Edit Suite 4 Corrected with ARC

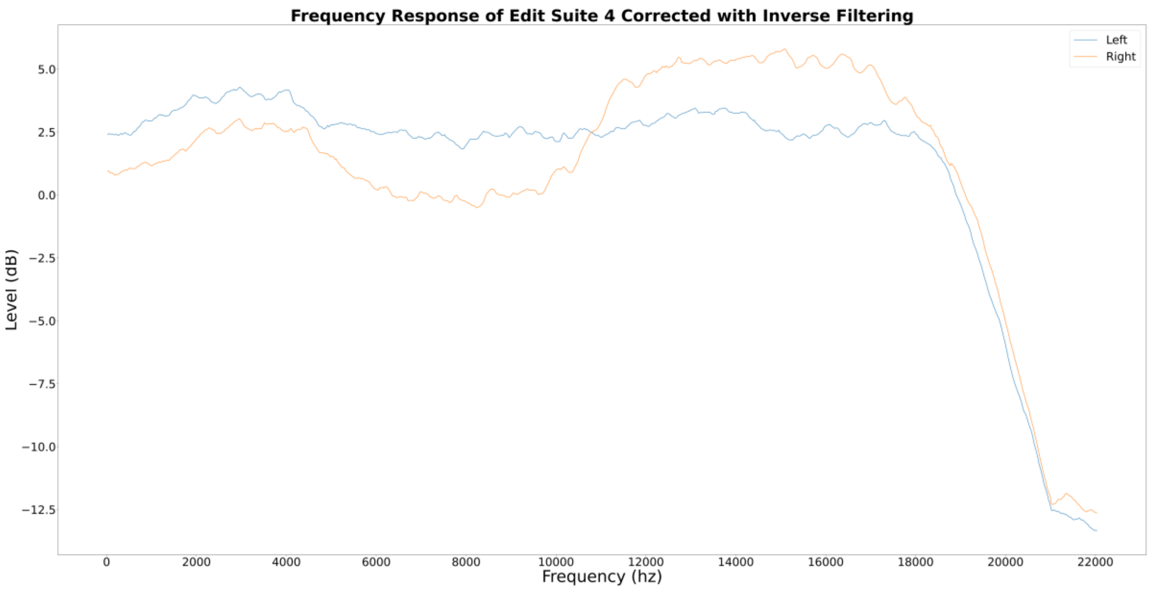


Figure 16. Frequency Response of Edit Suite 4 Corrected with Inverse Filtering

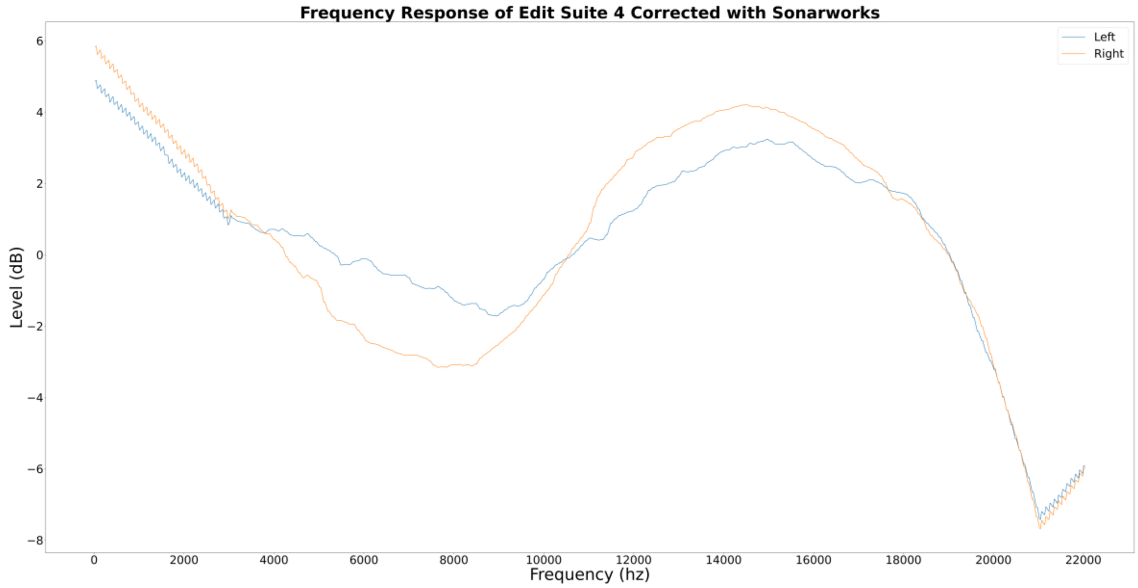


