A New Technique for Capturing True Coincidence in Ambisonic Microphone Arrays

James P. Keary
Thesis

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

New York University
Thesis Advisor: Professor Paul Geluso

2014/1/10

Table of Contents

ACKNOWLEDGMENTS

This project would not have been possible without the advisement and encouragement of my professors Paul Geluso and Agnieszka Roginska. Thanks also to my colleagues and classmates Harihana Mohanraj, Alexander Marse and Andrew Lection, for their help capturing impulse response measurements for this work.

Thank you to Prof. Angelo Farina from the University of Parma, Italy, for his helpful and detailed email correspondance on multichannel coincident array impulse response measurements and theory. Thank you to Mr. David McGriffy, creator of the Core Sound's VST plug-ins, for his advice on filtering and decoding Ambisonic signals. Both of these gentlemen are extremely influential and important in the field of modern day Ambisonics, without their work my thesis would not be possible.

Thank you to those anonymous audio professionals in the field who participated in the subjective listening sessions; your time and advice were an invaluable contribution.

For copyright permissions, thank you to the source publishers and authors of the figures used in this thesis document. In reference to their copyrighted materials used with permission in this thesis, the publishers and authors do not endorse any of NYU's products or services. Only internal or personal use of this material is permitted.

A special thanks to my fiancée, my family and friends for letting me talk their ears off about this subject for the past 7 months. They were more helpful than they know.

ABSTRACT

This thesis presents an in-depth investigation on the topics of correlation and spatial aliasing in Ambisonic microphone arrays. A measurement procedure that creates true coincidence using a multiple impulse response and a one capsule at a time approach is proposed. The goal of the proposed approach is the pursuit of more natural ambisonic 3D surround sound. The preface section of this document serves as an introduction to important notions and theories that are the background on which this work is established. The introductory section describes the project's goals and motivations. The measurement method is proposed and outlined. Analysis comparing the proposed technique to a traditional Ambisonic microphone array impulse response measurement was done. Measurements and analysis were also done on the Double MS Z array, an array that can be matrixed to the Ambisonic B format. It was discovered that the new approach did indeed succeed in achieving better correlation in the direct part of the signal and decorrelation in the reflection and reverb of the signal. However, the true breakthrough of this thesis work was in the discovery of medial and lateral reflections caused by the distance between microphone capsules. Professional listening sessions were done comparing the new technique to Ambisonic microphone arrays revealing interesting results. This thesis project is considered an extension in the field of Ambisonics.

0 PREFACE

In order to prepare the reader for the concepts in this thesis, it is thought important to go back to “the beginning” as it were; to tell the story of the sampling theorem. The author believes that the sampling theorem provides the best entry point to understanding this work scientifically and mathematically. And at the very least provides an understanding that something as little as a few milliseconds between samples, or a few millimeters between microphons, can make a big difference. That difference is the study of this thesis work.

0.1 On the Sampling Theorem

A general rule of thumb when digitally recording a sound signal is that the highest frequency a digitally encoded signal can represent is known as the Nyquist frequency. The Nyquist frequency of a Pulse Code Modulated (PCM) encoded signal is exactly half the sampling rate. Therefore the sampling rate directly determines the highest possible frequency in the digital signal. Depending on the signal that you need to record, and depending on the purpose for which you are making the recording, you can use different sampling rates.

Let’s assume that I want to record a kick drum, a sound which is primarily made up of low frequencies. The thud of a kick drum generally resonates somewhere around 60 to 80 Hz. So if I wanted I could digitally encode that signal using a fairly low sampling rate of 1600 PCM samples per second, or 16 Hz, which would be 2 times the highest frequency in the signal. However, the kick drum also contains some mid-range frequency content as it attacks somewhere around 3 kHz. So I would then think to raise my sampling rate to 6 kHz. Now, the recording comes to life and we hear a more natural kick drum. This example serves to point out that the higher the sampling rate is, the more the recorded signal will sound like the original, and this is termed *natural*. So if the purpose is for the kick drum to sound like a real kick drum, then it would be better to choose a higher sampling rate at which to encode the digital signal.

Let us instead say that the priority is to keep my sampling rate low, such is the case in modern day telephony. Sound can only move so fast, and 1 millisecond translates to 342.3 km through air or 200 km through a fiber optic cable[1]. It takes more time to transmit more samples, and humans start to notice that there is a delay in telephone conversation at roughly 25 ms. So, telephone companies want to keep the sampling rate as low as possible because one of their top priorities is real time conversation. Phone signals are generally comprised of the human voice, and human voices are generally comprised of frequencies lower than 3 kHz. So the sampling rate telephone companies use is 8 kHz.

With this information in mind, let's return to our kick drum recording, where transmission speed is not an issue. If the priority is the most natural sounding kick drum, what should our sampling rate be? There are 2 points at which sound waves become either too long or too short for the hardware of human ears to resonate; 20 Hz and 20 kHz respectively. This range is what we call the range of human hearing. So, if I encode my recorded signal at a sampling rate of 40 kHz, or 2 times 20 kHz, then I am encoding frequencies up to the highest we can hear, and I am in accordance with the sampling theorum. To be fair, this is a general rule, as many factors, such as listener age, hearing damage, air temperature, and medium through which the sound travels, can affect this range. However, if I assume that a young person with limited past exposure to loud noise will be listening to my recorded sound, in a relatively dry climate, at room temperature, then I can confidently set my sampling rate at 40 kHz.

However, most used sampling rates extend well beyond 40 kHz; 44.1 kHz for CD quality, 48 kHz for high quality audio with video, and further still to super high sampling rates of

[1] [Pohlmann:2010aa]

88, 96 and 192 kHz. If those frequencies are deemed inaudible, then what is the point to using these higher sampling rates? The reason is revealed in the second part of the sampling theorum, which warns us that if we do not stop frequencies above the Nyquist frequency with a filter before the sampling process, then they may "fold over" into the audible portion of the signal as aliased versions of their former self (figure 1). This effect is known as *temporal aliasing*, and it happens in the higher frequencies of a low sampling rate non-filter signal. To correct this we must use a low-pass filter to stop those frequencies from being encoded in the first place. However, it is quite difficult to build a low pass filter with an extremely sharp immediate cutoff. So, to avoid aliasing, higher sampling rates and low pass filters with more gradual cutoffs are used instead (figure 2).

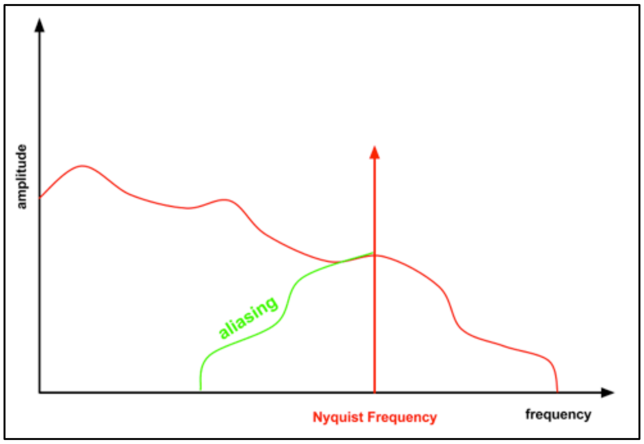


Figure 1. Visualization of aliased frequencies folding over above the Nyquist frequency.

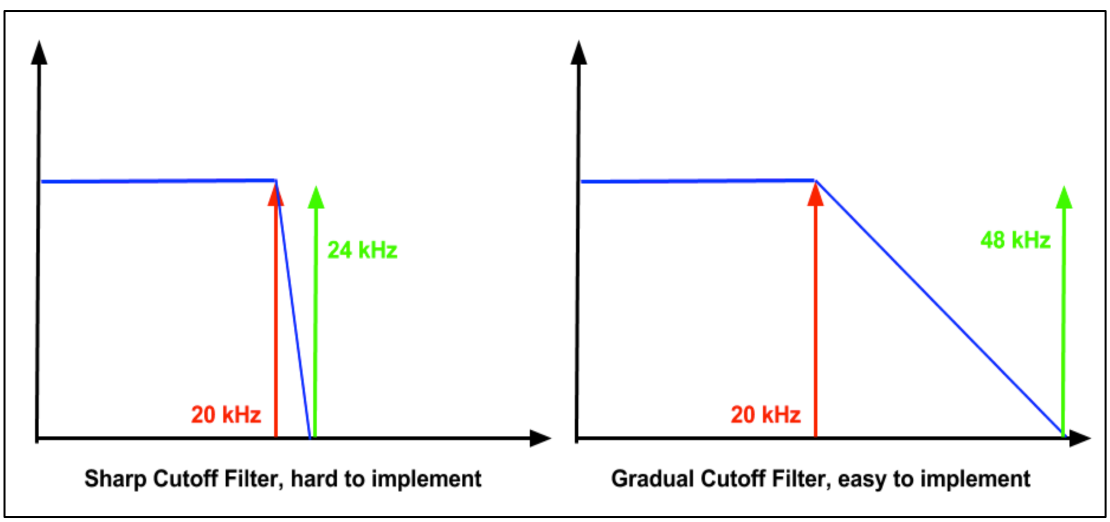


Figure 2. Two low-pass filters, x axis is frequency, y axis is amplitude.

0.2 "You need to learn to "Think spatial"…"[2]

Let us now apply the same principles as we did to time to the concepts of space. Imagine two microphones capturing a sound wave emanating from a source. They capture the sound pressure wave as it is at two different positions in space. It is not too difficult then to draw the following comparison: the space between these two microphone capsules is analogous to the time between two digital samples. They both give us the same information: the pressure of the sound wave as a discrete function. One breaks *time* into discrete units, and the other, *space*. So, the same rules of the the sampling theorum that we observed to in the digital time domain can now safely be applied in the new digital space domain.

It follows from this that, just as *temporal aliasing* occurs at frequencies faster than the time between two samples, *spatial aliasing* will occur at wavelengths shorter than the distance between the two capsules. To imagine this, picture the digital sampling process in the time domain. The wavelengths that are shorter than the time between two samples (green), will be reconstructed as a frequency of an entirely different wavelength (dotted red); a wavelength that is based on the amplitude of the true frequency at the time of the sampling interval. The same is true in space with two microphone capsules.

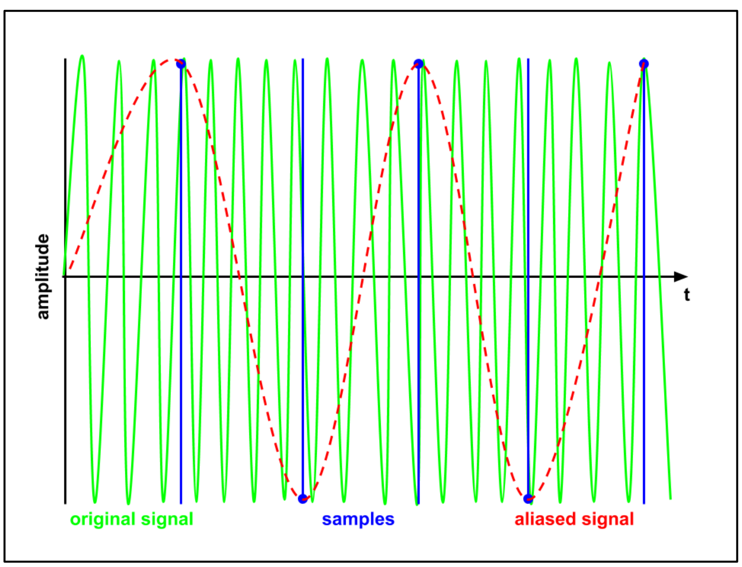


Figure 3. Temporal Aliasing in the Time Domain.

[2] Angelo Farina (Personal email communication, Nov 5th, 2013)

So, in order to avoid *spatially* aliasing higher frequencies, the space between two capsules must be less that half the smallest wavelength needed to be captured. Applying the sampling theorum to space lets us know that the distance between the acoustic center of two microphone capsules can, and will, determine the quality of the recorded signal. So the same questions we asked before apply once again. Only now instead of asking what sampling rate we want, we must ask: how close should our microphone capsules be? The answer to this question is explored in this thesis.

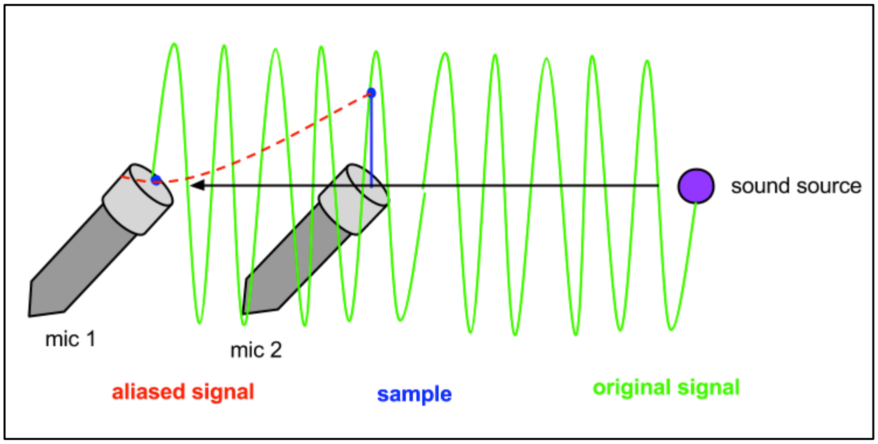


Figure 4. Spatial Aliasing caused by capsule distance.

1 Introduction

1.1 Basics Principles of Ambisonics

Ambisonics is a surround recording and playback system developed by Michael Gerzon in the 1970s. It was advertised as a system that could faithfully recreate the original 2D or 3D soundfield within a sweetspot through the use of four cardioid microphone capsules in a closely spaced array, output to as many speakers as one has available. It was considered to be the greatest advance in surround sound recreation when it was developed.

The basic matrixing of the system is to take the four cardioid signals, or A format, and matrix them into what is known as B format. The following equations are used to do so:

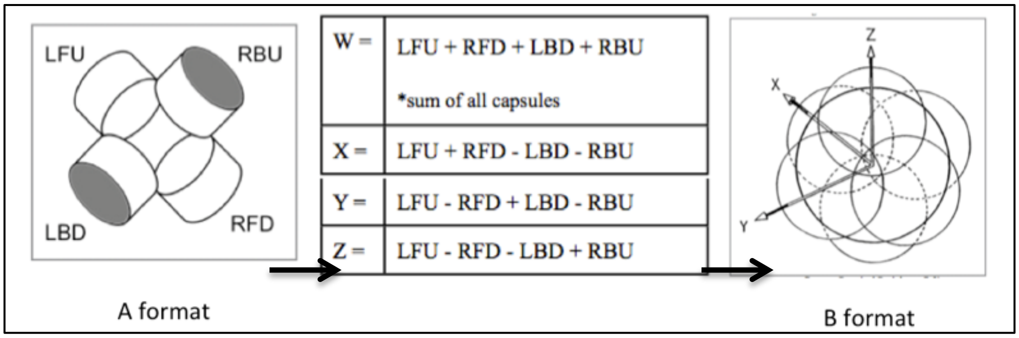


Figure 5. A format to B format equations. © 2009 Fulwider and © 1980 AES, used with permissions.

With these four B format signals, we have an omnidirectional signal, W, and three bidirectional signals: X for front and back, Y for left and right, and Z for up and down[3]. This is done because, as Montgomery[4] explains, "B format channels are transmission channels, not speaker feeds". With B format it is possible to create microphone responses pointing in various directions. The A format angles are all 45° off axis and 109.5° apart from each other. B format angles are on axial planes X front and back, Y left and right, and Z up and down. So, they can more easily be sent to their appropriate speakers depending on the speaker configuration you are using.

Once the signals are properly matrixed to B format, the next step is sending these signals to the appropriate speakers at the right amount through a process called *decoding*. In a general sense, decoding is done though a linear combination of a certain amount of each signal to each speaker feed. For every individual speaker, each of the four signals gets multiplied by a gain factor; the set of weighted gain factors for every speaker is determined by how many speakers are involved and their arrangement around the listener[5], usually creating some geometric polygon. The four B signals can be stripped down to a four speaker quadrophonic arrangment, a two speaker stereophonic or a one speaker monophonic arrangement. The system is forward compatible too in that the four signals can be distributed to as of yet undefined speaker arrangments with more than four speakers. By adding more speakers around the listener more

[3] [macCába:2002aa]
[4] [Montgomery:2013aa]
[5] [Gerzon:1973aa]

parts to the soundfield are added in, strengthening the surround sound image. Gerzon also describes the use of shelf filters in the decoding process to boost frequency ranges depending on psychoacoustic principles[6]. This step seems like an augmentation of the soundfield which is against the idea of natural soundfield representation, but perhaps this author is missing the intent of shelf filters. Regardless, the decoding system is more desirable than discrete channel mixing practices because it alleviates the problematic workflow that involves multiple mixing sessions per discrete channel playback system (stereo, 5.1, 7.1, etc). In comparison to discrete channel mixing, as long as Ambisonics is used throughout the entire process from recording to playback, there is no need to do multiple mixing sessions.

1.2 Ambisonics Today

Despite all the aparent advantages of Ambisonics, it has never been fully embraced commerically by the audio industry[7]. Traditionally, surround sound was confined to film, since such advanced and expensive audio systems were only available in movie theatres. Unfortunately for Ambisonics, the film industry chose discrete channel mixing practices over Ambisonics early on. It continues to do so today with Dolby's Atmos system, which combines discrete channel mixing practices and vector-based amplitude panning[8] for up to 128 discrete audio channels[9]. Despite this fact Ambisonics has always been a part of the conversation as producing an accurate and natural soundfield has always been desirable.

Despite Ambisonics's general non acceptance, Ambisonic software and hardware has become more mainstream. These days, Ambisonic programs and plug-ins are abundant, especially among smaller, less mainstream, and less movie studio based digital audio

---

[6] [Gerzon:1980aa]
[7] [macCába:2002aa], [Elen:2001aa]
[8] [Pulkki:2002aa]
[9] http://www.dolby.com/us/en/professional/technology/cinema/dolby-atmos.html

workstations (DAWs). DAWs such as Reaper[10] and Ardour support more I/O flexibility. This in turn supports Ambisonics; more channels and more speaker feed options gives Ambisonics the forward compatibility mentioned earlier. Furthermore, audio programming languages such as CSound and SuperCollider now have Ambisonic libraries[11] and toolkits[12]. These systems allow for Ambisonics to flourish with a multitude of encoder, decoder, panner and reverb plug-ins as well as standalone applications[13].

In hardware advances, the audio industry is starting to generate Ambisonic microphones with higher directionality and better localization capabilities. These microphones were previously considered impossible to create because of hardware limitations. However, due to advances in academia and technology[14], they are now becoming commerically. For example, the Eigenmike[15] is being used for everything from teleconferencing to academia to security and surveillance.

In summary, Ambisonics today is more widely supported than ever. As the hardare complexities of the system start to become a reality in the form of these better Ambisonic microphones, the elegant mathematical complexities are being tamed in these software applications for easy use. The technology once thought an historical deadend is being revived.

1.3 Motivation

There are several huge drawbacks to the Ambisonics system, drawbacks that the system has always faced and that, despite their best efforts, cannot seem to get around. It is due to these seemingly insurmountable drawbacks that Ambisonics has not yet gained wider acceptance in

[10] [Wiggins:2008aa]
[11] [Hofmann:2009aa]
[12] [TheAmbisonicToolkit:aa]
[13] [Adriaensen:2010aa], [Geeknet:2009aa], [VVAudio:2007aa], [Wiggins:2008aa]
[14] [Abhayapala:2002aa], [Manola:2012aa]
[15] [Eigenmike:aa]

film and music. As will be seen in the literature review, the system *was* developed for music recording, mixing and playback. But it has never advanced in that direction.

One major problem that Ambisonic microphones face is the symbiotic relationship between cost and quality. With any multi-capsuled microphone array, the increased number of capsules increases the cost. To cut costs, companies may use capsules that are lower in quality than desired. However, the use of better quality microphone capsules is a must if these microphones are to be used for professional music and film. This is exactly the point of the unestablished microphone array Double MS Z. The Double MS Z array is an Ambisonic-like array with 3 bidirectional microphones facing frontwards, leftwards and upwards[16]. It was designed with music production in mind, and it uses high-quality Senheiser Twin[17] microphones for the best frequency responses. With the more established Ambisonic arrays, i.e. arrays based on Gerzon's design, the quality of the microphone capsules have in the past been criticized for not being good enough for music recording. Companies that build Ambisonic microphone arrays, such as Core Sound (the Tetra Mic) and MH Acoustics (the Eigenmike), must tread the line between quality and quantity with their capsules.

The next major drawback in the Ambisonic system is *spatial aliasing* caused by capsule distance. All closely spaced microphone arrays operate on the principle that multiple microphone capsules work together as one to synthesize a natural soundfield. However, as illustrated in the preface, the capsules will, by their very nature, work against each other as well. Because it is physically impossibe for the capsules to occupy the same point in space and time, coincidence issues will arise, especially in the higher frequencies.

[16] [Geluso:2012aa]
[17] [Sennheiser:2010aa]

This project is a study of the distance between microphone capsules, the use of multiple unique capsules, the quality of those capsules, and the affect these factors have on sound quality. The drawbacks to the Ambisonics system as outlined above are the basic motivation behind this thesis work.

1.4 Goals

> ***This should convince skeptics of the fallacy of assuming that "nearly coincident" is good enough!*** - [Gerzon:1975ac]

The main goal of this project is to ask *what if*? What if we can simulate perfect physically coincident conditions? And, what if we can use better quality microphone capsules when doing so? The preface described how and why distance affects sound quality. This thesis proposes that there is a way to achieve improved sound quality. We can take impulse responses of the capsules of a closely spaced microphone array one at a time, at the same location in space, and then synthesize the different parts of the soundfield together in post. We can do this with a high quality microphone. It is hypothesized that upon convolving these impulse responses with a dry signal, we would even achieve a more accurate representation of a 3D soundfield than that of the signal received from the Ambisonic microphone array itself. The goal of this project is to avoid the bad artifacts that capsule distance and lower quality capsules produce, and to do so in a cost effective manner, i.e. only using one high quality microphone.

Furthermore, a smooth frequency response from capsule to capsule must also be taken into account, as it is in the Ambisonic microphone. With the proposed system, since we are using the same capsule at every Ambisonically related angle, the difference is seen in the *impulse responses*, not the mechanics of the microphone. As long as there is limited time in between impulse responses, the mechanics of the microphone will stay linear. It is hypothesized that this will give a smoother all-around frequency response from capsule to capsule. Furthermore, the differences in the room's reflections and reverberations of each impulse may even serve to give a better sense of sound *immersion*. And if we were to use a better quality microphone capsule, then the frequency response would be flatter all around, offering a smoother surround response when synthesizing the signals. This thesis sets out to determine if a more realistic soundfield can be achieved through this approach.

1.5 Contributions to the Field

What are the implications of doing such a study? It is theorized that the method can be used to create a realistic surround-sound convolution reverberation plugin; one that is perhaps more true to the 3D soundfield than an Ambisonic microphone itself. Furthermore, an exploration of this subject has apparently never before been attempted. To be clear, Professor Angelo Farina has done similar work with virtual ambisonic arrays[18], but his project placed the microphone at different locations as well as in different orientations. This project attempts only different orientations while keeping the same location. However, there are drawbacks to the proposed method. This section addresses those drawbacks.

First, the proposed system cannot *replace* Ambisonic microphones because it is a virtual implementation. It cannot be used to record music or speech. The only application of this work would be in the form of a convolution reverb plug-in. However, such a tool for producers and engineers to use in a musical application, would actually be a valuable outcome. Furthermore, Ambisonics has always had always faced barriers to breaking into the music production field; thus adding a high quality tool such as one produced through this process might be quite beneficial in the field.

[18] http://www.angelofarina.it/Public/RotorLego/

Second, there have been some researchers[19] who found the difference in capsule location to be acceptable. For instance, Benjamin proved that the interaural time differences produced by Ambisonic microphone arrays are similar enough to natural hearing in localizing sound, especially when more and more speakers are added in, with worse results 90° off axis. However, it also must be noted that tiny differences in time (and in this case space) will affect the overall timbre, or frequency response, of the signal in a process called temporal smearing[20]. Human ears are remarkably adept at noticing these tiny timbral differences in sound, so even though time cues may not be too misaligned, this thesis argues that the frequency content is still impacted. How dramatically it is impacted is studied in this work.

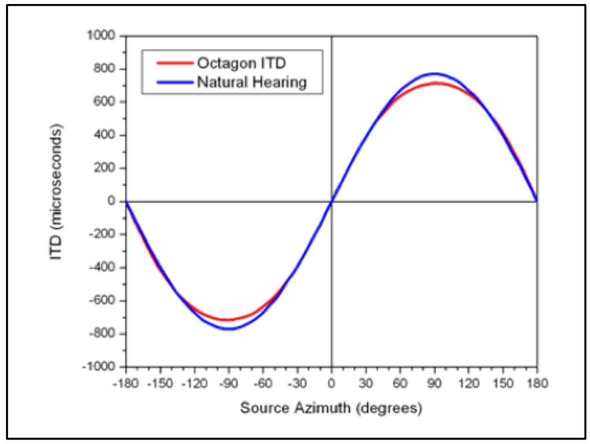


Figure 6. [Benjamin:2010aa]. p. 5. The Interaural Time Differences curve of full 360°. Greatest difference is noticed 90° off axis, i.e. side, otherwise there is no recognizable difference. © 2010 AES, used with permission.

Finally, it must be recognized that there are possibly some desireable artifacts that arise from capsule distance. When frequencies are longer in wavelength than the distance between capsules, the phase velocity of those frequencies, i.e. the speed of change of the phase between the two capsules, will give the listener directional information at low frequencies[21]. Which will in turn help the ears localize the signal. This has been noted as a positive, albeit, as this thesis argues again, unnatural, side effect to the Ambisonic design. Another intereting discovery was made by Dmochowski[22] when studying spatial aliasing. He noticed that in a broadband signal, such as music or speech, “our auditory system is able to remarkably resolve the phase ambiguity

[19] [Benjamin:2010aa]
[20] [Hou:1994aa]
[21] [Gerzon:1992aa]
[22] [Dmochowski:2009aa]

and localize quite precisely with ITD only." So, even thought there may be spatial aliasing in all the higher frequencies, our ears are still able to "wade through" that mess, and localize the sound. In this study it is not determined if this localization effect is present because of the distance between capsules. It must be noted that in another study[23] proper localization was found to be true only in the ear's of well-trained listeners, i.e. listeners with expereience in the the field of audio production and research. Only those well trained ears were able to stablize and localize the sound image, otherwise spatial aliasing would ruin localization cues in the high frequencies. As this author has been interjecting throughout this paragraph, it is clear that there are different schools of thought on the notion of what a natural listening experience implys.

Regardless, this work proposes solutions to the basic and inherent flaws that have held back the Ambisonic system since its inception. Despite all and any reason to the contrary to not explore coincidence issues in this manner, hardware-based solutions to these problems may never come in a cost effective way. This thesis proposes a unique contribution to the field in that we can simulate what actual true coincidence may sound like before it is techologically possible. If the application of Ambisonics is accurate soundfield representation, then if we were able to bypass the physical limitations described above, we would have achieved that goal. The true question that must be asked is this: If we are able to achieve these goals, would we recognize the outcome as more natural sound, and would we prefer it?

1.6 Outline of Thesis

The ideas discussed in Section 1 are the theoretical foundation upon which this thesis is built. Section 2 is a literature review on the study of coincidence. Ambisonics is examined, covering its history, theory, mathematics and physics. Section 3 is an outline of the experiment

[23] [Nettingsmeier:2010aa]

and a description of the measurement process. Section 4 analyzes the data retrieved from the measurements outlined in section 3. Section 5 describes a professional listening session and the results of these sessions. Section 6 offers concluding thoughts on the topic as a whole, on what was learned that was not known before, and suggesting further study to be undertaken. Although most of the literature review is in section 2, much literature is reviewed throughout these sections to help show how certain decisions were made, and the conclusions reached.

2 Literature Review

2.1 Quadrophonics and Discrete Channel Mixing Practices

Any proper introduction to Ambisonics should probably also have an introduction to the unsuccessful Quadrophonics system that directly preceeded it and led to Ambisonics's introduction. In the early 1970s, the Quadrophonic playback system was developed as the first commercial attempt at "surround sound" for home music listening[24]. It consisted of four speakers placed in a square configuration around the listener, and was soon deemed a failure because of poor localization issues. The problem with quadrophonics was actually present in stereo recordings as well, but less conspicuous: stereo only produces a natural sound image between the two speakers if the speakers are 60° apart and if the listener is seated in what is know as the "sweet spot". If the speakers are placed any further apart, such as 90° for quadrophonics, or if the listener is not seated in the sweet spot, then the constructed phantom sound image that appears between the two speakers will fracture, and the sound source would

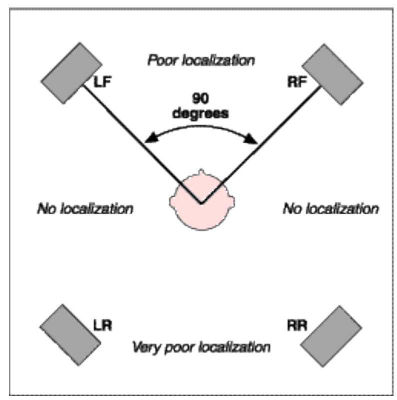


Figure 7. [Elen:2001aa]. p. 1. The original Quadrophonic system placed four speakers 90° apart: LF stands for Left Front, RF for Right Front, LR for Left Rear, and RR for Right Rear. The image shows where localization was deemed poor. Used with permission by Richard Elen ©.

[24] [Gerzon:1974aa], [Thornton:2009aa], [Elen:2001aa]

instead come from the closest speaker to the listener. Early quadrophonics only served to intensify this issue[25].

> ***"The basic fault of pairwise blending is that each musical instrument (or other source) considered by itself activates only one pair of channels or loudspeakers. This is not beyond stereo in concept or capability, but merely extends stereo to less and less suitable speaker-angles and hearing sectors as we go from front, to back, and to the sides."* - [**Fellgett:1975aa]

In the 1970s, Peter Fellgett, Peter Craven and Michael Gerzon, backed by the UK's National Research & Development Corporation (NRDC) set out to solve this problem. It was determined that the flaw with Quadrophonics was not in the four speaker arrangement approach itself, but rather in its misuse. The true reasons for poor sound localization were in the recording techniques and the discrete channel pair-wise mixing practices[26]. Fixing the problem would require a complete rethinking of the entire system from recording to playback. It would require a robust basis in psychoacoustic principles of how we as humans preceive sound. And it would require forward and backward compatibility so that adding more speakers would only serve to build more parts to the 360° sound field. It was an ambitious goal, made easier by the fact that someone else had already had the idea.

2.2 Stereophonics and Duplex Theory

> ***"Pan-potted stereo and quadraphonics gives only one spatial dimension, Blumlein stereo gives two dimensions (width and distance)."*** – [Gerzon:1976aa]

In 1931, the early days of stereo, Alan Blumlein patented a few microphone and circuit designs for Stereophonics based on the principles of sound perception[27]. One system was even a primative binaural dummy head. One of these designs was particularly important for the development of Ambisonics[28]. Although not named in the patent, the system has come to be known by several names – *mid-side pairing*, *the Blumlein pair*, *Blumlein stereo*, or various

[25] [Elen:2001aa]
[26] [Gerzon:1974aa], [Elen:2001aa]
[27] [Blumlein:1931aa]
[28] [Gerzon:1976aa]

derivations thereof. Important for the fact that it produced an extremely realistic image between 2 speakers, the patent outlines a system that involves two bi-directional microphones, side by side. One microphone is pointed directly toward the sound source and the other tilted 90° on the azimuth angle. Cardioids or omnidirectionals are often used in lieu of the front-facing bi-directional mic, the patent itself is not specific. The front-facing microphone would capture the distance by the direct part of the sound, thus making the *mid* signal. The left and right facing microphone would capture the sound's width, the reflections and reverberations of the sound, thus making the *side* signal. The system was mono compatible in that all one has to do is remove the side signal.

The second part of Blumlein's mid side design was in the matrixing and decoding circuit that would take the microphone outputs and create the appropriate left and right channel feeds for a stereo speaker configuration. Figure 8 shows Blumlein's original matrixing *sum and difference* circuit diagram from his 1931 patent, this author's notes illuminating the process are in blue. Blumlein's circuit uses transformers (indicated by the winding coils) to induce the voltages of the two signals. Reversing the leads on a transformer will reverse the polarity of the voltage[29]. The output of the above circuit is the following: mid plus side equals left, mid plus polarity-reversed side equals right. Today, such a configuration is simple to replicate in most DAWs and requires little knowledge of electronics to do so.

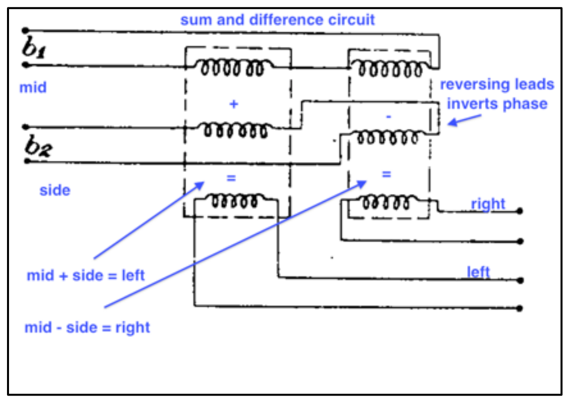


Figure 8. [Blumlein:1931aa]. p. 93. Sum and difference circuit. © 1958 AES, used with permission.

> ***"It can be shown however that phase differences necessary at the ears for low frequency directional sensation are not produced by phase differences at two loudspeakers (both of which communicate with both ears) but are produced by intensity differences at the speakers"***

29 [Gates:2011aa]

- [Blumlein:1931aa]

Blumlein knew that by placing the two microphones on the same axis, perpendicular to the sound source, the direct part of the sound wave, i.e. the first crest of the sound wave, would arrive at both microphones simultaneously. Thus, there would be no phase differences between the direct part of the two signals. What this meant is that his circuit would only have to account for magnitude differences. This is easily done by multiplying the mid and side signals by certain amounts, or gain factors. Specific gain coefficients are not listed in Blumlein's 1931 patent, but many producers and engineers just use their best judgement when mixing these signals together based on the desired sound, making it an artistic judgement call. Because the signals are phase matched, the two microphones are considered a *coincident* point in space, and thus the array is often considered to be one microphone.

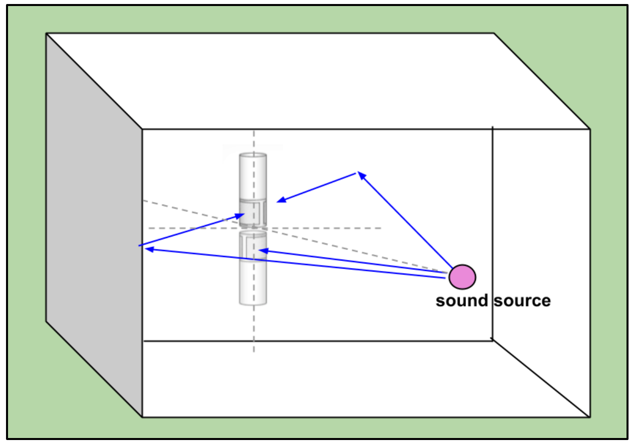


Figure 9. Mid-side paring with direct and reverberant sound paths in blue. X, Y, and Z axes indicated by the dotted grey lines. Sound source in pink.

The true genius of this design was in Blumlein's discovery that these *magnitude differences alone* between the two channels gave the perception of both *interaural level differences* (ILDs) in high frequencies and *interaural time differences* (ITDs) in low frequencies. At the turn of the 20th century, English physicist John William Strutt (Lord Rayleigh) proposed a two part theory on how humans perceive and localize sound, a

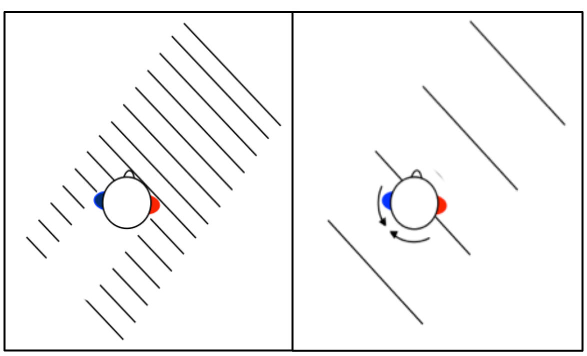


Figure 10. Unknown source. Duplex Theory. High frequencies, shorter in wavelength, on left, show head acting as a baffle. Low frequency longer wavelengths on right show sound wrapping around the head.

theory that is still recognized today[30]. At high frequencies, frequencies shorter than the diameter of the listener's head, the head acts as a baffle, blocking spectral content. So, location of high frequency sound is perceived by the difference in sound intensity between the ears, in what is known as *interaural level differences*. So, if the sound at my left ear is louder than at my right, then the sound is coming from my left side. If the frequencies are low, then the sound waves are longer than the diameter of the humn head, and these frequencies will wrap around the listener's head. Because of the length of longer waves, air pressure differences between the ears will be minimal. So the ears need more information than just the intensity difference to localize low frequency sound. In this case, the ears also use *interaural time difference*; i.e. if the sound arrives at my left ear first, the sound is coming from my left side. The mid-side pairing technique captures both of these spectral cues, placing the sound source at a location in space. The basic problem with early Quadraphonics, and its discrete channel mixing practices, is that it did not take these psycoacoustic principles into account. Duplex theory was observed by Blumlein in his 1931 patents, and by Gerzon when expanding upon Blumlein's ideas into the full 3D sound sphere.

2.3 Ambisonics and Tackling Precise Coincidence

> ***"In obtaining a complete directional "sound picture", i.e. both horizontal and vertical directional effects, the invention is not limited solely to the use of two microphones. A plurality may be employed and their outputs suitably collected, modified and separated to transmit suitable differences of impulses to a plurality of loud speakers. … Full three-dimensional location of the source is thus obtained by this arrangement."*** *- [Blumlein:1931aa]*

[30] [macCába:2002aa]

What Blumlein may not have realized when making the above statement, is that just adding new microphones capsules meant that eventually different axes would need to be accounted for. To demonstrate, let us start with just horizontal soundfield surround sound (figures 11 and 12), achieved with two bidirectional microphones and an omnidirectional microphone. Here, coincidence could be achieved in the same way as it is with a Blumlein pair: by placing the acoustic centers of the three microphones on a line perpendicular to the sound source. Just as it is in Blumlein stereo, the direct part of the sound signal would arrive at all three microphones at the same time. All that would need to be dealt with is the amplitude differences between the capsules; this is what is known as horizontal Ambisonics, and it works quite well.

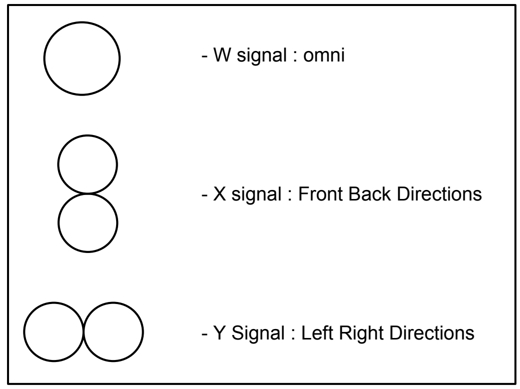


Figure 11. The microphone polar patterns of horizontal surround sound.

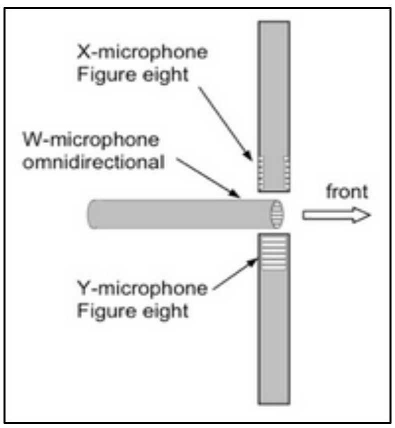


Figure 12. [Fulwider:2009aa]. p. 6. Microphone positioning of horizontal Ambisonics. Used by permission from author.

However, once we expand Ambisonics into the 3rd dimension, adding *Periphony*, as Gerzon called it, an off axis microphone must be added. Upon adding height, at the very least a third bi-directional microphone facing up and down, a Z signal if you will, is needed. But, by doing so, we are adding another axis, and this means that now in order to keep all capsules at a "coincident point", the capsules must be perpendicularly to each other on 2 axes. This configuration is called B format, and unfortunately, it is physically impossible to create. However, Gerzon knew that there was more than one way to achieve this goal.

2.3.1 Spherical Harmonics

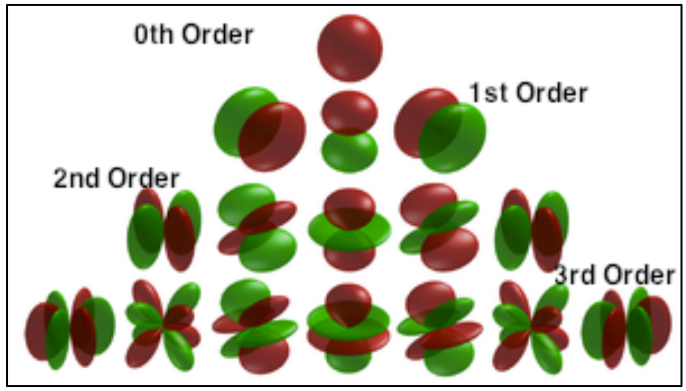


Figure 13. Spherical Harmonics up to 3rd Order.
Attribution: I, Sarxos. Used by permission under the Creative Commons Attribution-Share Alike 3.0 Unported license: http://creativecommons.org/licenses/by-sa/3.0/deed.en.

The approach that full sphere Ambisonics takes is through the use of *spherical harmonics* (figure 13). Just as a complex waveform can be broken down into a number of simple sinusoids, a spatial area within a sphere can be expressed as the sum of its spherical harmonics[31]. So, just as we can apply a Fourier transform in the time domain to get the resulting frequencies, we can apply a similar transform in the space domain, and get the resulting lobes of space that make up our 3D spherical area. This relates directly to microphone polar patterns, where a 0th order spherical harmonic is an omnidirectional microphone, and the 1st order spherical harmonics can be achieved by one bi-directional microphone creating a Figure-8 pick up pattern (figure 14).

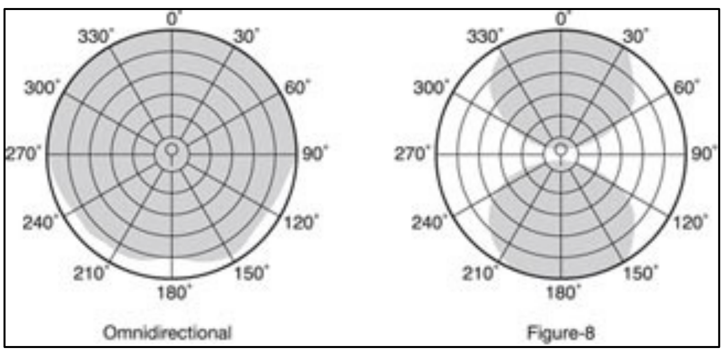


Figure 14. Unknown Source. Polar patterns of omni and bidirectional microphones.

In order to achieve spherical harmonics in microphone design, the capsules must be located on harmonically-related points on the surface of a sphere (figure 15). The points are considered harmonically related if they are the same distance from each other and the same distance to the center of the sphere. Mathematically this is known as an *isotropic* relationship, and creates 3D polygons. By placing the microphone capsules on the surface of a sphere at these isotropic points, we can accurately capture more and more parts of the soundfield.

[31] [Gerzon:1975ac]

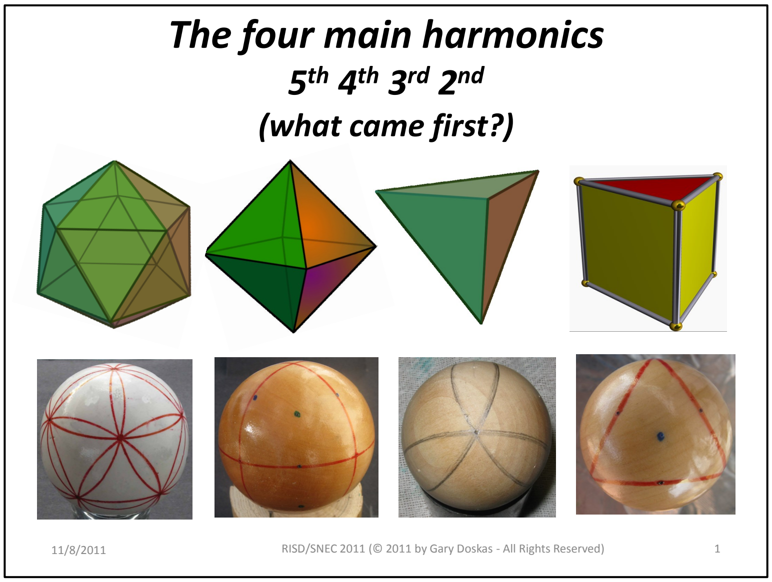


Figure 15. Shows harmonically related points on the surface of a sphere, and the polygons they are related to. Image © 2011 by Gary Doskas - All Rights Reserved from "Spherical Harmony" a presentation by Gary Doskas delivered at the Synergetics Collaborative's Third Biennial Design Science Symposium at Rhode Island School of Design 2011. Used with permission.

2.3.2 Phase Mismatch, a Limiting Frequency, and Why Some Say Ambisonics Does Not Work

Because the capsules are not on a line perpendicular to the sound source, each capsule would possess a different phase for the same frequency, meaning that upon playback the listener would hear comb filtering issues and have an unstable sound image. However, because the capsules were placed on harmonically related points on the surface of a sphere, this problem could be easily fixed. In his 1975 article, *The Design of Precisely Coincident Microphone Arrays for Stereo and Surround Sound*, Gerzon gives a mathematical proof that shows how to do so:

$$f(x,y,z)\ \exp\{+j(ux+vy+wz)\omega/c\}$$

(x,y,z) are direction cosines of direction from which sound arrives
c is the speed of sound (340 m/s)
coodrinates (u,v,w) are sphere's nominal center
ω is the angular frequency
$f(x,y,z)$ is the time delay due to extra distance

The equation makes the connection between isotropic points and phase matching. The isotropic points on a sphere can represent a coincident location in space that is the center of those isotropic points, through the use of complex points on the unit circle. All that is needed to account for the phase mismatch between spherically harmonic points on the surface of a sphere is the radius to the shared center of the sphere. Thus, one can approximate the phase that a frequency should be when it arrives at the true coincident location that is the center of the sphere, and by doing so flatten out the frequency response. Furthermore, one can design a filter in the circuitry of the microphone itself that keeps out those unstable frequencies. The equation creates a physical model of what the frequencies should be. It is how Gerzon dealt with phase mismatch between capsules in his Ambisonic microphone, and it is how the Ambisonic microphone became the next step in coincident microphone recording techniques after the Blumlein pair.

The second major discovery in Gerzon's article, which his equation hints at, was that this system only worked up to what he called a *limiting frequency*. Above this limit, spatial aliasing will occur. So, Gerzon's microphone only works at frequencies below that limiting frequency. Gerzon uses the following equation to determine limiting frequency:

$$F \cong \frac{c}{\pi * r}$$

Where F is limiting frequency measured in kHz, r is the radius of the sphere, and c is the speed of sound in the medium through which the sound travels, which in most applicable cases, is air (343.20 m/s). Based on the spatial sampling theorem, we can modify this equation to find the limiting frequency.

> ***"… in order to reconstruct a monochromatic signal from a set of spatial samples (i.e., with uniform sampling occurring along one spatial dimension), the sampling period must be equal to less than half of the wavelength corresponding to the monochromatic wave."***
> **-** [Dmochowski:2009aa]

As is confirmed in the above quote, and as was established in the preface, spatial aliasing will occur at wavelengths shorter than the distance between the two capsules. So, there is an optimal distance between the 2 capsules for every frequency, that is the wavelength of that frequency. Thus, we can determine the capsule-to-capsule distance by using the equation that gives us a frequency's wavelength:

$$\lambda \cong \frac{c}{F}$$

where λ is the wavelength of frequency *F*, which is the distance traveled by 1 full period of that frequency. This distance must be the length between two capsules at that frequency.

With the above equations, let us now determine what the distance between the capsules must be if the limiting frequency is 20 kHz (i.e. the full range of human hearing), and the medium is 68° dry air. In those conditions the distance between the capsules must be:

$$\lambda \cong \frac{34{,}320}{20{,}000}$$
$$\lambda \cong 1.716cm \cong 17.16mm$$

And, with Gerzon's equation, the radius of the sphere must be the following:

$$20{,}000 \cong \frac{34320}{3.14 * r}$$
$$20{,}000(3.14 * r) \cong 34320$$
$$62800 * r \cong 34320$$
$$r \cong 0.546cm \cong 5.46mm$$

Gerzon notes it to be 5 mm (0.5 cm) or smaller[32]. At this time, such a radius and capule-to-capsule distance is technologically impossible to achieve, due to the physical size of the microphone capsules. Furthermore, the overall goal of Ambisonics is to add more capsules at the harmonically related points to get better 3D spatial resolution. But, in doing so the distance

[32] [Gerzon:1975ac]

between these points must become smaller and small if the radius is to remain the same. The only solution for more harmonics is a larger radius, which lowers the limiting frequency.

The simplest solution to this problem was to have fewer capsules, and sacrifice the spatial resolution, but gain a broader overall frequency span. Thus, the four cardioid tetrahedral array was developed. In this design, the capsules are placed 45° off the X, Y or Z axes and 109.5° apart from each other. This was the first Ambisonic microphone prototype, and it is known as the *soundfield* microphone, seen here in figure 16. At the time of its invention, the capsules were placed 1.47 cm from the sphere's center. According to the spatial sampling theorum, this meant that high frequency spatial aliasing occurred above 7.5 kHz:

Figure 16. The soundfield microphone. Used by permission from SoundField ©. http://www.soundfield.com/soundfield/soundfield.php

$$F \cong \frac{34320}{3.14 * 1.47}$$
$$F \cong 7435.33Hz$$

Even in the most modern A format microphones, such as Core Sound's Tetra Mic[33] (figure 17), the limiting frequency has not improved much. With a capsule-to-capsule distance of 3.5 cm, the limiting frequency is determined to be:

$$3.5 \cong \frac{34320}{F}$$
$$F \cong \frac{34320}{3.5} \cong 9805.7Hz$$

This is only slightly better than the original soundfield microphone. It seems that the best we can physically accomplish is to only capture half the range of human hearing correctly. The only reason this is commercially acceptable is

Figure 17. [Core:2012aa] Core Sound's Tetrahedral Microphone. Used by permission from Core Sound ©.

[33] [Core:2012aa]

because frequencies higher than that only make up a small percentage of the sounds we hear.

2.3.3 How to Deal with the Spatial Aliasing Problem

> ***"Fortunately the common errors at high frequencies are less audible than one might expect. The ear ignores some incorrect information if the majority of frequencies from a source are clear"*** – [Griesinger:1987aa]

Griesinger observed that, unlike temporal aliasing, spatial aliasing is not as audibly detrimental to the signal. Rather, the result of spatial aliasing is simply an unstable sound image. Griesinger's view in this statement is in line with Dmochowski who told us that our ears do the work of wading through the spatially aliased information. However, as was noted prior, this only true for professional, skilled listeners in the audio industry[34]. Furthermore, it has been concluded that this deterioration in the high frequencies could be one of the major factors that led to the film and music industry choosing discrete mixing practices over Ambisonics[35]. Bamford and Vanderkooy support this claim in thier 1995 conference paper, *Ambisonics for Us*[36], in which they compared the soundfield mic quad speaker system with the 5.1 discrete channel mixing and playback system. They concluded that even though Ambisonics has better sound imaging over 360° and a more stable and wider listening area, the benefits deteriorate in the higher frequencies giving discrete channel mixing the upper hand. And finally, Griesinger himself has also researched the importance of clear high frequency content for sound source localization[37]. So, even though spatial aliasing does not produce a piercing sound effect as it does with temporal aliasing, and even though it is only on rarely heard frequencies in the range of human hearing, the effect has directly resulted in Ambisonic's underuse in the industry. Gerzon never had a solution for spatial aliasing, and unfortunately it seems that it was perhaps necessary.

[34] [Nettingsmeier:2010aa]
[35] [Nettingsmeier:2010aa]
[36] [Bamford:1995aa]
[37] [Griesinger:1987aa]

However, with modern day computing, some new computationally heavy solutions have been attempted. Core Sound uses a filter plug-in designed by David McGriffy known as VV Tetra[38] (figure 18). The details of this are proprietary so it is unclear exactly what the program does.

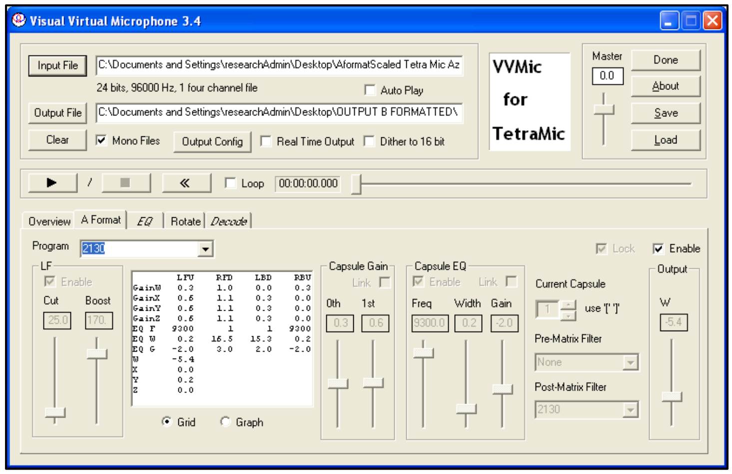


Figure 18. Screen Shot of the VVTetra Filter, showing several functionalities.

However, the process can be surmised from their calibration files and through an examination of their frequency response plots. It seems that instead of using the physical model solution that Gerzon came up with, McGriffy measured impulse responses of the Core Tetra microphone at different azimuth and elevation angles, and then designed inverse corrective filters based on those IRs of each capsule. These filters are in the form of the calibration files that come with the software and flatten out the frequency responses of the capsules. This process works well as long as you know the direction of the incoming signal to match with the proper IR filter at that orientation. Farina[39] developed a similar process in which the signals of the A format mic are pre-filtered in this fashion and then a reference mic is used to calibrate all the microphones to an even flatter frequency response in a post filter. These inverse IR filters as used by both Farina and McGriffy give a finer control over phase than Gerzon's theoretical physical modeling filters. Such "brute force" solutions are only possible in this day and age because of faster computing and larger hard drives. This author has found not documentation on how well this method affects performance in the range above 10 kHz, if it does at all.

[38] [VVAudio:2007aa]
[39] [Farina:2006aa]

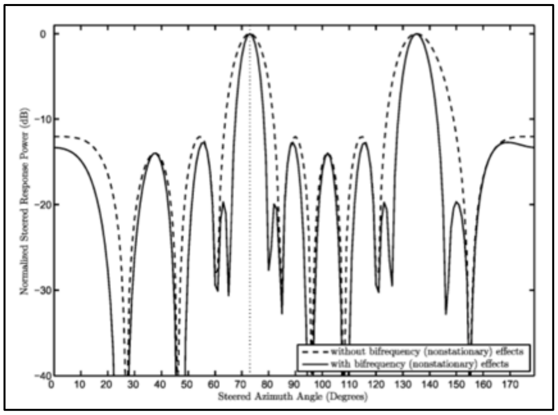


Figure 19. [Dmochowski:2009aa]. p. 1390.
Shows the lobes that the range can be broken into.
Used by permission © 2009 IEEE.

In the 1978 paper, *Processing Spatially Aliased-Arrays*, author Melvin J. Hinich outlined a solution to the spatial aliasing problem, called beamform filterbanks[40]. In frequency domain beamforming, the signal is broken down into frequency bands or lobes, shown in figure 19, and individual delays are set for each band depending on the distance between the capsules. This is also known as *delay and sum filtering*. It is uncertain if Gerzon or if Core Sound ever used this technique, however, the Eigenmike, created by MH Acoustics, does.

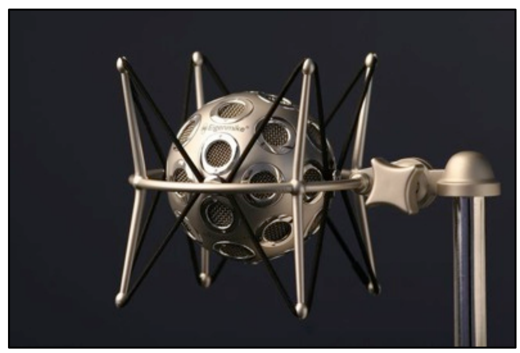


Figure 20. The Eigenmike. The only higher order Ambisonic microphone on the market. Used by permission from MH Acoustics LLC ©. http://www.mhacoustics.com/products.

The Eigenmike, shown in figure 20, is the only commercially available higher order ambisonic microphone on the market. Higher order ambisonics is based on the principle that just as one can achieve higher resolution in the frequency domain with more samples of the time signal, more harmonically related points on the surface of the sphere can give higher spatial resolution[41]. In the 3D soundfield, greater resolution translates to greater direction, greater location of sound sources, and a more stable 3D image. Gerzon theorized, and mathematically proved in detail, the blueprints for higher order ambisonics; but because of the spatial aliasing problem, it was seen as impossible for the time. The way the Eigenmike deals with the spatial aliasing problem is through the use of *adaptive beamforming.* Adaptive

[40] [Hinich:1978aa]
[41] [Gerzon:1973aa], [Gerzon:1978aa], [Gerzon:1980aa], [Gerzon:1992aa]

beamforming not only adds delays to correct spatial aliasing but also gains factors to augment certain frequency bands. This adds greater control over directionality; each individual capsule can be controlled with the aid of the Eigenmike software. With 32 capsules at 41 mm apart and adaptive beamforming, the Eigenmike is reported to capture perfectly coincident 0^{th} and 1^{st} order Ambisonic components across the full range of human hearing and up to 3^{rd} order Ambisonic components with a limiting frequency of 10 kHz[42].

However, as of a 2009 article[43], higher order microphone capsule quality of the Eigenmike was in question. It was reported that high noise floors resulted from all of the capsules added together, which led to the ultimate conclusion that these microphones were not ready for high quality music recording. More recently, Eigenmikes with higher quality capsules have been built by MH Acoustics and have been used for music recording purposes, but again the greater number of high quality capsules must come at a great cost.

It is in this current climate of quality versus quantity that the experiments for this thesis were done.

3 The Experiment

3.1 On the Methodology of the Experiment

The basic premise of this proposed experiment is a comparison of a virtually true coincident array to an actual, what we will call, *physically near-coincident* microphone array. *Physical near coincidence* is the system already established by such microphone arrays as Blumlein Pair, Ambisonic Soundfield, Core Tetra, the Eigenmic and Double MS Z, among others. *Virtual true coincidence* is the system outlined in the goals section of this thesis. It is a

[42] Farina
[43] [Nettingsmeier:2010aa], [Craven:2009aa]

measurement system of taking multiple impulse responses of the capsules of the aformentioned arrays one after the other at the different orientations of said arrays, but in the same location. Because a room's reverberation is a linear and time invariant system, the point of coincidence can be synthesized in post-production simply by putting together the start times of the impulse responses. This is the general methodology of the measurement procedure.

The hypothesis is that this *virtual true coincidence* approach will give a better quality signal in two ways: gain and phase. Signals will be gain-matched because we use the same microphone. There will be better phase correlation in the direct part of the signal because of the coincident location. And there will be better phase de-correlation in reflections and reverberations because of the use of different impulse responses. The measurement process itself was based on the most recent prior work in the field of high quality multichannel surround sound impulse response measurement[44]. For clarification purposes, the two terms - *physical near coincidence* and *virtual true coincidence* will be used to differentiate in the description of the measurement proceedures. See **Appendix A** for pictures of all the measument sessions.

3.2 On the Measured Rooms and Their Acoustic Parameters

Impulse response measurements were made in three available NYU spaces of different reverberant and acoustic qualities: a large performance hall, a medium live room, and a small semi-anechoic listening room. By testing different-sized rooms, we can analyse exactly how reverberance affects coincidence. The dimensions of the rooms were as follows:

	Large Performance Hall – Loewe Theatre	**Medium Live Room – Dolan Studio Live Room**	**Small Semi Anechoic Listening Room – Research Laboratory**
Length	23.97 m	8.79 m	5.89 m
Width	19.03 m	4.65 m	3.44 m
Height	22.03 m	9.05 m	2.53 m

[44] [Farina:2000aa], [Farina:2000ab], [Farina:2003aa], [Ben-Hador:2004aa]

Temp	69° f	69.7° f	69.8° f

Table 1. Room dimensions.

3.3 On the Microphone Arrays

The impulse response measurement process was done with two different Ambisonic microphone array configurations: Core Sound's A format tetra microphone, and Paul Geluso's Double MS Z array[45]. The tetra microphone's arrangement has already been discussed at length. The Double MS Z array is as follows; three Sennheiser MK 800 Twin microphones facing front/back (X), left/right (Y), and up/down (Z). Sennheiser Twin microphones are unique in that they are back-to-back cardioids, with control over both signals separately. The two signals added together give an omni pattern, if the back is phase-inverted you have a figure 8 pattern will null on the sides. Control over both signals separately gives unique advantages in building whatever pick-up pattern is desired[46]. It is possible with this array to matrix the signals to the universal ambisonic B format. No filters have been designed for the array to correct coincident issues. Until the array specifications and design are established, filter design is not possible: however, recent work is being done at NYU to do just that. In the meantime, adjustments can be made in a DAW to match the signals. The Double MS Z array was developed for adding height to classical music recordings[47].

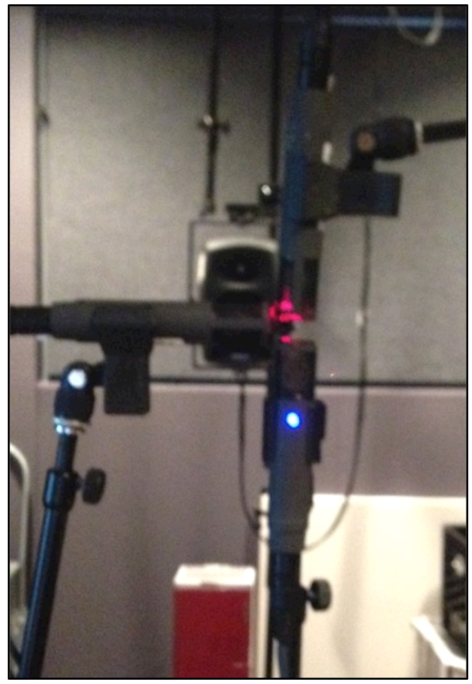


Figure 21. The Doube MS Z Array.

[45] [Geluso:2012aa]
[46] [Sennheiser:2010aa]
[47] [Geluso:2012aa]

3.3.1 On a Microphone's Quality

Much discussion has already been devoted to the term *quality* in reference to microphones. Exactly what that means is discussed here. Microphone quality is genereally defined by a few determining factors: how flat the frequency response is off axis, and how much the gain drops overall off axis[48]. In both regards, there will generally be some drop-off: in the high end of the frequency spectrum there is a genereal roll-off, and the signal as a whole is generally a few dB lower. If that drop is consistent, then the overall frequency response of the microphone is considered smooth. One can look at response curves such as in figure 22, or a polar diagram to see on and off-axis sound together, such as in figure 23.

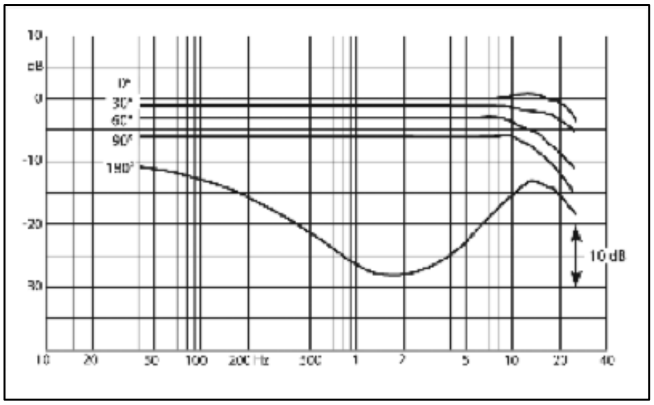


Figure 22. [DPA:2013aa]. The frequency response curves of the DPA Cardioid 4011A, at different on and off axes positions. Used by permission from DPA Microphones ©.

When dealing with microphone arrays you will be dealing with off-axis sound. In the A format microphone, because of the geometric proportions of a tetrahedron, all the microphone capsules are placed 45° off axis. The point is to not place any of the capsules facing forward. The genius of this design is that it sacrifices sound quality a bit overall (caused by off axis sound) but gives a smoother frequency response curve throughout the entire 360° soundfield. Thus, it is important to use high quality microphone capsules in order to have a flatter off-axis response with less of a gain reduction. Core Sound's website does not list off axis frequency response. It is assumed for this experiment, that Core Sound's capsules are of a lower quality than DPA's 4011A.

[48] [Hibbing:1989aa]

As for the Double MS Z array, only the frontward facing microphone captures on-axis sound of the direct part of the signal, the backward, left/right and up/down microphones capture off-axis sound in a variety of directions. The polar diagram of the Sennheiser twin, displayed in their manual and shown here in figure 23, shows good on and 90° off-axis sound[49]. There will however be gain mismatch in the signals because of the on and off axis sound. They do not however list off-axis gain drop.

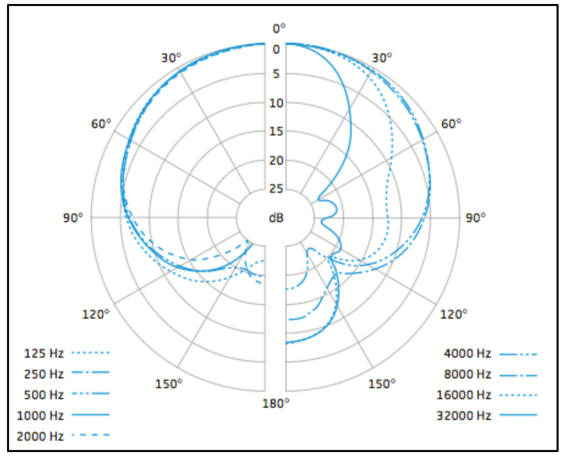


Figure 23. [Sennheiser:2010aa]. Polar diagram of one of the dual cardioids. Used by permission from Sennheiser Electronics Corporation ©.

3.4 On the Sound Source

Just as the receiver needs a flat response, the source does as well. So, the emitted sound's energy level needs to have a flat spectrum across the frequency range. This means that speakers utilized must have a flat response so as to not affect the dynamics or timbre of the excitation signal. The speakers used in this experiment are the Genelec 8030A. Figure 24, taken from the Genelec manual[50], shows a fairly flat frequency response, the power response of the system is shown in the lowest green curve. In Farina's experiments, a speaker with an omnidirectional output was used, which is good for exciting all parts of the room equally. The Genelec is directive in its output (akin to a

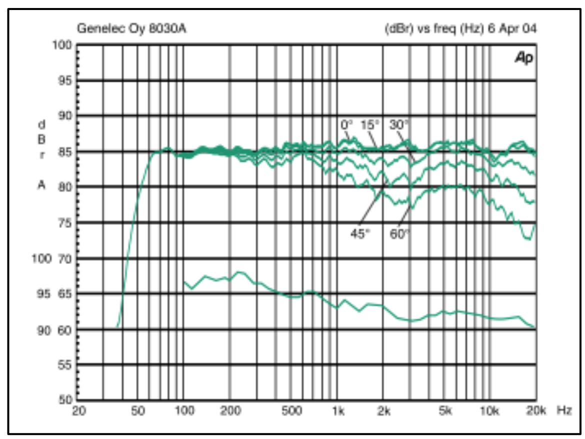


Figure 24. [Genelec:2007aa]. Genelec 8030A frequency response curves and power curve. Used by permission from Genelec ©.

[49] [Sennheiser:2010aa]
[50] [Genelec:2007aa]

cardioid microphone's input pattern). However as Farina notes, such a directive characteristic is more like real world sources[51].

3.5 On the Setup and Configuration

For the excitation signal, only one Genelec speaker was used and it was oriented directly towards the microphone array. This may not be the best configuration for capturing the most reverberant qualities of the room; however this is a study of microphones not room reverberation, so a direct signal is desirable. In the large performance hall and the medium sized live recording room, the space was divided into thirds length-wise: the speaker was placed 1/3rd of the room's length to the front wall, and the microphone array was placed 1/3rd of the rooms length to the back. The large performance hall is a proscenium style theatre that seats three hundred (see figures). The speaker was placed on the stage in the middle front, and the microphone array was placed in the middle of the audience. The floor of the large hall was on a slope. In the small semi anechoic room the microphone array was set in the middle of the room and the speaker emitted from the middle of the front wall. The dimensions were as follows:

	Large Performance Hall – Loewe Theatre	**Medium Live Room – Dolan Studio Live Room**	**Small Semi Anechoic Listening Room – Research Lab**
Speaker to Mic	7.99 m	3.1 m	1.9 m
Speaker to Ground	4.9 m (on stage)	1.19 m	1.41 m
Mic to Ground	1.6 m	1.19 m	1.41 m

Table 2. Configuration dimensions

The interface used for connecting the computer to the speaker and microphones was the Metric Halo ULN-8 I/O console[52]. It acted as the system's I/O routing, provided phantom power for the microphones and monitoring control for the calibration of the microphones. Before any excitation signals were captured, the microphones and the speakers needed to be calibrated. The

[51] [Farina:2003aa]
[52] http://mhsecure.com/metric_halo/products/hardware/uln-8.html

speaker was calibrated with an SPL meter and a pink noise signal to 70 SPL in all 3 rooms. Because each microphone capsule in the tetra and double MS Z arrays has its own unique set of transducers, resistors and other mechanics, the microphone array capsules needed to be calibrated to each other. This was done by sending a pink noise signal through the speaker and placing each microphone capsule at the front one at a time. Then, on the Metric Halo interface the signals were matched to roughly the same peak level, plus or minus a few decibels. In Loewe Theatre (the Large Hall), it was roughly -25 dB, and in the medium sized room (Dolan studio) and the small listening room (the Research Lab), it was roughly -19.8 dB.

3.6 On the Measurement Procedure

In order to create a coincident location in space, two laser pointers, or laser line tools, were used. Then, the process of creating *true virtual coincidence* was moderately simple. First, the correct angles of the A format microphone needed to be found. In order to do so, the actual Tetra mic itself was used as a tool. The Tetra mic was placed with its center at the laser's intersection, shown here in figure 25.

Figure 25. Tetra mic with the two laser pointers meeting at its center.

Then, each angle was found using a laser pointer and an angle locator tool. An angle locator tool can be found at your local hardware store, the one we used was unique in that it had a rounded base that would easily sit on top of the microphone capsules. A laser pointer was glued to the other end of the rounded base, as shown here in figure 26.

Figure 26. Angle locator tool with laser pointer glued to the end pointing at the ground. Rounded base sitting on top on the DPA 4011a microphone.

This created a direct line extending from the capsule to the laser pointer, to a point on the ground or ceiling, which would be marked with a piece of tape, seen here in figure 27.

Figure 27. Tape on floor marking the point at which the laser pointer points to. LBD stands for the capsule Left Back Down that that laser is representing.

In order to reinforce the vector line between the coincident point in space and these new angular points of the capsules, a laser line tool was used, shown in figure 28 on a mic stand.

Figure 28. Laser line tool creating an axis that intersects at the microphone capsule.

The laser line would run from the point on the tape, throught the coincident point in space, and up the shaft of the microphone capsule. This process was done to find all of the tetra mic's angles at 0° and 90° azimuth from the sound source. With these points, the correct orientations could be recreated for the *virtual true coincidence* technique. This process of finding correct angles was easier for the Double MS Z array because the angles of this array are directly on the X, Y, and Z axes.

Once the angles and points were marked, the true virtual coincidence impulse response measurements could begin. First, the face of the capsule would be placed at the location of the laser's intersection, as seen in figure 29.

Figure 29. Lasers intersecting on grill of capsule.

Then with the help of the laser line and the angle locator tool and laser pointer tool, the correct angle would be found. The measurement process would be as follows: find the angle, take IRs, rotate to next orientation and repeat. The process at the A format angles was done with a capsule of the Core Sound Tetra mic, and then with a DPA cardioid 4011A[53]. The DPA cardioid gave at least one professional music studio quality microphone for the A format configuration. The process for the Double MS Z array was much simpler, since the angles of the Double MS Z array are all fairly simple to find. Also, because rotating a Double MS Z array 90° still keeps the microphones at the same orientations, these angles were not taken (figures 30 through 32).

[53] [DPA:2013aa]

Figure 30. Dual Capsule Senheiser Twin capturing front and back

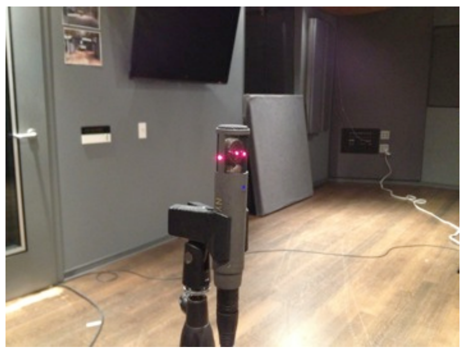


Figure 31. Dual Capsule Senheiser Twin capturing left and right

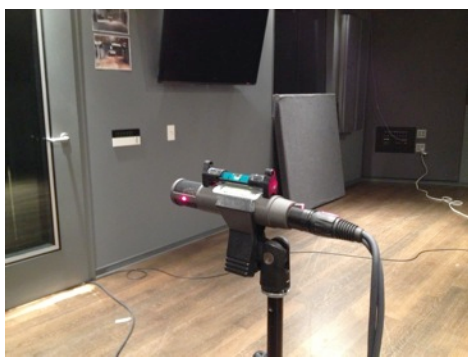


Figure 32. Dual Capsule Sennheiser Twin capturing up and down.

Direct comparisons needed to be made between the *virtual true coincident* measurements and the *physical real coincident* array one which these techniques are based. So impulse responses for these arrays were taken as well. These IRs were taken all at once, instead of one capsule at a time. Table 3 shows all configurations that were measured.

Capsules	Configuration	Axes
Core Sound Tetra Microphone	Physical Near Coincident	Az 0° and Az 90°
Core Sound Tetra Microphone	Virtual True Coincident	Az 0° and Az 90°
DPA 4011 Cardioid – Tetra Angles	Virtual True Coincident	Az 0° and Az 90°
Sennheiser Twins Double MS Z Array	Physical Near Coincident	Az 0°
Sennheiser Twins Double MS Z Array	Virtual True Coincident	Az 0°

Table 3. Microphone Array configurations measured.

3.7 On the Excititation Signal

Recently, much work has been done in the field of multichannel impulse response measurement[54]. Following Farina, the test signals used were exponential sine sweeps (ESS). ESS is a sine wave of a band-limited frequency range that varies exponentially over time. This technique has been recognized as the most successful technique for multichannel spatial IR measurements[55]. It offers low signal-to-noise ratio, and the frequency response can be easily manipulated in post-production. A low signal to noise is of the utmost importance when dealing with multiple capsules, as their noise floors add up.

A long sweep time is considered desirable, since more time equals greater resolution in the frequency domain and reduces the signal to noise ratio. However, Farina found that at a certain point, the longer the sweep time is, the more distorted the IR can become, due to smearing caused by slight air pressure changes[56]. He suggested a compromise length of 15 seconds regardless of the space. This gives a noise floor of 100 dB[57]. Farina also recommended a sampling rate of 96 kHz and a 24-bit rate[58]. Several sweeps were taken for each microphone measurement, and then averaged in post to smooth out any further distortions caused by the speaker and the air pressure variations, especially in the high frequencies[59]. Also, the pressure variations between the capsules caused by the virtual true coincident technique are also

[54] [Guy-Bart:2002aa], [Farina:2000ab], [Kessler:2005aa], [Muller:2001aa], [Ben-Hador:2004aa]
[55] [Kessler:2005aa], [Ben-Hador:2004aa]
[56] [Farina:2003aa]
[57] [Ben-Hador:2004aa]
[58] [Farina:2003aa]
[59] [Farina:2000ab]

smoothed out through the above-proposed technique of averaging multiple responses[60]. There is a five second lag between each test signal to account for the decay of the reverberations in each room. The sweep type was logarithmic. This technique creates a smooth linear response especially in the higher frequencies, which are important for psychoacoustic spatialization. Table 4 shows the important information of the excitation signal.

Starting Frequency	20 Hz
Ending Frequency	20 kHz
Length of Sine Sweep	15 seconds
Sweep Type	Logarithmic
Sampling Rate	96 kHz
Bit Rate	24 bit
# of sweeps averaged	5

Table 4. Characteristics of the excitation signal.

3.8 On the Software and Post-Processing the Captured Signal

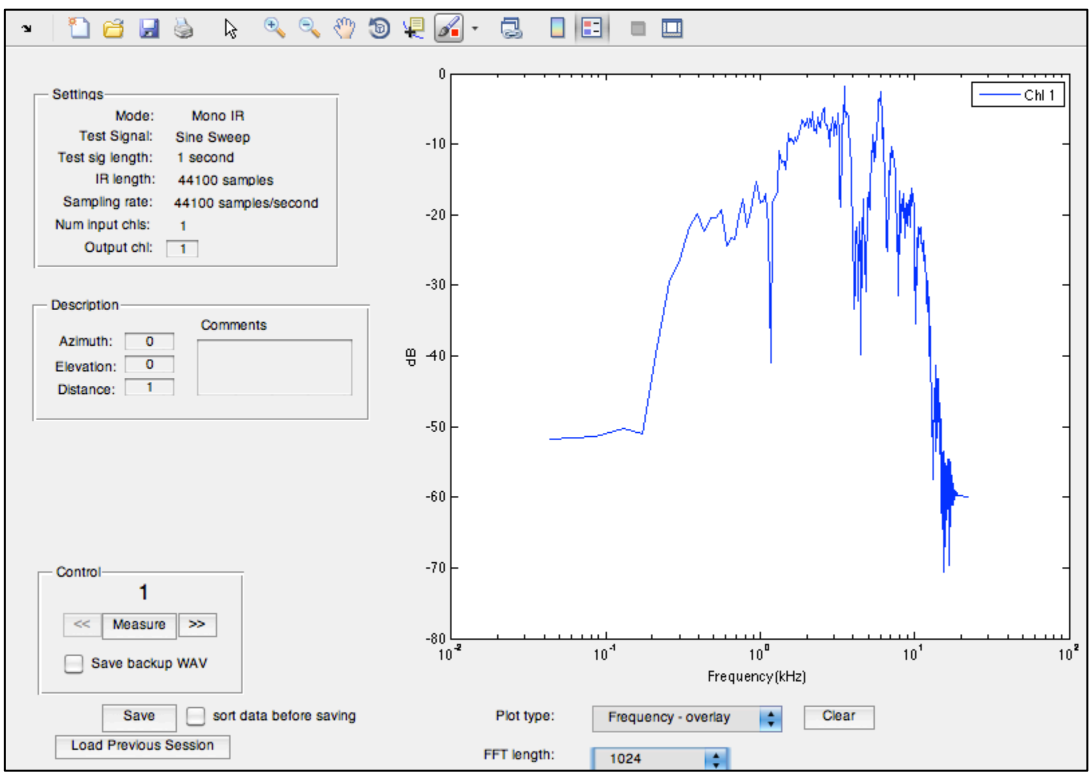


Figure 33. Screen Shot of the ScanIR Utility GUI Software in Matlab.

[60] [Farina:2000ab]

The sine sweeps were generated, captured and processed with the Matlab utility and GUI Scan IR (figure 33), created here at NYU by Braxton Boren[61]. This software was an invaluable time saver as the algorithms for the processes were already written[62]. The following is an outline the process.

After capturing the five logarithmic sweeps, the Scan IR algorithm first normalizes, then averages, the sweeps. This lowers the signal to noise ratio by smoothing out any larger peaks caused by harmonic distortions or strange room reflections and reverberations[63]. Then the Scan IR algorithm de-convolves the response of the system through a linear convolution with an inverse of the sine sweep. The inverse sweep is not an inverse of the recorded signal, but an inverse of the logarithmic sine sweep, as generated by Scan IR. The linear convolution involves the following steps: zero padding the recorded sweep and the reverse sweep, doing a Fast Fourier transforms to the frequency domain (fft), then a linear convolution of two fft-ed signals with a window length that is the length of the sweep, then an inverse fft back to the time domain signals, and finally a removal of the zero pad. The output signal is an impulse response. The signals are now ready for analysis.

4 ANALYSIS

4.1 On the Process of Analysis

The signals were analyzed via three different factors: correlation, cross correlation between each channel, and frequency response content. There were a few stages at which these factors were analyzed. A pre-analysis stage examined the raw output signals from Scan IR. Core Sound's VVTetra Plugin was used for filtering. In this way, a direct comparison could be

[61] [Boren:2011aa]
[62] [Farina:2000ab]
[63] [Farina:2000ab]

done between the *virtual true coincident* technique with the core sound capsules and the actual tetra mic, with their own filtering. These signals are shown in Table 5.

Capsules	Configuration	Format	Axes
Core Sound Tetra Microphone	Physical Near Coincident	A format	Az 0° and Az 90°
Core Sound Tetra Microphone w/ VVTetra Filtering	Physical Near Coincident	A format	Az 0° and Az 90°
Core Sound Tetra Microphone	Virtual True Coincident	A format	Az 0° and Az 90°
DPA 4011 Cardioid – Tetra Angles	Virtual True Coincident	A format	Az 0° and Az 90°
Senheiser Twins Double MS Z Array	Physical Near Coincident	Double MS Z format	Az 0°
Senheiser Twins Double MS Z Array	Virtual True Coincident	Double MS Z format	Az 0°

Table 5. All configurations in which IR measurements were taken.

Then all ten cases (including the 90° off axis cases of the A format) were matrixed to B format for the actual analysis stage. The matrixing for Double MS Z to B format is shown in Table 6.

W =	Front + Back + Left + Right + Up + Down
X =	Front – Back
Y =	Left – Right
Z =	Down – Up

Table 6. Matrixing of Double MS Z channels to B format.

The W channel can just be Front plus Back or Front plus Back plus Left plus Right for horizontal mixing. Or one can also add Up plus Down when dealing with height channels. The VVTetra Mic plug-in did the A to B format matrixing in the cases where it was used. Then, all the B format signals were analyzed on the three factors: correlation, cross correlation between each channel, and frequency response content of all the signals, separately. All programming was done in MatLab, see **Appendix B** for code. Figure 34 is a block diagram of the workflow.

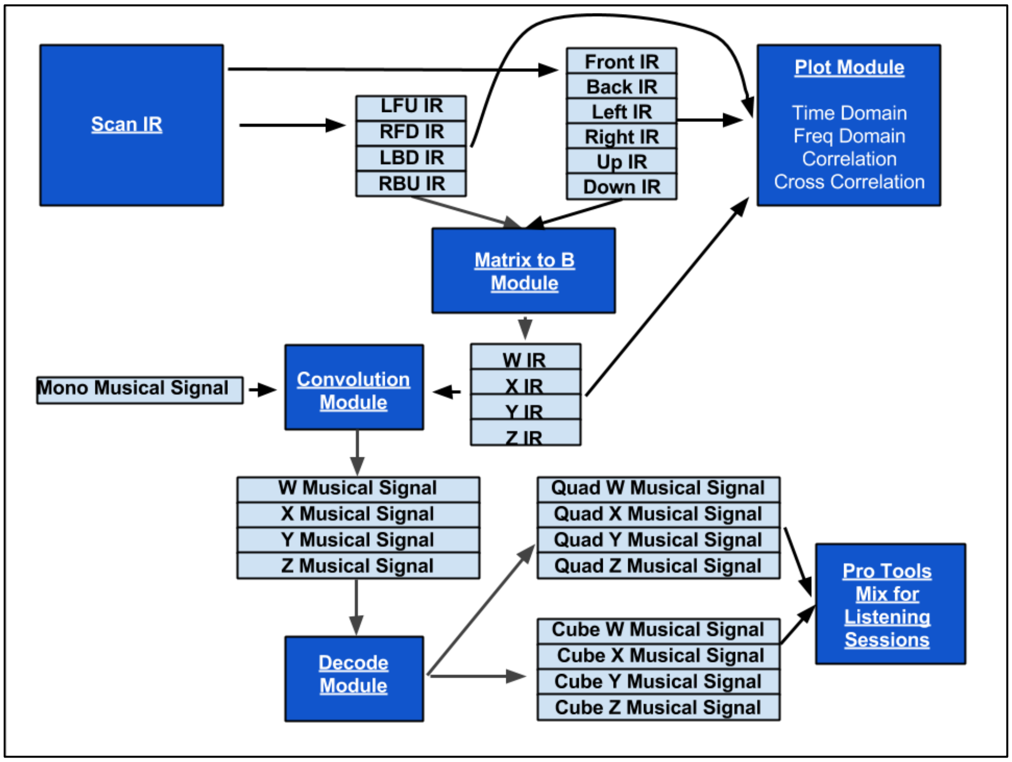


Figure 34. Block diagram of all the programming modules all done in MatLab.

4.2 On the Techique and Methodology of Analysis

The basic technique of the analysis portion was a straightforward search for the areas of greatest correlation in these cases and determining how time lags and more or less correlation related to and affected frequency response content. Adding in the methodology to this was simple as well; psychoacoustics on how the human auditory system precieves sound was seen as a guiding reference for how the frequency responses were expected to be acting at the different orientations. For example if the microphone was facing front, then a more flat spectrum was expected than a microphone facing back from the sound source which usually capture more low end content in the reverberations and reflections. This is just one example of how the analysis process was performed. Certain assumptions were made, some proving to be correct and some incorrect, however, they served as a starting point for the analysis to progress.

Correlation is the degree to which signals vary together over time. It is dependent on the content, meaning that because the W signal contains the X, Y and Z signals in it, W will be highly correlated to those transmission channels. X, Y and Z will generally show less correlation to each other because they are picking up the reflections and reverberations of different parts of the soundfield. Correlation has been studied extensively over the past 40 years, including how it relates to psychoacoustics. It has been generally found that low correlation during late reverberation is desirable as it gives a greater sense of *spaciousness*. Greisinger, writing on the issue of decorrelation, stated:

> ***"Perception of both spaciousness and depth in loudspeaker reproduction depends critically on having low correlation."*** - [Griesinger:1987aa]

So, for a greater sense of spaciousness and sound envelopment, decorrelation is desirable, especially in the low end. Decorrelation has also been established as desirable in the early reflections of a signal. Early lateral reflections give distance, and early medial reflections provide loudness[64].

> ***"Early lateral reflections have been investigated and related to the perception of spaciousness. A strong single early lateral reflection added to a direct sound creates a change in the perception of spaciousness."*** – [Conceicao:2013aa]

However, in the case of our closely spaced microphone arrays there is a point where correlation is desirable in surround sound. Correlation is frequency-dependent. For instance, low frequencies that have a longer wavelength vary more slowly over time (and space). And higher frequencies vary more quickly. So, if we think of our spaced microphone capsules, then we can assume there will be greater correlation between the signals in the low end than there is in the high. This is fine for reverb and reflections, but in the direct part of the signal, we, of course, get spatial aliasing in the high frequencies. In this study it is assumed that higher correlation is desirable during the direct part of the signals.

[64] [Griesinger:1999aa]

See **Appendix C** for B format cross correlation plots. See **Appendix D** for B format correlation data sets. And see **Appendix E** for B format frequency spectral plots. The A format plots and the Double MS Z format data sets and plots were left out of the Appendices, due to space limitations and since the B format transmission channels are the most crucial comparisons. Any important points on A format or Double MS Z format will show plots and data in this section of the document. All plots were made in Matlab.

4.3 Correlation Analysis

Using Matlab, the correlation data of the *virtual true coincidence* techinque was directly compared to that of the *physical near coincidence* of the Soundfield and the Double MS Z arrays. Correlation data was gathered for both the direct part of the signal and the full signal. First, it was assumed that the *virtual true coincidence* technique would achieve higher correlation in the direct part of the signal than the *physical near coincidence* technique. Second, it was hoped that there would be equal or greater decorrelation in the early reflections and late reverb of the signal than that seen in *physical near coincidence*. Correlation is measured on a scale from ±1 to 0, where ±1 is maximum correlation, and 0 is minimum correlation. Negative correlation occurs when a positive and negative peak are correlated, and positive correlation when two positive or two negative peaks are correlated. Low correlation (0) generally corresponds to a high peak in one signal and a null crossing in the other.

In a direct comparison between the *virtual true coincidence* with Core Tetra's capsules and *physical near coincidence* with Core Tetra's array, the *virtual true coincidence* correlation of the direct signals was shown to perform moderately better than of that of *physical near coincidence*. However, when converting to B format, those signals that were once strongly correlated due to the *virtual true coincidence* technique became less correlated in the direct part of the signals.

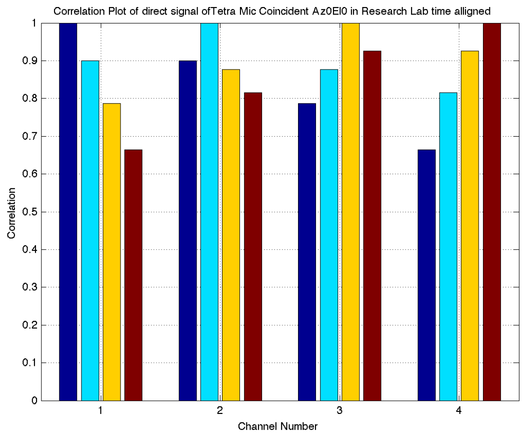


Figure 35. A format Core's Tetra Virtual Coincidence Correlation Bar Plots. Channel 1 is in dark blue, ch 2 is in light blue, ch 3 in yellow, and ch 4 in red. So, for example the first set of 4 bars from left to right are the correlation between channels 1 and 1, 1 and 2, etc.

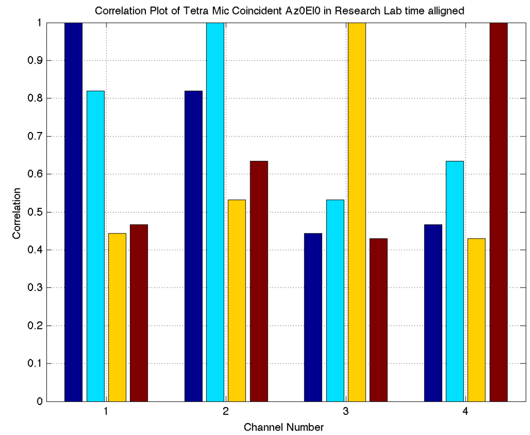


Figure 36. B format Core's Tetra Virtual Coincidence Correlation Bar Plots. Channel 1 is in dark blue, ch 2 is in light blue, ch 3 in yellow, and ch 4 in red. So, for example the first set of 4 bars from left to right are the correlation between channels 1 and 1, 1 and 2, etc.

Above, in figure 35, we have the A format of the Tetra Microphone at virtual coincidence, showing high correlation across the board. In figure 36, the signals have been matrixed to the transmission channels, the B format signals, showing lower correlation between channels. This is true across the board of all A to B format signals. However, with the DPA 4011A cardioid capsules at the A format angles, correlation is higher in both A and B format, thus it is noted that the decorrelation occuring during the matrixng process is not as destructive.

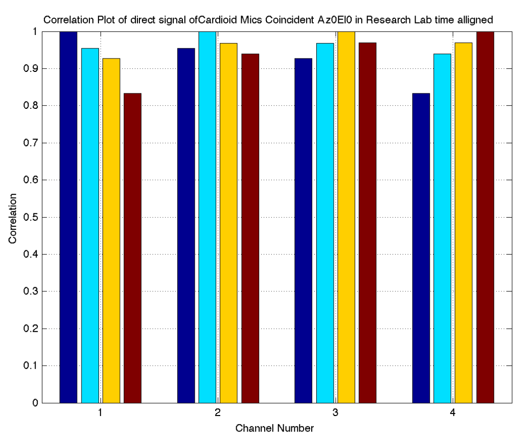


Figure 37. A format DPA 4011s at Tetra Angles Virtual Coincidence Correlation Bar Plots. Channel 1 is in dark blue, ch 2 is in light blue, ch 3 in yellow, and ch 4 in red. So, for example the first set of 4 bars from left to right are the correlation between channels 1 and 1, 1 and 2, etc.

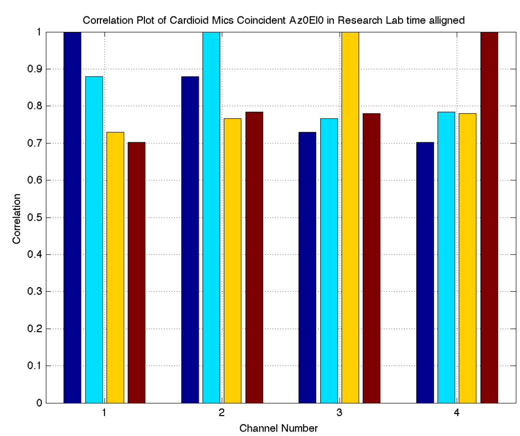


Figure 38. B format DPA 4011s at Tetra Angles Virtual Coincidence Correlation Bar Plots. Channel 1 is in dark blue, ch 2 is in light blue, ch 3 in yellow, and ch 4 in red. So, for example the first set of 4 bars from left to right are the correlation between channels 1 and 1, 1 and 2, etc.

As we see in figures 37 and 38, the correlation in the direct part of the signal is as good as it was for the A format of the tetra capsules. This even holds true at 90° off-axis, and even in larger, more reverberant spaces.

The drawback to the DPA cardioid however, is that it offers mid-level, or "good", correlation (figure 39) in the reflective and reverberant parts of the signal, precicely where we want low correlation.

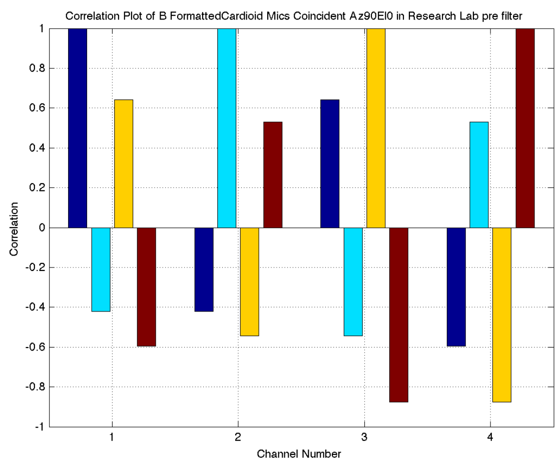


Figure 39. B format of DPA Cardioids at Virtual True Coincident location. Correlation Bar plots just of reverberation, shows "good" or mid level correlation, where we want to be seeing low to no correlation.

It seems that the *physical near coincident* array of the Core Tetra mic offers a blend of correlated and uncorrelated signals. The best of both worlds however is the Double MS Z array, offering extremely high correlation in the direct parts of the signals (figure 40) and extremely low correlation in the reflections and reverberations (figure 41).

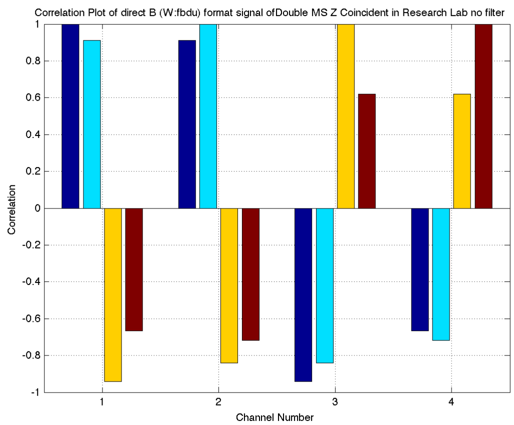


Figure 40. Direct part of the signal of the coincident Double MS Z in B Format in Research Lab.

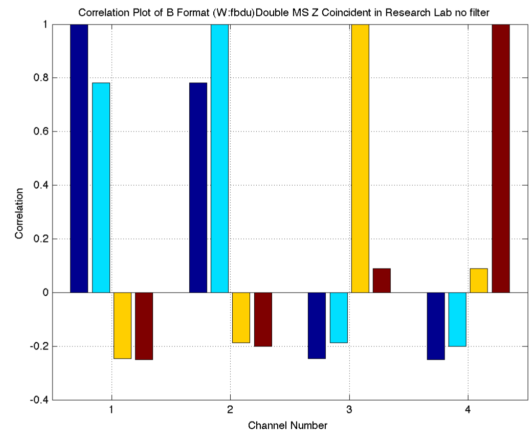


Figure 41. Reflections and Reverb of the signal of the coincident Double MS Z in B Format in Research Lab.

4.4 Cross Correlation Analysis

Cross correlation is the measurement of the correlation between signals as a function of a time lag. Simply put, it answers the question how many samples apart in time must the two signals be for maximum correlation between those two signals to occur? For example, if the cross correlation between two channels of the A format IR returns a value of 10 samples, then if we shifted one signal 10 samples up, the two signals would then be aligned so that the area of best correlation happens simultaneously. In the space domain, 10 samples of a 96 kHz sampling rate signal is 3.575 cm, or the distance between two adjacent capsules of Core's Tetra Mic.