Figure 17. Frequency Response of Edit Suite 4 Corrected with Sonarworks

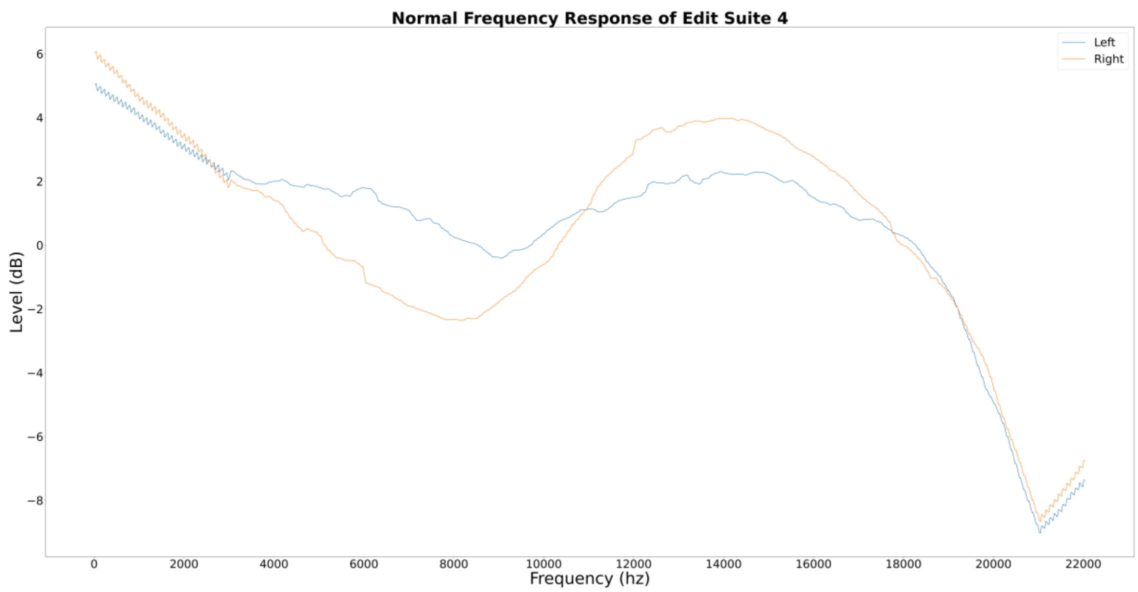


Figure 18. Normal Frequency Response of Edit Suite 4

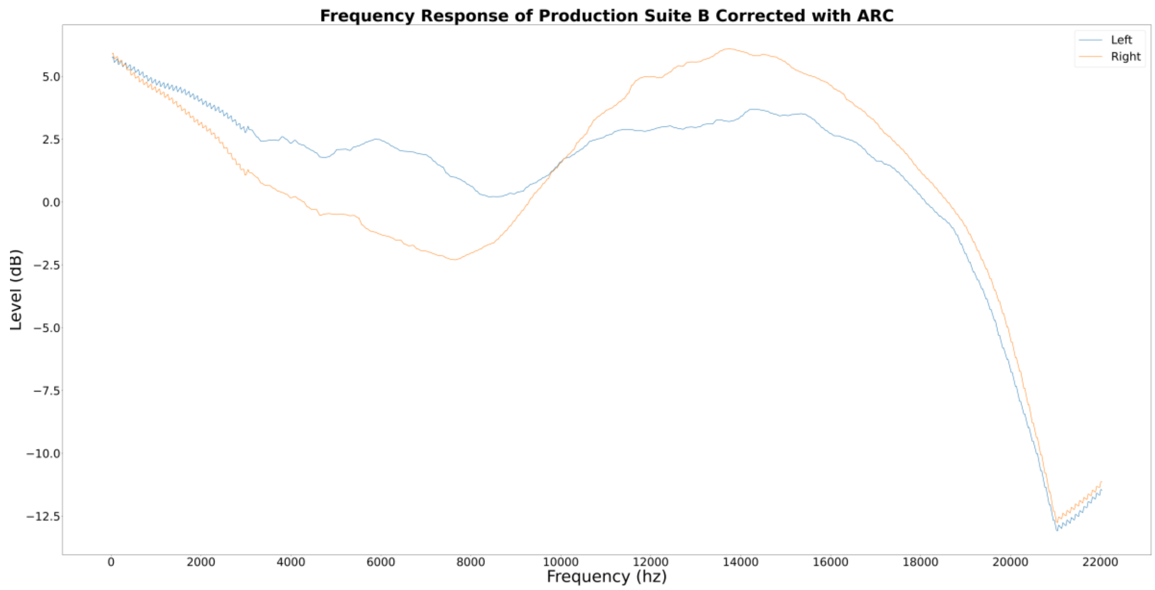