In this experiment, it was hoped that the cross correlation between the signals at the same coincident location would be 0 samples apart. However, as realized during preliminary analysis, this was not always the case. First, getting the exact location of the coincident point in space is actually quite difficult when you are dealing with millimeters and the A format's angles. But, the more important point is that finding the true acoustic center of the microphone's capsule is very difficult as well. In fact, the true acoustic center of the Core Sound's Tetra Microphone is three to four mm below the face of the grid. This is also true for the DPA 4011A cardioid and the Sennheiser twin. The advantage of the twin microphone is that its grid can be seen through, so the diaphragm and the true acoustic center of the microphone is visible. Thus, virtual coincidence was easiest to achieve when using the Double MS Z twin microphones, because it was easier to find the coincident point, and also because the 90° angles were simpler to find.

Another factor that could have had something to do with the unexpected time lag in the *virtual true coincident* cases is the size of the capsule's diaphragm itself, i.e. the capsule diameter. As noted, Core Sound's tetra capsules showed the greatest time lag in the *virtual true coincidence* cases, these capsules have a 12 mm diameter[65], which is generally categorized as a small diaphragm[66]. The DPA 4011a[67] cardioid and the Sennheiser twin[68], which were less likely to have any time lag present in the *virtual true coincident* cases, both have a diameter of 19 mm, making them fall under the medium to large diaphragm size. It is generally understood that smaller diaphragm microphones are thought to be more accurate, and especially because we are dealing with a scale of a few millimeters, accuracy is crucial. It is hypothezied that the DPA 4011a and Sennheiser twin yielded a smaller time lag precisely because of their larger diaphragms which produced a less accurate signal.

Regardless of the reason for why the time lags occurred, this mean that methodological choice had to be made on what to do with them, whether or not to time-align these signals in post-production. In many cases the time lag was equal to, or greater than, the time lag caused by the distance between the capsules in the *physical near coincidence* cases. This was not true for all cases in all the rooms, for example there was one DPA A format cardioid case where *virtual*

[65] [Core:2012aa]
[66] [DPA:2013aa]
[67] [DPA:2013aa]
[68] [Sennheiser:2010aa]

true coincidence was in fact achieved through the measurement process itself. However, this one result could be considered an outlier. It is noted that even with the time lag in the *virtual true coincidence* system, there were some cases where the correlation data was better in the *physical near coincident* cases. However, this was not consistent. It was determined that because the proposed technique of this thesis work is in testing a coincident location in space, then it must be followed to the logical conclusion, that time aligning the signals in post production is necessary. Time aligning these signals happened before matrixing the signals to B format.

Remarkably, the *physical near coincident* Double MS Z array signals were incredibly well time aligned before any post-processing - so much so that the *virtual true coincidence* technique seemed not much different. This is understandable since the three microphones were placed on the same axis, perpendicular to the sound source. So, just as Blumlein discovered, the time of arrival will be the same. Thus, the three microphones of the Double MS Z are *already phase matched* regardless of which method *virtaul true coincidence* or *physical near coincidence,* is used. And, just like the Blumlein pair, all that needs to be adjusted for in post is gain mismatch.

However, upon matrixing to B format, a mysterious artifact showed up in the cross correlation plots. Time lags began to appear in the Double MS Z *physical near coincidence* array signal's cross correlation plots. They ranged from 10 to 100 samples. At first it was thought there was an issue with the algorithm used to plot the signals. However, upon re-examining which channels were lagging, and examining the

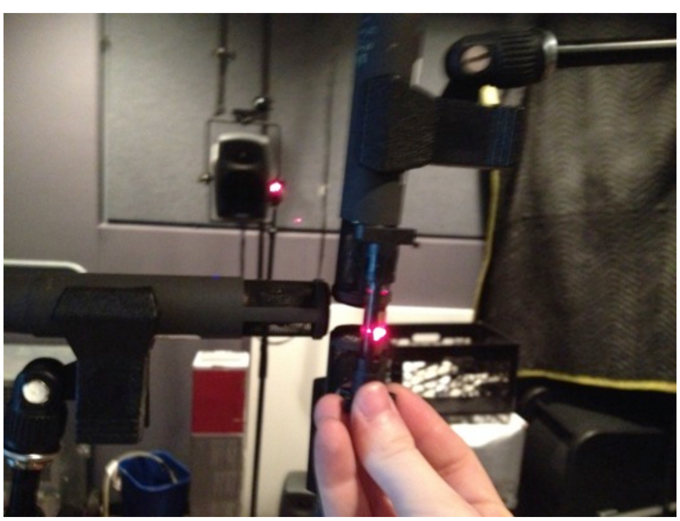


Figure 42. Closeup of Double MS Z array in the Research Lab.

pictures of the measurement sessions, an interesting realization occured. Figure 42 is a typical arrangement of the Double MS Z array, with the three microphones all on the same perpendicular axis to the sound source. In this arrangement, it is necessary to place one microphone off axis from the others. As is evident in figure 43, the Z microphone (down/up) is positioned on the y axis, and the X and Y microphones on the X axis. The Z channel is precicely the channel where see the time lag in the cross correlation plots, figure 44.

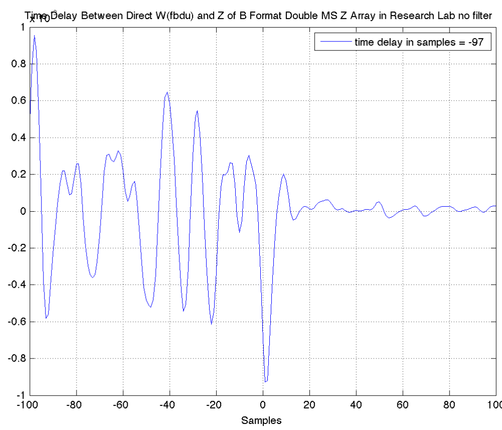


Figure 43. Double MS Z array in the Research Lab. Cross Correlation plot between W and Z channel. 97 sample lag.

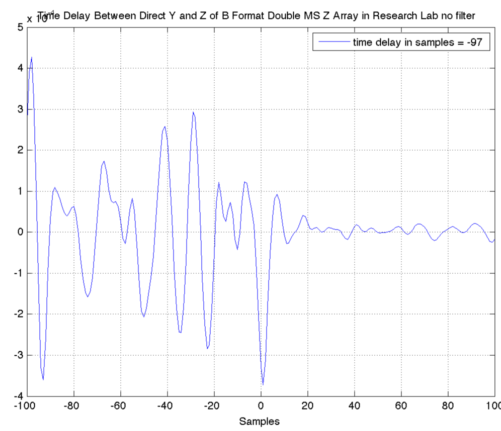


Figure 44. Double MS Z array in the Research Lab. Cross Correlation plot between Y and Z channel. 97 sample lag.

It is hypothezed that the lag forms because of the fact that the height channel is off axis from the other two capsules. Psychoacoustically, this should cause strong early *medial reflections* towards the middle of the sound stage which is reported to give a sense of depth[69]. Lateral reflections are reported to cause a sense space because they are akin to relections coming from the sides. The Double MS Z array is unique in that you can select which microphone is the off axis one, thus giving either a sense of depth with an off-axis Z microphone, or a sense of space with an off-axis Y microphone.

As noted prior, the matrixing from A to B caused a little bit of time lag in the A format cases as well as the Double MS Z cases. It became aparent that how far that time lag was was

[69] [Griesinger:1999aa]

determined by how far apart the capsules were, not just on the perpendicular axis to the sound source, but on the other axes as well. With the *virtual true coincident* technique cases, the time lag was by only a few samples. Larger lags were seen in the *physical near coincident* A format tetra mic array. And the largest was noticed in the Double MS Z array. Sending the Tetra Mic A format signals through their software, the VVTetra filter, would smooth out the time lags a bit on the Tetra array signals: not all the way to 0, but more so than without.

Another phenomenon of note can be seen in the cross correlation plots of the B format DPA 4011A capsules (figure 48) compared to all other A format to B format signals (figures 45 through 47). The B format signals of the DPAs show no questionable side peaks and very strong correlation at 0 or 1 lag. Core's Tetra mic always showed definite side peaks competing with the index of strongest correlation, regradless of *physical near coincident* or *virtual true coincident* cases. This was an interesting and consistent phenomenon across each of the rooms and at 0 and 90°. It is hypothesized that this most likely has to do with the quality of the capsules. Because the DPA 4011As has a strong off axis frequency response curve, more so than the Core Tetra's, it is hypothesized that this led to a strong correlation without any questionable distortions.

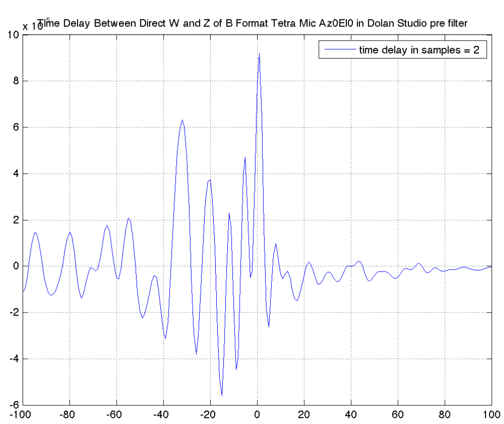


Figure 45. Tetra Mic physical near coincidence array in Dolan Studio. Cross Correlation plot between W and Z channels.

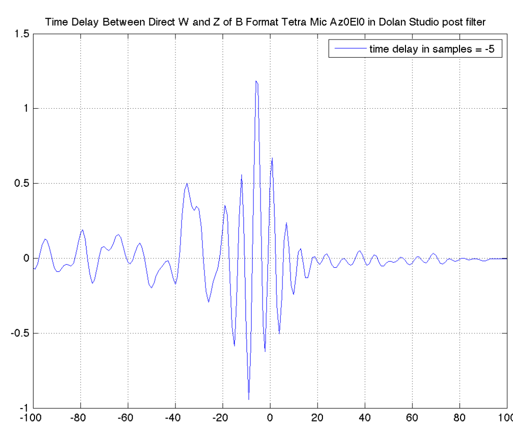


Figure 46. Tetra Mic physical near coincident array in Dolan Studio. Cross Correlation plot between W and Z channels, with VVTetra Filter.

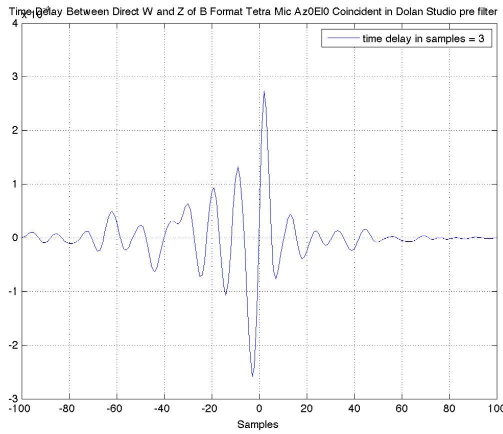


Figure 47. Tetra Mic virtual true coincidence array in Dolan Studio. Cross Correlation plot between W and Z channels.

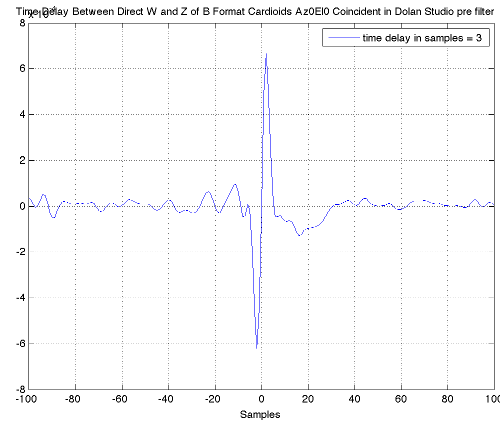


Figure 48. DPA 4011a's virtual true coincidence array in Dolan Studio. Cross Correlation plot between W and Z channels.

4.5 Freqeuncy Response Analysis

Finally, the direct part of the signals was analyzed in the frequency domain. The signals with the most consistently smooth response all around were the *physical true coincident* DPA 4011s (figure 49) and the *physical true coincident* Core Tetra capsules (figure 50). The reason for this smooth response is because the DPAs and the Core Tetra mic capsules are all set at 45° angles off axis. Such a design aids in creating a smooth frequency response in surround. It was hoped that this would result in a high performance in the professional listening tests.

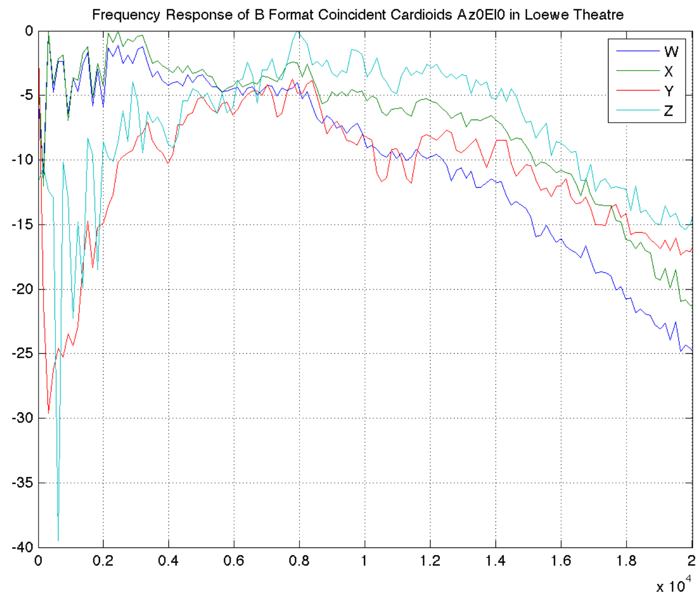


Figure 49. B format frequency responses of the DPA Cardioid *Virtual True Coincident* signals in Loewe Theatre. Notable for the similarly smooth sloping higher frequencies above 1 kHz.

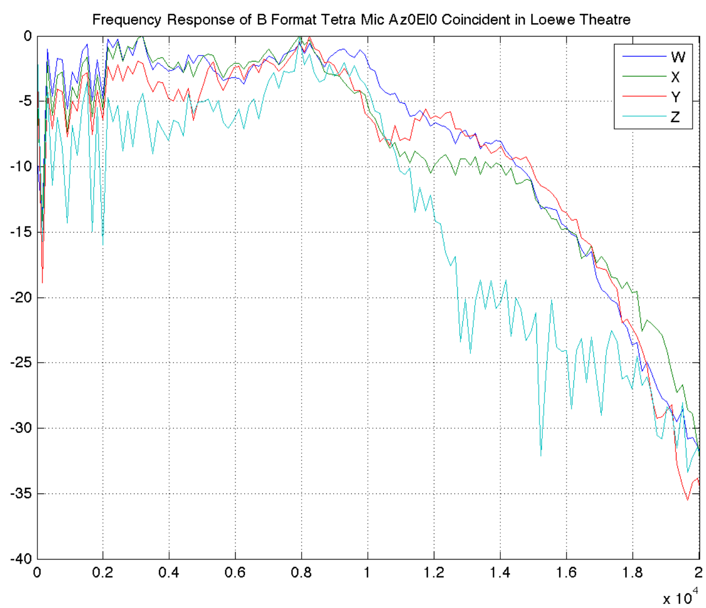


Figure 50. B format frequency responses of the Core Tetra Mic in *Virtual True Coincidence* signals in Loewe Theatre. Also has smooth slopes in the higher frequencies above 1 kHz.

The *physical near coincident* Core Tetra array itself caused a peculiar dip at the 1 kHz range of the Y signal. This is notible for the fact that 1 kHz is exactly where spatial aliasing starts in the Core Tetra microphone. The dip is consistent in all rooms, at all axes, and it is not present in the *virtual true coincident* signals of the Core Tetra capsules. Thus, it is assumed that the array itself is causing this strange artifact in the direct frequency response. This artifact is not corrected in the VV Tetra filter either (see figure 51).

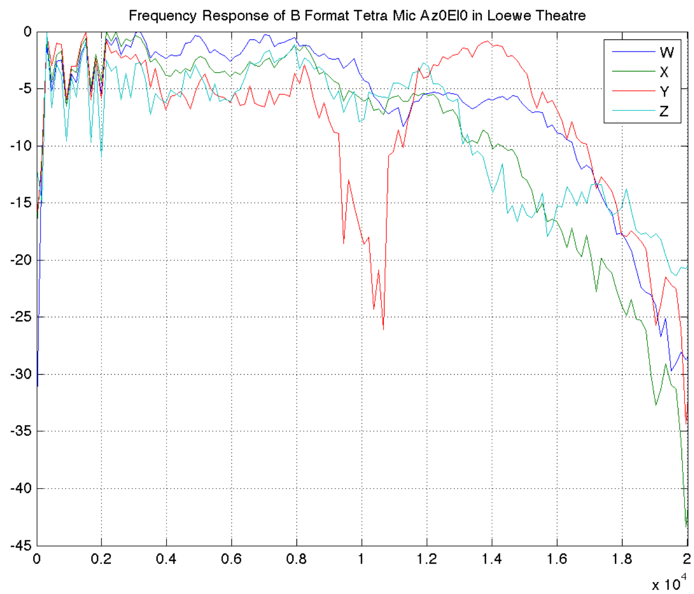


Figure 51. B format frequency responses of the Core Tetra Mic in *Physical Near Coincidence* signals in Loewe Theatre, post VVTetra Filtering. Note the still present dip in the Y signal at 1 kHz.

The responses of the Double MS Z arrays are unique in that they don't consistently roll off in the high frequencies all together. This is because they are placed 90° off axis from each other, so their responses are going to be different. The Y and the Z responses are a little more jagged, since they are capturing reflections and reverb. And the W and X are more smooth since they are capturing the direct part of the signal. The frequency response of the coincident (figure 53) and non-coincident (figure 52) arrays are practically the same.

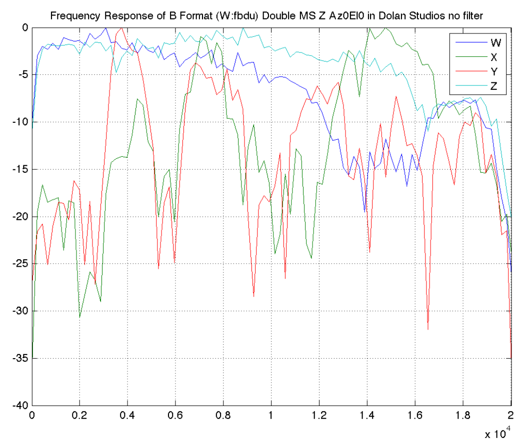


Figure 52. Double MS Z B format physical near coincidence in Dolan Studio.

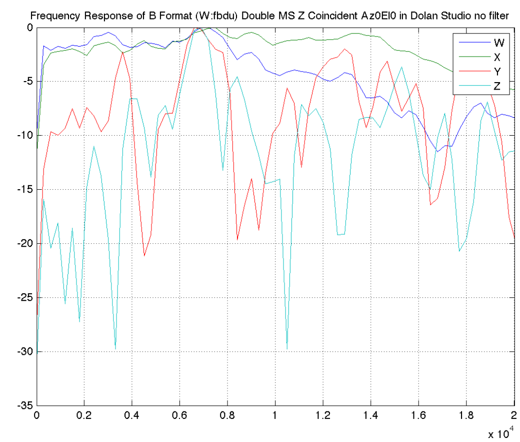


Figure 53. Double MS Z B format virtual true coincidence in Dolan Studio.

For more details on analysis, please see **Appendices C** through **E**.

4.6 Summary of Analysis

The array with the best combination of correlation in the direct and decorrelation in the reverb and reflections was the *virtual true coincident* Double MS Z. The *physical near coincident* Core Tetra Mic Array had good qualities of both. The array with the smoothest surround frequency response curve was the *virtual true coincident* A format DPA 4011As. The *physical near coincident* Double MS Z produced an very interesting arteficial early medial reflection. This is reported to cause a greater sense of spaciousness and depth. This was an unexpected side effect. The phenomenon was also recognized to a lesser extent in the A format Tetra spaced array, although in the A format spaced array it did not create a strong early lateral or medial reflection, but rather a group of weak ones seen as side lobes around the main maximum index of cross correlation. It will be interesting to see how these various characteristics translate to psychoacoustic qualities in a subjective listening test.

5 SUBJECTIVE LISTENING SESSIONS

5.1 On the Methodology and Purpose of the Listening Sessions

The overarching approach behind the listening test was to see what professionals in the audio industry would choose for optimal musical content, *virtual true coincidence* or *physical near* coincidence? It was hypothesized early on that the *virtual true coincidence* technique would cause a more natural soundscape, with stronger higher frequency content and no high frequency spatial aliasing. The theory is that the technique would give high frequency content localization, which would create a more natural 3D soundfield. It is hypothesized that the *virtual true coincidence* technique is the more natural of the two approaches. The goal of the listening session is to see if this is first even noticable and if so if it is more or less preferable to the listener.

5.2 On the Parameters of the Listening Sessions

Four professionals in the field of audio and music production were asked to participate in subjective listening sessions. The selected subjects were invited precicely because of their professional background and experience with understanding minute spectral differences for musical, acoustic and natural characteristics. The sessions were designed based on ITU recommendations for subjective sound quality assessments[70]. The basic proceedings of the test was a blind listening session of XAB comparisons, X being the orginal, A the first, and B the second. The subjects were told that they would be comparing different surround sound impulse response convolution reverbs. They were not told about the differing techniques used. They were asked to compare on the following qualities as defined by the ITU (table 7):

[70] [ITU:1997aa]

Characteristic	ITU Definition
Sense of Space	Does the performance appear to take place in an appropriate spatial environment? Can you tell features such as apparent room size, and distance between you and the sound source?
Envelopment	Does the listener feel completely immersed in sound, laterally and backwards? Can you tell speaker detent?
Presence of Digital Distortions	Does the listener percieve any electrical artifacts, such as high noise floors, aliasing, or anything that makes the sound unnatural.
Sound Quality Evaluation	Over all timbre of the sound. Boomy, sharp, dark, light, warm and cold are appropriate attributes. Would the listener need to eq the signal upon mixing, and if so, how?
Overall Preference	Disregarding the content of the programme, which did you prefer, A, B, or no preference?

Table 7. [ITU:1997aa]. Characteristics and ITU Definitions.

The sessions were done in a conversational manner, with the author asking questions on the above qualities. All answers and discussion were recorded.

The signal used was a short five second solo piano excerpt originally recorded close mic-ed in stereo. This particular signal was chosen for its musical broadband quality. The instrument has tonal characteristics as well as brighter transient characteristics. The impulse responses that were chosen to be convolved with the musical signal for the comparisons in these listening sessions were the ones taken in Dolan Studio. The Loewe theatre impulses were too reverberant and since the convolutions were set to 100% wet, this was seen as an unneccesary distraction to the above questions. The impulse response set from Dolan studio offered a good acoustically treated live room feel that was amenable for testing. There were four XAB comparisons sessions, as follows (table 8):

	AB1	AB2	AB3	AB4
A	*Virtual true coincidence,* Core Tetra Mic Capsules, A format	*Virtual true coincidence,* Core Tetra Mic Capsules, A format	*Virtual true coincidence,* Sennheiser Twin, Double MS Z format	*Virtual true coincidence,* DPA cardioids, A format
B	*Physical near coincidence,* Tetra Mic array, A format, unfiltered	*Physical near coincidence* Tetra Mic A format filtered	*Physical Near coincidence,* Sennheiser Twin, Double MS Z format	*Virtual true coincidence,* Sennheiser Twin, Double MS Z format

Table 8. The four AB comparisons for the listening sessions

The subjects were notified that they could listen to the signals as many times as they wished. The were seated in a listening spot in the center of the room and speaker array, but they were told they could move around to see if the sound image would hold true. The listening room chosen for the test was the Research Lab. The listening sessions were controlled by the author through a Pro Tools mixing session. The signals were output through a 16 speaker array; quad height and quad low, and octagon on ear level (figure 54).

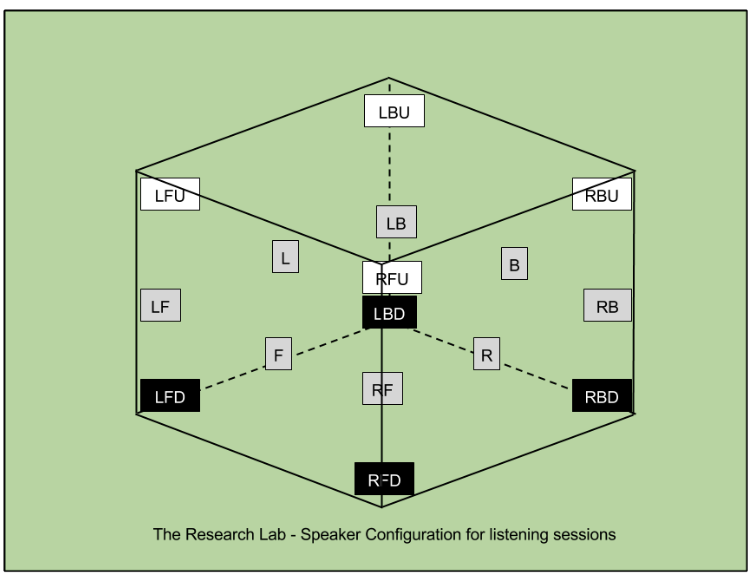


Figure 54. Speaker arrangement for the listening sessions.

This particular speaker arrangement is not a widely established Ambisonic layout. In fact no set of decoding coefficients were found for this particular layout, which indicates that no decoding coefficients have been set for this unique layout. However, this arrangement was chosen to get the best affects of all the microphone arrays. It is also true that this arrangement is not typical for music playback systems; but this was a test for listener preference and naturalness, so a playback configuration was chosen to serve that purpose. The decoding coefficients used were those of a similar established layout, one that had established coefficients, the cubioid and quad configurations combined. Also, a mixing session was done prior to all the subjective listening sessions, to make sure the sound was properly balanced in a "natural" way throughout the speakers. The author would not suggest using the quad and cubioid decoding coefficients as a standard for this unique layout, since equalization and gain mixing was done prior to the listening sessions.

5.3 On the Listeners for the Listening Sessions

Listener 1 is a mastering engineer with 40 plus years experience in the music industry and over 500 credits to his name in the genres of rock, jazz, classical, pop, and R&B. Listener 2 is a professional producer, mix engineer, musician and sound artist with over 10 years professional experience as a master producer and hunderds of commercial music credits to his name in classical, jazz and rock. He has worked extensively with surround sound and surround with height for music recording. Listener 3 is an audio professional and scientific researcher with background in 3D spatial sound, acoustics, and audio reproduction. Listener 4 is a composer and producer of electronic music with over 15 years experience in the field.

5.4 Results of the Listening Sessions

The most noteworthy results are discussed in the following sections.

5.4.1 A to B 1

AB1
A : *Virtual true coincidence,* Core Tetra Mic Capsules, A format
B : *Physical near coincidence,* Tetra Mic array, A format, unfiltered

Table 9. AB comparison set 1.

On spaciousness, listener 1 said signal B felt more spacious with more sense of depth. Listener 2 felt that A provided a better sense of space, noting that A felt correct in the space realm. Listener 4 felt a good sense of space for both. Listener 3 had some interesting comments saying that A felt "inside the piano" and much more of a "first person experience". B had more depth, and felt a little more inside a room.

On envelopment, listener 1 said both were poor, preferring B a little more. Listener 2 preferred A commenting that B felt tilted. Listener 4 preferred the immersion of B. Listener 3 said A was more sensitive to movement, the image would change as the listener moved through the room. With A the speakers seemed more "discrete": as the listener moved to each loudspeaker that part of the instrument would come out more creating an "inside the instrument" feel. Listener 3 noted that B had a "robust" spatial image that would not move from its location as the listener moved around the room. This listener made the interesting distinction between instrument envelopment and room envelopment, clarifying the more tradition sense of envelopment is the notion of being in a room with a strong spatial image, and in that sense, listener 3 prefered B for envelopment.

On the matter or distortions, listeners 1 and 2 noted that there were no distortions. Listener 3 said A had slight phasing issues as the listener moved around the room, and B did not. Listener 4 noticed slight digital noise artifacts in both, with A not as noticeable.

On quality, listener 1 said neither were pleasant, calling both “chilly”, “unattractive” and “unnatural”. Listener 2 commented that A was of good quality, and B was slightly annoying. Listener 4 said A is slightly better quality because of the artifacts.

On preference, listeners 1 and 3 preferred B. Listener 3 stated that this was because of the robust spatial image. Listeners 2 and 4 preferred A, stating that A felt more natural and lusher.

5.4.2 A to B 2

AB2
A : *Virtual true coincidence,* Core Tetra Mic Capsules, A format
B : *Physical near coincidence* Tetra Mic A format filtered

Table 10. AB comparison set 2

On the sense of space, listener 1 said A felt more realistic, and that B still had the sense of a room, but overall sounded too “processed” and “unnatural”, i.e. the listener knew the reverb was fake. Regardless, listener 1 felt “almost in the room” on both. Listener 2 also noted a good sense of space in both, but preferred B over A in this regard. The listener noted that, even though B had “a hole in the middle” and was “perceptually unbalanced”, the listener would still choose B over A and fix the unbalanced signal in production.

On immersion, listener 2 preferred B, stating that “A felt too mono”, and “at least I could work with B.” Listener 3 agreed, saying A sounded too “omnipresent”, noting again that the sound of the instrument was enveloping. However this time the “inside the instrument” immersion happened in both A and B, and the instrument sound was “too big” with “no focal point, coming from everywhere.”

On the subject of distortions, listener 1 thought B felt processed. Listener 4 noted that as the listener moved, the balance was affected in A. Listener 2 and 3 did not notice distortions.

On quality, listener 1 said that the sense of room in A was "moderately pleasant" and that the listener "didn't think about the sound". Listener 1 also noted that B sounded "dryer" than A. Listener 2 liked both. Listener 3 noted that they both sounded boomy and dark. Listener 4 said A felt "brighter" and "clearer", but B felt "warmer" with "more mids and high bass" content. Listener 4 expaned on this saying there was "less transient info in B" making it "a bit warmer and smoother overall."

On Preference, listener 1 preferred A. Listeners 2, 3 and 4 preferred B.

5.4.3 A to B 3

AB3
A : *Virtual true coincidence,* Sennheiser Twin, Double MS Z format
B : *Physical Near coincidence,* Sennheiser Twin, Double MS Z format

Table 11. A to B comparison set 3

On the sense of space, listener 1 noted B sounded unnatural. But listeners 2 and 3 said B was wider, more roomy, and a bigger space. Listener 3 noted again that the space seemed to be inside the piano, as if B had a "bigger resonant chamber". Listener 4 noted B had "spatial cues" and a "sense of depth."

On sense of envelopment, listener 2 said A had "a little more intense direct sound", but B was "less speaker-centric", giving overall envelopment to B. As a side note, both listeners 3 and 4 noted better envelopment and more reverberance in this AB comparison than in the prior 2 AB sessions. Listener 3 did once again note the "inside the instrument feel", and "discrete sounds coming from loudspeakers." Listener 4 said B had a better sense of envelopment, noting that with A, the listener felt more envelopment in the height, front, and back areas, but with B envelopment was more in the front.

On quality, listener 1 said A felt more natural, and that B felt “contrived” and “mechanical”. Listeners 2 and 4 noted that A felt “a little comb filtery”, meaning some phasing issues, but nothing too annoying though, ultimately giving quality preference to B. Listener 3 noted that both of them had a lot of mid frequency content, but B was a bit less annoying in this regard. Listener 4 said A felt fuller, but not as natural as B.

On preference, listener 1 preferred A, stating “I could imagine that having some sort of presence in a mix. The attacks and reverb seemed to be in consonance.” However, listeners 2, 3 and 4 preferred B.

5.4.4 A to B 4

AB4
A : *Virtual true coincidence,* DPA cardioids, A format
B : *Virtual true coincidence,* Sennheiser Twin, Double MS Z format

Table 12. A to B comparison set 4

On distortions, listener 3 noted A was “omnipresent”, “suceptible to movement and coloration”. Listener 4 noted comb filtering, phase issues, occuring in A upon moving around.

On quality, even though B had better envelopment and space, listener 2 preferred the sound quality of A, noting that it sounded more natural. Listener 4 felt that B was brighter, “with better attacks of the transients.”

On preference, Listener 2 was split, saying it depended on the context. Listener 3 prefered B hands down over A, noting that B was “clearer”, “crisper” and “airier”, and A sounded “muffled” in comparison. Listener 3 noted that with B there was slightly more of a sound stage present, and B had a better sense of space. Listener 4 preferred A over B.

5.5 Conclusions of the Listening Sessions

In general, there was no clear consensus when it came to any of the categories. In most cases there was generally a split between the four listeners with some very good reasons on either side. In a general sense this means that the *virtual true coincident* cases performed as good as the *physical near coincident* cases. To be fair, it did seem that the overall preference for musical content went slightly more for the *physical near coincident* cases, i.e. those cases using the microphone arrays instead of the *virtual true coincident* technique. However, *physical near coincident* cases were never 100% preferred across the board for all test subjects, for all AB comparison sets, and for all ITU characteristics. There were some notable observations made by the professional listeners, ones that that do not necessarily undermine the hypothesis of this thesis.

It seems that the space between the capsules provided a better sense of space, and a stronger spatial image of the sound source in a room at a particular location in that room. This is in line with the discovery made during the analysis stage that early lateral and medial reflections were caused by the distance between the capsules. Past research has shown that medial reflections are important for establishing a sense of depth, which sense of being in a room, and lateral reflections are important for establishing as sense of reverb, which defines a location of the sound source in that room. And this research shows that these phenomena are present in the Ambisonic arrays.

However, in contrast to this, several listeners noted that the sense of immersion in the *virtual true coincidence* array was more enveloping in some regards. Listener 3 had the noteworthy observation that the envelopment of the *virtual true coincidence* was a different kind of envelopment. It is hypothesized that in these listening sessions, the combination of a close mic-ed piano original signal and the use of the *virtual true coincident* technique gave a sense of immersion on a different level; that of being "inside the instrument". Listener 2, who noted that there was more intense direct sound causing envelopment with the *virtual true coincident* technique, echoed this view.

It was also noted that the *virtual true coincident* technique was more prone to phasing issues upon head movements. Both listeners 2 and 3 noted these issues in the *virtual true coincidence* technique upon moving their heads around. However, such comments were followed with remarks on the naturalness of the changing soundfield upon head movement, and that the *sound* was more "discrete", and not the speakers being more pronounced. This leads to a conclusion that perhaps this technique could be good for natural immersive environments, where the sound stage is not necessarily in front, if the phasing issues could be dealt with.

Although it seemed clear that preference usually went to the spaced microphone arrays in the listening tests, the author is not convinced that the technique proposed by this thesis should not be used for music. The listening test itself had a few inherent problems that need to be addressed. First, the sound source itself was problematic, the closed microphone technique used on the piano was noted to cause an unnatural disconnect between the reverb and the instrument. Listener 1 noted that the sample of the solo piano, "was not a very good source recording which makes it hard to judge quality." Listener 3 noted similarly that the piano is "huge". However, if the sound sample itself was not close mic-ed, then perhaps the convolved signal would not have been as unnatural overall. Furthermore, it is important to note that "proper" Ambisonics decoding was not done for this test. Many defenders of Ambisonics say that in order for Ambisonics to work properly, the decoding needs to be done correcly[71] with shelf filters and correct decoding coefficients for the layout. This author agrees that the decoding step is very

[71] [Heller:2008aa]

important because it has been noted in the listening sessions that gain affects the perception of aliasing. Despite this, documentation on ambisonic decoding (especially for larger, more irregular arrays such as the one used in these listening sessions) is either illegible[72], incomplete[73] and just plain lacking. It is hypothesized that proper decoding with proper filtering may provide a better sense of space and depth to the *virtual true coincident* technique. Further research in the area of completing and documenting proper decoding will be an important area for the improvement of Ambisonics.

6 CONCLUSIONS

6.1 Summary

In this work, a theory of complete coincidence was tested. A measurement and DSP process for simulating true coincidence was originated, measured and then tested against established coincident microphone arrays. Upon analysis, there were many interesting discoveries about coincidence. It was revealed that the space between the capsules in a closely spaced microphone array actually serve to create very early lateral or medial reflections. These reflections are more pronounced the further apart the capsules of the array. On subsequent professional listening sessions, it was discovered that the medial and lateral reflections of the closely spaced microphone arrays resulted in a perceived sense of space and localization for the listener. This led to the conclusion that perhaps this technique is not preferable for music production, since often a strong frontal sound stage and sense of space is desired over such things as naturalness or immersion.

[72] [Lee:2008aa]
[73] [Benjamin:2008aa]

6.2 Discussion

As noted in the literature review section, much has been written about the problems of spatial aliasing and phasing issues of high and mid frequency content in coincident microphone arrays. This led the author to attempt to address these flaws in this thesis work. However, what was unknown at the time, and what is still perplexing to this author is why these issues of spatial aliasing, phasing problems and poor image localization are so well documented as the problme with Ambisonics. To explain, let us take a look again at the listening session results. Spatial aliasing, phasing and localization issues were not recognized by any of the listeners in the very same microphone arrays that claim to have these issues, i.e. the *physical near coincident* cases. This fact alone is a baffling mystery to the author since so much has been written and documented about this topic. It has been said by other researchers in the field that skilled listeners have an easier time with locating low frequency sound sources with Ambisonics, and this can account for this this as skilled listeners were indeed used for this project. However, it seems unlikely that skilled listeners would not notice phase issues, especially when they were asked to listen for it as they were in these listening sessions. In fact, in the cases of *physical near coincidence*, phase and spatial aliasing issues were rarely noted.

Furthermore, it was actually easier for the listeners to locate the sound stage because of the space between the capsules. This artificial augmentation of the soundfield produced by the capsule spacing, actually serves to trick our brains into perceiving a sense of space and location of the sound source. Additionally, it seems that the early medial and lateral reflections were combating the psychoacoustic spatial aliasing affect in the higher frequencies and seeming masking the comb filitering phasing issues in the mid frequencies enough that it did not matter that these issues were present at all. This leads to the ultimate conclusion that the true problem with Ambisonics has nothing to do with the spatial aliasing issue so overdiscussed, but rather has to do with the poor spatial resolution caused by the use of a small number of capsules. And this is precisely why higher order ambisonic microphones are necessary for the further development of Ambisonics.

Finally with the *virtual true coincident* techniqe established by this thesis work, phase issues were noted on multiple occasions. Furthermore, what is interesting about this observation is that those phase issues noted by the listeners, were generally coupled with a statement on how natural the soundfield was. At the beginning of this project it was hypothesized that phase issues would not be present because of the fact that all microphones would be at the same location in space at different times. It is hypothezised that the reason phasing issued did persist is because of the fact that multiple impulse responses were used, which meant that the air pressure fluctuations caused by the impulse response would be different every time. Due to research in the field done by Farina, it was thought that by averaging multiple impulse responses, those phase issues in the higher frequencies would be smoothed out. However, this did not seem to be the case. There are however some ways to fix these phasing issues in the future.

6.3 Future Work

6.3.1 Improving the Technique

What was always asserted is that this technique would indeed create a more natural soundfield. What is clear is the technique suceeded in creating a natural soundfield, however, what was not expected, and what took away from the naturalness of the soundfield was the phasing issues. An hypothesis on how this issue occurred was presented in the previous section. There are several ways to combat the issue to hopefully create the most natural 3D soundfield without phasing issues. One solution would be to use a shorter sine sweep room response during the measurement session. The length of the signal was noted by several researchers[74] as an area of debate, a longer sweep gives better frequency resolution, however, also may cause phasing issues. Farina suggests 15 second sweeps where as Ben Hardor used 10 second sweeps. This project used 15 second sweeps, which may have been too much. Another solution would be to do the same type of inverse filtering that the VVTetra software does. With this solution the frequency response of the system could be flattened by using the inverse of the impulse responses as the system's own filters. This would get rid of phasing issues as it does for Core Sound's Tetra microphone. Finally, the system could be modified with shelt filters boosting the low end, to cut the phasing issues as well as perhaps creating a warmer sound with a better sense of space. This exact method is used in "proper" Ambisonic decoders, which as noted earlier, the author did not use for the listening sessions. Much potential work can be done to fix the phasing issue.

6.3.2 Modifying the Technique

If we think about the spatial sampling theorum once again, we realize that there is an optimal distance for every frequency, a distance that perfectly captures the peaks and valleys of the frequency of that corresponding wavelength. However, once we get down to a certain spacing we are out of the range of human hearing. Once we have achieved that distance, achieving true coincidence does not matter anymore, since frequencies of wavelengths shorter than that distance are inaudible. As has been true since Gerzon's time, there are no microphones small enough to create such an array (although we are getting pretty close).

However, we could modify the *virtual true coincident* technique proposed by this thesis. Instead of capturing the truly coincident location, we could capture the *virtual optimal distance*

[74] Farina and Ben Hardor

instead; the distance at which wavelengths are no longer audible. Furthermore, we could break the frequency spectrum up into different bands and use multiple optimal distances for each band in order to keep the lateral and medial reflections in the lower frequencies.

6.4 Conclusions

It is not clear when, but Gerzon was aware of the beneficial aspect of the capsule to capsule distance, he called it phase velocity information in his 1992 paper *General metatheory of auditory localisation*. What was originally believed at the start of this thesis work that the techological holdback of Gerzon's time was that he could not place the capsules close enough together to fix the spatial aliasing problem. But if Gerzon knew the benefit of the capsule spacing and if Gerzon knew that spatial aliasing was not that big of an issue, then perhaps this was not the case. Perhaps the real holdback was the inability of first order ambisonics to impress the public.

Indeed augmented reality is more facinating that real life when it comes to sound. When hearing a natural soundscape there is a disconnect if the visuals of that soundscape are not also present. So perhaps we prefer the unnatural 3D soundscape precisely because of the augmentations that help our ears make sense of everything around us. In music mixing practices the frequency spectrum is unnaturally modified all the time. What we precieve as natural, is actually augmented to give the sense of natural.

This thesis was built on the hopes that true coincidence would sound "better" than near coincidence; that no spacing would be "better" than any spacing between the capsules. So, early on, the methodological choice was made to time align all the signals, however, it is concluded that this may not be the best way to go, and at best this really depends on the application. With this impulse response technique, capsule distance has the potential to become a maliable feature of a convolution reverb plugin. Future work with this idea has much potential.

The *virtual true coincident* techinque proposed by this thesis does works at providing realistic natural convolution reverb: even if the *physical near coincident* cases tested slightly better among four audio professionals in terms of preference, it is not clear if it was considered more or less natural. And perhaps if the noted phasing and unstable soundfields issues were fixed this would become more clear. Regardless, the testing results show that the new techique works at creating a natural sense of immersion of good quality with not audible distortions. Improvements can be made on the work, and further testing can be done. This thesis work offers a unique contribution to the field of research by establishing an under published and little understood fact: that capsule distance provides a sense of depth and sound stage. With these new understandings in mind, the author sees limitless virtual microphone possibilities.

References

1. T. D. Abhayapala and D. B. Ward. *Theory and Design of High Order Sound Field Microphones Using Spherical Microphone Array*. In: International Conference on Acoustics, Speech and Signal Processing. © 2002 IEEE. [Abhayapala:2002aa].

2. Fons Adriaensen. *AmbDec - 0.4.2 User Manual*. At: http://kokkinizita.linuxaudio.org/linuxaudio/downloads/ambdec-manual.pdf. © 2010 Fons Adriaensen. [Adriaensen:2010aa].

3. J. S. Bamford and J. Vanderkooy. *Ambisonic Sound for Us*. In: 99th Audio Engineering Society Convention, New York, October 1995. © 1995 Audio Engineering Society. [Bamford:1995aa].

4. R. Ben-Hador and I. Neoran. *Capturing Manipulation and Reproduction of Sampled Acoustic Impulse Responses*. In: 117th Audio Engineering Society Convention, San Francisco, CA, October 2004. © 2004 Audio Engineering Society. [Ben-Hador:2004aa].

5. Eric Benjamin. *Ambisonic Loudspeaker Arrays*. In: 125th Audio Engineering Society Convention, San Francisco, CA, USA, October 2008. © 2008 Audio Engineering Society. [Benjamin:2008aa].

6. Eric Benjamin, Richard Lee, and Aaron Heller. *Why Ambisonics Does Work*. In: 129th Audio Engineering Society Convention, San Francisco, CA, USA, November 2010. © 2010 Audio Engineering Society. [Benjamin:2010aa].

7. Blue Ripple Sound. *HOA Technical Notes – Decoding*. At: http://www.blueripplesound.com/decoding. © 2013 Blue Ripple Sound, Ltd. [BlueRippleSound:2013aa].

8. Alan Blumlein. *Improvements in and Relating to Sound-Transmission, Sound-Recording and Sound-Reproducing Systems*. British Patent Specification, (394,325), Application Date: Dec 1931. In: Journal of the Audio Engineering Society, 6 (2), April 1958. © 1958 Audio Engineering Society. [Blumlein:1931aa].

9. Braxton Boren and Agnieszka Roginska. *Multichannel Impulse Response Measurement in Matlab*. In: 131st Audio Engineering Society Convention, New York, NY, USA, October 2011. © 2011 Audio Engineering Society. [Boren:2011aa].

10. Marco Conceição and Dermot Furlong. *Influence of Different Test Room Environments on IACC as an Objective Measure of Spatial Impression or Spaciousness*. In: 131st Audio Engineering Society Convention, New York, NY, USA, October 2011. © 2011 Audio Engineering Society. [Conceição:2011aa].

11. Marco Conceição and Dermot Furlong. *Influence of different microphone arrays on IACC as an objective measure of spaciousness*. In: 134th Audio Engineering Society Convention, Rome, Italy, May 2013. © 2013 Audio Engineering Society. [Conceição:2013aa].

12. Core Sound. *Tetra Mic Specifications*. At: http://www.core-sound.com/TetraMic/2.php. October 2012. © 2012 Core Sound, LLC. [Core:2012aa].

13. Peter Craven, M. J. Law, C. Travis. *Microphone Arrays Using Tangential Velocity Sensors*. In: Proceedings of the Ambisonics Symposium, Graz, Austria, 2009. © Unknown. [Craven:2009aa].

14. Jacek Dmochowski, Jacob Benesty, and Sofiène Affès. *On Spatial Aliasing in Microphone Arrays*. In: IEEE Transactions on Signal Processing, 57 (4), April 2009. © 2009 IEEE. [Dmochowski:2009aa].

15. DPA. *Microphone University - the Essentials*. http://www.dpamicrophones.com/en/Mic-University/The-Essentials.aspx. © 2013 DPA Microphones. [DPA:2013aa].

16. MH Acoustics LLC. http://www.mhacoustics.com/products. © Used MH Acoustics LLC. [Eigenmike:aa].

17. R. Elen. *Ambisonics: The Surround Alternative*. 2001. © 2001 Richard Elen. [Elen:2001aa].

18. A. Farina. *Guidelines for Acoustical Measurements Inside Historical Opera Houses*. In: Journal of Sound and Vibration, 232 (1), pages 281-301, April 2000. © 2000 Elsevier. [Farina:2000aa].

19. A. Farina. *Simultaneous Measurement of Impulse Response and Distortion with a Swept-Sine Technique*. In: 108th Audio Engineering Society Convention, Paris, France, February 2000. © 2000 Audio Engineering Society. [Farina:2000ab].

20. A. Farina and R. Ayalon. *Recording Concert Hall Acoustics for Posterity*. In: 24th Audio Engineering Society, International Conference on Multi-Channel Audio, Banff, Canada, June 2003. © 2003 Audio Engineering Society. [Farina:2003aa].

21. A. Farina. *A-Format to B-Format Conversion*. http://www.ramsete.com/Public/B-format/A2B-conversion/A2B.htm. October 2006. © 2006 Angelo Farina. [Farina:2006aa].

22. Peter Fellgett. *Ambisonics. Part One: General System Description.* Reproduce from Studio Sound, 17 (40): pages 20-22, August 1975. © IPC Media Ltd. [Fellgett:1975aa].

23. Geeknet. *The Ambisonic Reflector*. http://thereflector.sourceforge.net/. © 2009 GNU General Public License version 2.0 (GPLv2). [Geeknet:2009aa].

24. Earl Gates. *Introduction to Electronics, 6th Edition.* © 2011 Delmar Cengage Learning. [Gates:2011aa].

25. Paul Geluso. *Capturing Height: The Addition of Z Microphones to Stereo and Surround Microphone Arrays*. In: 132nd Audio Engineering Society Convention, Budapest, Hungary, April 2012. © 2012 Audio Engineering Society. [Geluso:2012aa].

26. Genelec. *Data Sheet Genelec 8030A Active Monitoring System*. © 2007 Genelec. [Genelec:2007aa].

27. Michael Gerzon. *Experimental Tetrahedral Recording: Part Two*. Reproduced from: Studio Sound, Vol. 13: pages 472, 473 and 475, September 1971. © IPC Media Ltd. [Gerzon:1971aa].

28. Michael Gerzon. *Periphony (With Height Sound Reproduction)*. In: Journal of the Audio Engineering Society, 21 (1), pages 2–10, 1973. © 1973 Audio Engineering Society. [Gerzon:1973aa].

29. Michael Gerzon. *What's Wrong with Quadraphonics*. Reproduced from Studio Sound, May 1974. © IPC Media Ltd. [Gerzon:1974aa].

30. Michael Gerzon. *The Design of Precisely Coincident Microphone Arrays for Stereo and Surround Sound*. In: 50th Audio Engineering Society Convention, March 1975. © 1975 Audio Engineering Society. [Gerzon:1975ac].

31. Michael Gerzon. *More on "The Presentation of Statistical Information on a Circle"*. In: Journal of the Audio Engineering Society, 23 (6), July/August 1975. © 1975 Audio Engineering Society. [Gerzon:1975ad].

32. Michael Gerzon. *Blumlein Stereo Microphone Technique*. In: Journal of the Audio Engineering Society, 24 (1): pages 36–38, February 1976. © 1976 Audio Engineering Society. [Gerzon:1976aa].

33. Michael Gerzon. *Mathematics and Sound Perception*. In: Journal of the Audio Engineering Society, 26 (1/2), January/February 1978. © 1978 Audio Engineering Society. [Gerzon:1978aa].

34. Michael Gerzon. *Practical Periphony: The Reproduction of Full-Sphere Surround*. In: 65th Audio Engineering Society Convention, London, England, February, 1980. © 1980 Audio Engineering Society. [Gerzon:1980aa].

35. Michael Gerzon. *General Metatheory of Auditory Localization*. In: 92nd Audio Engineering Society Convention, Vienna, Austria, March 1992. © 1992 Audio Engineering Society. [Gerzon:1992aa].

36. D. Griesinger. *New Perspectives on Coincident and Semi Coincident Microphone Arrays*. In: 82nd Audio Engineering Society Convention, London, England, March 1987. © 1987 Audio Engineering Society. [Griesinger:1987aa].

37. D. Griesinger. *The Science of Surround*. A Power Point Presentation. Presented at: 107th Audio Engineering Society Convention, New York, New York, September 1999. © 1999 D. Griesinger. [Griesinger:1999aa].

38. S. Guy-Bart, J. J. Embrechts, and D. Archambeau. *Comparison of Different Impulse Response Measurement Techniques*. In: Journal of the Audio Engineering Society, 50 (4), April 2002. © 2002 Audio Engineering Society. [Guy-Bart:2002aa].

39. Aaron J. Heller, Richard Lee and Eric M. Benjamin. *Is My Decoder Ambisonic?* In: 125th Audio Engineering Society, Oct 2008. © 2008 Audio Engineering Society. [Heller:2008aa].

40. Manfred Hibbing. *XY and MS Microphone Techniques in Comparison*. In: Journal of the Audio Engineering Society, Volume 3, October 1989. ©1989 Audio Engineering Society. [Hibbing:1989aa].

41. Melvin J. Hinich. *Processing Spatially Aliased-Arrays*. In: Journal of the Acoustical Society of America, 64 (3), Sept 1978. © 1978 Acoustical Society of America. [Hinich:1978aa].

42. Jan Jacob Hofmann. *Ambisonics Tutorial / Spatialization with Set of the Csound Spatializing Instruments*. At: http://www.sonicarchitecture.de/. © Unknown. [Hofmann:2009aa].

43. Zezhang Hou, Chaslav V. Pavlovic. *Effects of Temporal Smearing on Temporal Resolution, Frequency Selectivity, and Speed Intelligibility*. In: Journal of the Acoustical Society of America, vol. 96, no. 3. (1994), pages 1325-1340. © 1994 Acoustical Society of America. [Hou:1994aa].

44. R. Kessler. *An Optimized Method for Capturing Multi-Dimensional "Acoustic Fingerprints"*. In: 118th Audio Engineering Society Convention, Barcelona, Spain, May 2005. © 2005 Audio Engineering Society. [Kessler:2005aa].

45. Richard Lee. *Shelf Filters for Ambisonic Decoders*. http://www.ambisonia.com/Members/ricardo/shelfs.zip/view. © 2008 University of York. [Lee:2008aa].

46. C. J. mac Cába. *Surround Audio That Lasts: Future-Proof Ambisonic Recording and Processing Technique for the Real World*. In: 112th Audio Engineering Society Convention, Munich, Germany, May 2002. © 2002 Audio Engineering Society. [macCába:2002aa].

47. D. G. Malham. *Second and Third Order Ambisonics - the Furse-Malham Set*. http://www.york.ac.uk/inst/mustech/3d_audio/secondor.html. © 2005 University of York. [Malham:2005aa].

48. Fabio Manola, Andrea Genovese, and Adriano Farina. *A Comparison of Different Surround Sound Recording and Reproduction Techniques Based on the Use of a 32 Capsules Microphone Array, Including the Influence of Panoramic Video*. In: 25th UK Audio Engineering Society Convention - Spatial Audio in Today's 3D World and 4th International Symposium on Ambisonics and Spherical Acoustics, March 2012. © 2012 Audio Engineering Society. [Manola:2012aa].

49. T. K. Matsudaira, T. Fukami, and J. Sugati. *On "The Presentation of Statistical Information on a Circle"*. In: Journal of the Audio Engineering Society, vol. 22, page 536, September 1974. © 1974 Audio Engineering Society. [Matsudaira:1974aa].

50. Christopher Montgomery. *Xiph.org - Ambisonics*. At: http://wiki.xiph.org/Ambisonics. Page last modified June 18th, 2013. © Xiph.org Foundation. [Montgomery:2013aa].

51. S. Muller. *Transfer-Function Measurements with Sweeps*. In: Journal of the Audio Engineering Society, 49 (6), June 2001. © 2001 Audio Engineering Society. [Muller:2001aa].

52. Jörn Nettingsmeier. *Higher Order Ambisonics - A Future-Proof 3D Audio Technique*. In: 26th Tonmeistertagung - VDT International Convention, November 2010. © 2010 VDT. [Nettingsmeier:2010aa].

53. Ken Pohlmann. *Principles of Digital Audio*, 6th edition. © September 2010 McGraw-Hill Professional Publishing. [Pohlmann:2010aa].

54. V. Pulkki. *Microphone Techniques and Directional Quality of Sound Reproduction*. In: 112th Audio Engineering Society Convention, Munich, Germany, May 2002. © 2002 Audio Engineering Society. [Pulkki:2002aa].

55. The Ambisonic Toolkit. http://www.ambisonictoolkit.net/wiki/tiki-index.php?page=Downloads. General Public License version 2.0 (GPLv2). [TheAmbisonicToolkit:aa].

56. Stephen Thornton. *http://www.michaelgerzonphotos.org.uk/ambisonics.html*. © 2009 Stephen Thornton. [Thornton:2009aa].

57. Sennheiser Electronic. *MKH 800 TWIN Instructions for Use.* © 2010 Sennheiser Electronic. [Sennheiser:2010aa].

58. The ITU Recommendation Assembly. *Recommendation ITU-R BS. 1284-1. General Methods for the Subjective Assessment of Sound Quality*. © 1997-2003 ITU-R. [ITU:1997aa].

59. VV Audio. *Tetra Mic TM / VV Mic TM Users Guide*. © 2007 VV Audio. [VVAudio:2007aa].

60. B. Wiggins. *Has Ambisonics Come of Age?* In: Institute of Acoustics, 30 (6), 2008. © 2008 Institute of Acoustics. [Wiggins:2008aa].

APPENDIX A – MEASUREMENT SESSIONS

Here are pictures taken during the three measurement sessions in the three different sized spaces. These pictures show, in detail, how the measurement processes were done in the Research Lab, Dolan Studio and Loewe Theatre. Giving these pictures their own section serves the purpose of hopefully answering any remaining questions that may have been unanswered about how the measurement process was done.

Research Lab

Figure 55. Core's Tetra Mic in Research Lab.

Figure 56. DPA 4011a Cardioid with Angle Locator and laser pointer resting on end pointing to piece of tape on the ground, marking the angle. Unpictured, are the two laser pointers intersecting on the face of the DPA cardioid anchoring the micophone at the coincident location in space.

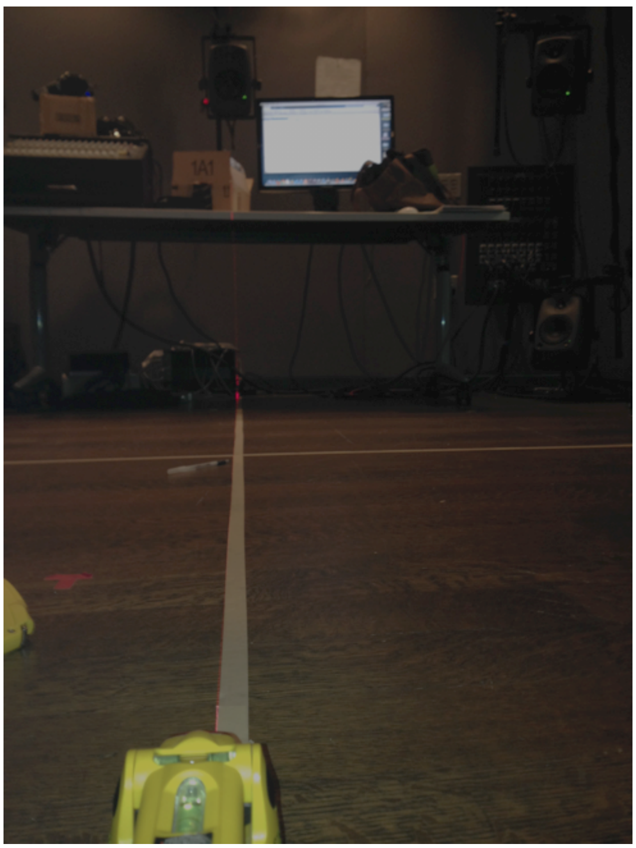


Figure 57. Tape was used to measure the room dimensions and mark the spot where the microphoen should be located. Laser line tool used to create a visible axis through the room.

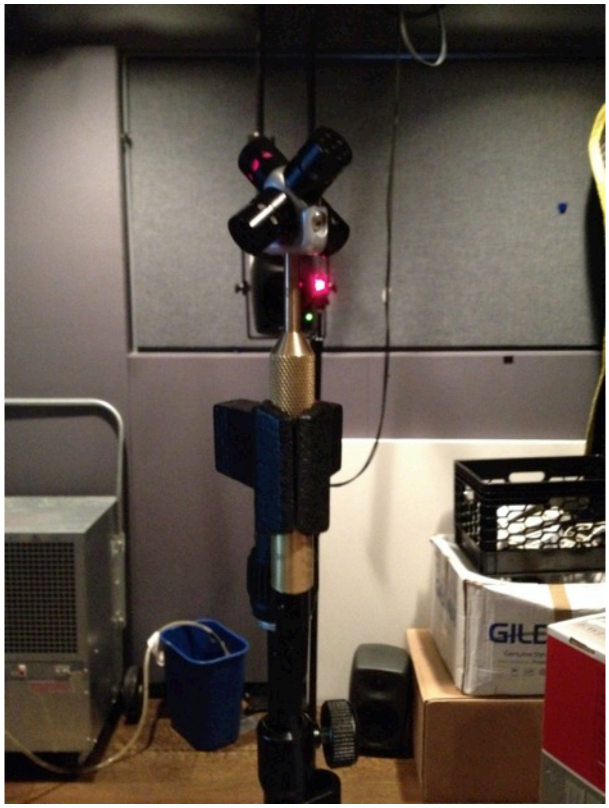


Figure 58. Core Tetra Mic at coincident location.

Figure 59. Laser pointer on mic stand with level on top to create right angle. Creating coincident location.

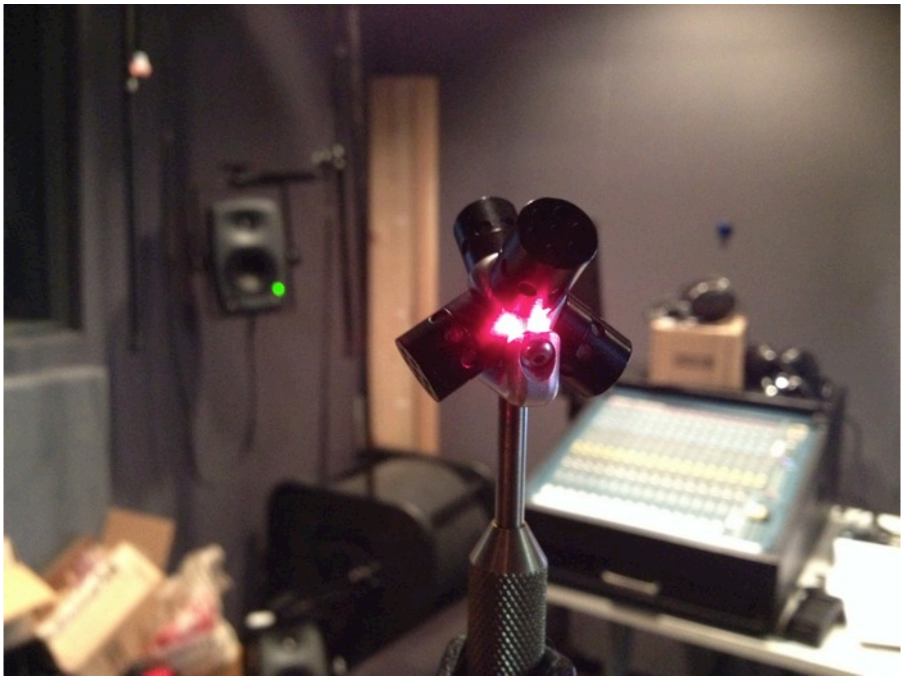


Figure 60. Core Tetra with 2 laser points at the nominal center. This setup was used when measuring the IRs of the physical near coincident cases. That way the nominal center would be the same for both physical near coincident and virtual true coincident cases.

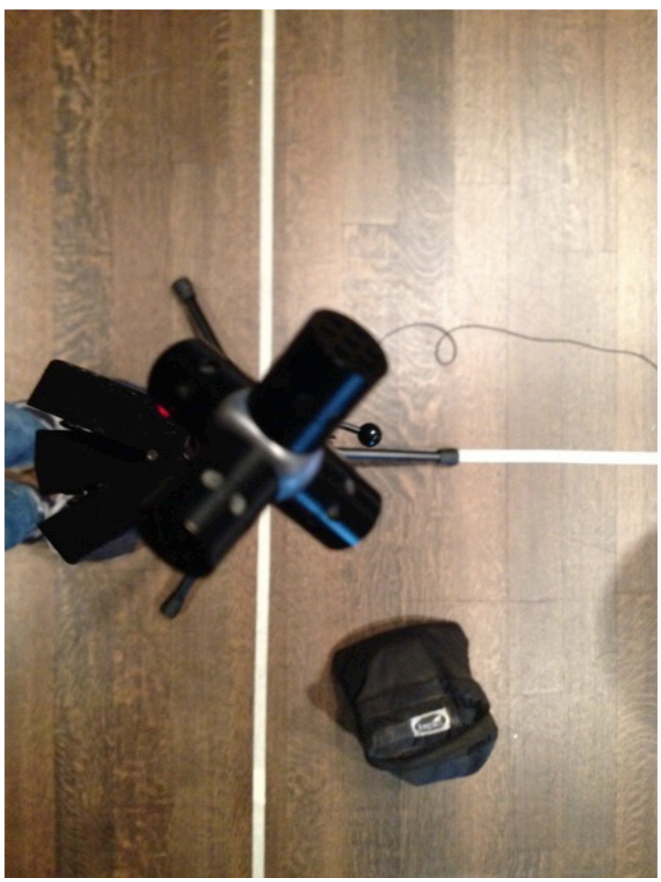


Figure 61. Birds eye view of Core Tetra at coincident location in space.

Figure 62. Core Tetra Virtual True Coincident measurement. Laser points intersecting on grill's face.

Figure 63. DPA 4011a Cardioid Virtual True Coincident Measurement. Laser's intersection on face of grill center.

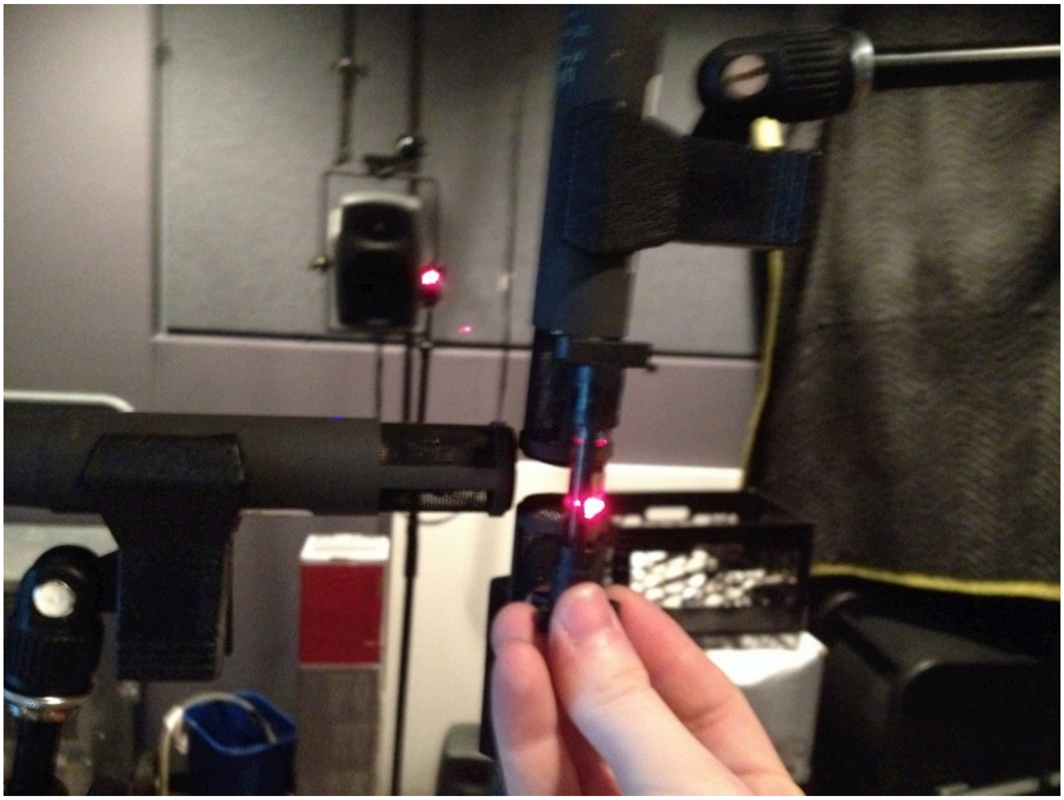


Figure 64. Double MS Z Array, holding up a level line tool to show laser point passing through the center of the array.

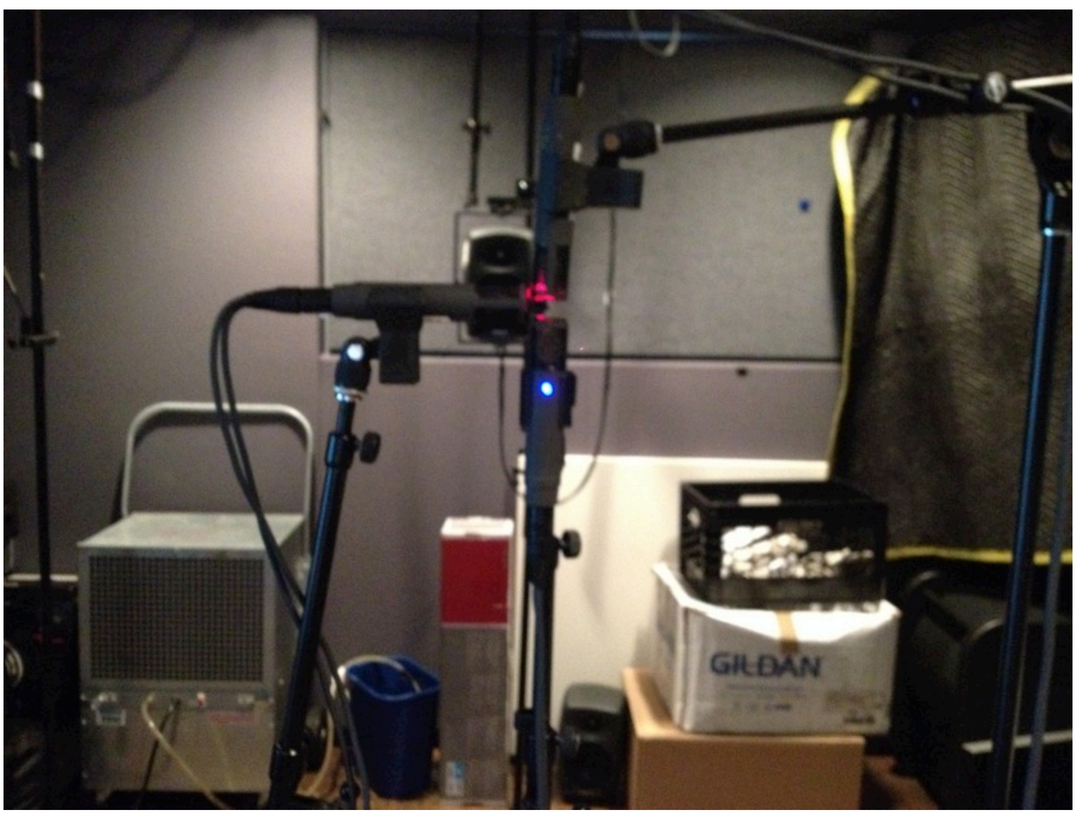


Figure 65. Double MS Z array.

Figure 66. Double MS Z array.

Dolan Studio

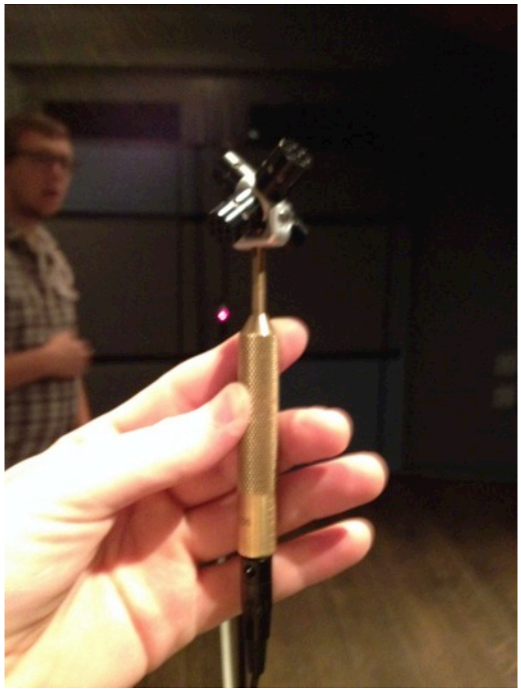


Figure 67. Core Tetra Mic, Alex Marse in background.

Figure 68. Dolan Studio setup.

Figure 69. 2 laser pointers intersecting at Core Tetra Mic.

Figure 70. Speaker and Microphone relationship.

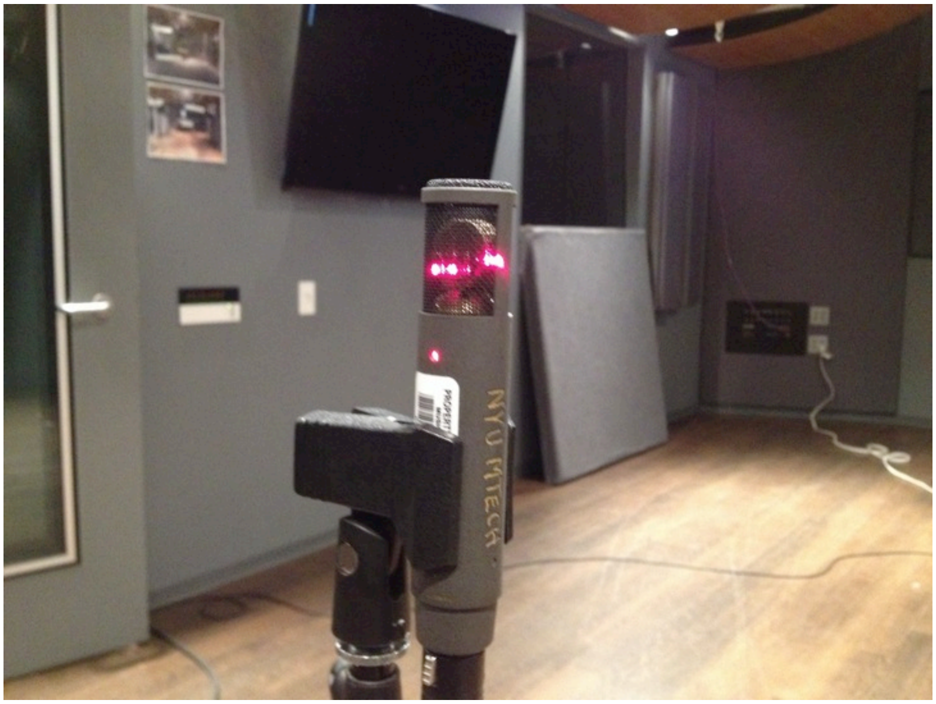


Figure 71. Sennheiser twin, Laser point marking center of diaphragm. Capturing front back of Double MS Z.

Figure 72. Sennheiser twin, Laser point marking center of diaphragm. Capturing left right of Double MS Z.

Figure 73. Sennheiser twin. Capturing up down of Double MS Z.

Loewe Theatre

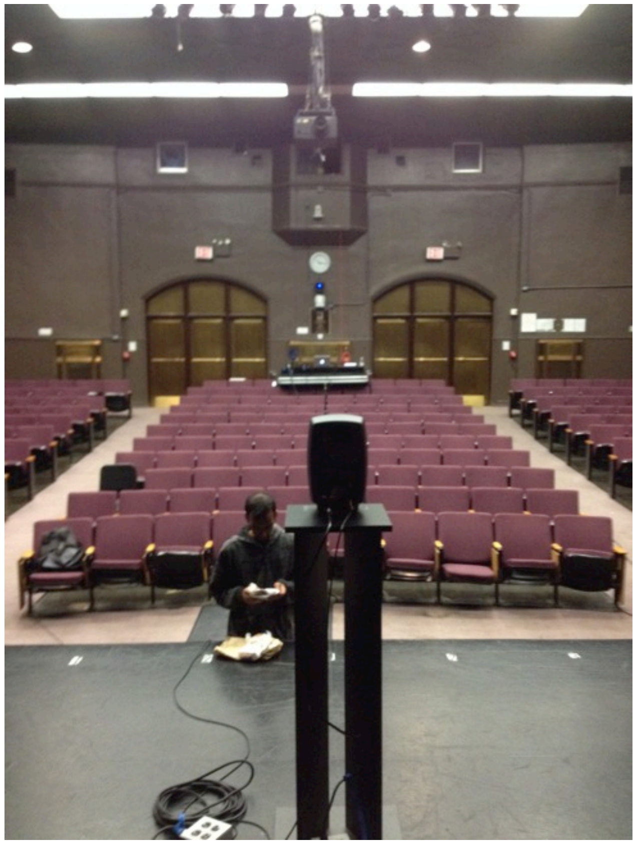


Figure 74. Speaker center stage, Hari Mohanraj in the background

Figure 75. Microphone on stand 5 rows back. Laser pointer and speaker also pictured.

Figure 76. Genelec Speaker center stage, 1/3 of length of theatre and centered.

Figure 77. The entire setup. In foreground, Computer with ScanIR session open, connected to the Metric Halo interface. Right behind is a laser pointer pointing to a microphone, behind which is the speaker on stage.

Figure 78. Microphone and speaker relationship.

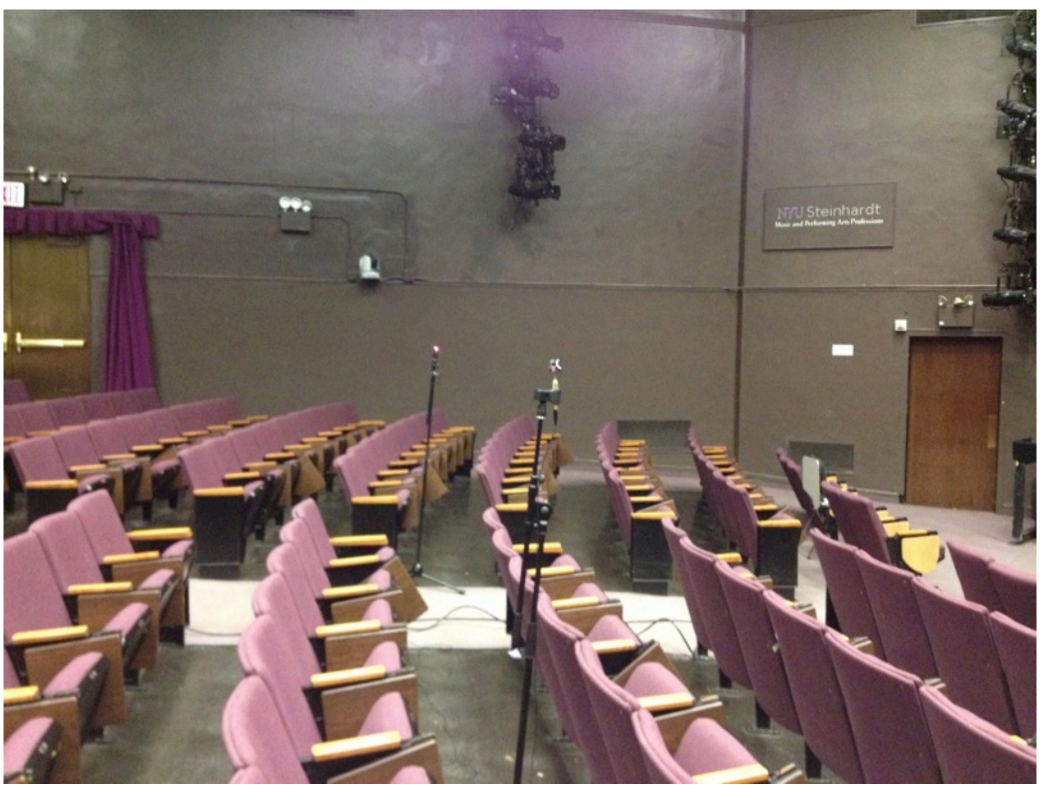


Figure 79. Laser pointer pointing to microphone at coincident location.

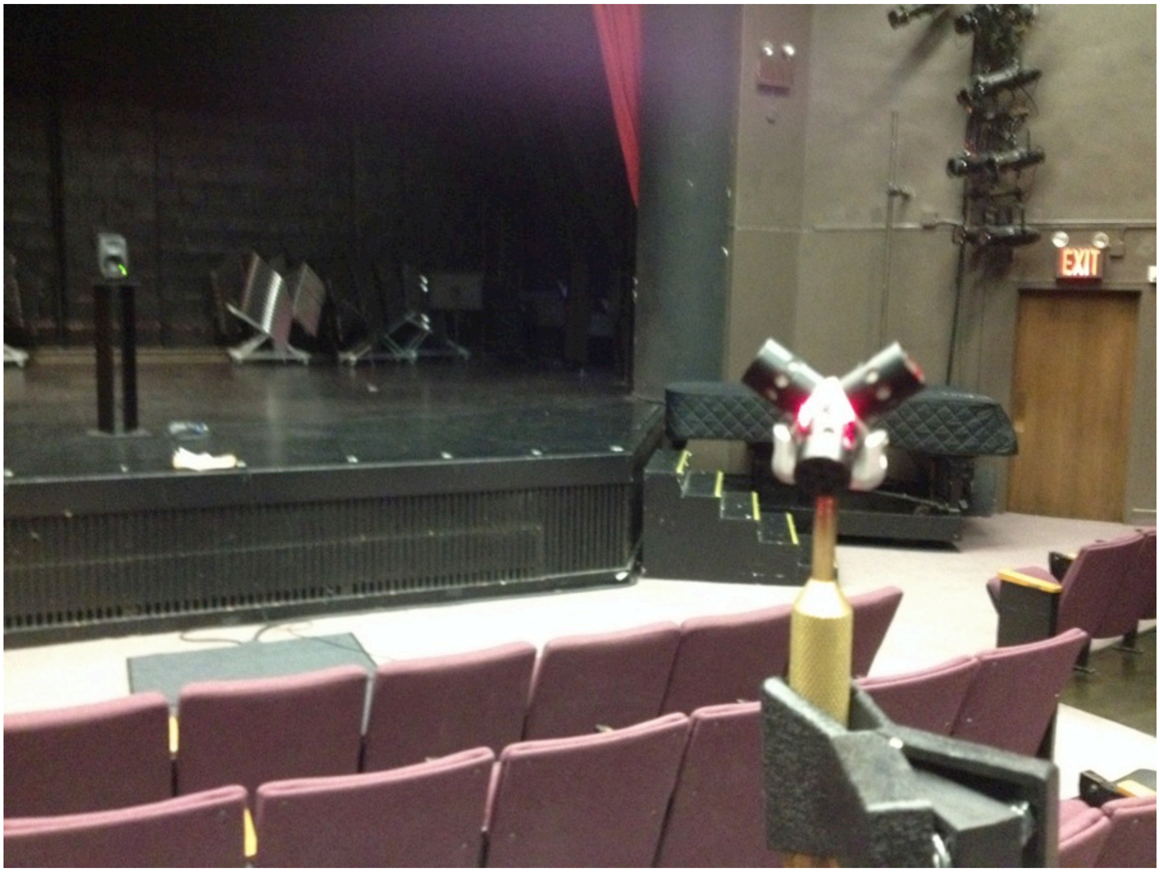


Figure 80. Closeup of Core Tetra with laser points, speaker in background.

Figure 81. Laser line tool on mic stand creating axis at the A format angle axis.

Figure 82. Laser line tool on mic stand.

Figure 83. Core Tetra Mic with laser points and laser line intersecting on grill center.

Figure 84. DPA 4011a Cardioid at A format Angle with Laser points intersecting at coincident location.


Figure 85. The Setup in Loewe Theatre.

Appendix B – Code Excerpts

Appendix B contains some excerpts from the important parts of the code done for this thesis work. Not all code is included, since there were several thousands of lines of code. A brief descriptions of each of the functions are given pointing out the important work being shown.

Plotting Modules

MasterPlot.m

This is an excerpt from the masterplot.m file, the main coding module for this work upon which most other coding modules branched from. The do_my_work function is its own .m file, explained later in detail; basically it is the plotting module that plots all input files. Then, there are the time align modules per channel for the *virtual true coinicdent* cases, that time align all the channels so that the start times line up. Then the do_my_work function is called again on the time aligned channels of the *virtual true coincident* cases. The channels are sent through a normalizing function. Then the *physical near coincident* channels get sent through a make multiple wave file function, so that the normalized *physical near coincident* channels can now be sent to the VV Tetra Mic software plugin for A to B matrixing and inverse filtering. There is an A to B function for those *virtual true coincident* cases that need to be matrixed in Matlab. The A to B function is its own .m file with other modules within such as another plotting module for the B format signals.

```
function masterplot(matrix, z, t)

[time_delayD_LFU_RFD, time_delayD_LFU_LBD, time_delayD_LFU_RBU,
...
    time_delayD_RFD_LBD, time_delayD_RFD_RBU,
time_delayD_LBD_RBU, ...
    ch1LFU, ch2RFD, ch3LBD, ch4RBU, corrALL, corrDirectAll] ...
    = do_my_work(matrix, z);
```

```
curChans = [ch1LFU ch2RFD ch3LBD ch4RBU];

% time align a format signals if
if nargin == 3 % then time align the signals, only do for
coincident sigs
    % ch1 and ch2
    if time_delayD_LFU_RFD >= 1
        direction = 'push';
        ch2RFD_TA = timeallign(ch2RFD,
time_delayD_LFU_RFD,direction);
    elseif time_delayD_LFU_RFD <= -1
        direction = 'pull';
        ch2RFD_TA = timeallign(ch2RFD,
time_delayD_LFU_RFD,direction);
    elseif time_delayD_LFU_RFD == 0
        ch2RFD_TA = ch2RFD;
    end
    putvar(ch2RFD_TA);

    % ch1 and ch3
    if time_delayD_LFU_LBD >= 1
        direction = 'push';
        ch3LBD_TA = timeallign(ch3LBD,
time_delayD_LFU_LBD,direction);
    elseif time_delayD_LFU_LBD <= -1
        direction = 'pull';
        ch3LBD_TA = timeallign(ch3LBD,
time_delayD_LFU_LBD,direction);
    elseif time_delayD_LFU_LBD == 0
        ch3LBD_TA = ch3LBD;
    end
    putvar(ch3LBD_TA);

    % ch1 and ch4
    if time_delayD_LFU_RBU >= 1
        direction = 'push';
        ch4RBU_TA = timeallign(ch4RBU,
time_delayD_LFU_RBU,direction);
    elseif time_delayD_LFU_RBU <= -1
        direction = 'pull';
        ch4RBU_TA = timeallign(ch4RBU,
time_delayD_LFU_RBU,direction);
    elseif time_delayD_LFU_RBU == 0
        ch4RBU_TA = ch4RBU;
    end
    putvar(ch4RBU_TA)

    Aformat_TA = [ch1LFU ch2RFD_TA ch3LBD_TA ch4RBU_TA];
    putvar(Aformat_TA);
    save(matrix);

    [time_delayD_LFU_RFD_TA, time_delayD_LFU_LBD_TA, ...
        time_delayD_LFU_RBU_TA, time_delayD_RFD_LBD_TA, ...
        time_delayD_RFD_RBU_TA, time_delayD_LBD_RBU_TA, ...
```

```
        ~, ~, ~, ~, corrALL_TA, ...
        corrDirectAll_TA] = do_my_work(matrix, z, t);

    % set the var curChans to your time aligned set so that gets
passed
    % into the matxrixing to B module.
      curChans = [ch1LFU ch2RFD_TA ch3LBD_TA ch4RBU_TA];

end

% normalize raw possibly time alligned A, create wav files for
VVTetra
Aformat_normalized = curChans/max(max(abs(curChans)));
makeWav(Aformat_normalized, matrix, 'Aformat_normalized');

% matrix to B signal in Matlab to plot without filters
[time_delayD_W_X, time_delayD_W_Y, time_delayD_W_Z,
time_delayD_X_Y, ...
    time_delayD_X_Z, time_delayD_Y_Z, corrBALL, corrDirectBALL]
...
    = AtoB(matrix, z, curChans);

% plot with Filters from VVTetra Plugin
[time_delayD_W_X_filtered, time_delayD_W_Y_filtered, ...
    time_delayD_W_Z_filtered, time_delayD_X_Y_filtered, ...
    time_delayD_X_Z_filtered, time_delayD_Y_Z_filtered, ...
    corrBALL_filtered, corrDirectBALL_filtered] ...
    = AtoB(matrix, z, curChans, 'VV');

…
```

TimeAlign.m

This is an excerpt from the timealign.m file, the function is called in the masterplot.m file.

```
% pulls channel forward by z samples
function [channel] = timealign(channel, z, direction)

if strcmp(direction, 'pull') == 1
    channel = channel(abs(z)+1:length(channel));
    zeropad = zeros(abs(z),1);
    channel = [channel' zeropad'];
    channel = channel';
end
if strcmp(direction, 'push') == 1
    zeropad = zeros(z,1);
    channel = [zeropad' channel'];
    channel = channel';
    channel = channel(1:480000);
end

end
```

DoMyWork.m

This is an excerpt from the do_my_work.m file, the function is called in the masterplot.m file as well. Here we see time domain plotting calls, magnitude spectra plotting calls, correlation and cross correlation plotting calls.

```
% create 4 channel matrix prior to running this function and save
it in a
% .mat file.  File name must include mic, coincident point, room,
AzEl. Load that
% '.mat file' into the function.  need to check start time for
where first
% order reflections come in too.  z is the number of samples to
the first
% order reflections.

function [td1, td2, td3, td4, td5, td6, ch1, ch2, ch3, ch4,
corrVal, ...
    corrDVal] = do_my_work(matrix, z, t)

% extract variables
channels = load(matrix);
x = -479999:479999;
y = (0:(z/2-1))/(z/2)*(96000/2);

if nargin == 2
    ch1LFU = channels.ch1LFU;
    ch2RFD = channels.ch2RFD;
    ch3LBD = channels.ch3LBD;
    ch4RBU = channels.ch4RBU;
    chALL = channels.chALL;
end
if nargin == 3 % then the signals are time alligned
    ch1LFU = channels.ch1LFU;
    ch2RFD = channels.ch2RFD_TA;
    ch3LBD = channels.ch3LBD_TA;
    ch4RBU = channels.ch4RBU_TA;
    chALL = channels.Aformat_TA;
    matrix = [matrix, ' time alligned'];
end

    %plot time domain
    plot(chALL);
    xlim([0 z]);
    grid on;
    legend('ch1LFU', 'ch2RFD', 'ch3LBD', 'ch4RBU');
    plottitle = ['Start Times of Signals in', matrix];
    title(plottitle);
    fsave = ['./plots/Figs/', plottitle, '.fig'];
    saveas(gcf, fsave, 'fig');
    psave = ['./plots/Pics/', plottitle, '.png'];
```

```
saveas(gcf, psave, 'png');

% Plot Magnitude Spectra of the direct signal
%LFU
directLFU = ch1LFU(1:z);
fftLFU = fft(directLFU);
MagLFU = abs(fftLFU(1:z/2));
normMagLFU = MagLFU/max(MagLFU);
dBMagLFU = 20*log10(normMagLFU);
plot(y, dBMagLFU);
xlim([0 20020]);
grid on;
xlabel('Frequency (Hz)');
ylabel('Magnitude (dB)');
legend('ch1LFU');
plottitle = ['Frequency Response of Direct LFU of', matrix];
title(plottitle);
fsave = ['./plots/Figs/', plottitle, '.fig'];
saveas(gcf, fsave, 'fig');
psave = ['./plots/Pics/', plottitle, '.png'];
saveas(gcf, psave, 'png');
```

…

```
% compute correlation data
corrCommand = sprintf('corrALL = corr(chALL);');
eval(corrCommand);
% bar plot correlation data, save plots
bar(corrALL);
xlabel('Channel Number');
ylabel('Correlation');
plottitle = ['Correlation Plot of ', matrix];
title(plottitle);
grid on;
fsave = ['./plots/Figs/', plottitle, '.fig'];
saveas(gcf, fsave, 'fig');
psave = ['./plots/Pics/', plottitle, '.png'];
saveas(gcf, psave, 'png');
```

…

Convolution Modules

Convolver.m

This is an excerpt from the convolver.m file. There are two types of convolution available, fast convolution for a quicker convolution, and direct convolution for a more brute force approach. The convolution function offers a choice of window length as well for the user's preference.

```
% convolver ( IRfilename, SIGfilename, OUTfile, convMethod,
segLength )
%
% The function will convolve an impulse response with a signal
% using either direct or fast convolution. The resulting
% signal will be written to a .wav file.
%
% Inputs:
%
%   1) IRfilename : Name of .wav file containing the impulse
response
%   2) SIGfilename : Name of .wav file containing a signal
%   3) OUTfile : Name of .wav file to which the resulting
(convolved) signal
%       will be written
%   4) convMethod : Convolution method:
%       1: direct convolution
%       2: fast convolution
%   5) segLength : (optional) Length of input segments (blocks)
to convolve
%       with impulse response, using the overlap-add method.
Default is
%       entire length of input file.
%
% Output:
%
%       convolver ( IRfilename, SIGfilename, OUTfile, convMethod,
segLength )

function convolver( IRfilename, SIGfilename, OUTfile, convMethod,
location, IRPath, SIGPath, segLength)

%-------- ERROR CHECKING OF INPUTS ---------

% number of arguments check (remember segLength is optional).

%  if nargin < 4,
%     error('need at least 4 inputs, guy')
%  end
%
% if nargin > 5,
%     error('too many inputs, buddy')
% end

% filename check, make sure your files are strings

if ischar(IRfilename) == 0,
    error('IRfilename must be a string, friend')
end
```

```
if ischar(SIGfilename) == 0,
    error('SIGfilename must be a string, bro')
end

% ---------- CONSTANTS AND EQUATIONS ----------

% The function reads the .WAV files, sig and imp, converts to
mono and
% normalizes to prevent clipping

IRfilename = [IRPath IRfilename];
[impulseresponse,fs1] = wavread(IRfilename);
%impulseresponse = mean(impulseresponse,2); ALL ARE MONO SO
SUPERFLUOUS
%impulseresponse = impulseresponse / 1.001 *
max(abs(impulseresponse));

% need to cut the IR so that it ends after the late reflections
if strcmp(location, 'RL') == 1
    impulseresponse = impulseresponse(1:9000,:);
end
if strcmp(location, 'Dolan') == 1
    impulseresponse = impulseresponse(1:9000,:);
end
if strcmp(location, 'Loewe') == 1
    impulseresponse = impulseresponse(1:479996,:);
end

irLength = length(impulseresponse);

SIGfilename = [SIGPath SIGfilename];
[signal,fs2] = wavread(SIGfilename);
%signal = mean(signal,2);  ALL ARE MONO SO SUPERFLUOUS
signal = signal / 1.001 * max(abs(signal));
sigLength = length(signal);

outputVec = zeros((irLength + sigLength -1), 1);

if fs1 ~= fs2
    error('sampling rate of the signal is not the same as the
impulse response');
end

% If no segLength is provided, default segLength is entire length
of the
% signal

if nargin == 7
    segLength = sigLength;
end

% If user provides segLength input make sure it is not longer
than the
% total number of samples of the signal.
```

```
if segLength > sigLength
    segLength = sigLength;
    warning('your segLength is longer than the number of samples
in the SIGfilename, using signal length instead, foo');
end

numSeg = ceil(sigLength / segLength);

if numSeg * segLength > sigLength
    numZeros = (numSeg * segLength) - sigLength;
    signal = [signal ; zeros(numZeros, 1)];
end

for k = 1:numSeg
    startIndex = (k - 1) * segLength + 1;
    segment = signal(startIndex:startIndex + segLength - 1);

    switch convMethod
        % direct convolution
        case 1
            segLength = length(segment);
            output = zeros(irLength + segLength - 1,1);

            for i = 1:segLength
                dConv = segment(i) .* impulseresponse;
                output(i:i+length(dConv)-1) =
output(i:i+length(dConv)-1) + dConv;
            end
        % fast convolution
        case 2
            irFFT = fft(impulseresponse, (segLength + irLength) -
1);
            segFFT = fft(segment, (segLength + irLength) - 1);

            output = ifft(irFFT .* segFFT);
    end

    % ----- OUTPUT -----
    idx = (1:length(output)) + startIndex-1;
    outputVec(idx) = outputVec(idx)+output;
end

outputVec = outputVec / (1.0001 * max(abs(outputVec)));

wavwrite (outputVec, fs1, OUTfile);

end
```

First Order Decoder Modules

FirstOrderDecode.m

This is an excerpt from the firstorderdecode.m file.