Figure 19. Frequency Response of Production Suite B Corrected with ARC

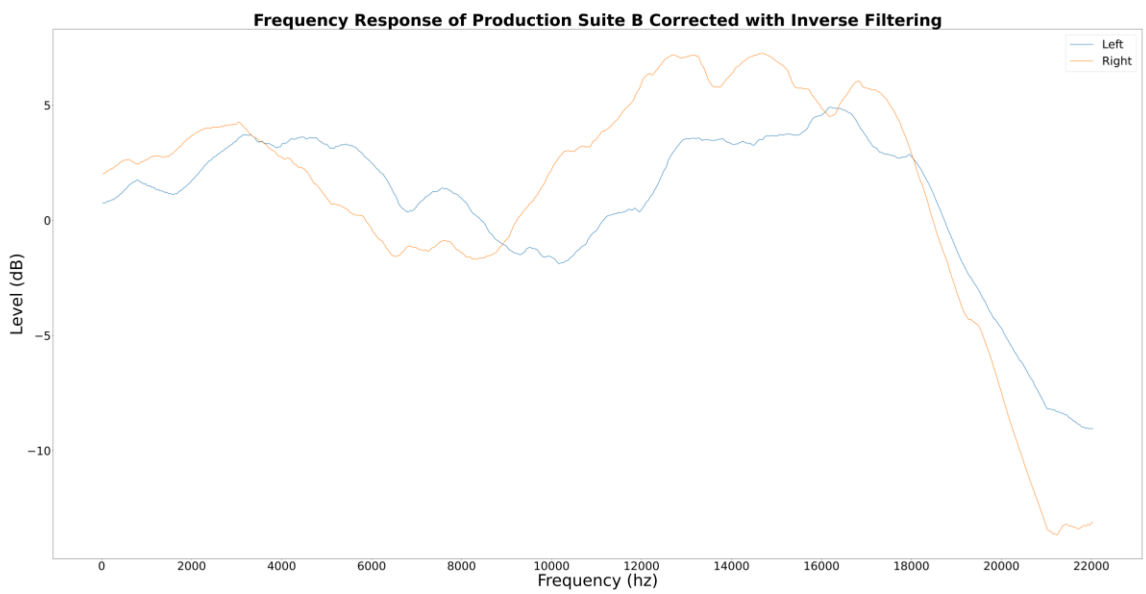


Figure 20. Frequency Response of Production Suite B Corrected with Inverse Filtering