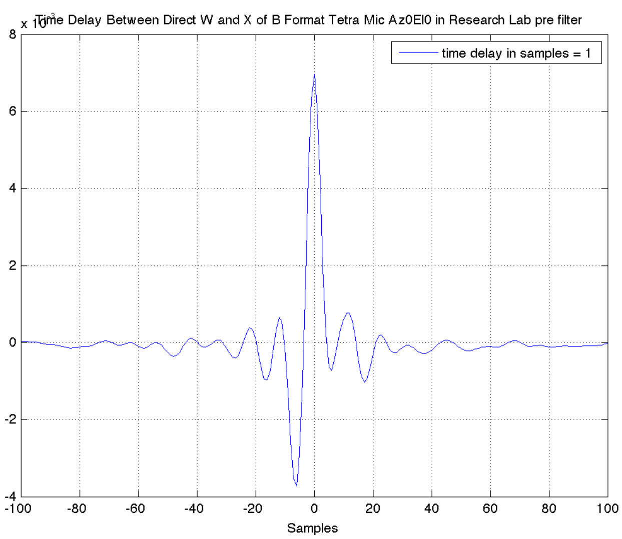
```
% decodes convolved signals into quad and cube first order
signals with proper coefficients.  Not shelf filtered for
different high and low frequency filtering though.  That seems a
bit more complex, and can be done in DAW.

function firstorderdecode

…
                % W Signal
                    % first order quad decode
                    wQuad = w .* (0.3536);
                    wavwrite(wQuad, 96000, 24, ...
                        [OUTPathQuad 'WQuad_' curSigName]);
                    % first order cube decode
                    wCube = w .* (0.1768);
                    wavwrite(yCube, 96000, 24, ...
                        [OUTPathCube 'WCube_' curSigName]);
                end

                % X Signal
                    % first order quad decode
                    xQuad = x .* (0.1768);
                    wavwrite(xQuad, 96000, 24, ...
                        [OUTPathQuad 'XQuad_' curSigName]);
                    % first order cube decode
                    yCube = x .* (0.0722);
                    wavwrite(xCube, 96000, 24, ...
                        [OUTPathCube 'XCube_' curSigName]);
                end

                % Y Signal
                    wavwrite(yQuad, 96000, 24, ...
                        [OUTPathQuad 'YQuad_' curSigName]);
                    % first order cube decode
                    yCube = y .* (0.0722);
                    wavwrite(yCube, 96000, 24, ...
                        [OUTPathCube 'YCube_' curSigName]);
                end

                % Z Signal
                    % first order quad decode
                    zQuad = z .* 0;
                    wavwrite(zQuad, 96000, 24, ...
                        [OUTPathQuad 'ZQuad' curSigName]);
                    % first order cube decode
                    zCube = z .* (0.0722);
                    wavwrite(zCube, 96000, 24, ...
                        [OUTPathCube 'ZCube_' curSigName]);
                end
            …
end
```

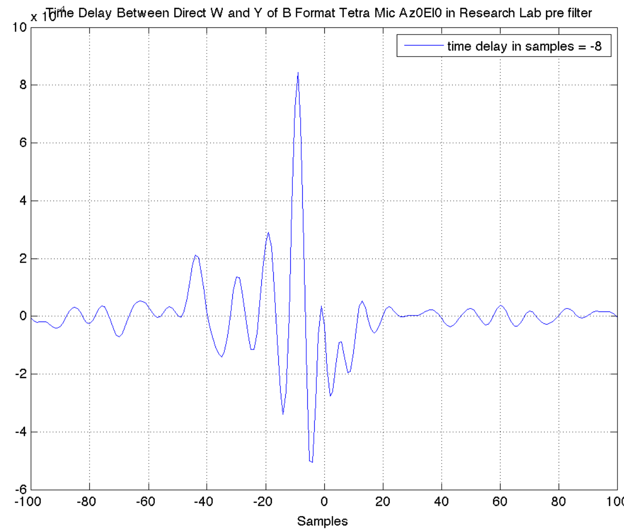
For Appendices C through E, all plots and tables are included for B format signals as proof of the concepts and hypotheses made in the thesis document itself. All plots and tables are clearly labeled at the top of each plot or chart. Not all data was included in the appendices; all the A format plots was left out since B format gave the best comparisons across all cases.

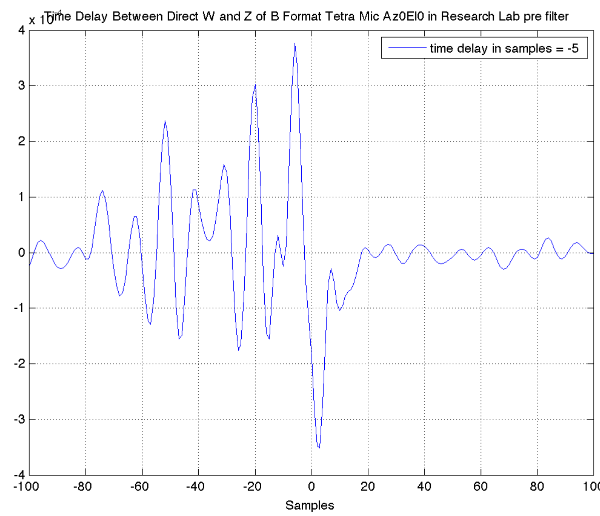
APPENDIX C – B FORMAT CROSS CORRELATION PLOTS

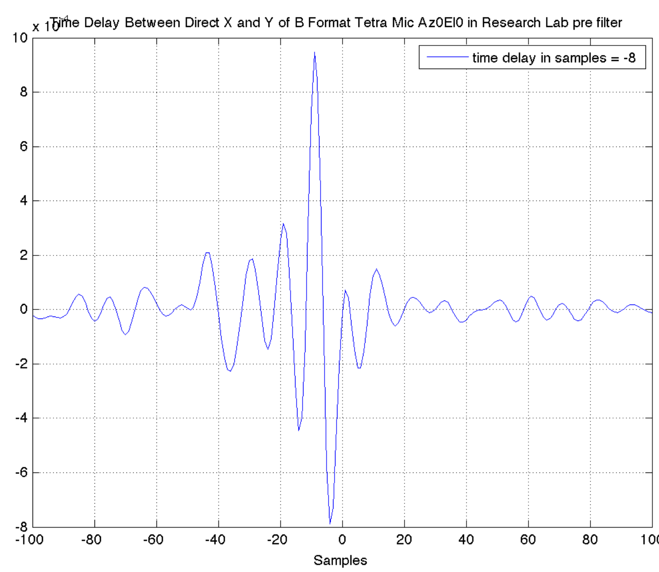
Research Lab

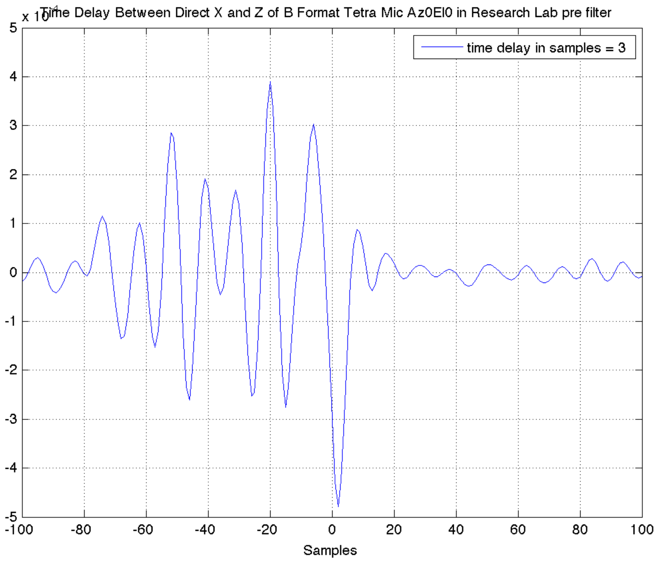
Physical Near Coincident Core Sound Tetra Mic Az0

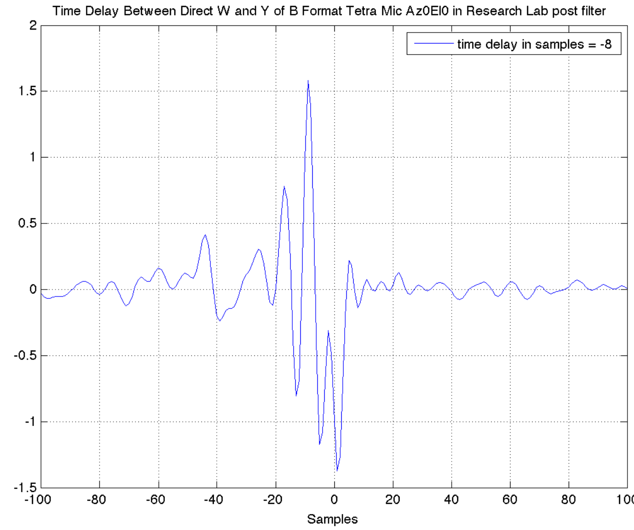


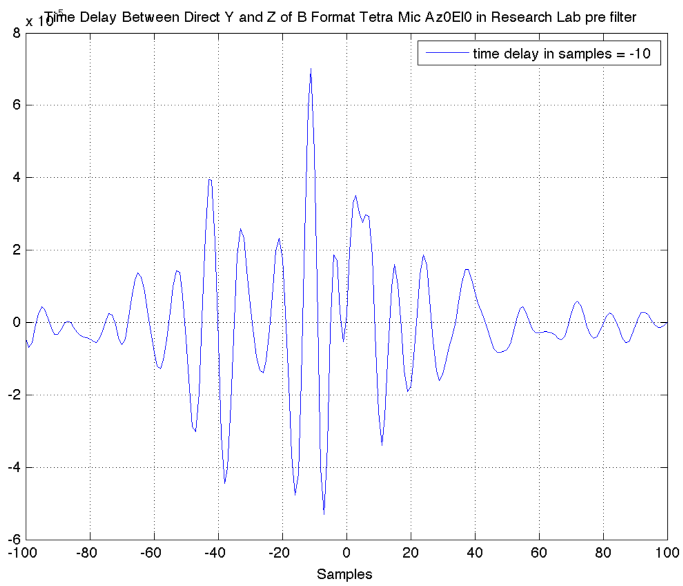


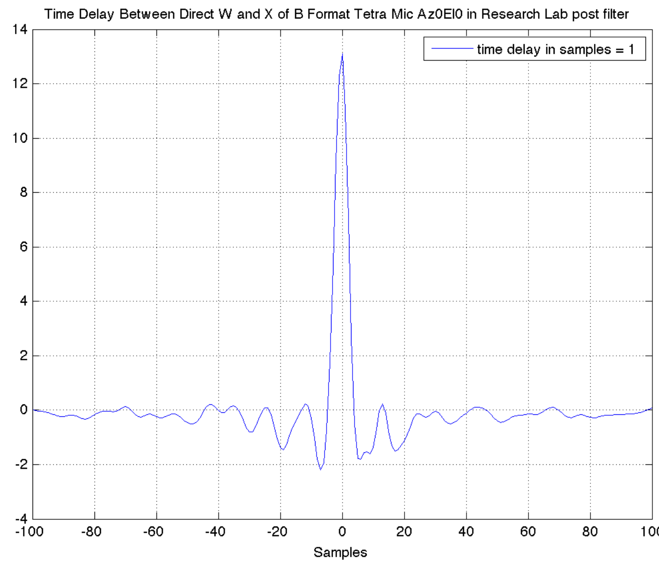


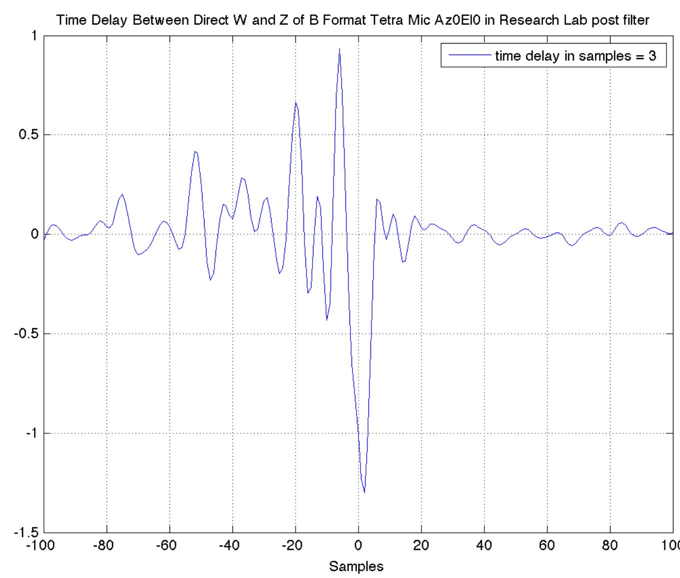



Time Delay Between Direct X and Z of B Format Tetra Mic Az0El0 in Research Lab pre filter
time delay in samples = 3
Samples

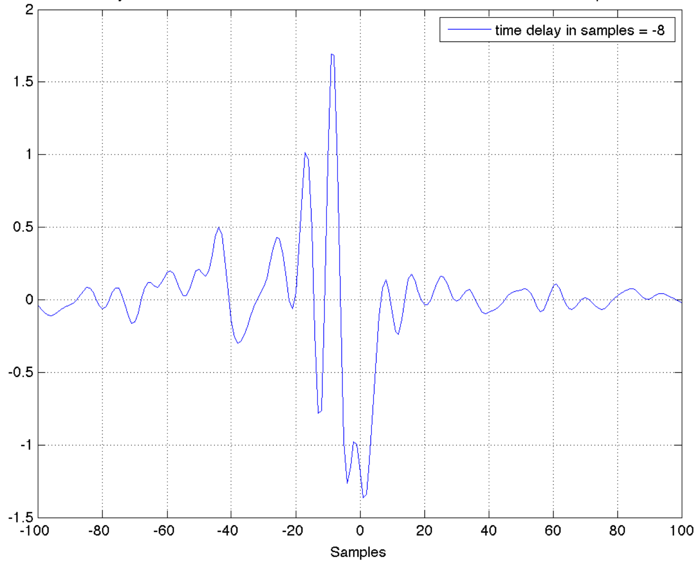

Time Delay Between Direct W and Y of B Format Tetra Mic Az0El0 in Research Lab post filter
time delay in samples = -8
Samples


Time Delay Between Direct Y and Z of B Format Tetra Mic Az0El0 in Research Lab pre filter
time delay in samples = -10
Samples

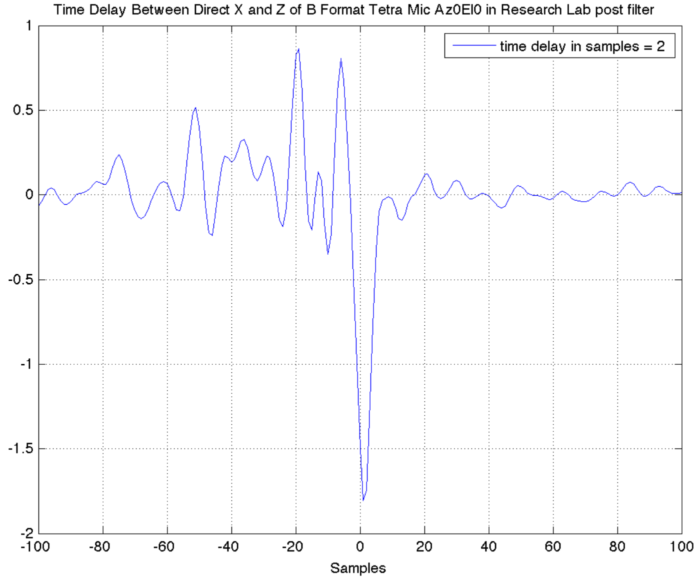

Time Delay Between Direct W and X of B Format Tetra Mic Az0El0 in Research Lab post filter
time delay in samples = 1
Samples


Time Delay Between Direct W and Z of B Format Tetra Mic Az0El0 in Research Lab post filter
time delay in samples = 3
Samples

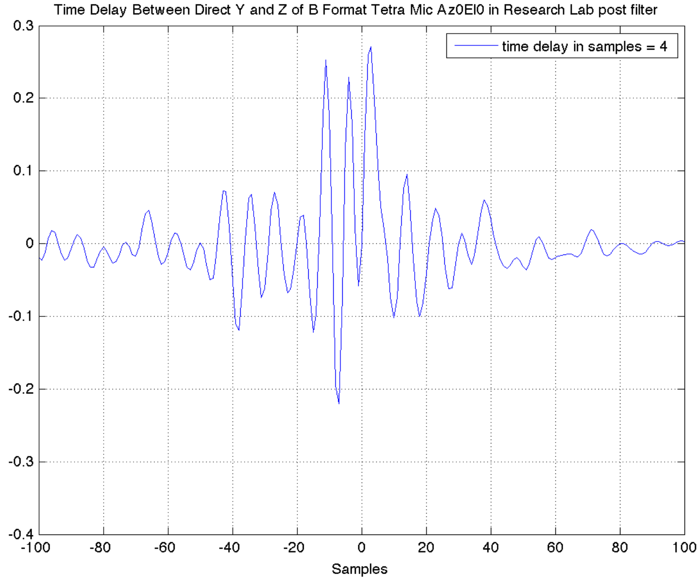
Time Delay Between Direct X and Y of B Format Tetra Mic Az0El0 in Research Lab post filter



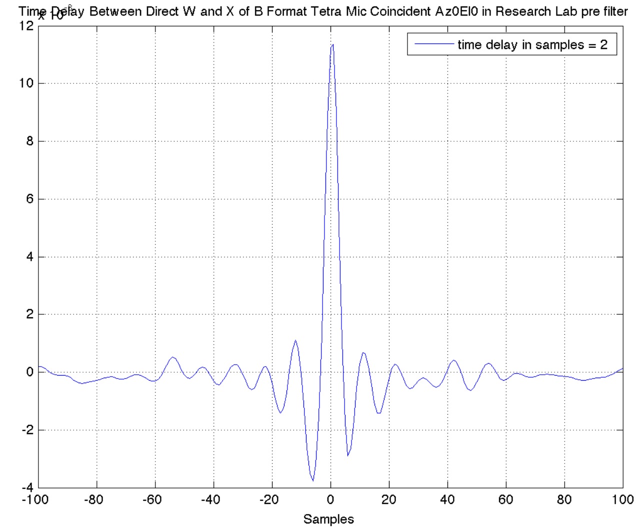
Time Delay Between Direct X and Z of B Format Tetra Mic Az0El0 in Research Lab post filter



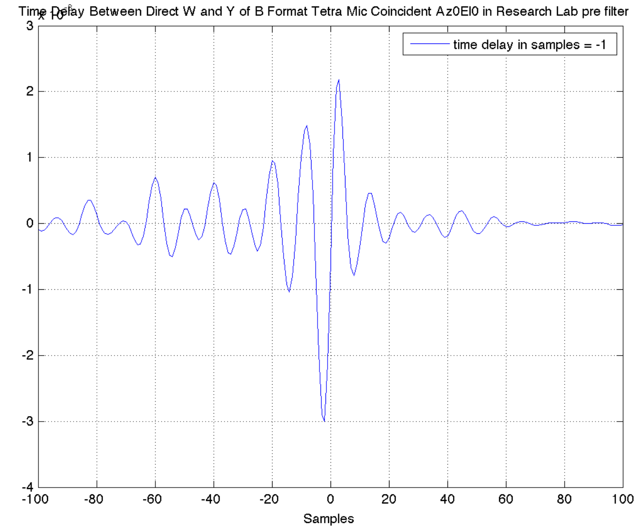
Time Delay Between Direct Y and Z of B Format Tetra Mic Az0El0 in Research Lab post filter



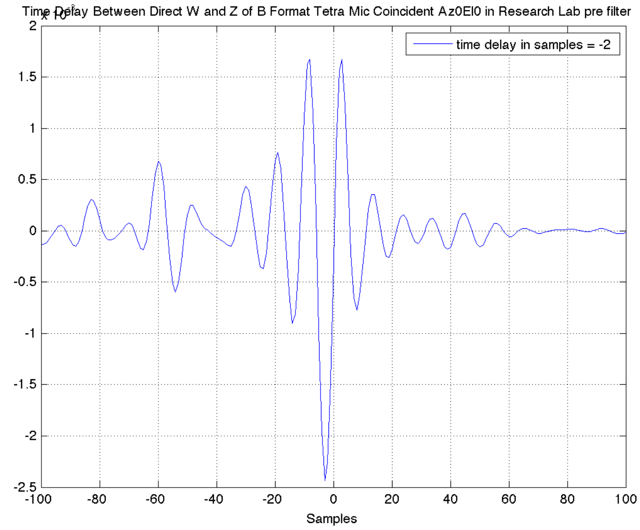
Time Delay Between Direct W and X of B Format Tetra Mic Coincident Az0El0 in Research Lab pre filter

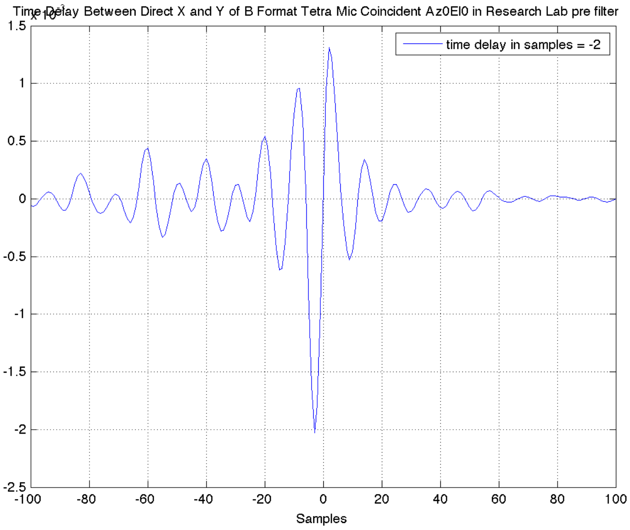


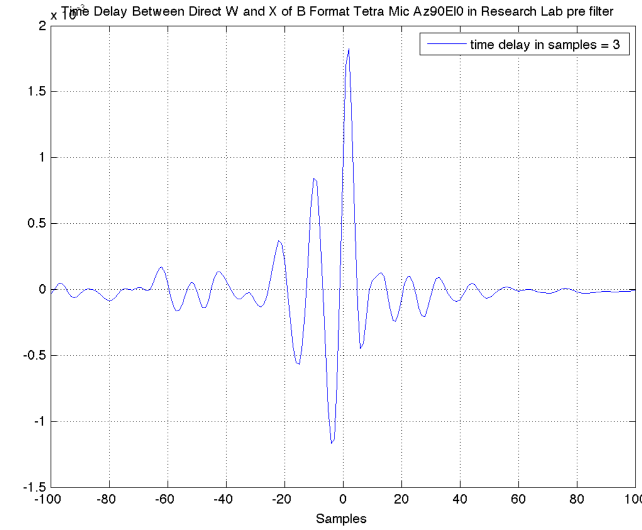
Time Delay Between Direct W and Y of B Format Tetra Mic Coincident Az0El0 in Research Lab pre filter

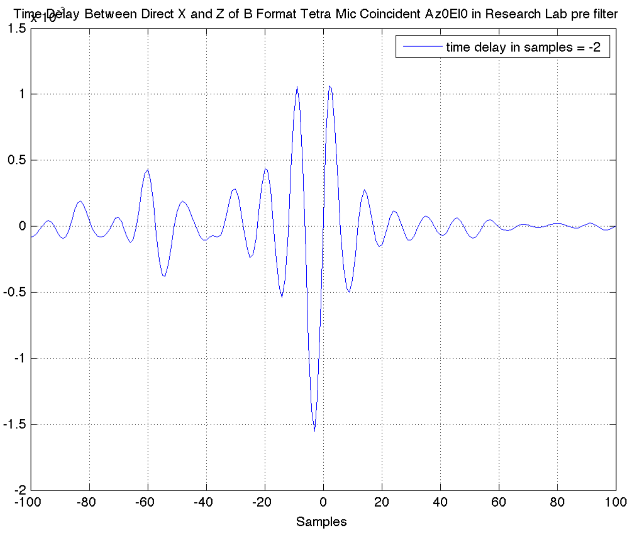


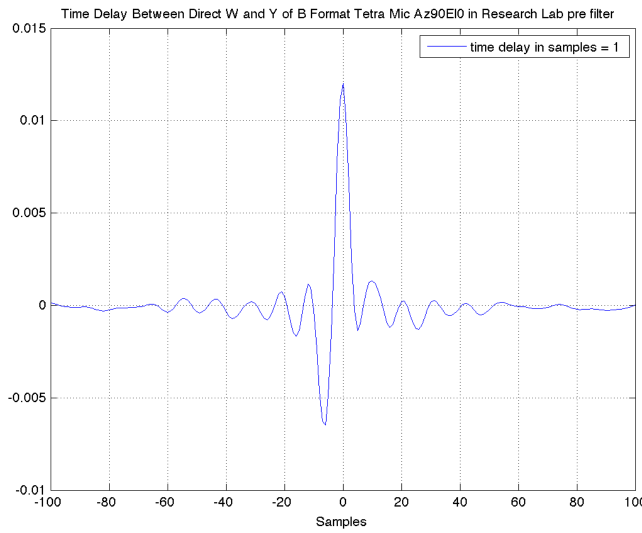
Time Delay Between Direct W and Z of B Format Tetra Mic Coincident Az0El0 in Research Lab pre filter

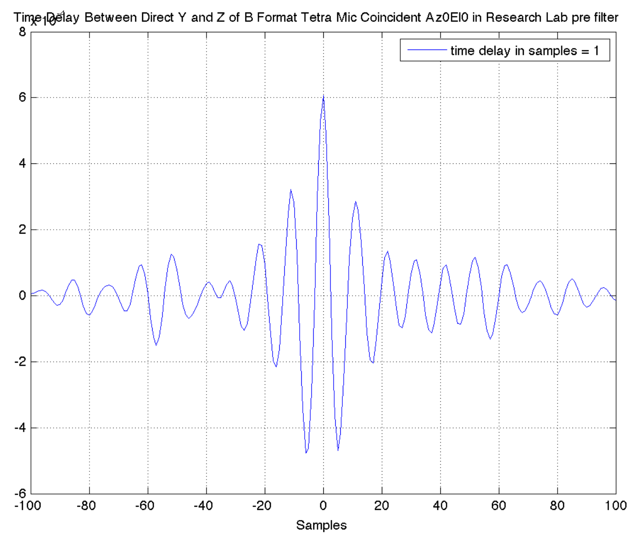


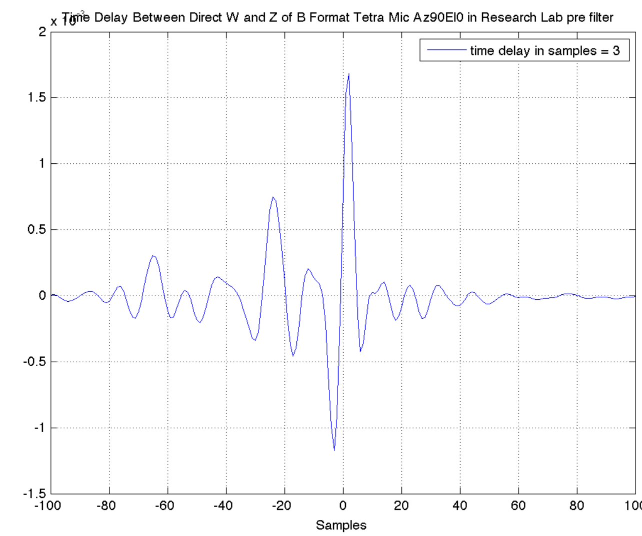




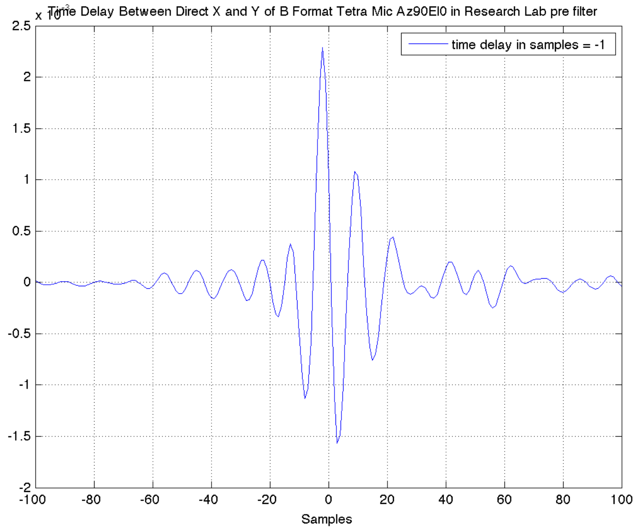


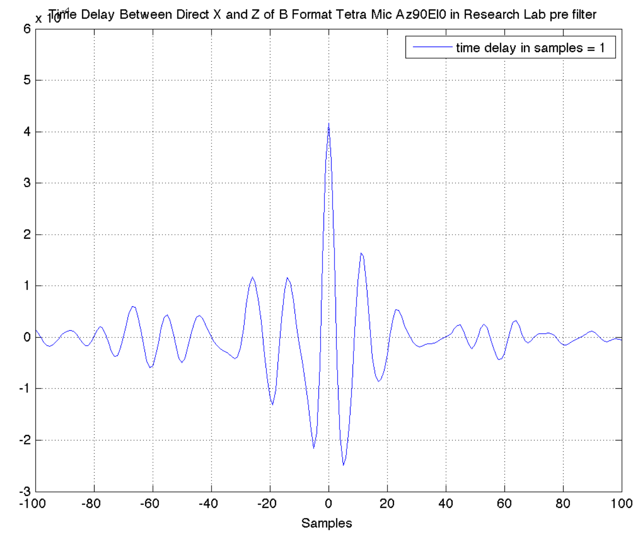


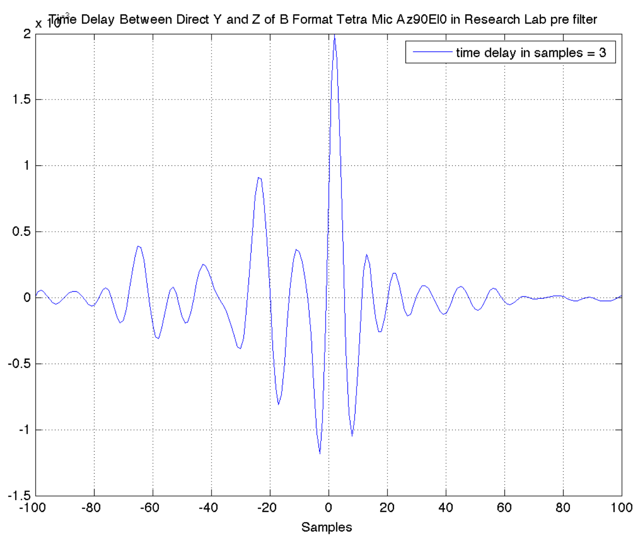


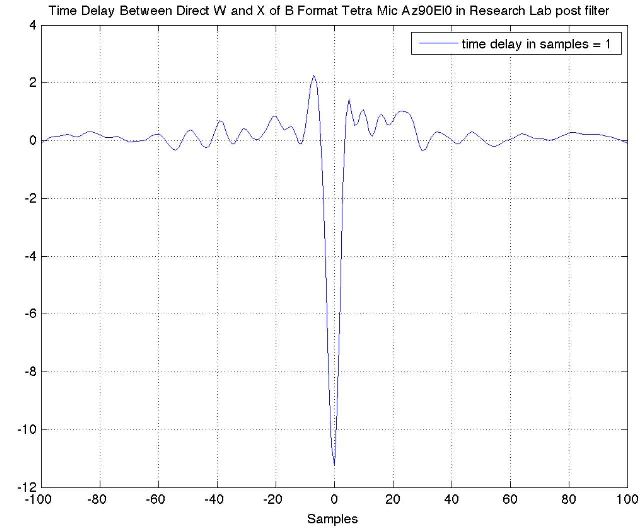


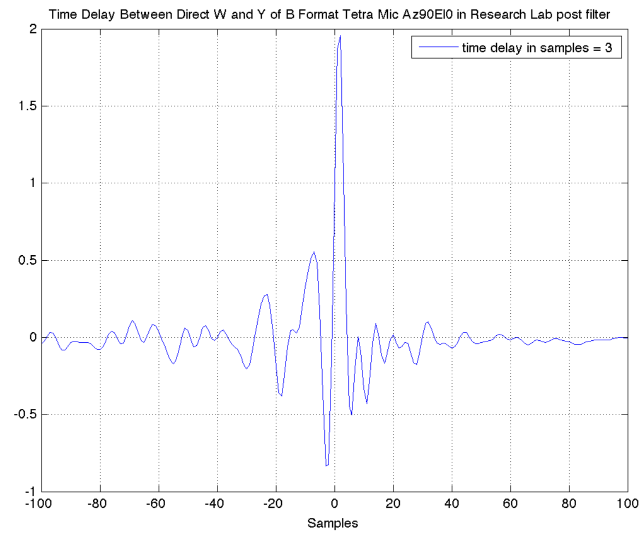
Physical Near Coincident Core Sound Tetra Mic Az90

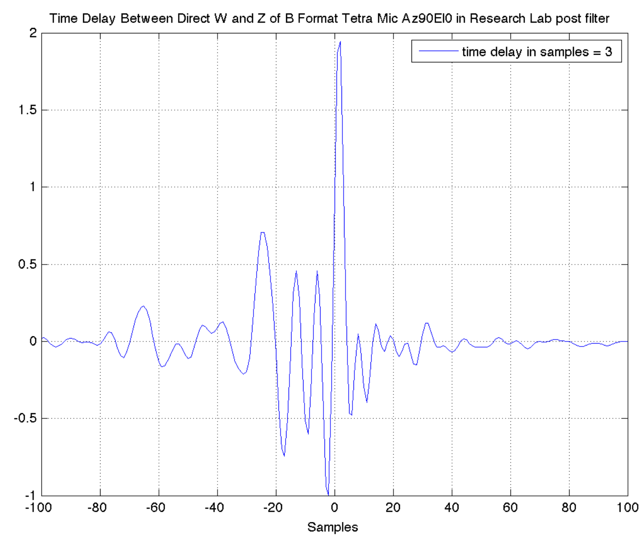

Time Delay Between Direct X and Y of B Format Tetra Mic Az90El0 in Research Lab pre filter
time delay in samples = -1
Samples

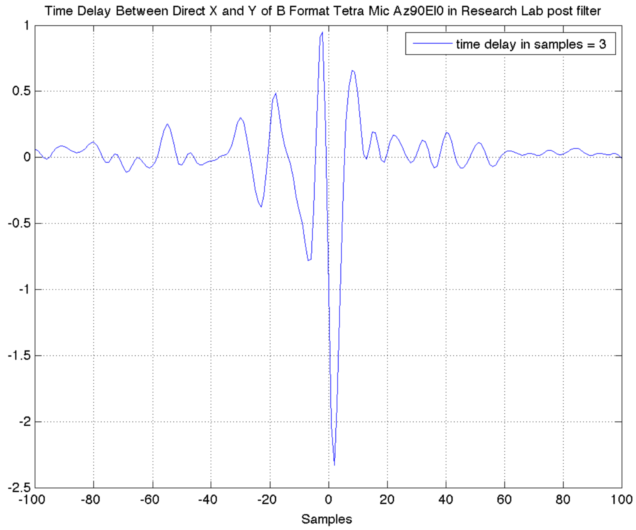

Time Delay Between Direct X and Z of B Format Tetra Mic Az90El0 in Research Lab pre filter
time delay in samples = 1
Samples

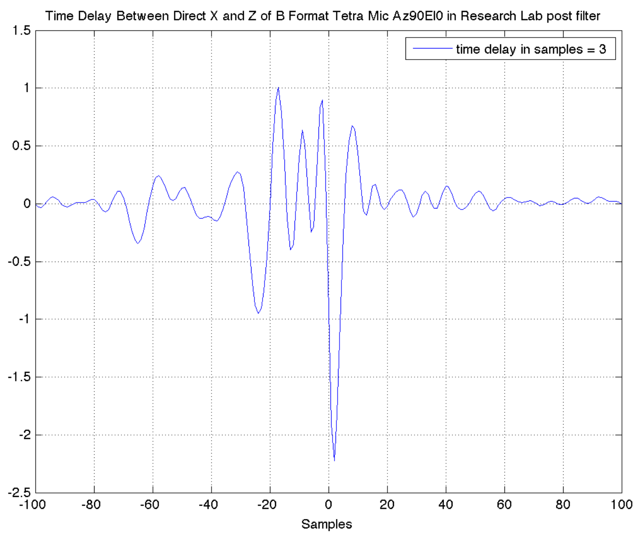

Time Delay Between Direct Y and Z of B Format Tetra Mic Az90El0 in Research Lab pre filter
time delay in samples = 3
Samples

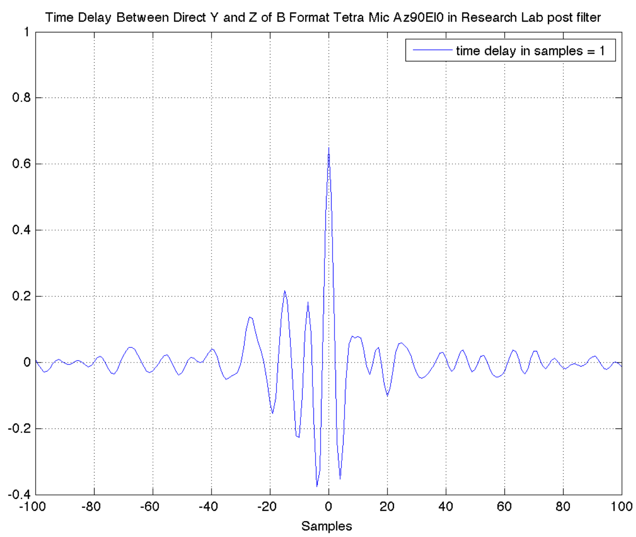

Time Delay Between Direct W and X of B Format Tetra Mic Az90El0 in Research Lab post filter
time delay in samples = 1
Samples

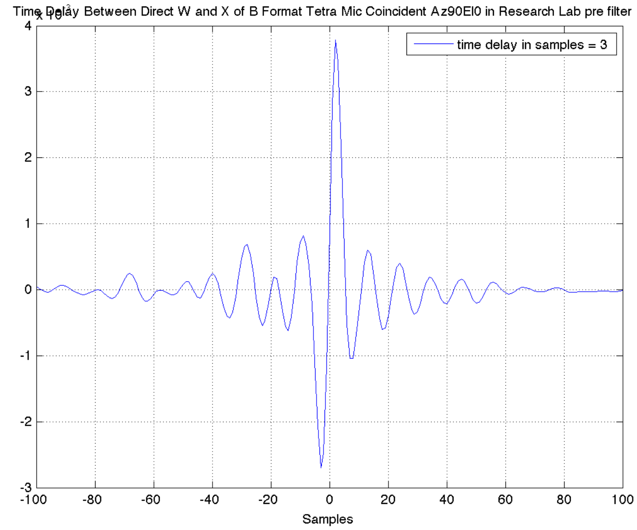

Time Delay Between Direct W and Y of B Format Tetra Mic Az90El0 in Research Lab post filter
time delay in samples = 3
Samples

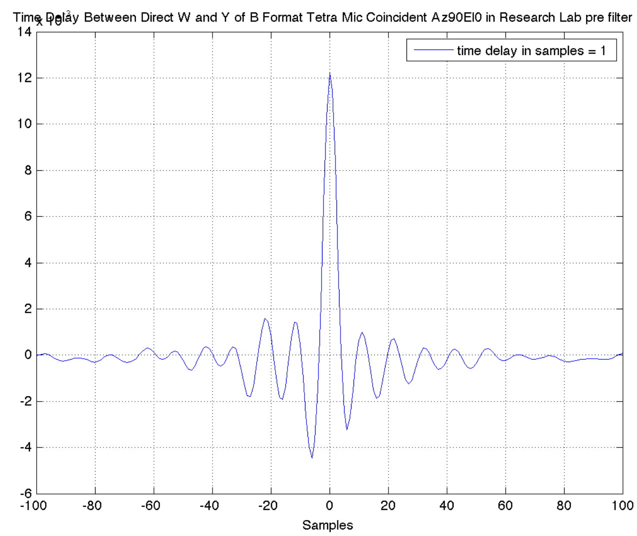

Time Delay Between Direct W and Z of B Format Tetra Mic Az90El0 in Research Lab post filter
time delay in samples = 3
Samples

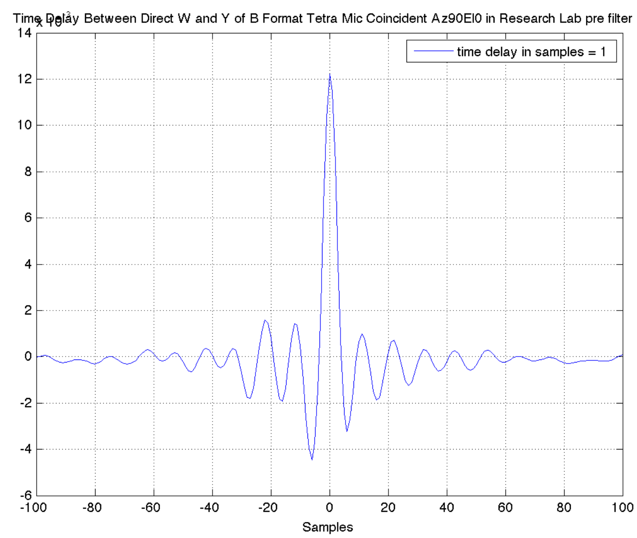




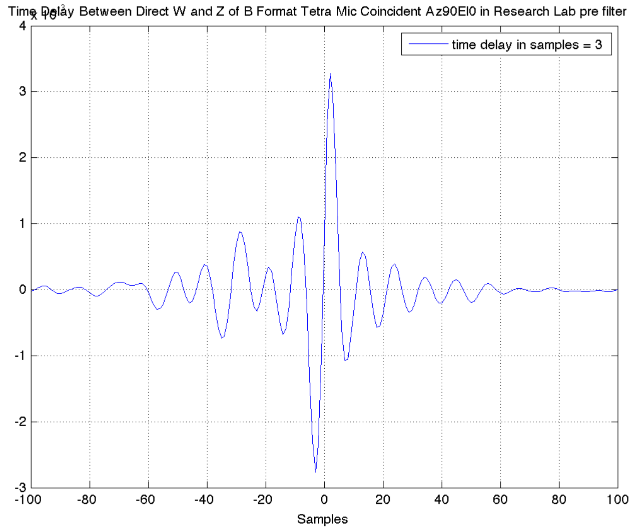


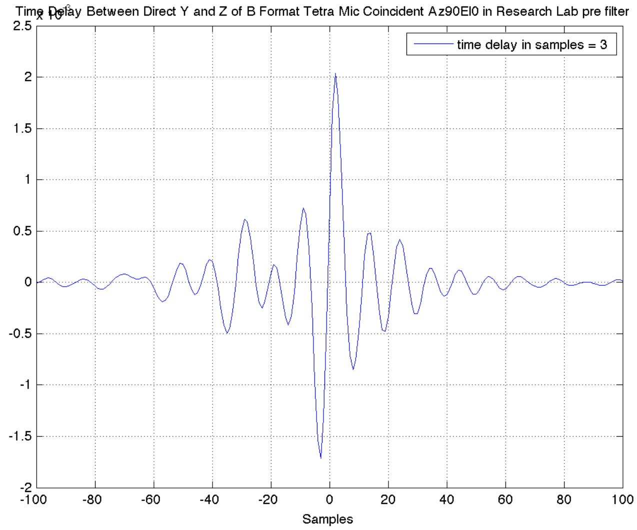


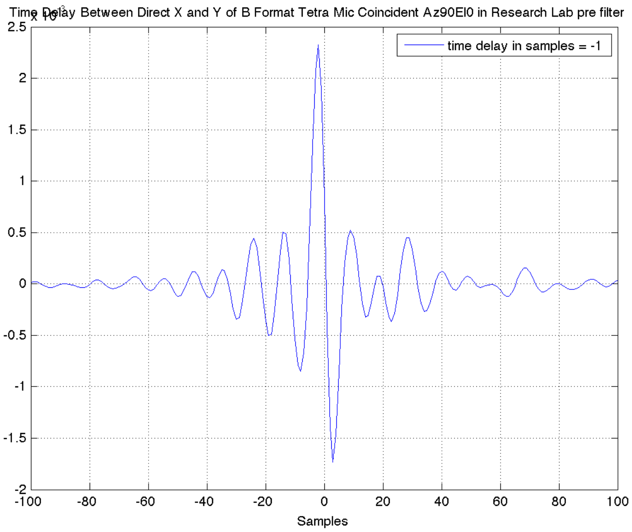


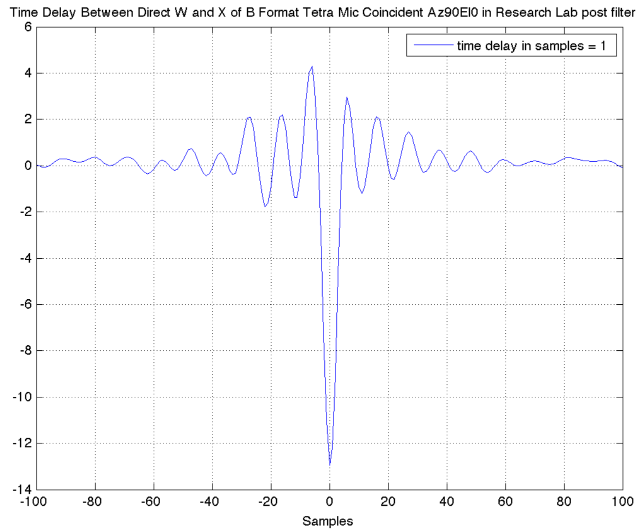


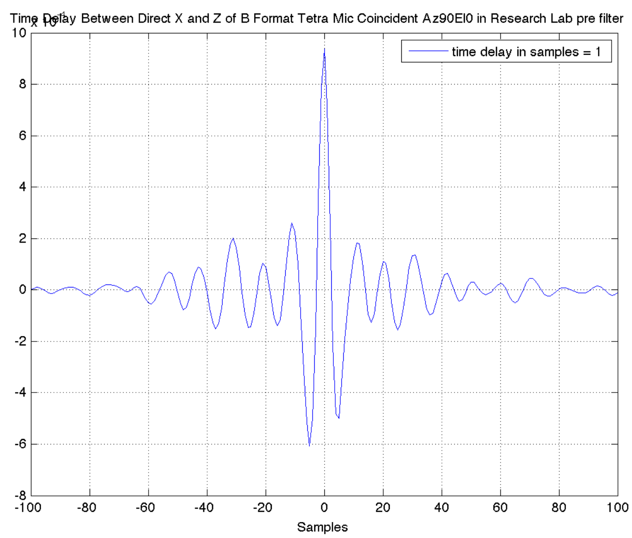
Virtual True Coincident Core Sound Tetra Mic Az90

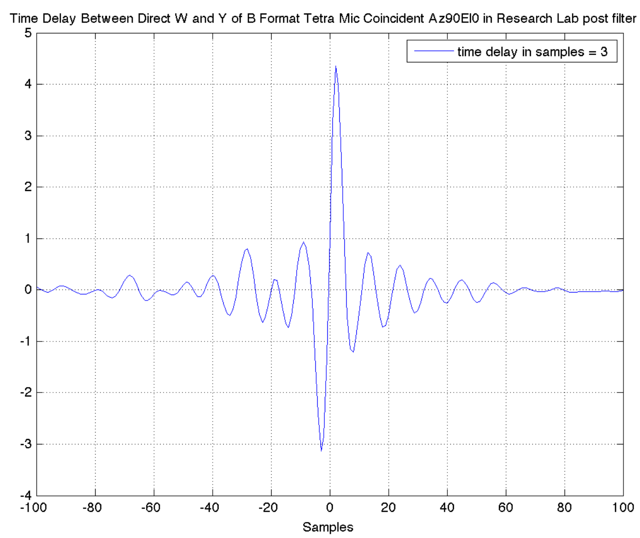

Time Delay Between Direct W and Z of B Format Tetra Mic Coincident Az90El0 in Research Lab pre filter
time delay in samples = 3
Samples


Time Delay Between Direct Y and Z of B Format Tetra Mic Coincident Az90El0 in Research Lab pre filter
time delay in samples = 3
Samples

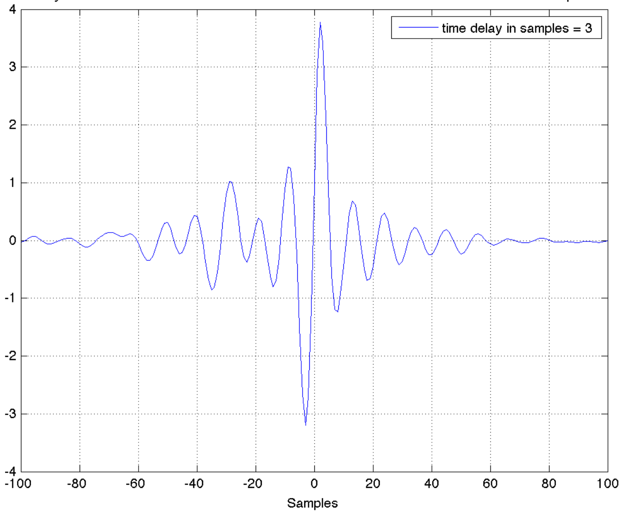

Time Delay Between Direct X and Y of B Format Tetra Mic Coincident Az90El0 in Research Lab pre filter
time delay in samples = -1
Samples


Time Delay Between Direct W and X of B Format Tetra Mic Coincident Az90El0 in Research Lab post filter
time delay in samples = 1
Samples

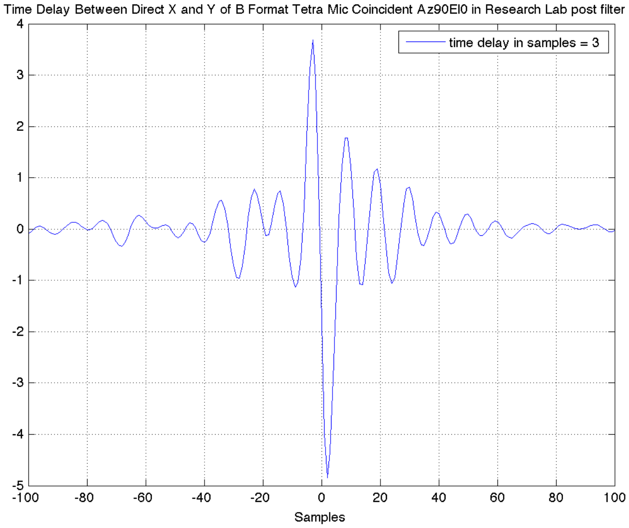

Time Delay Between Direct X and Z of B Format Tetra Mic Coincident Az90El0 in Research Lab pre filter
time delay in samples = 1
Samples


Time Delay Between Direct W and Y of B Format Tetra Mic Coincident Az90El0 in Research Lab post filter
time delay in samples = 3
Samples

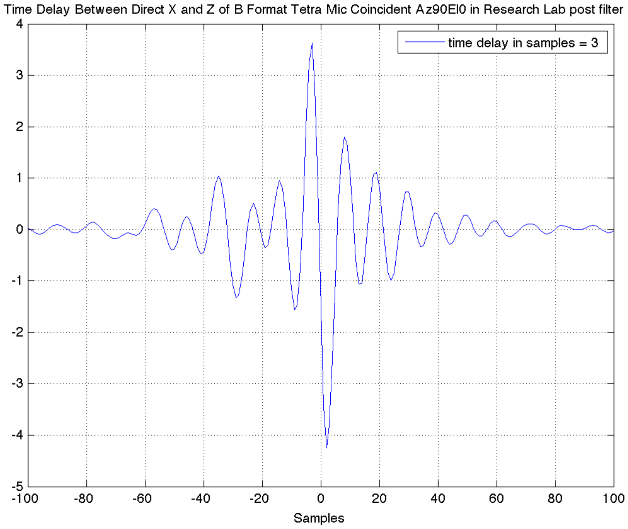
Time Delay Between Direct W and Z of B Format Tetra Mic Coincident Az90El0 in Research Lab post filter



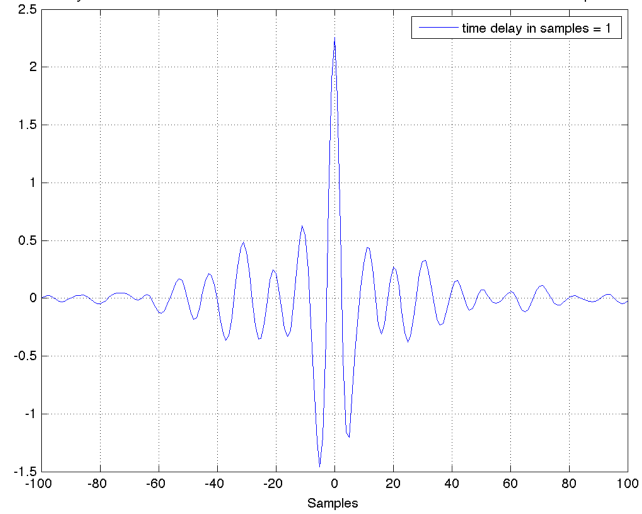
Time Delay Between Direct X and Y of B Format Tetra Mic Coincident Az90El0 in Research Lab post filter



Time Delay Between Direct X and Z of B Format Tetra Mic Coincident Az90El0 in Research Lab post filter

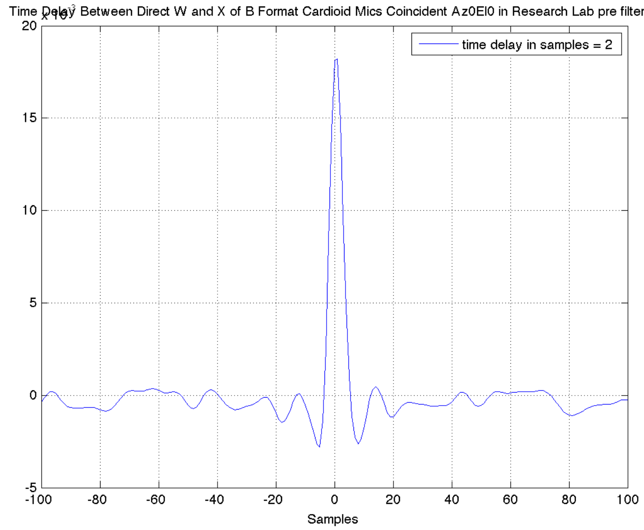


Time Delay Between Direct Y and Z of B Format Tetra Mic Coincident Az90El0 in Research Lab post filter

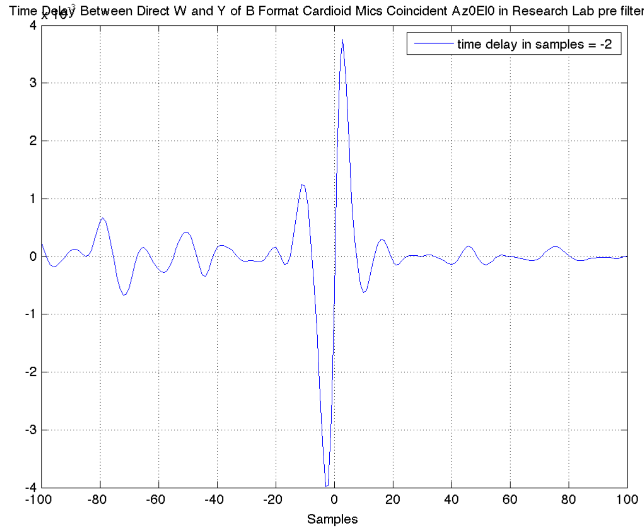


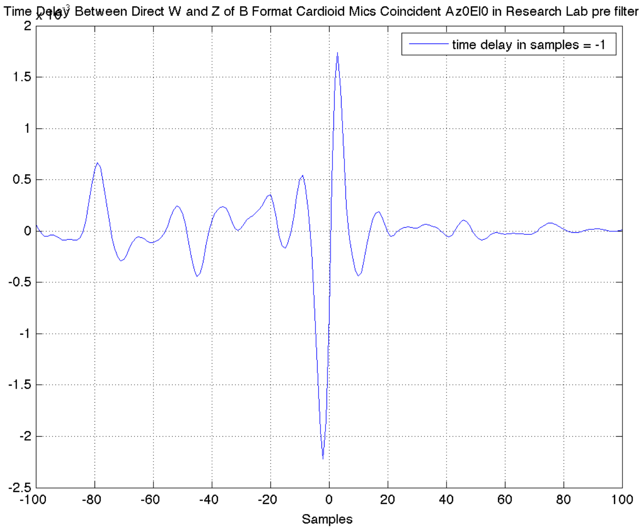
Virtual True Coincident DPA 4011A Tetra Configuration Az0

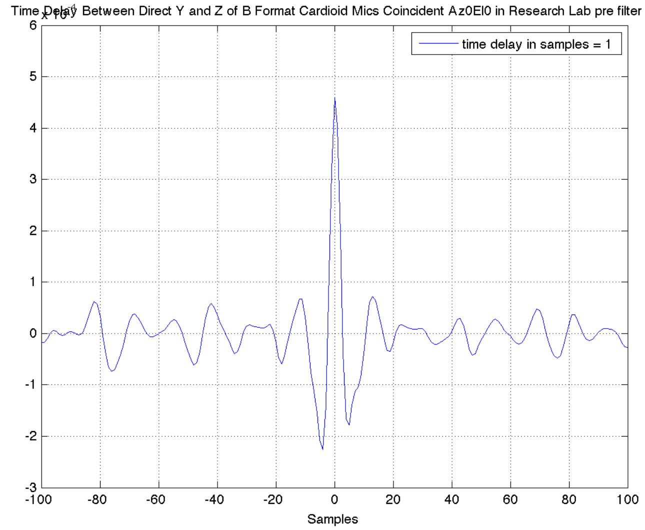
Time Delay Between Direct W and X of B Format Cardioid Mics Coincident Az0El0 in Research Lab pre filter

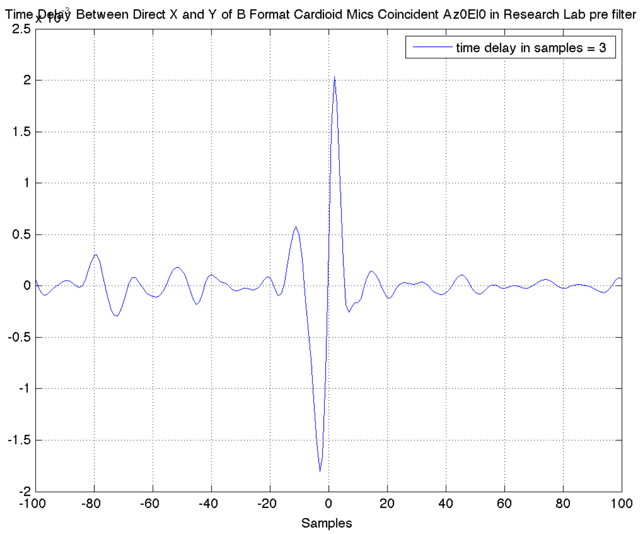


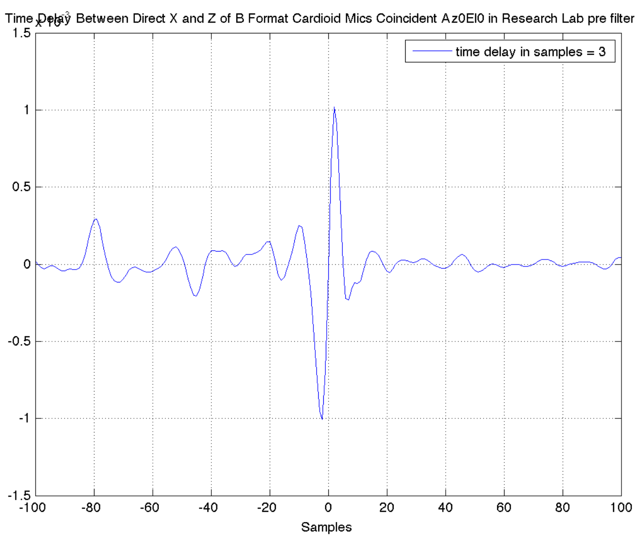
Time Delay Between Direct W and Y of B Format Cardioid Mics Coincident Az0El0 in Research Lab pre filter



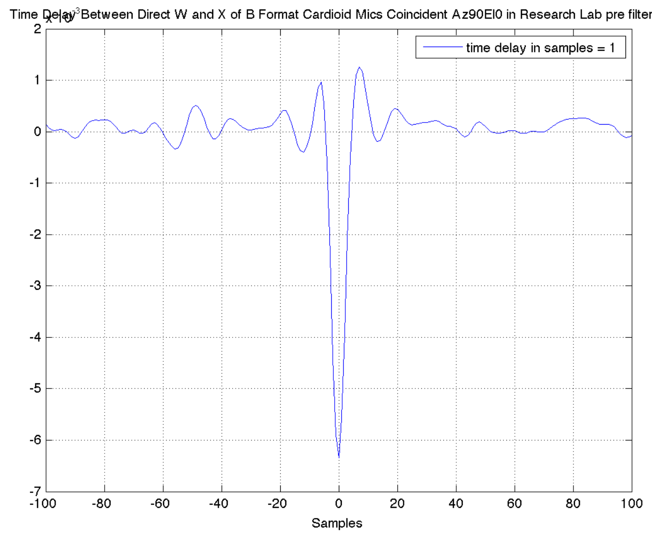


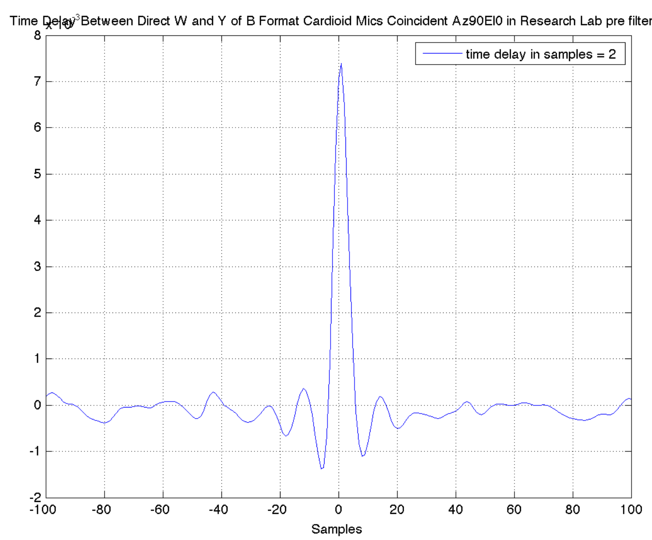


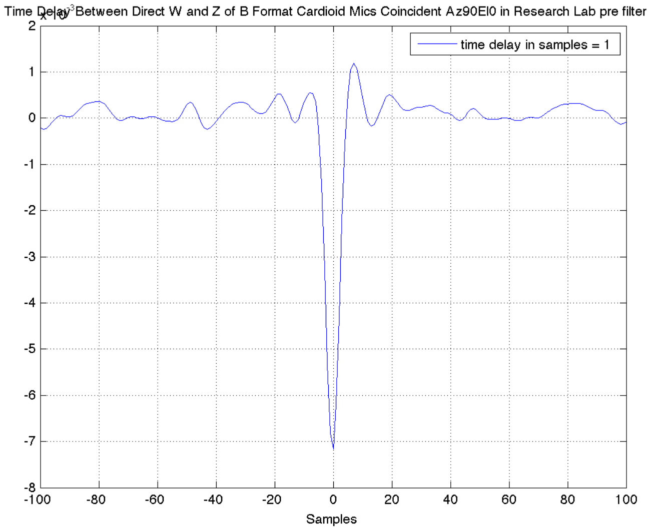


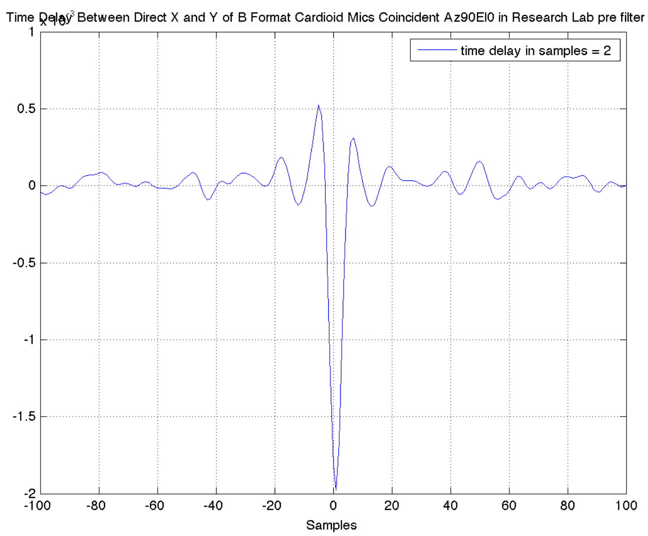


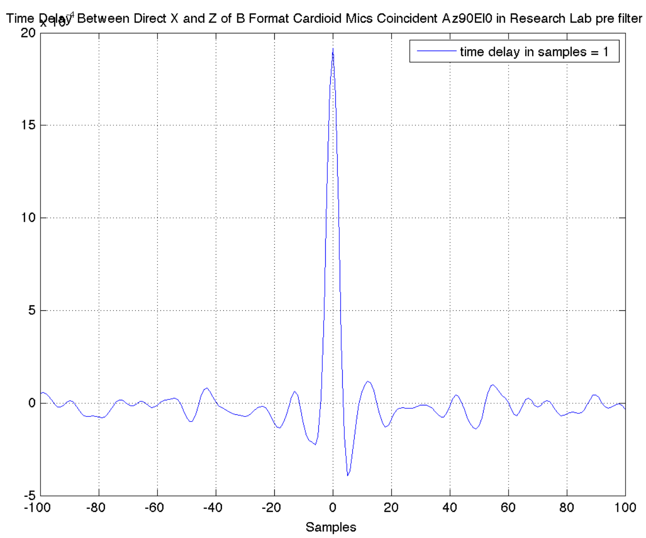
Virtual True Coincident DPA 4011A Tetra Configuration Az90

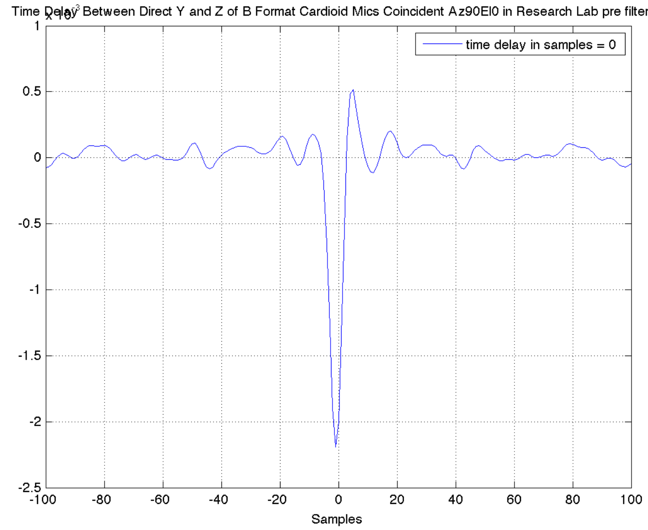




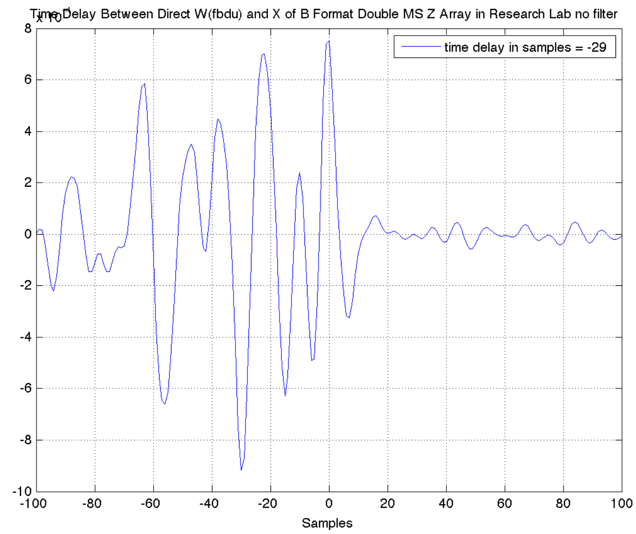


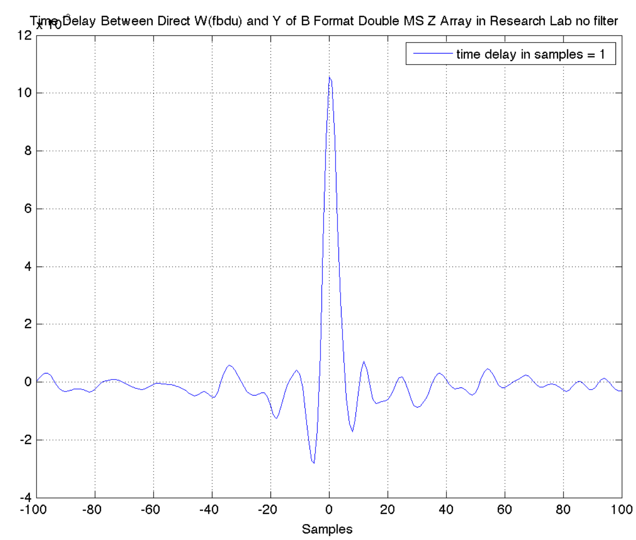


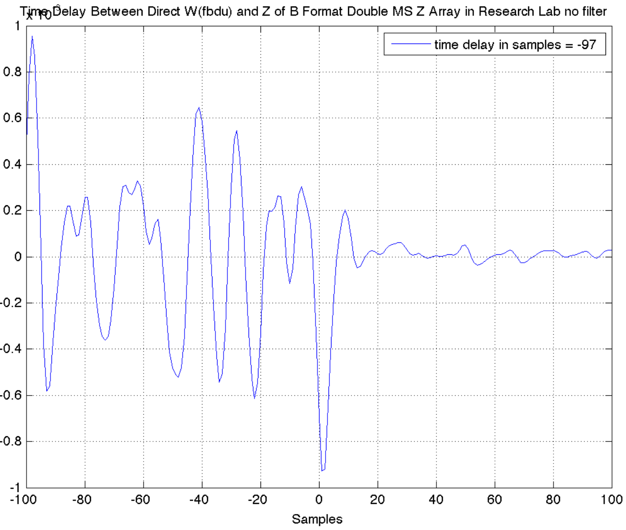


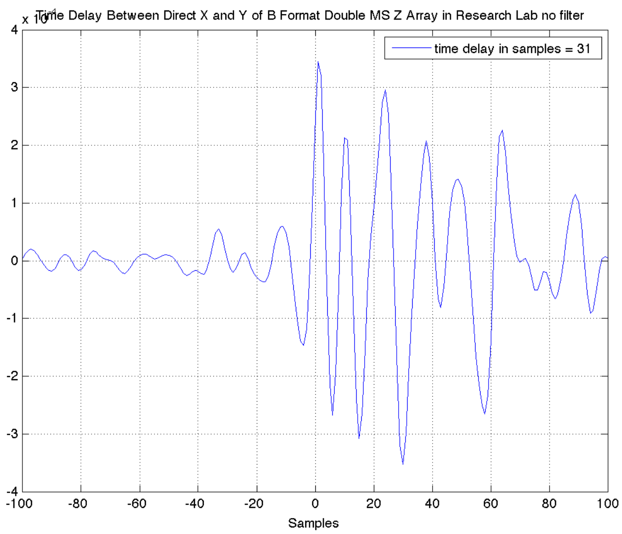


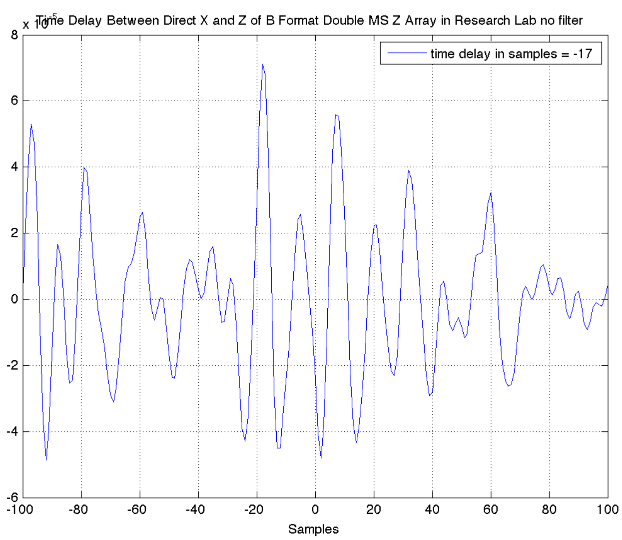
Physical Near Coincident Double MS Z Configuration Sennheiser Twins Az0

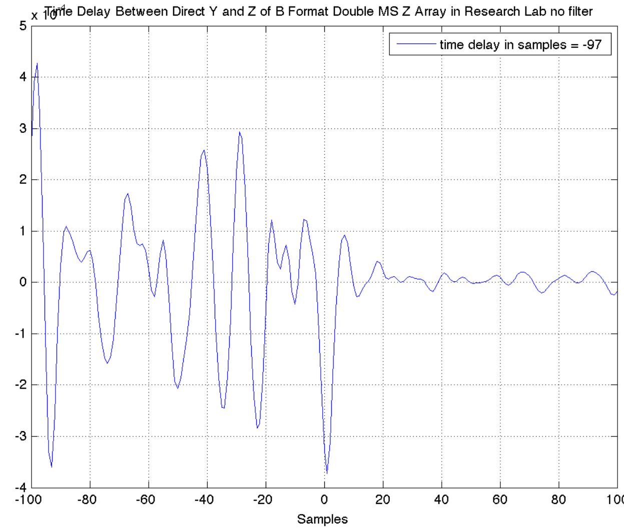




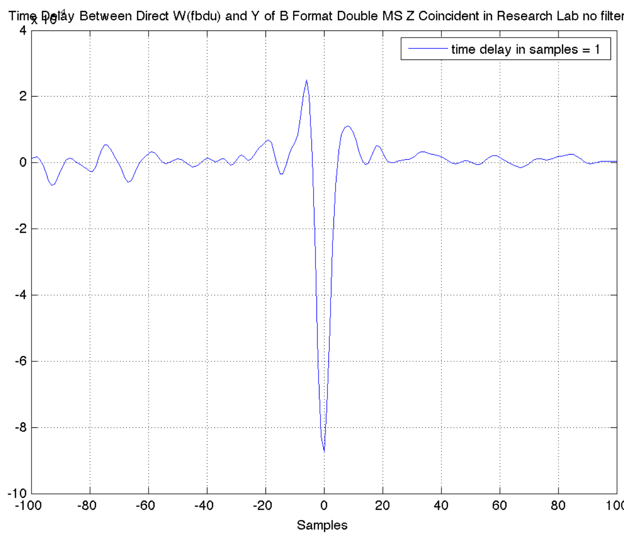


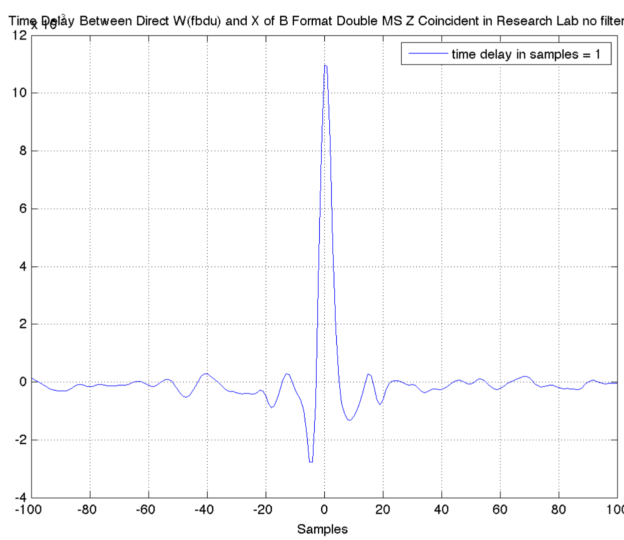


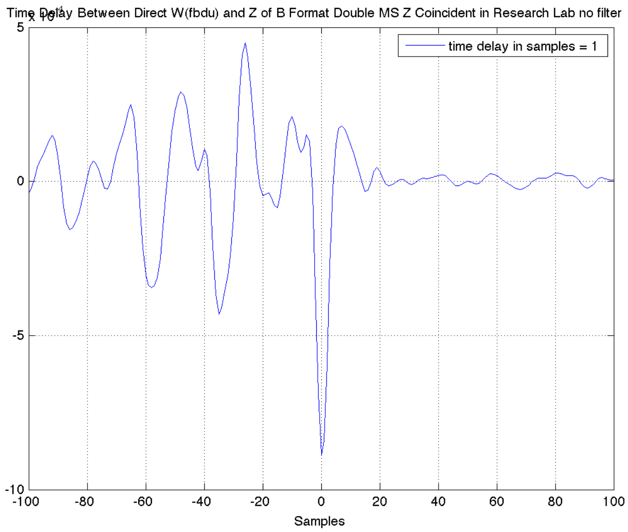


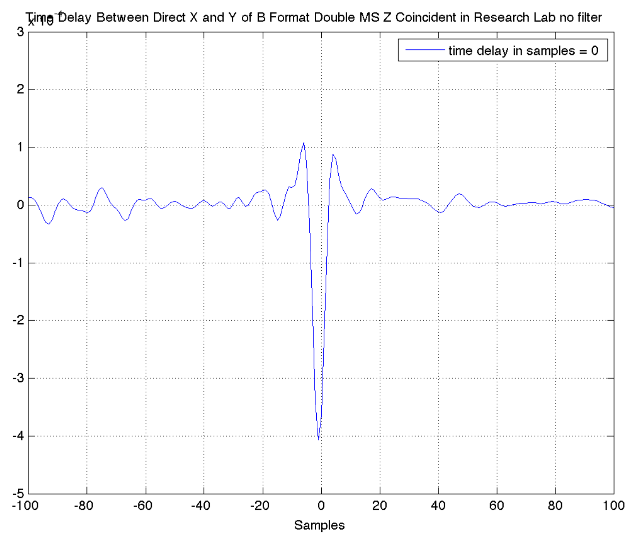


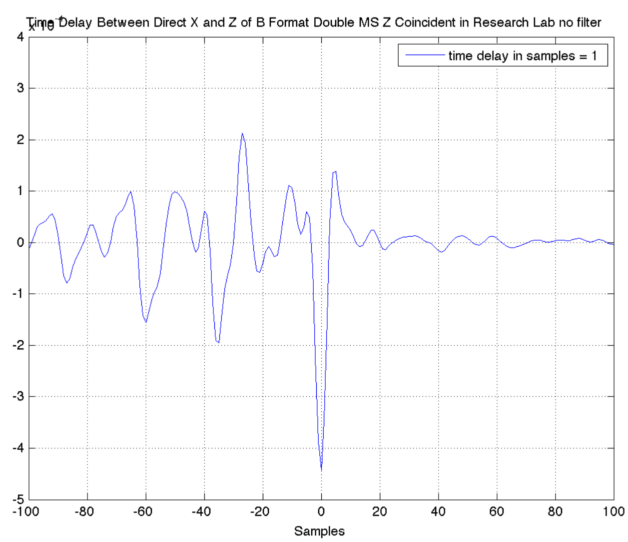
Virtual True Coincident Double MS Z Configuration Sennheiser Twins Az0

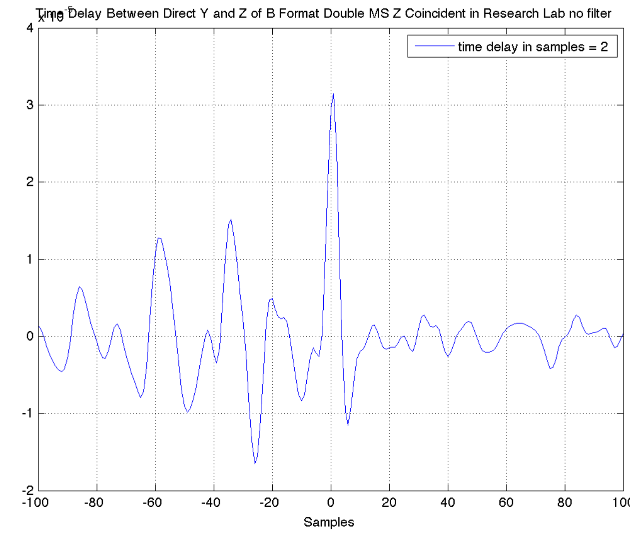






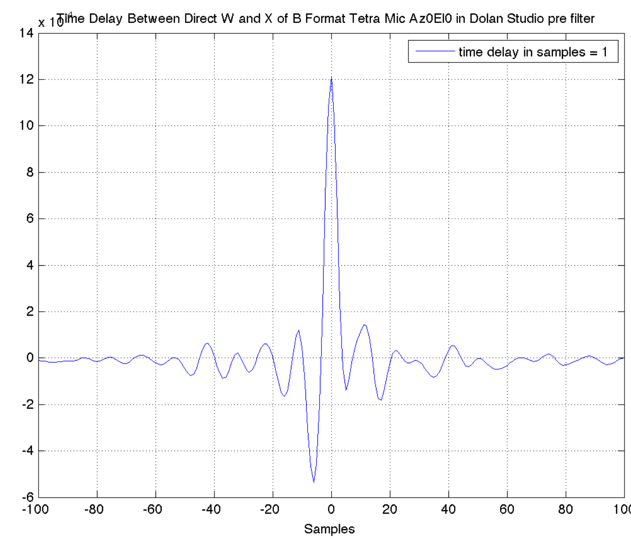


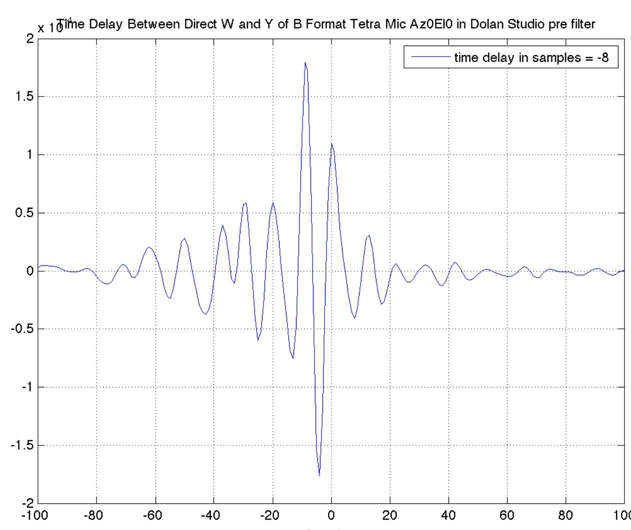


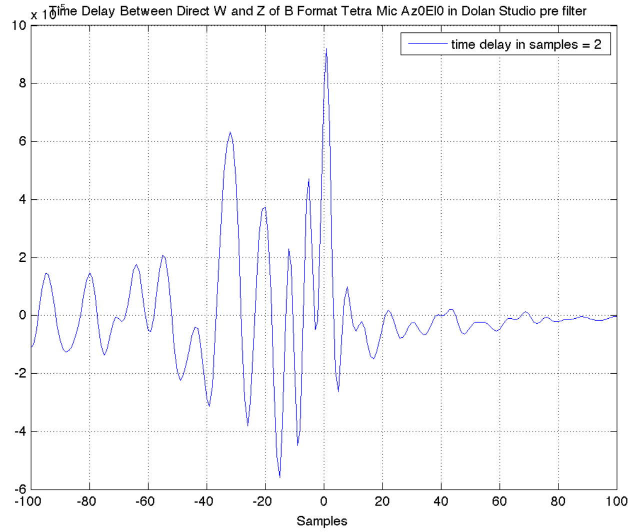


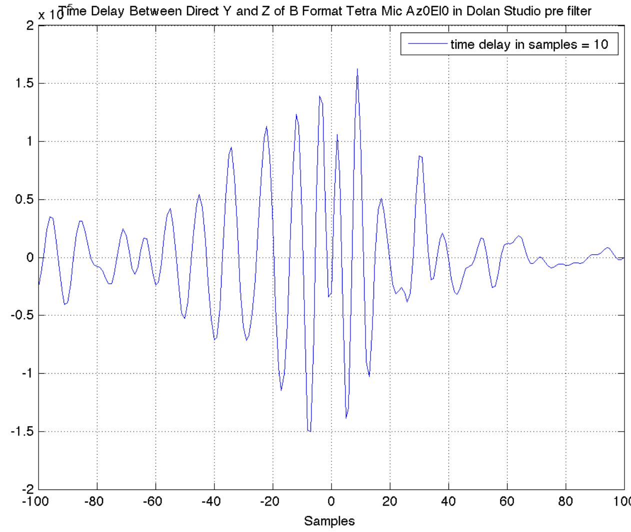
Dolan Studios

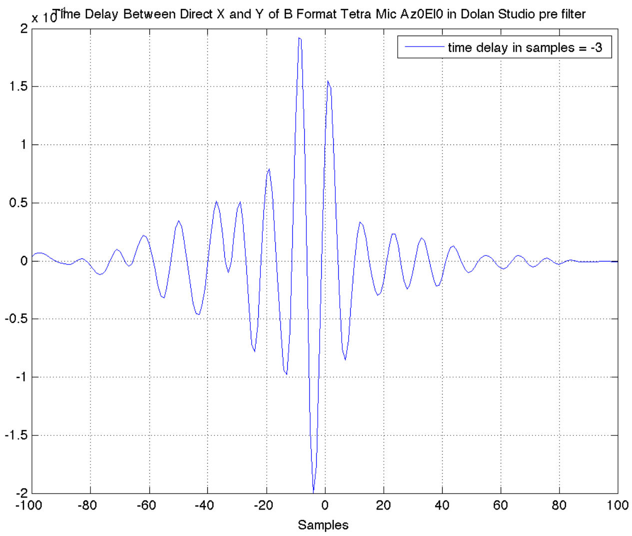
Physical Near Coincident Core Sound Tetra Mic Az0

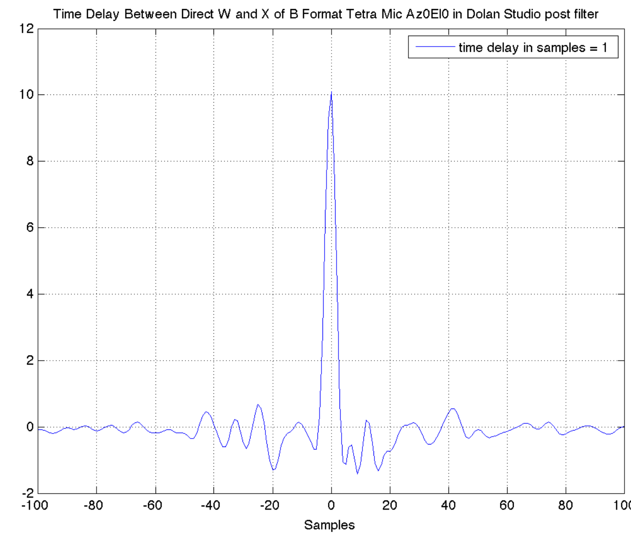


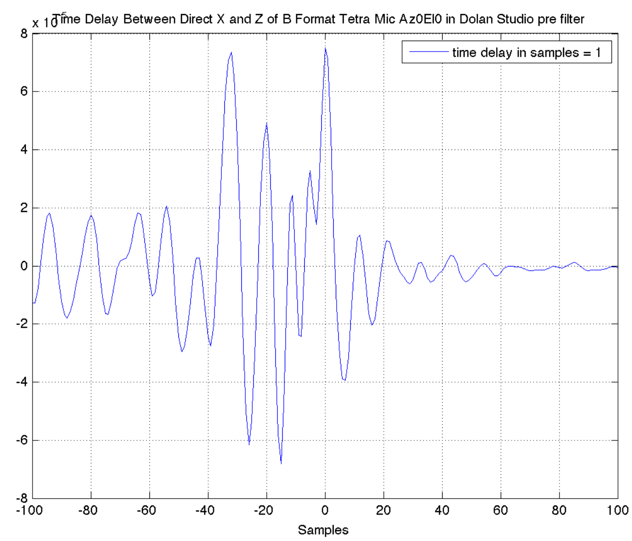


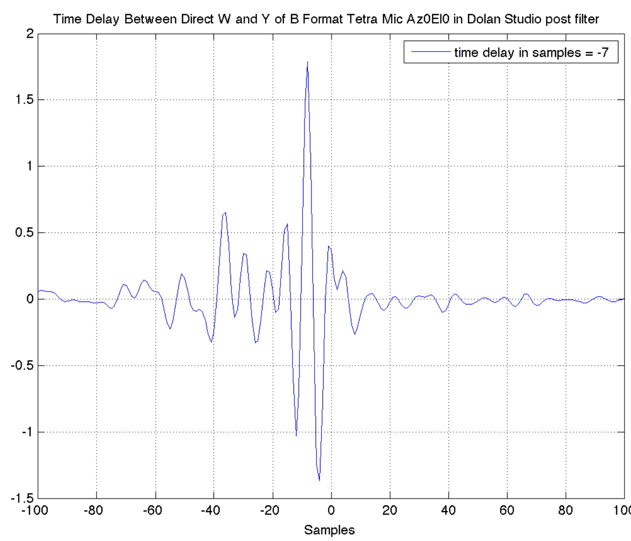

Time Delay Between Direct W and Z of B Format Tetra Mic Az0El0 in Dolan Studio pre filter
time delay in samples = 2
Samples

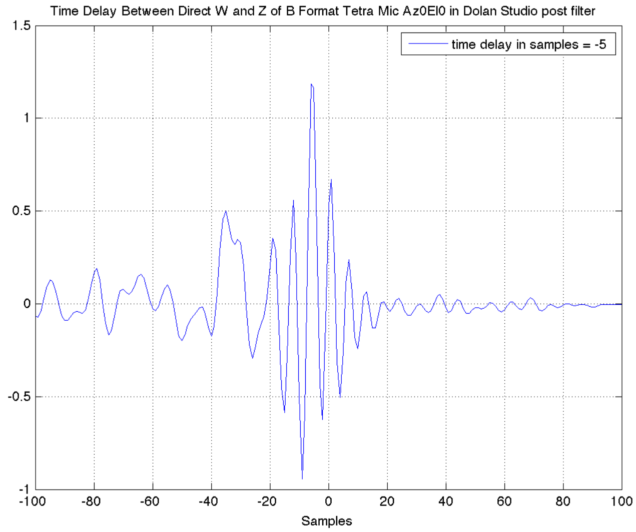

Time Delay Between Direct Y and Z of B Format Tetra Mic Az0El0 in Dolan Studio pre filter
time delay in samples = 10
Samples

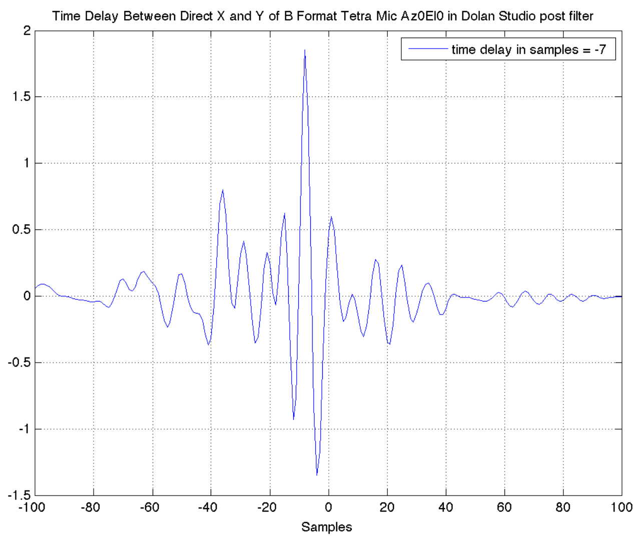

Time Delay Between Direct X and Y of B Format Tetra Mic Az0El0 in Dolan Studio pre filter
time delay in samples = -3
Samples

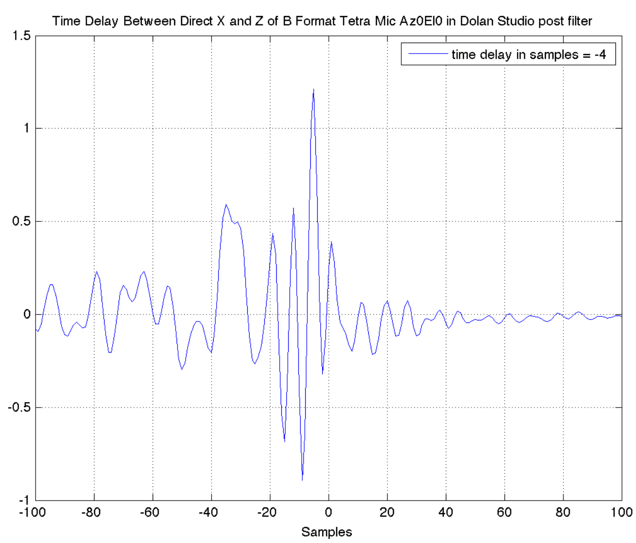

Time Delay Between Direct W and X of B Format Tetra Mic Az0El0 in Dolan Studio post filter
time delay in samples = 1
Samples

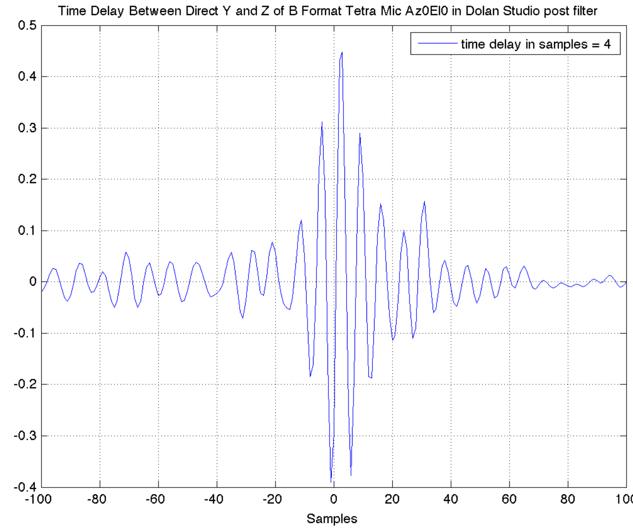

Time Delay Between Direct X and Z of B Format Tetra Mic Az0El0 in Dolan Studio pre filter
time delay in samples = 1
Samples


Time Delay Between Direct W and Y of B Format Tetra Mic Az0El0 in Dolan Studio post filter
time delay in samples = -7
Samples

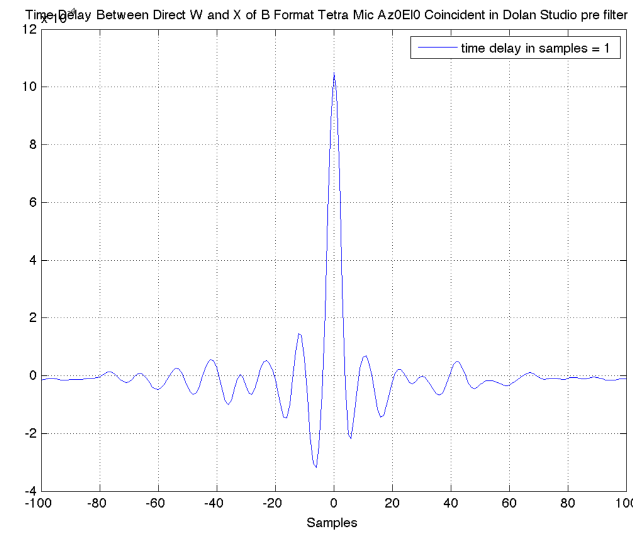


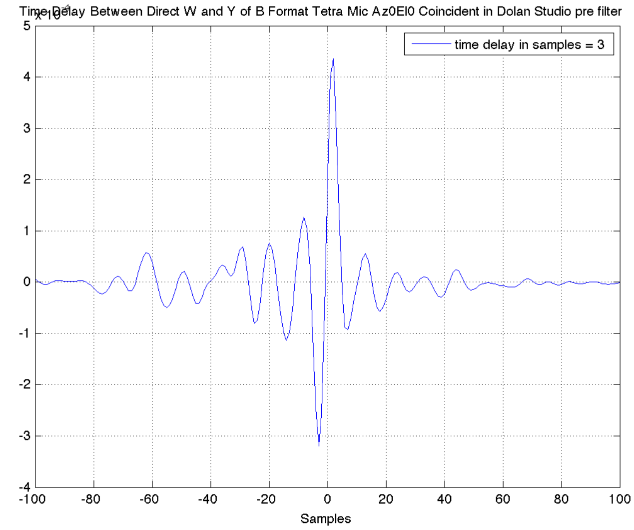


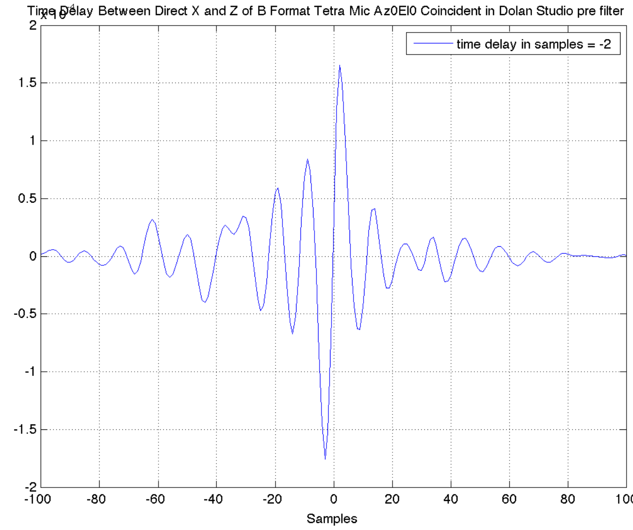


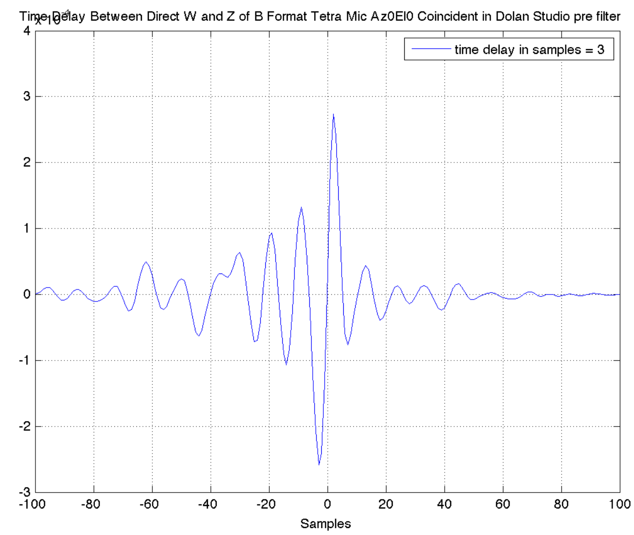


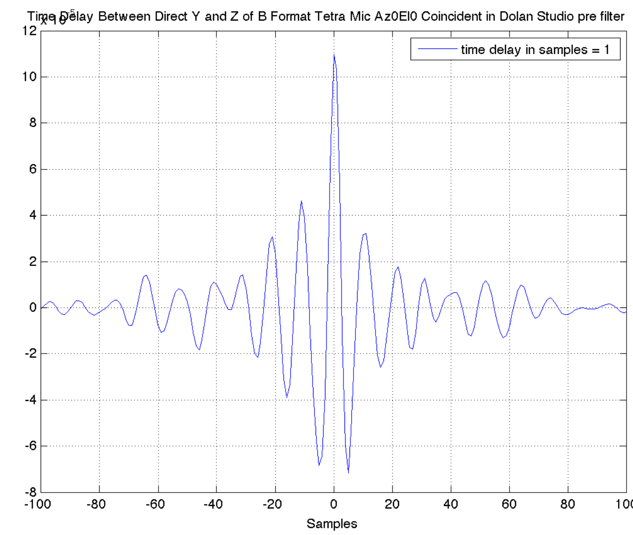
Virtual True Coincident Core Sound Tetra Mic Az0