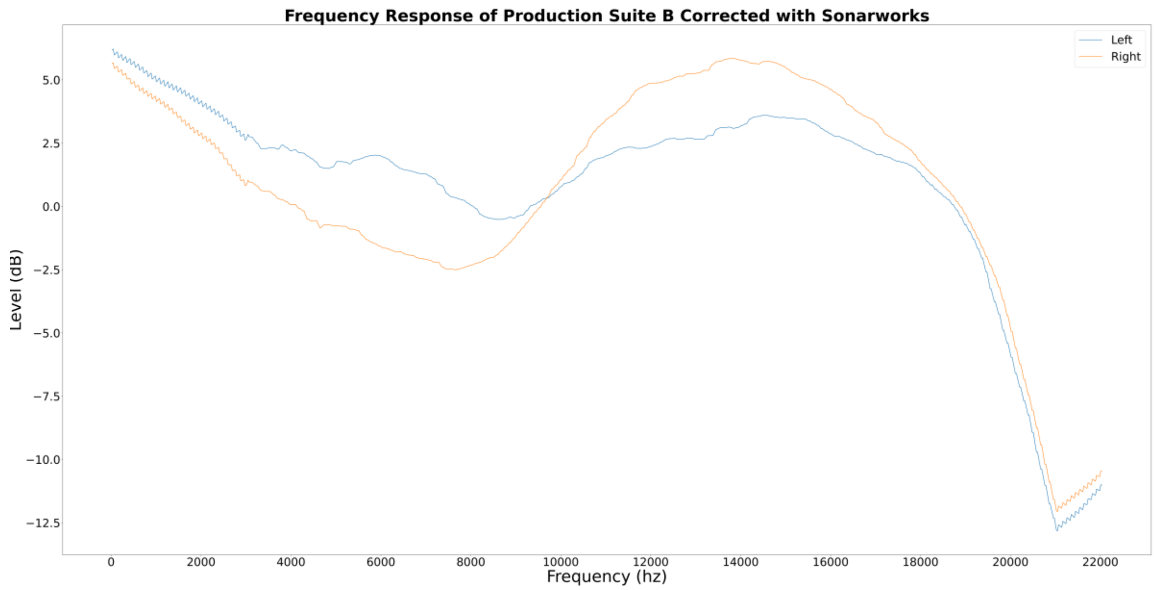


Figure 21. Frequency Response of Production Suite B Corrected with Sonarworks

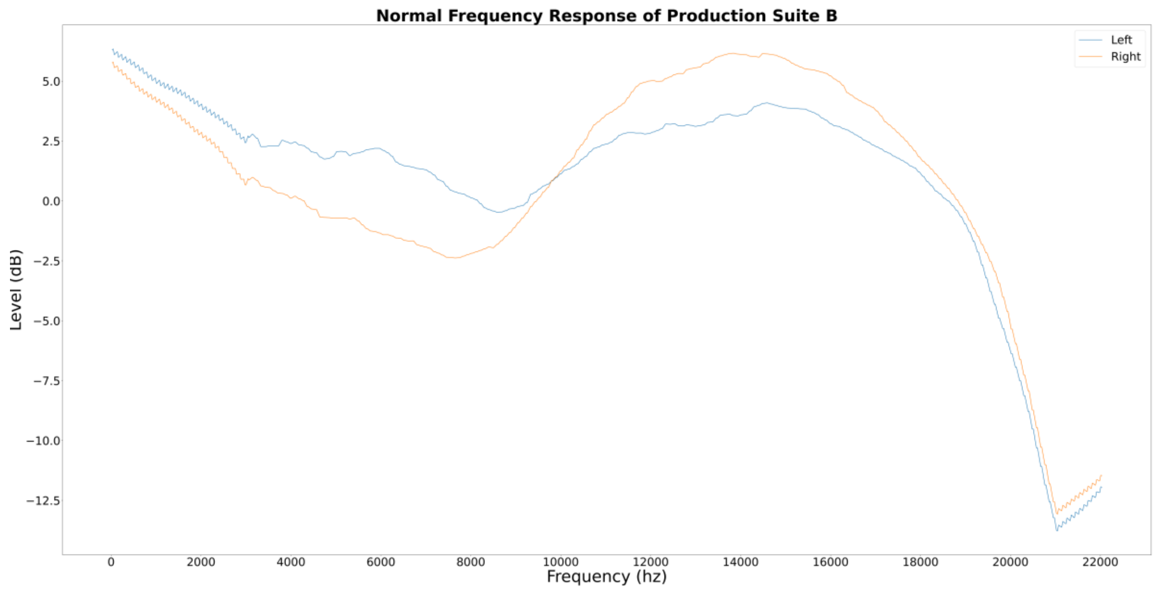


Figure 22. Normal Frequency Response of Production Suite B

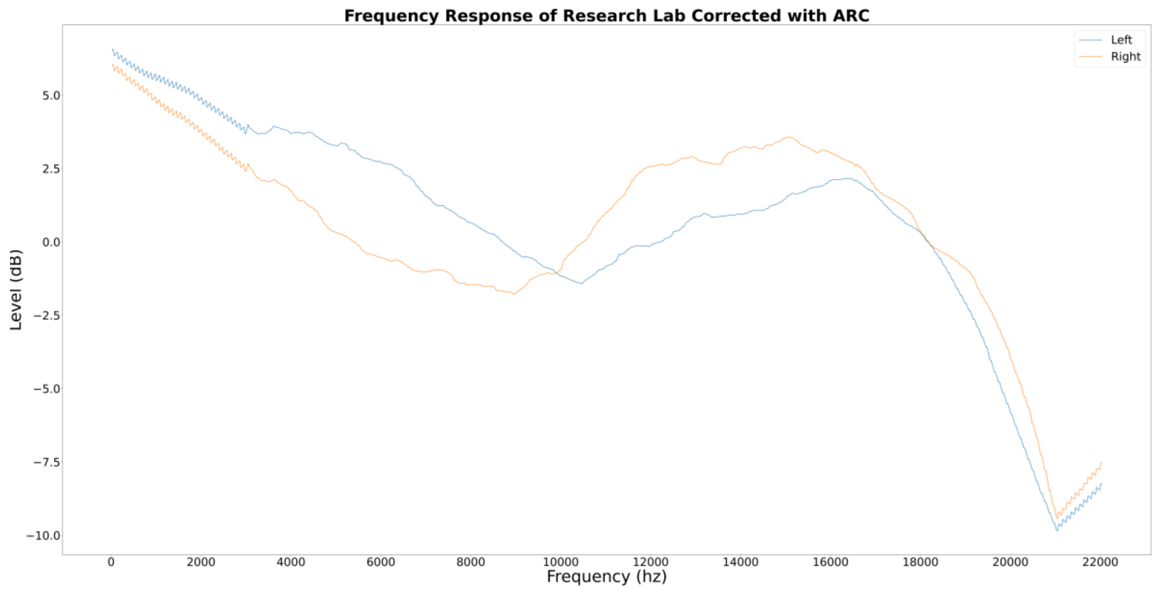


Figure 23. Frequency Response of Research Lab Corrected with ARC

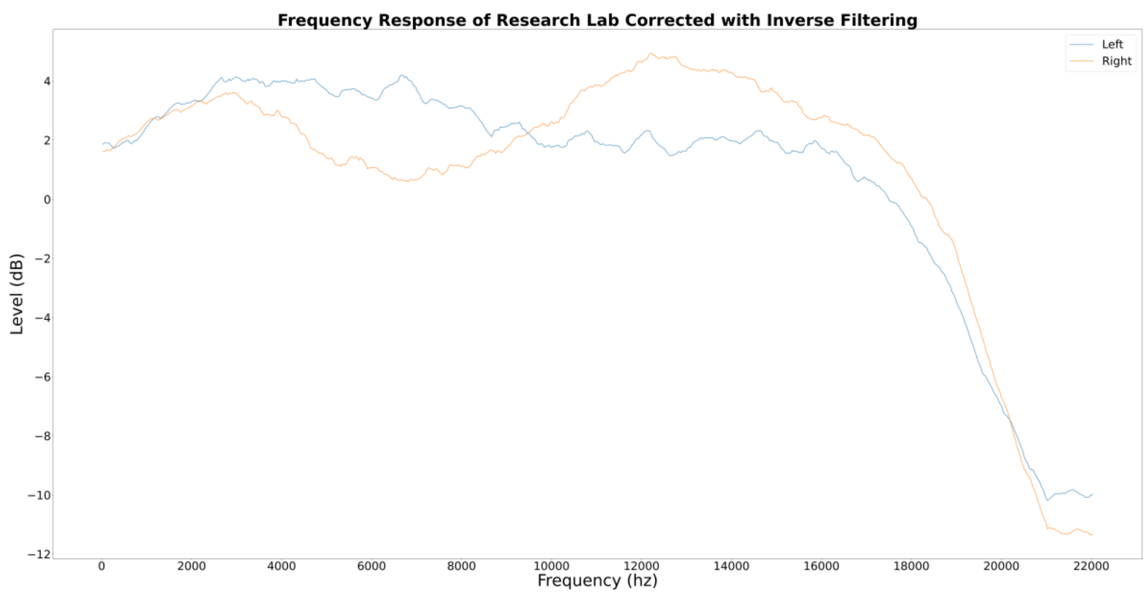


Figure 24. Frequency Response of Research Lab Corrected with Inverse Filtering

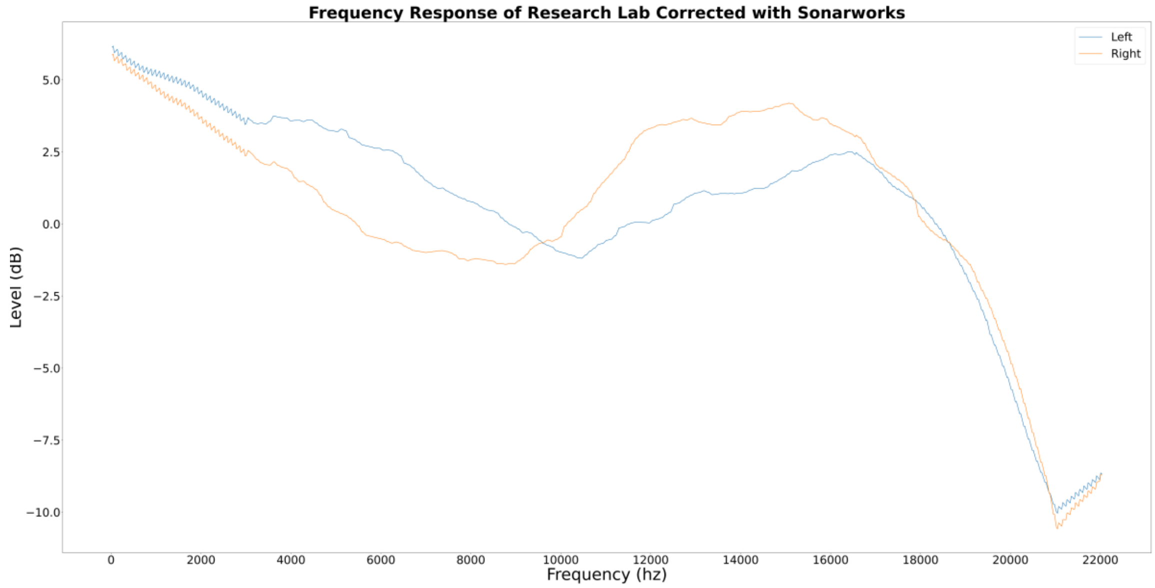


Figure 25. Frequency Response of Research Lab Corrected with Sonarworks

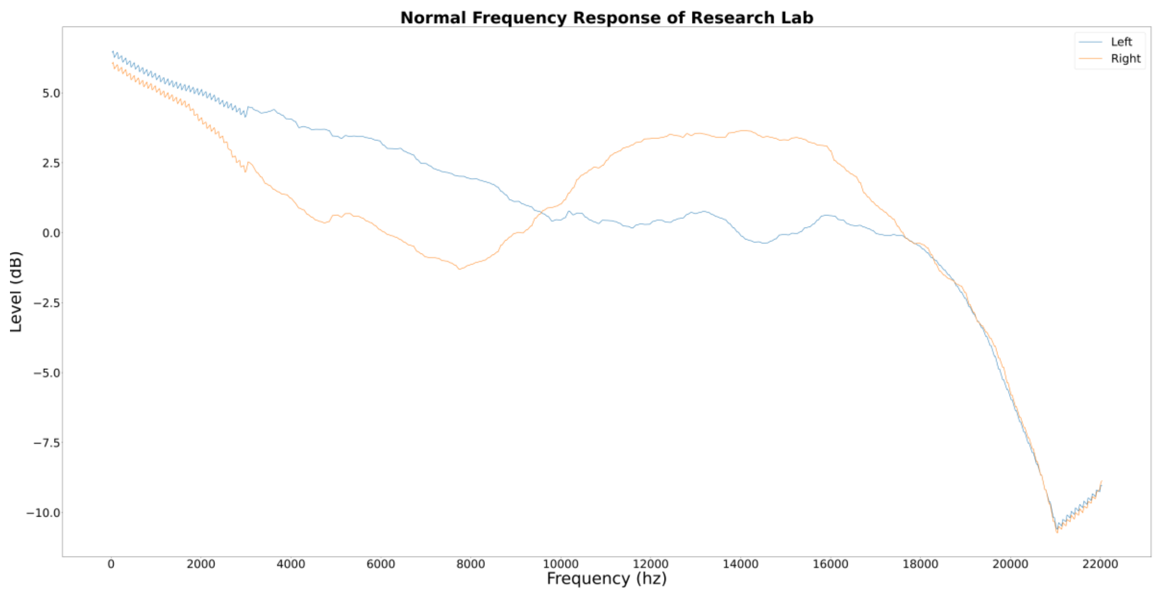