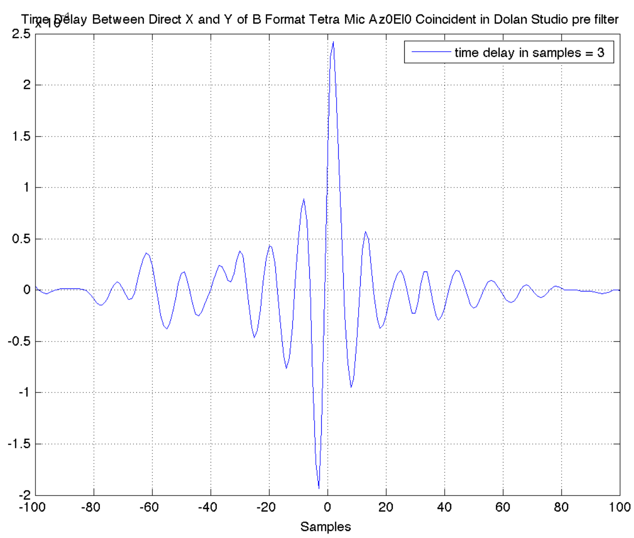




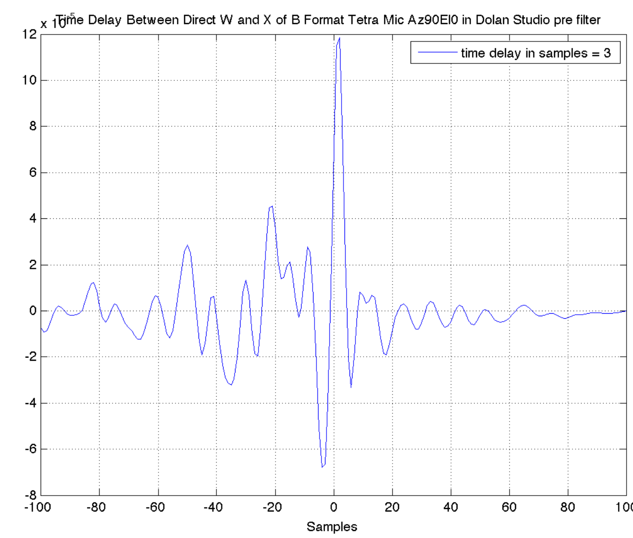


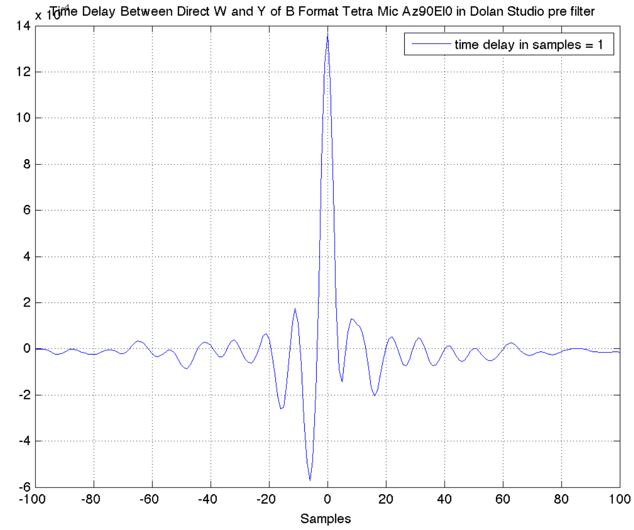


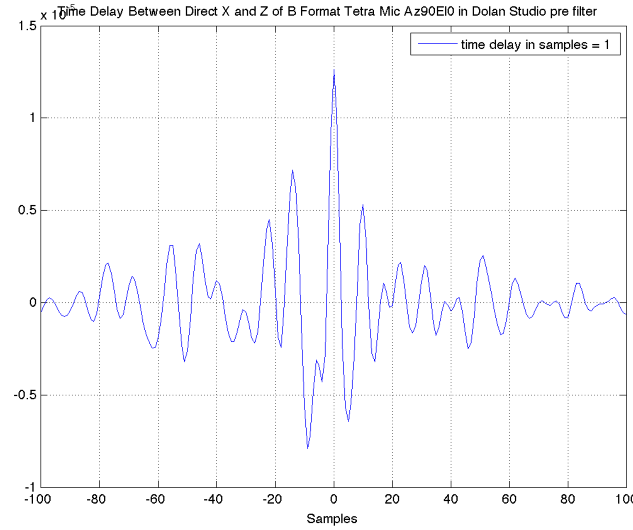


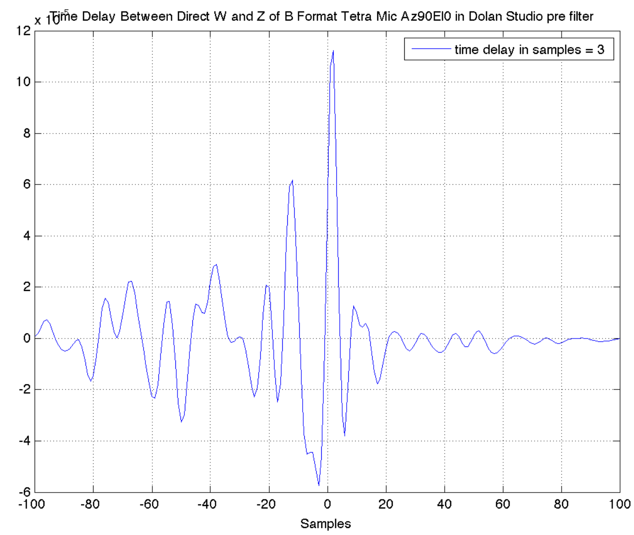


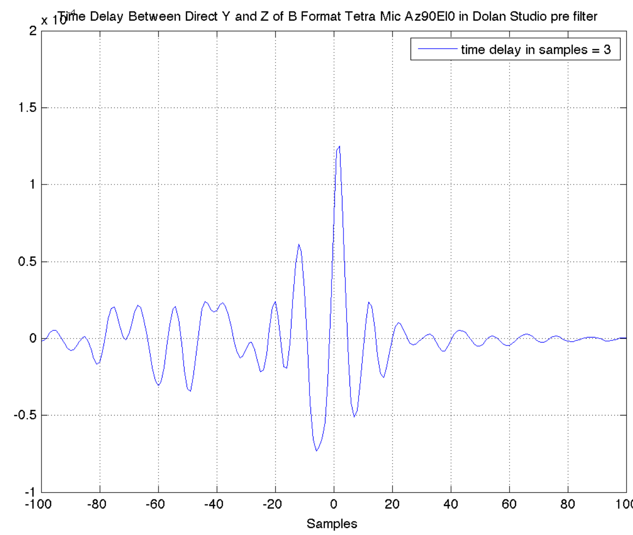
Physical Near Coincident Core Sound Tetra Mic Az90

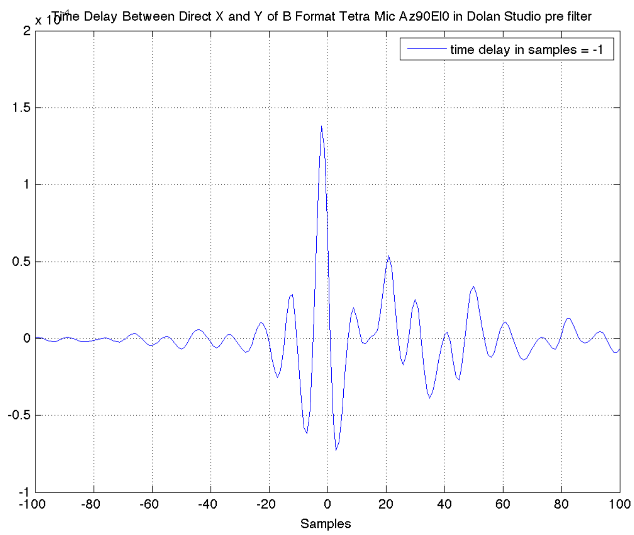


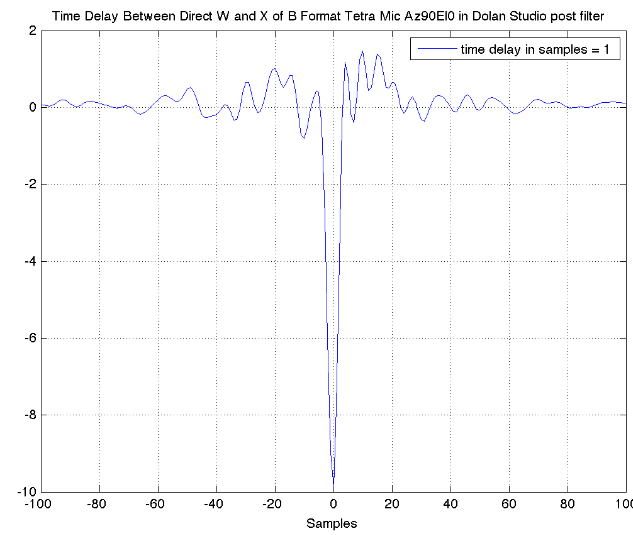

Time Delay Between Direct W and Y of B Format Tetra Mic Az90El0 in Dolan Studio pre filter
time delay in samples = 1
Samples

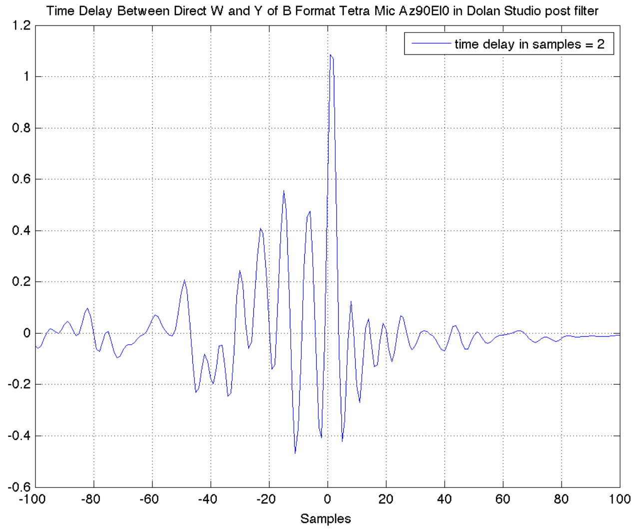

Time Delay Between Direct X and Z of B Format Tetra Mic Az90El0 in Dolan Studio pre filter
time delay in samples = 1
Samples

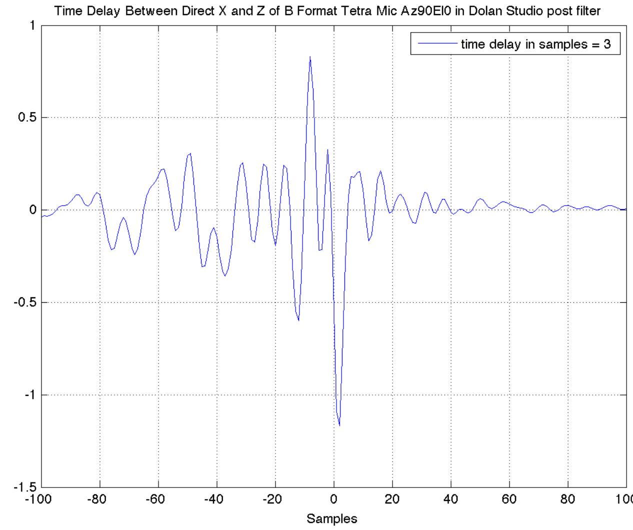

Time Delay Between Direct W and Z of B Format Tetra Mic Az90El0 in Dolan Studio pre filter
time delay in samples = 3
Samples

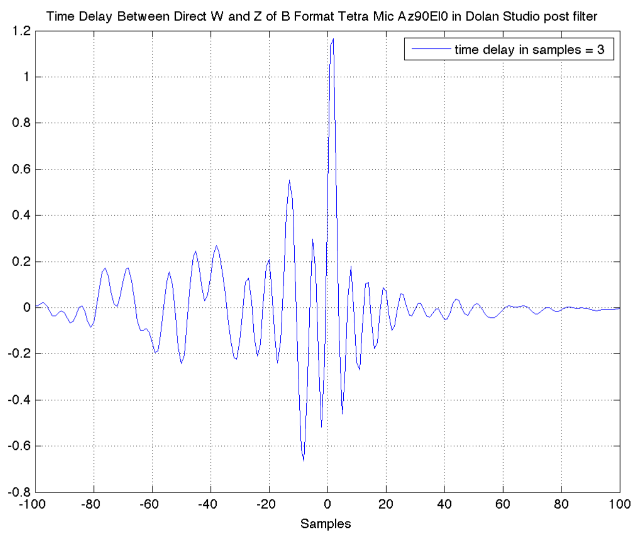

Time Delay Between Direct Y and Z of B Format Tetra Mic Az90El0 in Dolan Studio pre filter
time delay in samples = 3
Samples

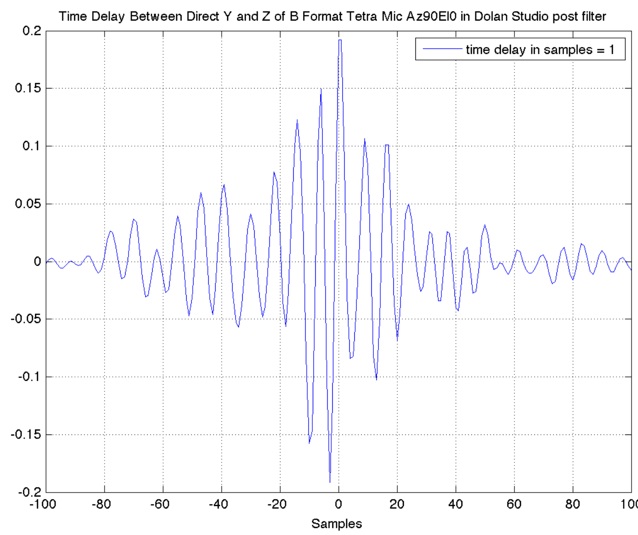

Time Delay Between Direct X and Y of B Format Tetra Mic Az90El0 in Dolan Studio pre filter
time delay in samples = -1
Samples

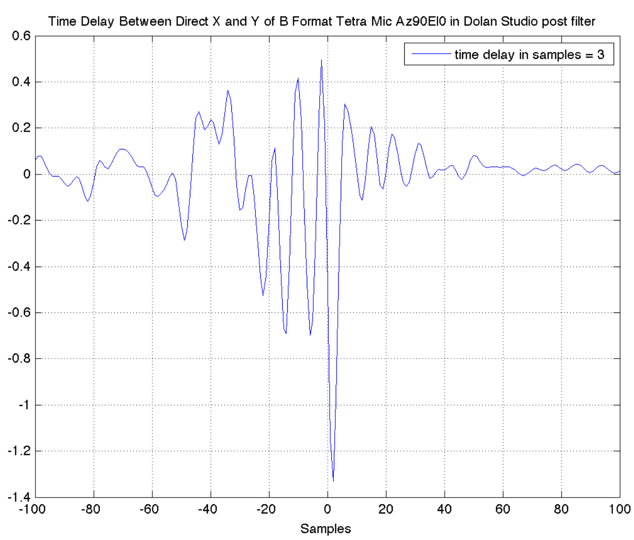

Time Delay Between Direct W and X of B Format Tetra Mic Az90El0 in Dolan Studio post filter
time delay in samples = 1
Samples



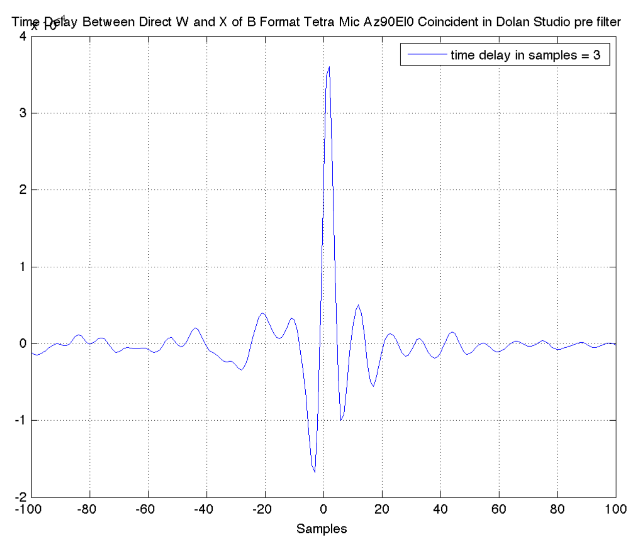


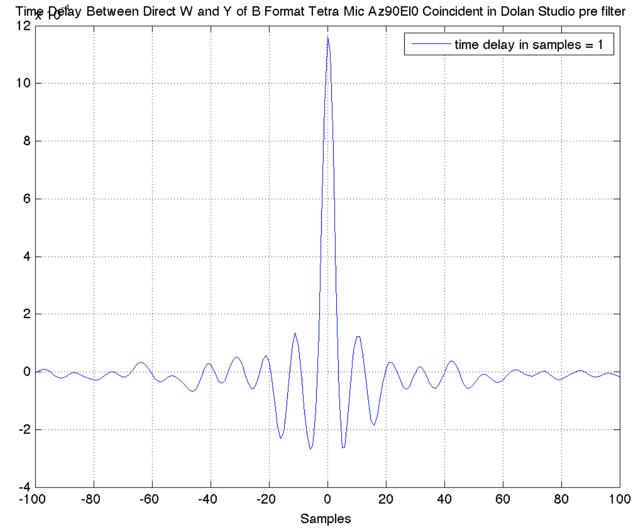


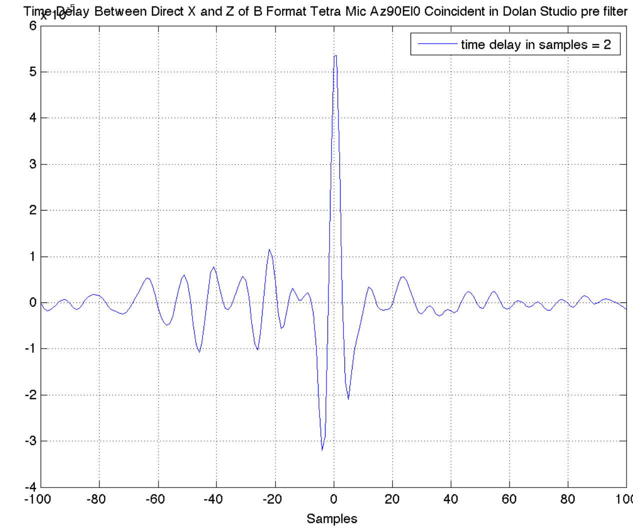


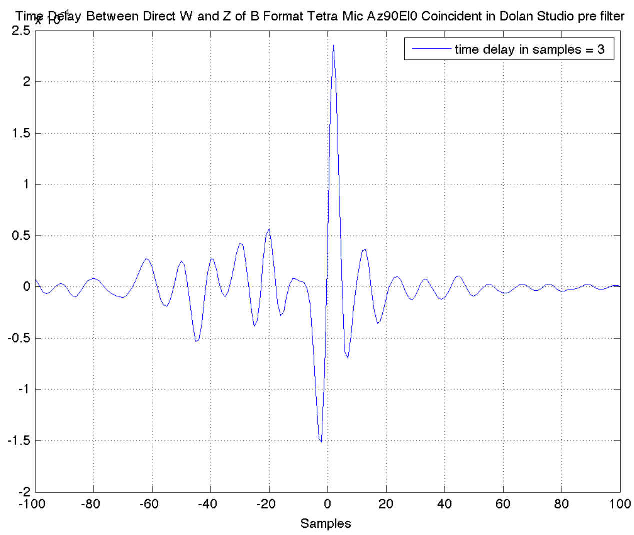


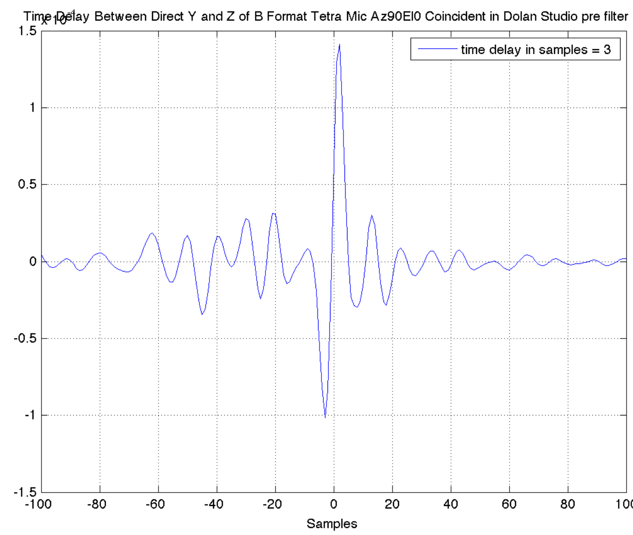
Virtual True Coincident Core Sound Tetra Mic Az90

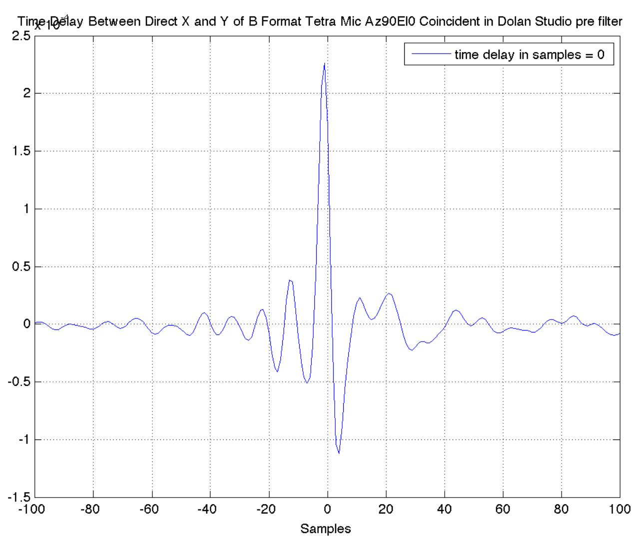




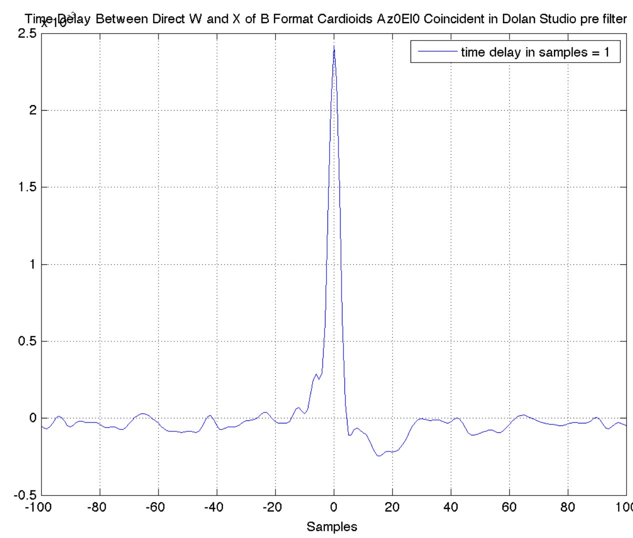


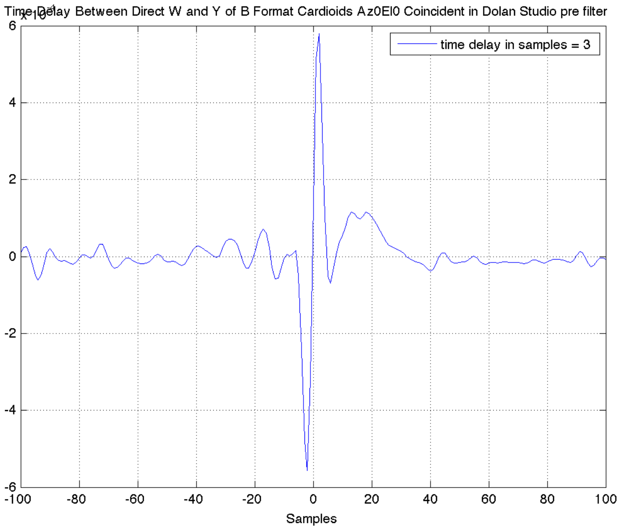


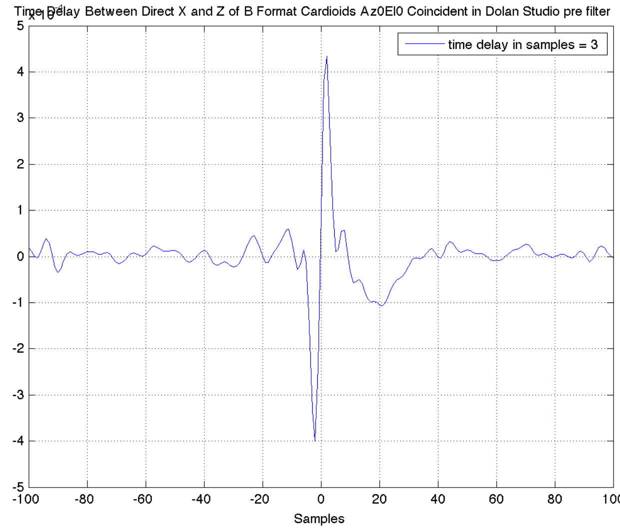


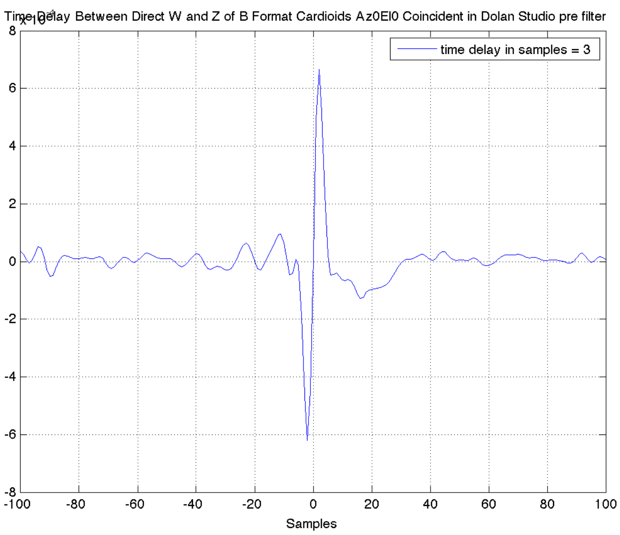


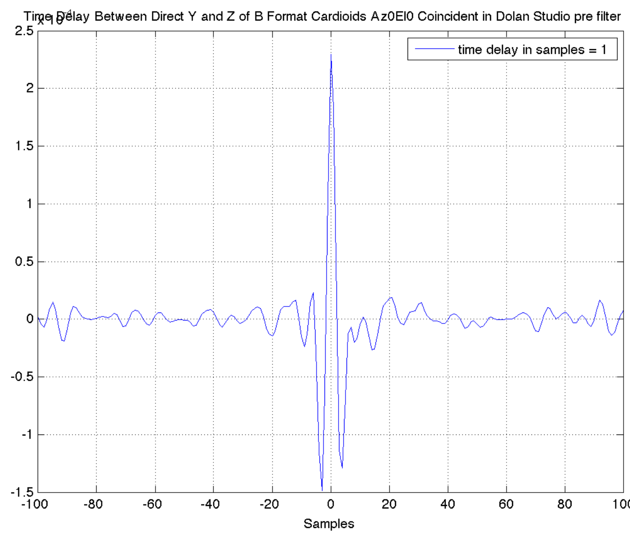
Virtual True Coincident DPA 4011A Tetra Configuration Az0

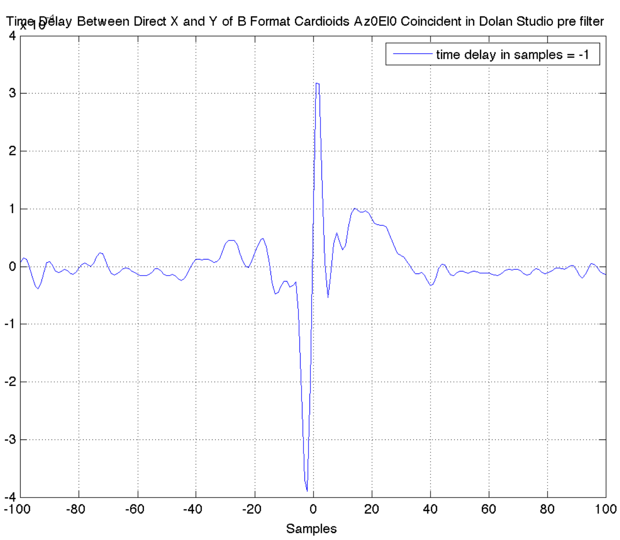




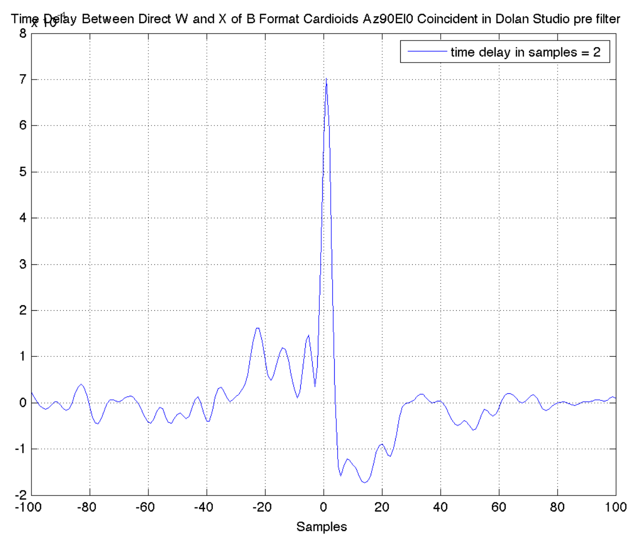


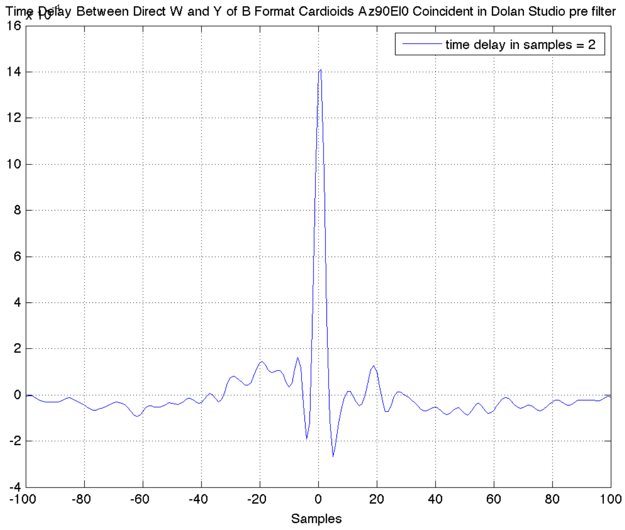


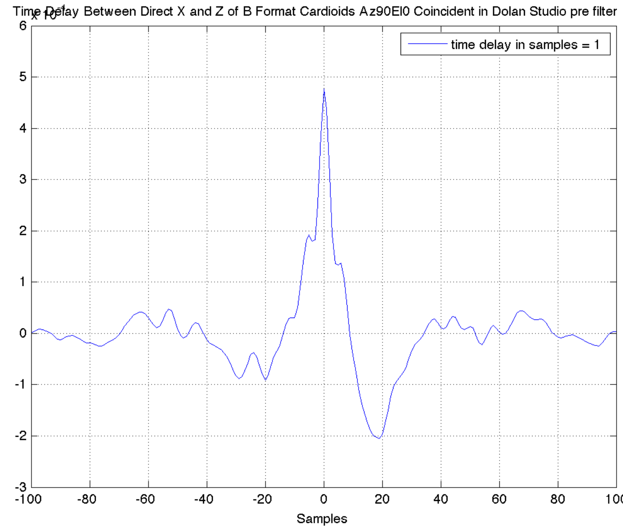


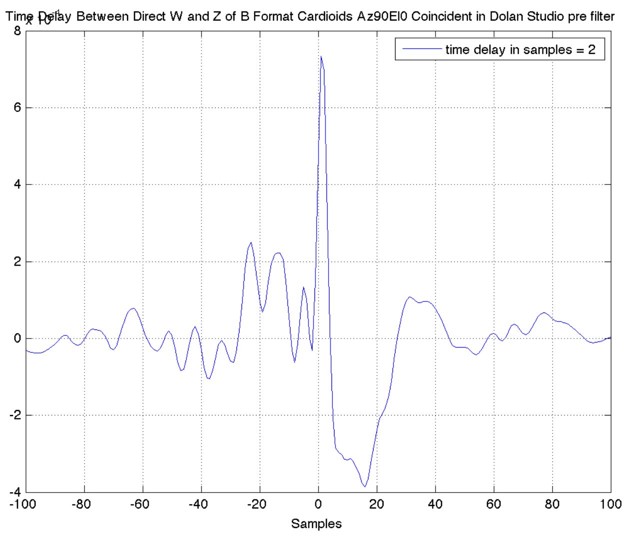


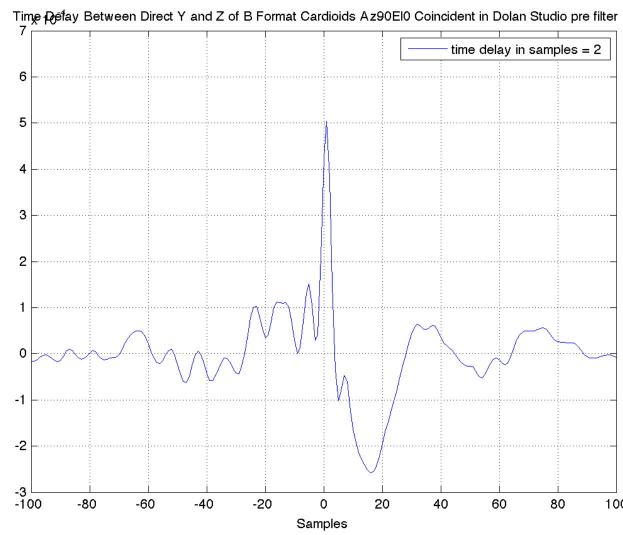
Virtual True Coincident DPA 4011A Tetra Configuration Az90

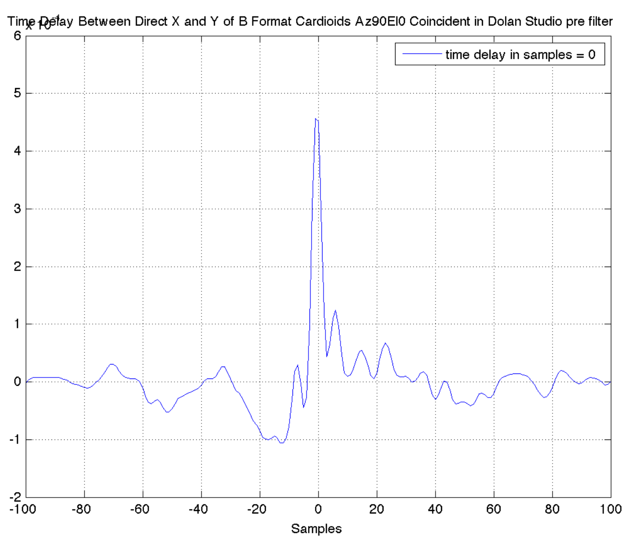




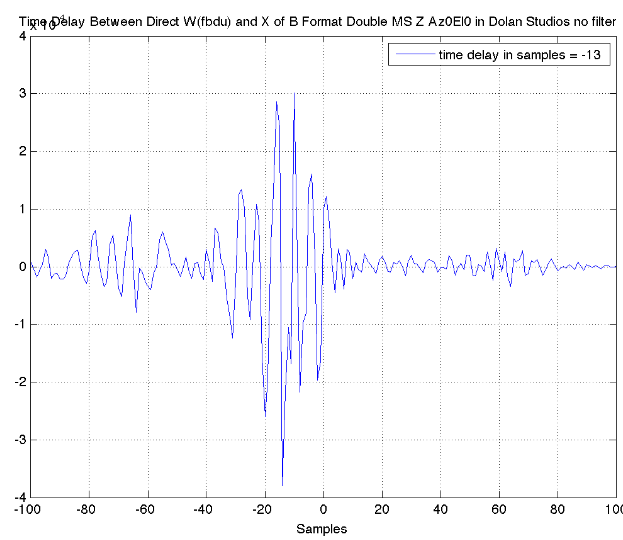


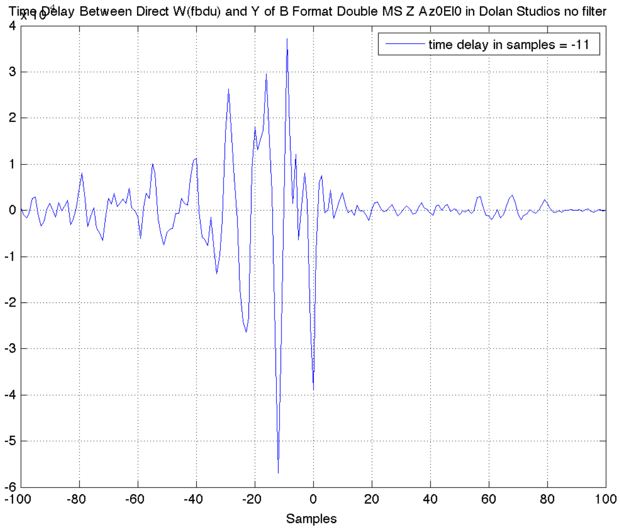


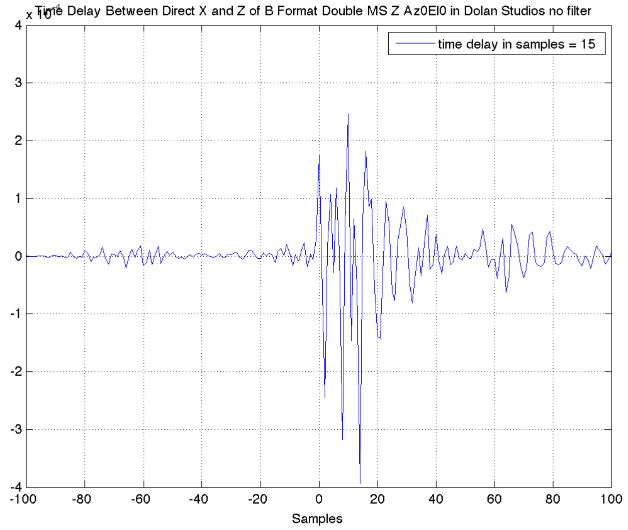


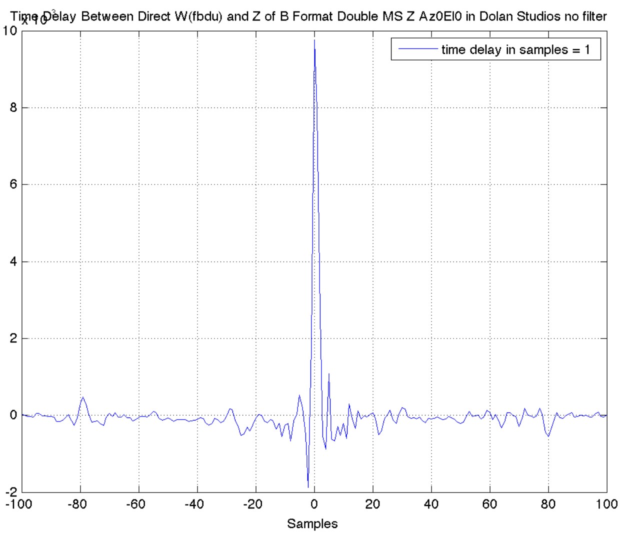


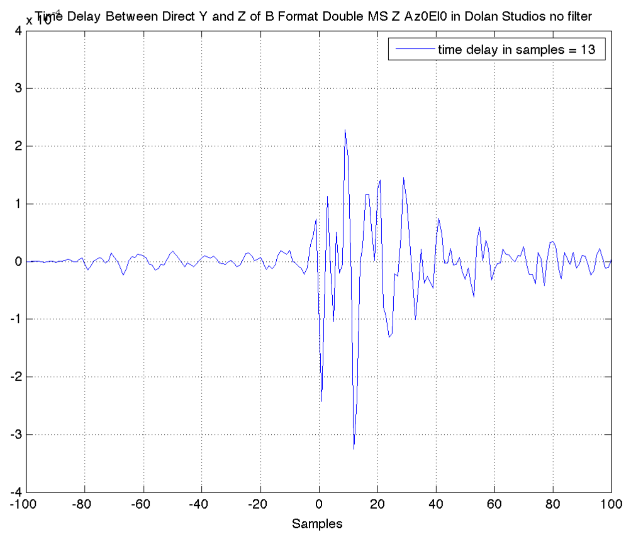
Physical Near Coincident Double MS Z Configuration Sennheiser Twins Az0

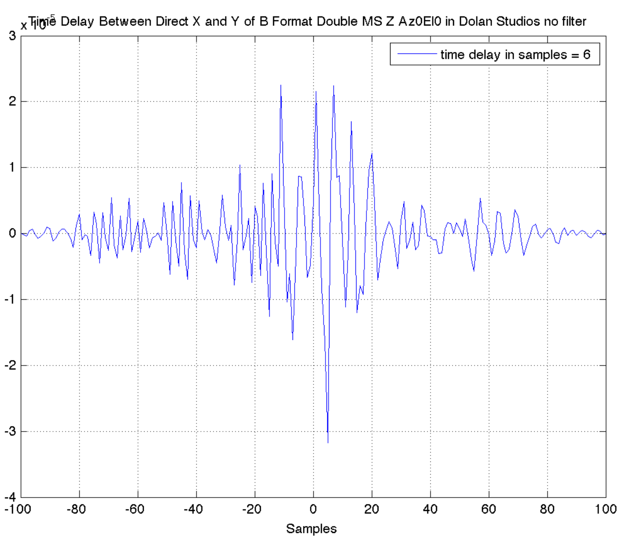




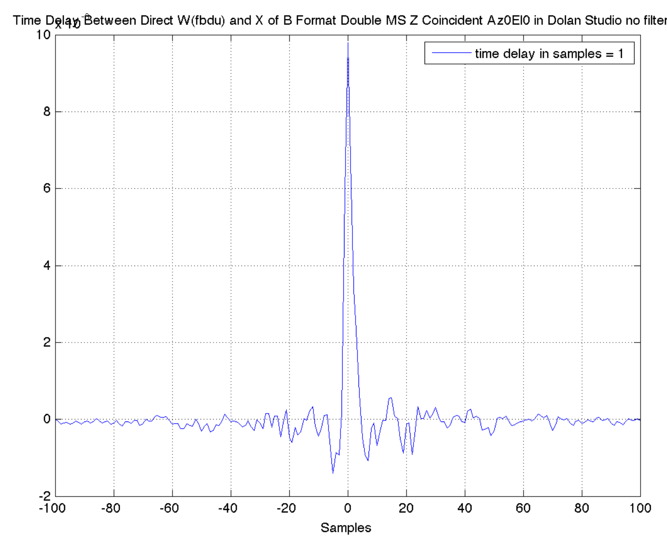


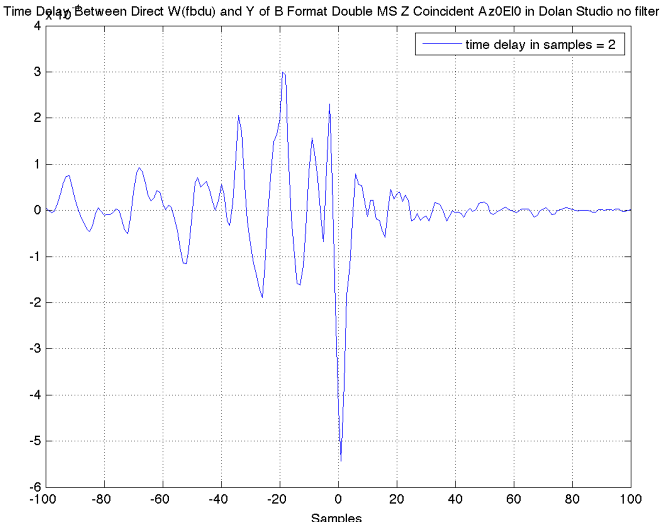


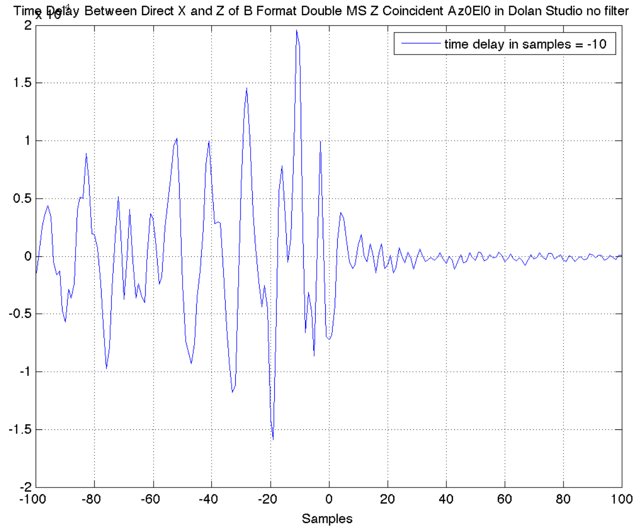


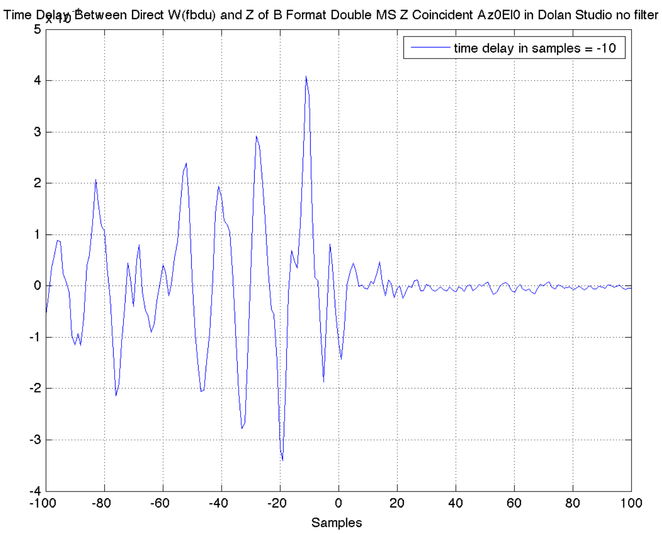


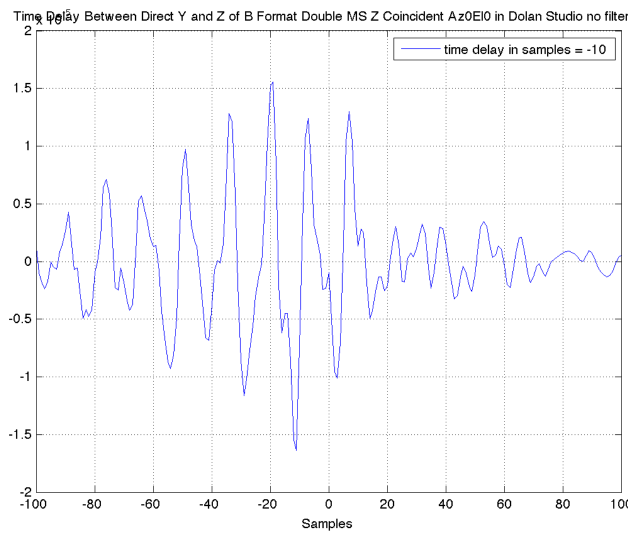
Virtual True Coincident Double MS Z Configuration Sennheiser Twins Az0