Figure 26. Normal Frequency Response of Research Lab

APPENDIX C – Code and Scripts

All Code, scripts, and Jupyter/Google Colaboratory Notebooks used during the process of this thesis can be found at https://github.com/michaelphagen/DIGITAL-CORRECTION-IN-SMALL-ROOMS]

The Smoothed Deconvolution code is included below for quick reference:

```
function[out]=mph354_deconvolution(sig,ir,smoothing)
%MPH354_DECONVOLUTION function made by Michael Hagen- MPH354
%
%   This function is primarily for use in the mph354_deconvolver
script, which is used for practical, flexible, and fast deconvolution
of two signals. This script completes the deconvolution of two signal
vectors, assuming that they have identical sample rates and vector
lengths.
%
%   Arguments should be the signal vector, the impulse response vector,
and the smoothing factor to be used.
%
%   sig: the vector of the main signal to have components removed via
%        deconvolution.
%   ir: the vector of the impulse response that will be what the signal
is deconvolved by.
%   smoothing: the strength of the smoothing (a higher smoothing value
will reduce the effect of the deconvolution process

%Compute the FFTs
SIG = fft(sig);
IR = fft(ir);

% Compute deconvolution
%generate the compensation vector
compensationFactor=smoothing*mean(IR.*conj(IR));
%create the IR vector compensated for phase manipulations
compensatedIR=(IR.*conj(IR));
%create the signal vector compensated for phase manipulations
compensatedSig=(SIG.*conj(IR));
%element-wise divide the compensated signal by the sum of the
compensated IR and the compensation factor to get the compensated
frequency domain output
OUT = compensatedSig./(compensatedIR+compensationFactor);
%convert the output into the time domain using ifft
out=ifft(OUT);
%take the real output of the time domain signal.
out=real(out);
%normalize the output
out= out / abs(max(out));
return
```