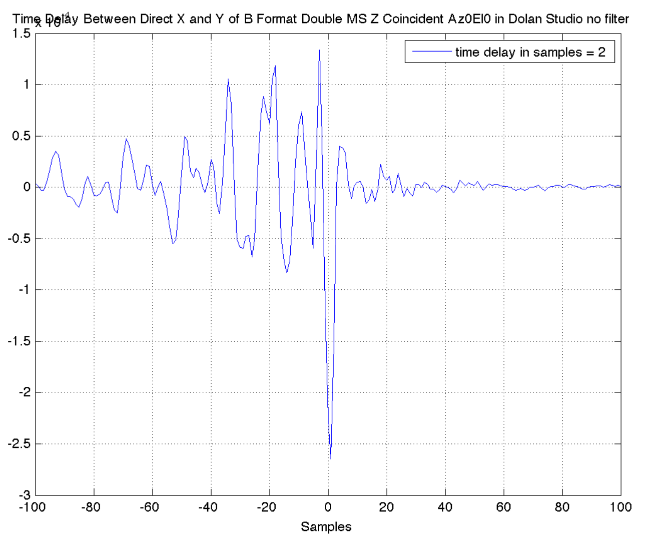






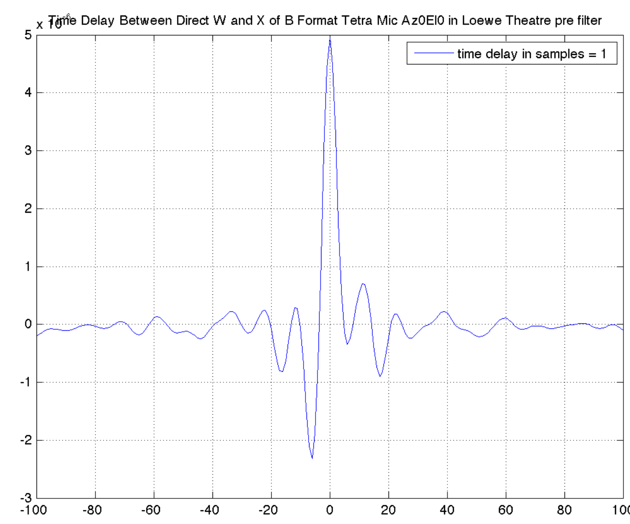


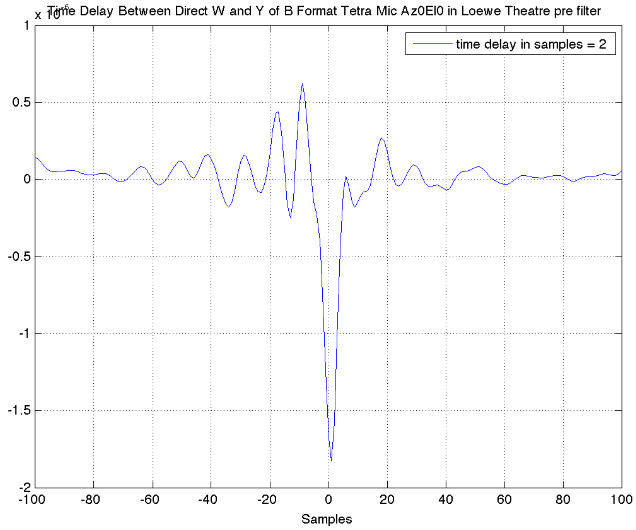


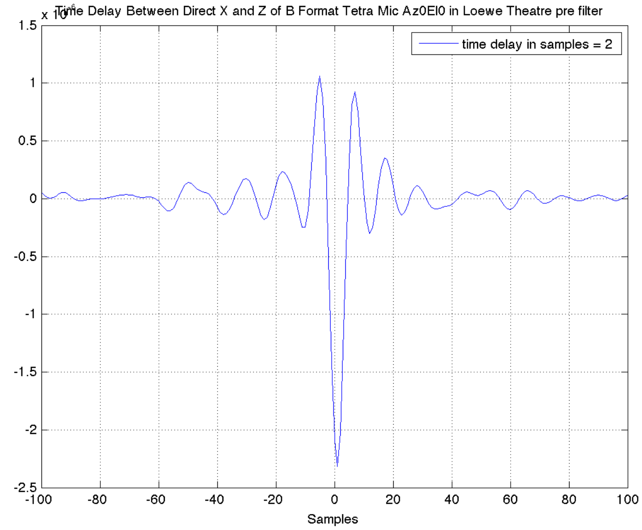


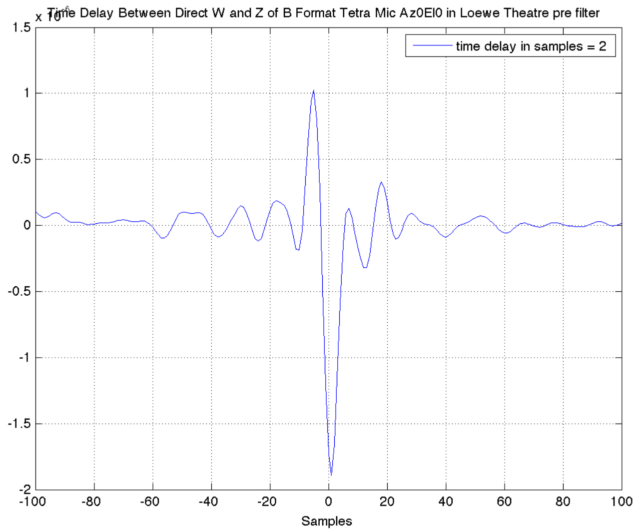
Loewe Theatre

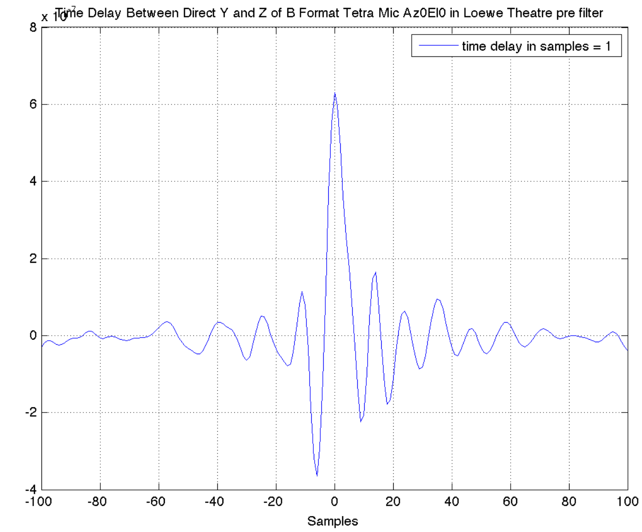
Physical Near Coincident Core Sound Tetra Mic Az0

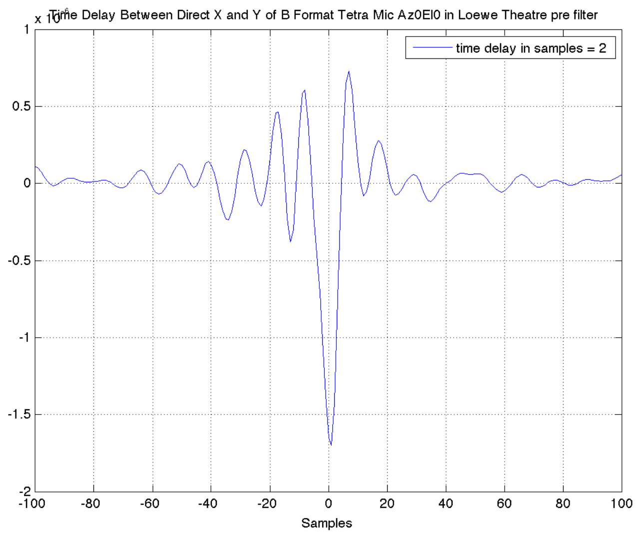


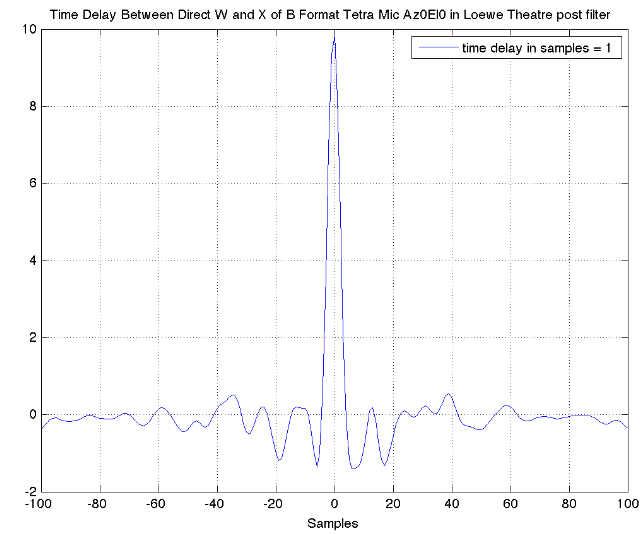

Time Delay Between Direct W and Y of B Format Tetra Mic Az0El0 in Loewe Theatre pre filter
time delay in samples = 2
Samples

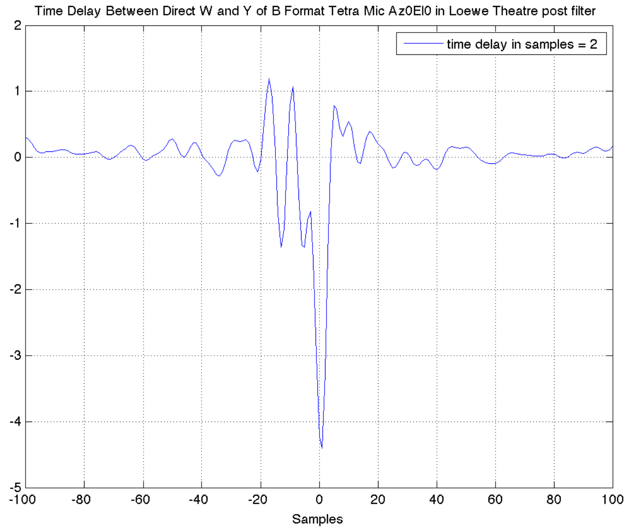

Time Delay Between Direct X and Z of B Format Tetra Mic Az0El0 in Loewe Theatre pre filter
time delay in samples = 2
Samples

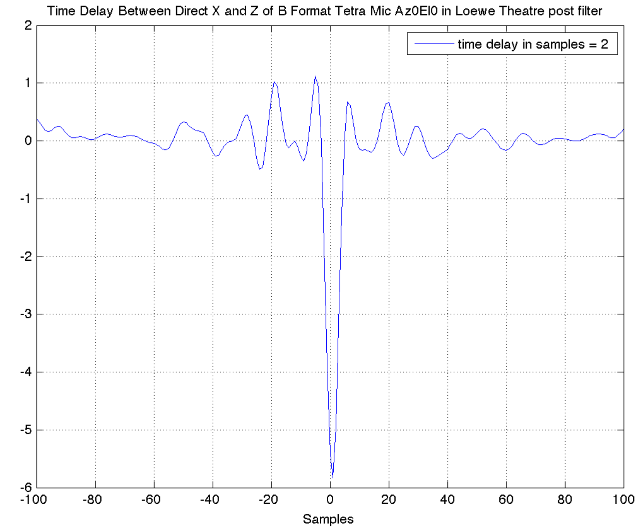

Time Delay Between Direct W and Z of B Format Tetra Mic Az0El0 in Loewe Theatre pre filter
time delay in samples = 2
Samples

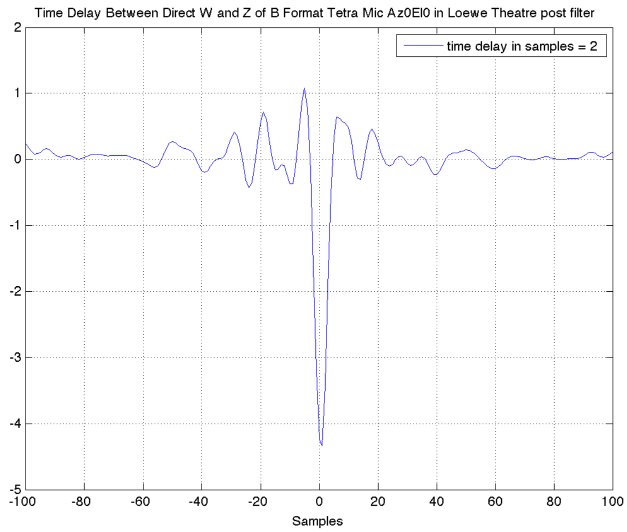

Time Delay Between Direct Y and Z of B Format Tetra Mic Az0El0 in Loewe Theatre pre filter
time delay in samples = 1
Samples

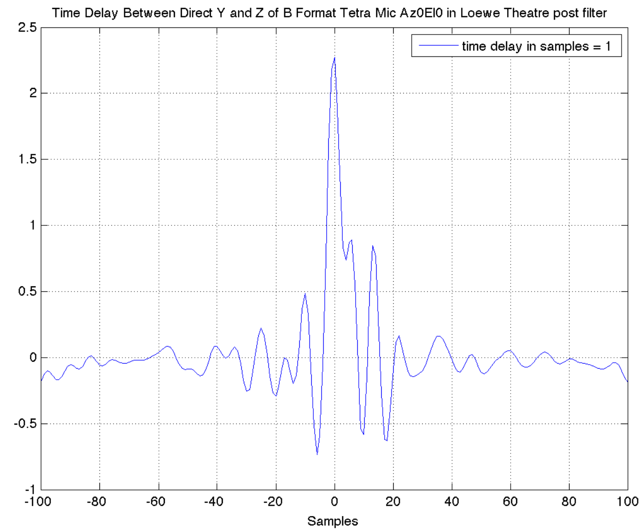

Time Delay Between Direct X and Y of B Format Tetra Mic Az0El0 in Loewe Theatre pre filter
time delay in samples = 2
Samples

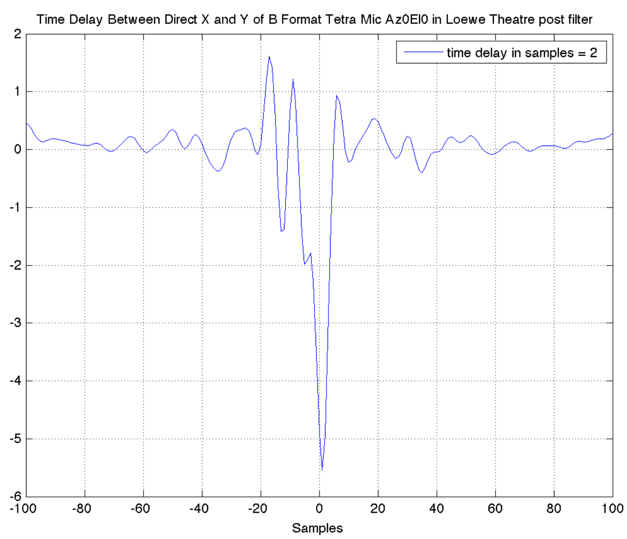

Time Delay Between Direct W and X of B Format Tetra Mic Az0El0 in Loewe Theatre post filter
time delay in samples = 1
Samples



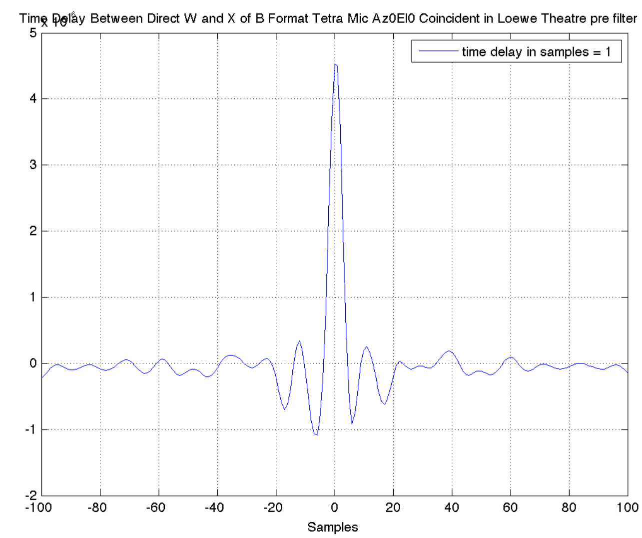


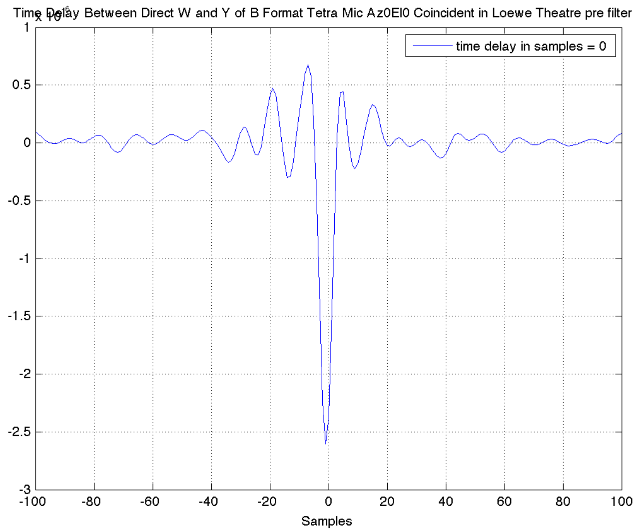


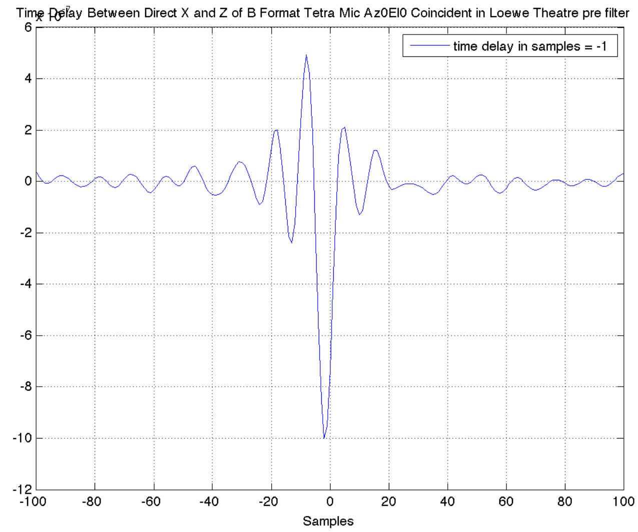


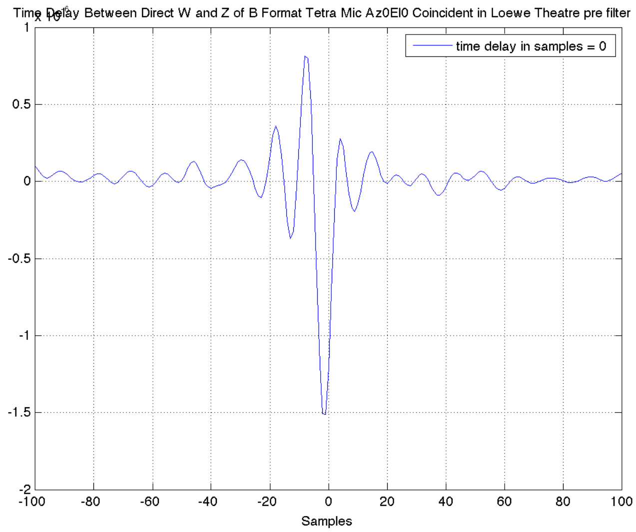


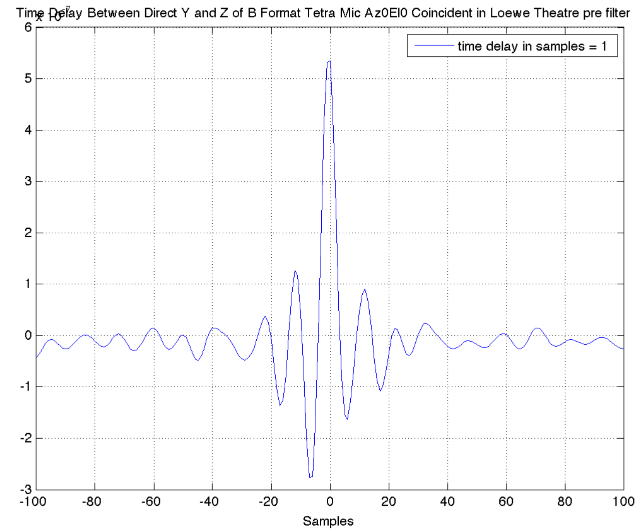
Virtual True Coincident Core Sound Tetra Mic Az0

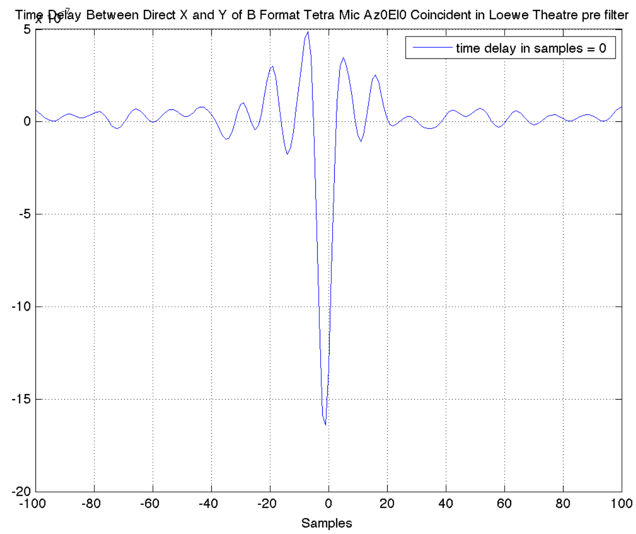




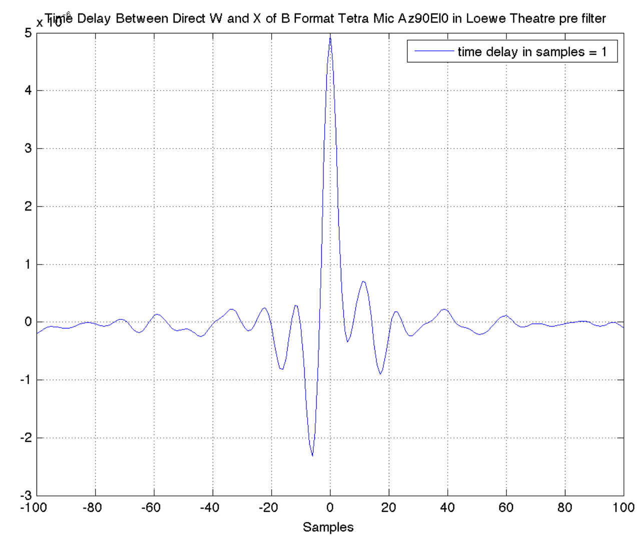


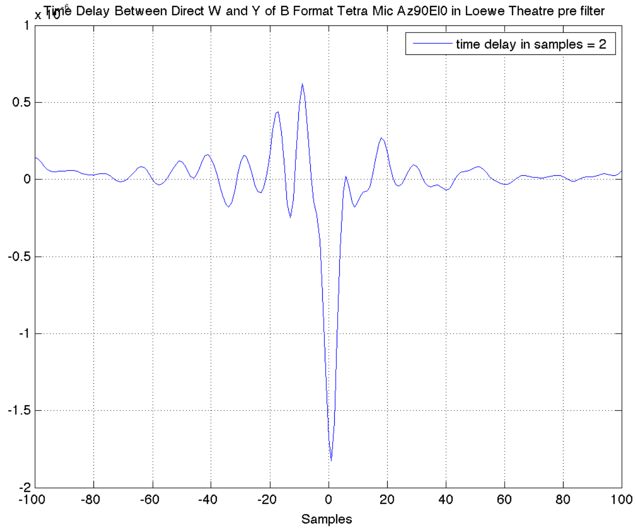


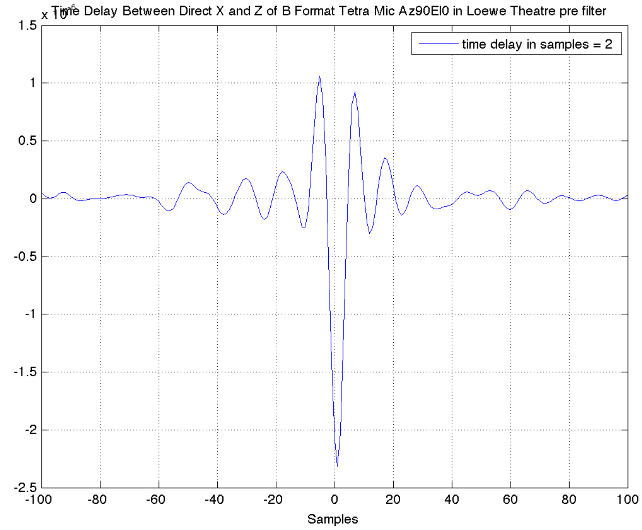


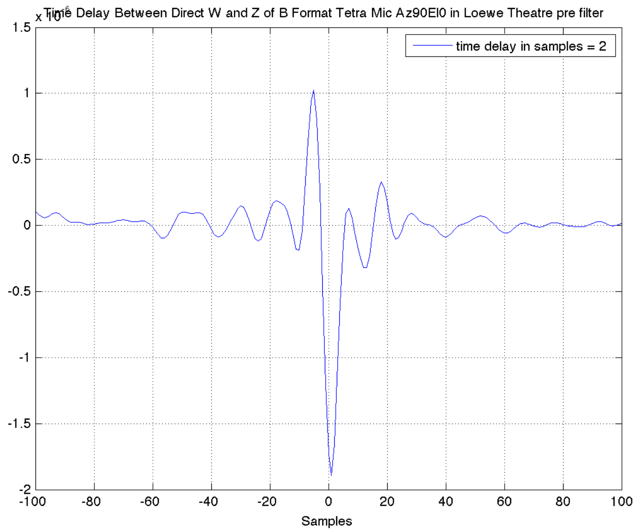


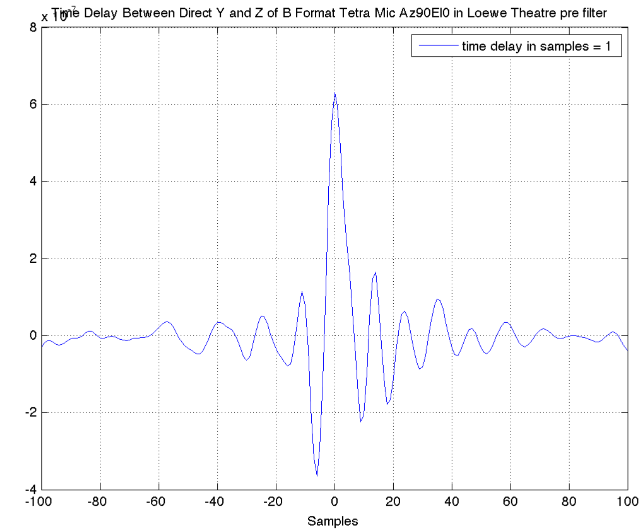
Physical Near Coincident Core Sound Tetra Mic Az90

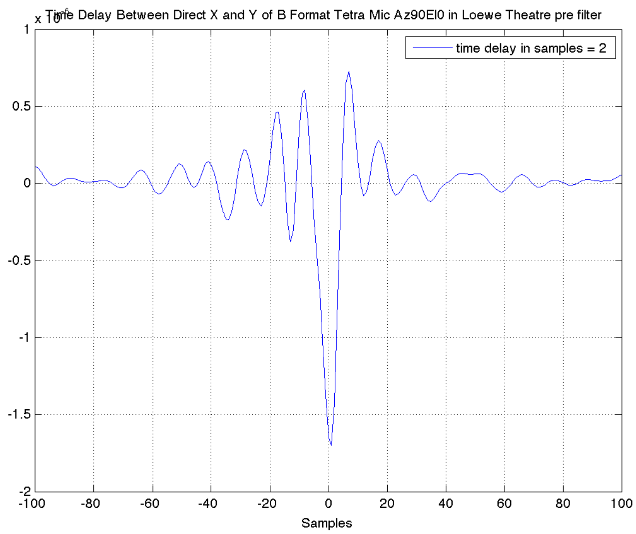


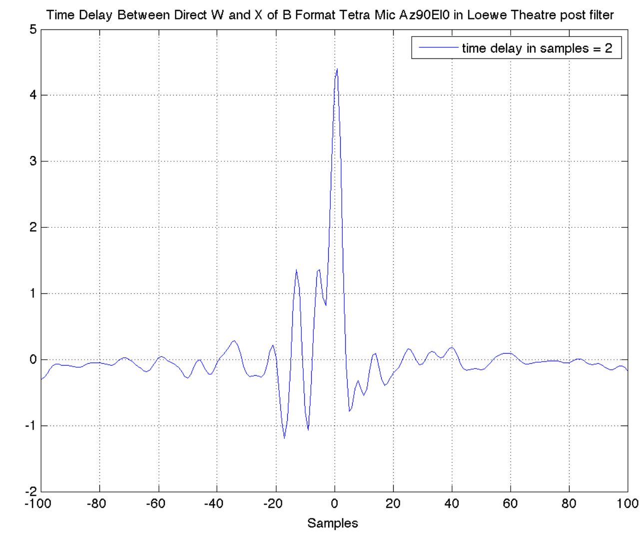

Time Delay Between Direct W and Y of B Format Tetra Mic Az90El0 in Loewe Theatre pre filter
time delay in samples = 2
Samples

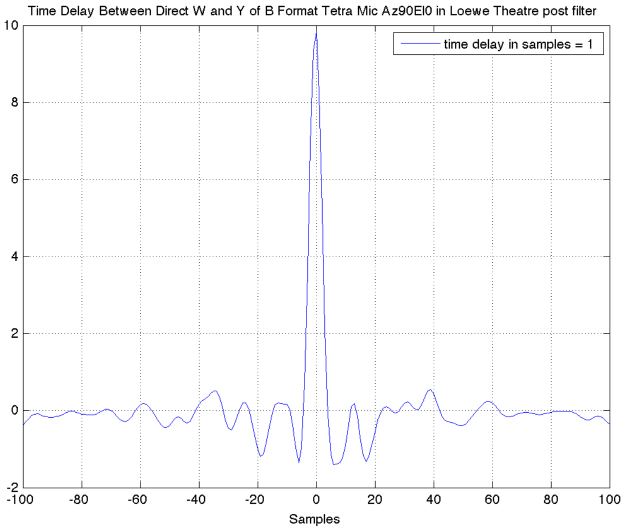

Time Delay Between Direct X and Z of B Format Tetra Mic Az90El0 in Loewe Theatre pre filter
time delay in samples = 2
Samples

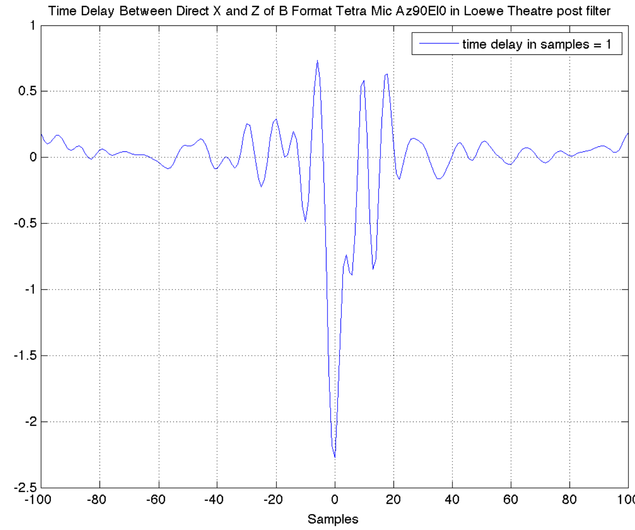

Time Delay Between Direct W and Z of B Format Tetra Mic Az90El0 in Loewe Theatre pre filter
time delay in samples = 2
Samples

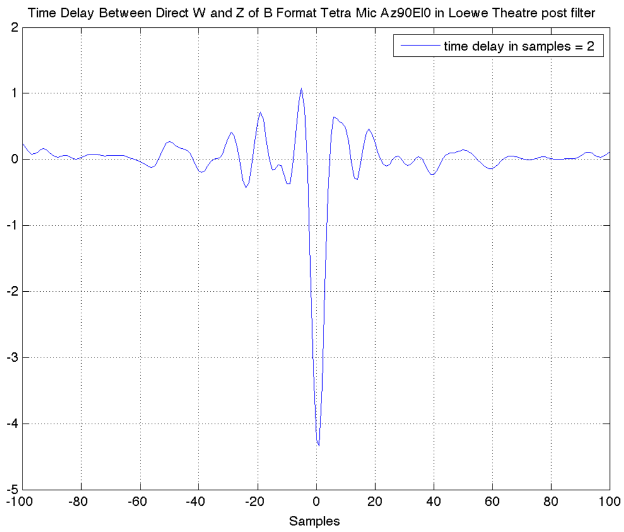

Time Delay Between Direct Y and Z of B Format Tetra Mic Az90El0 in Loewe Theatre pre filter
time delay in samples = 1
Samples

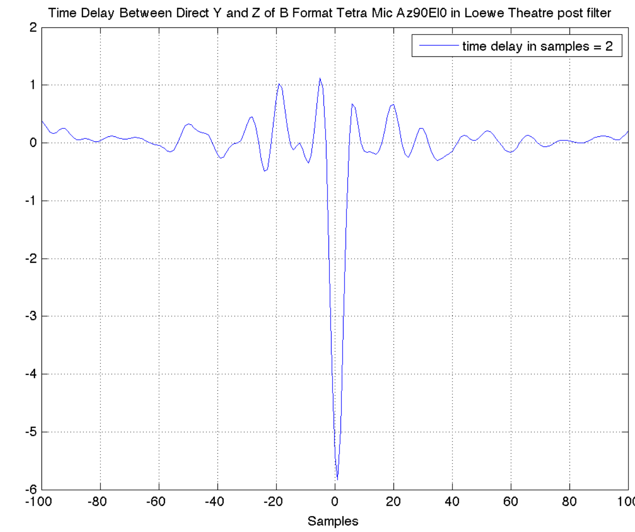

Time Delay Between Direct X and Y of B Format Tetra Mic Az90El0 in Loewe Theatre pre filter
time delay in samples = 2
Samples

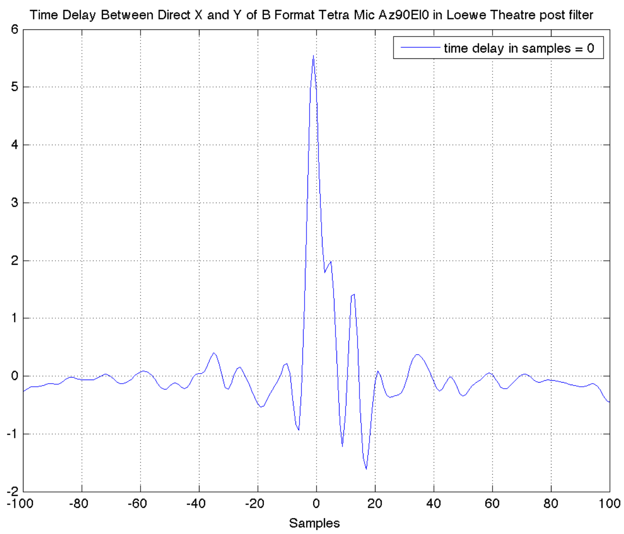

Time Delay Between Direct W and X of B Format Tetra Mic Az90El0 in Loewe Theatre post filter
time delay in samples = 2
Samples



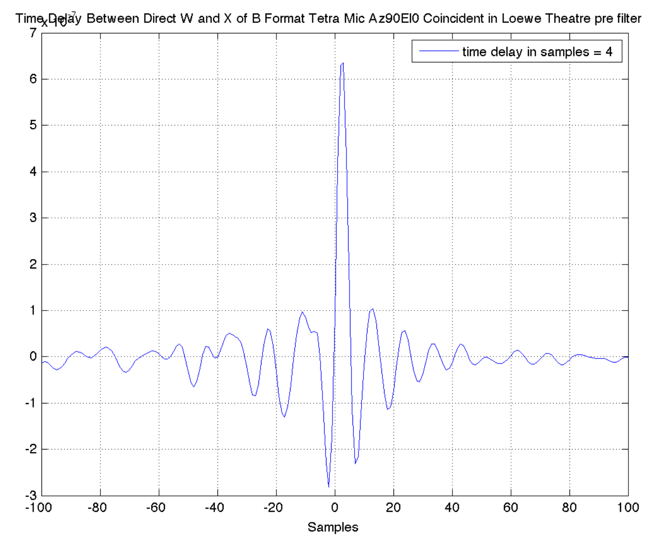


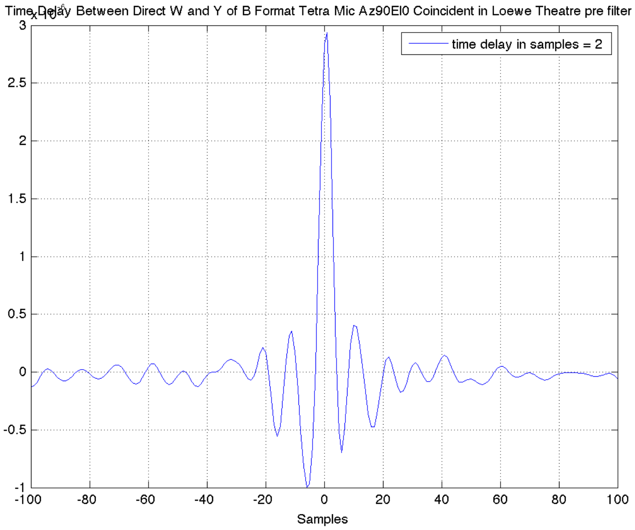


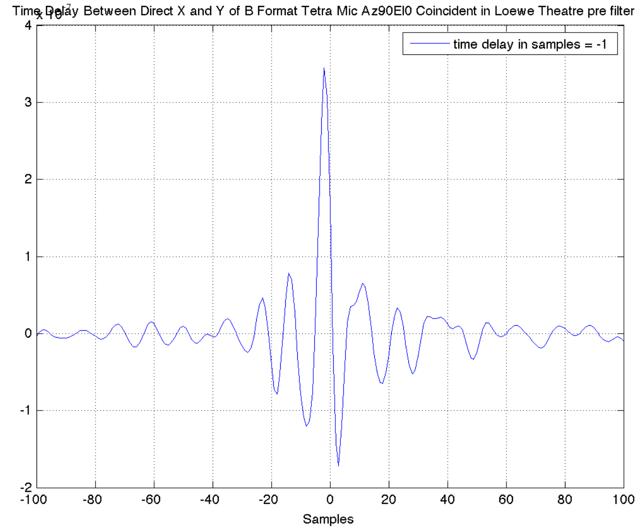


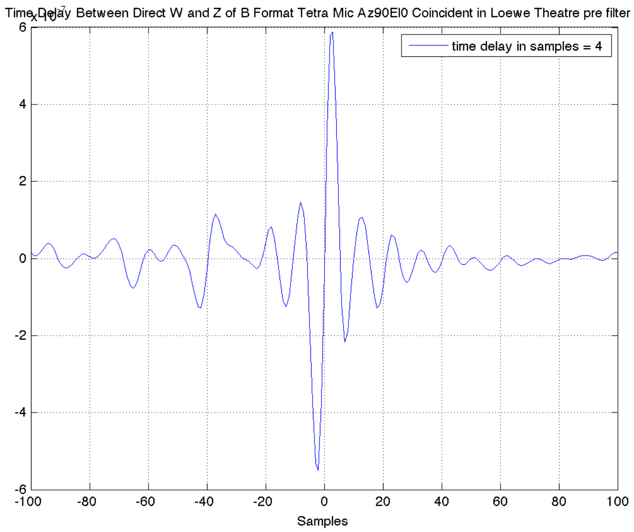


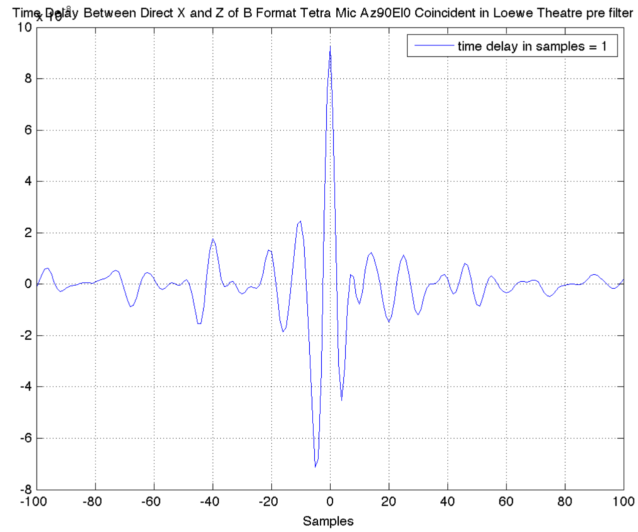
Virtual True Coincident Core Sound Tetra Mic Az90

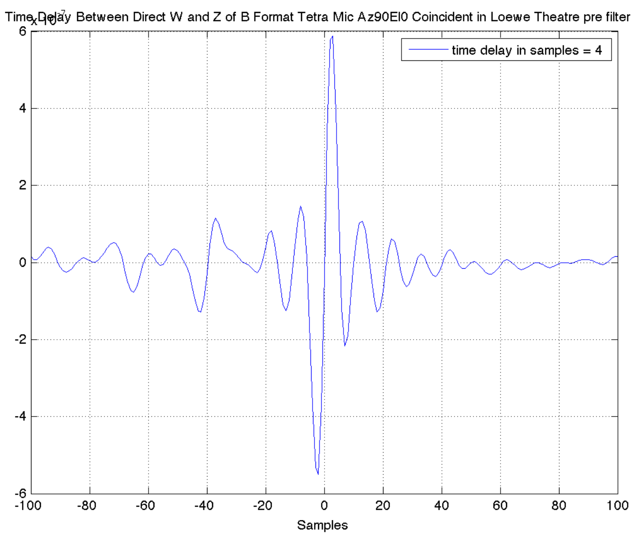


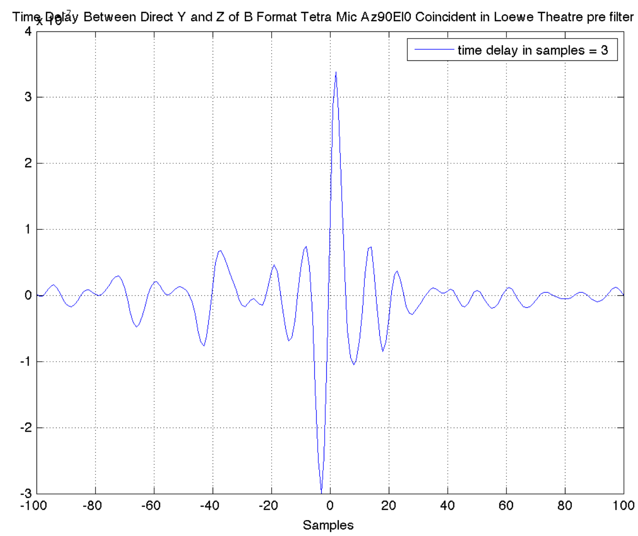




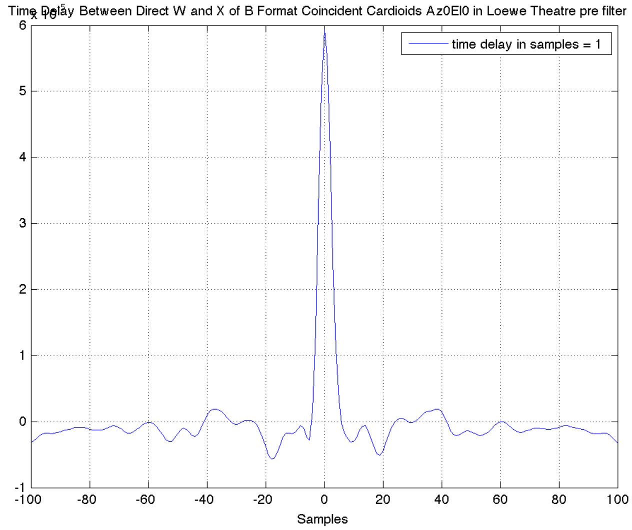


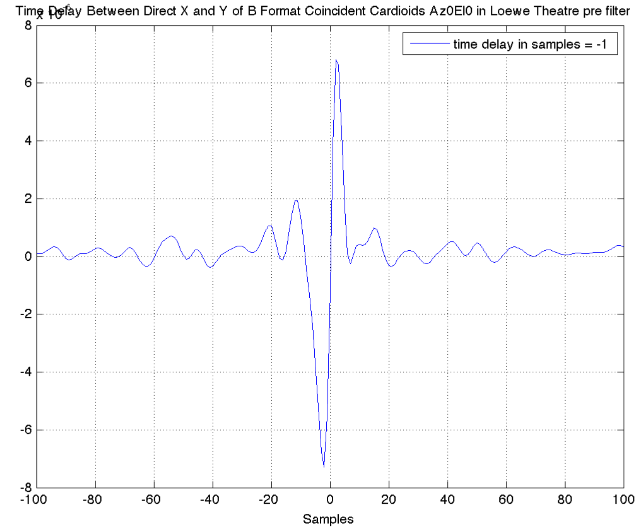


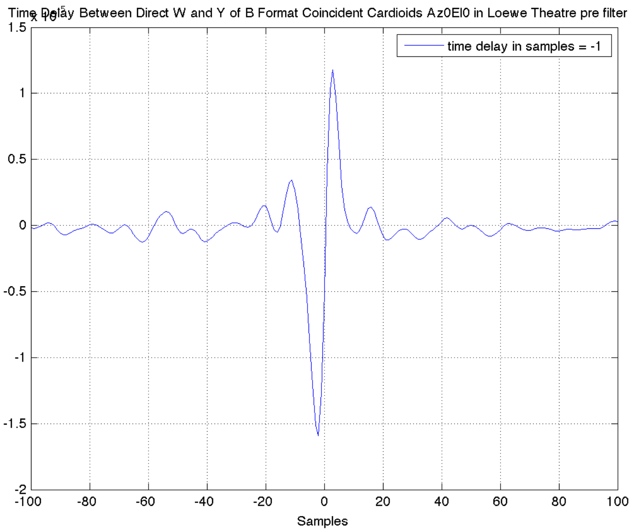


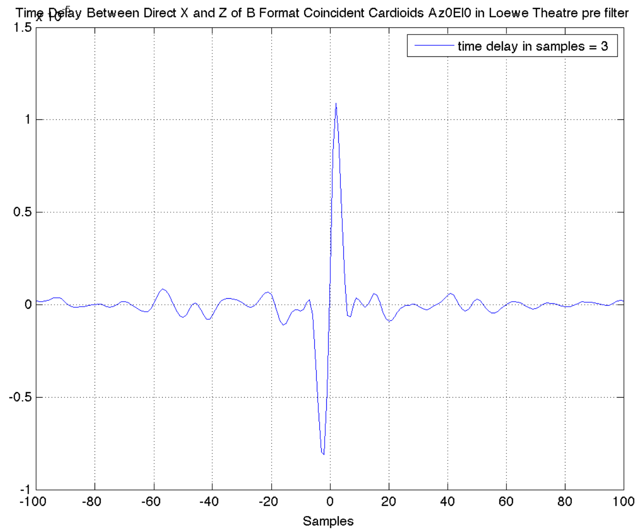


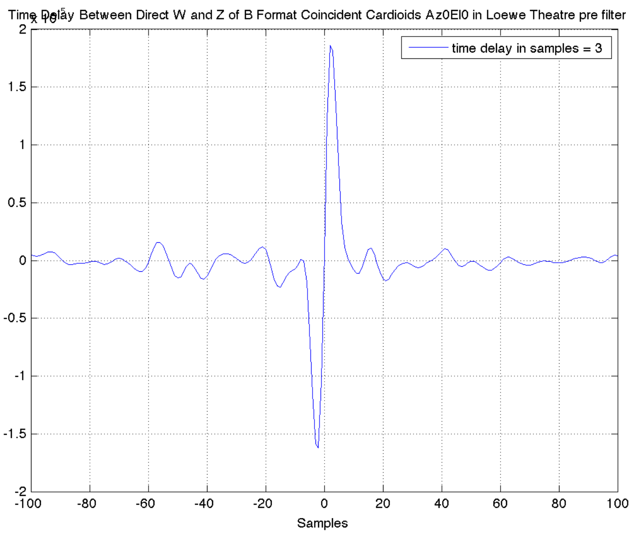
Virtual True Coincident DPA 4011A Tetra Configuration Az0

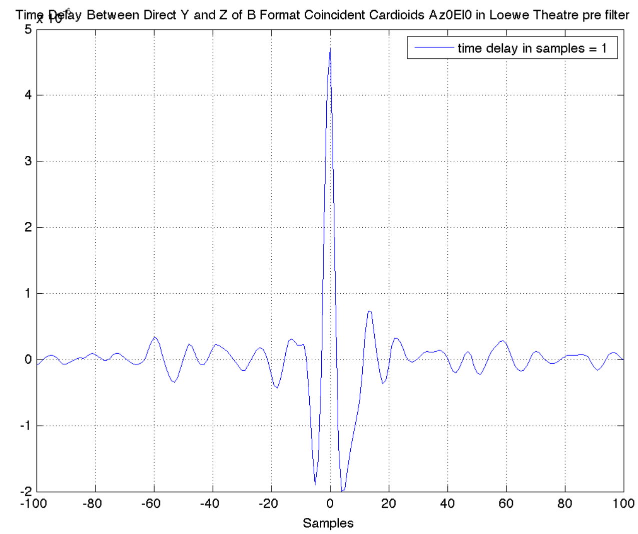




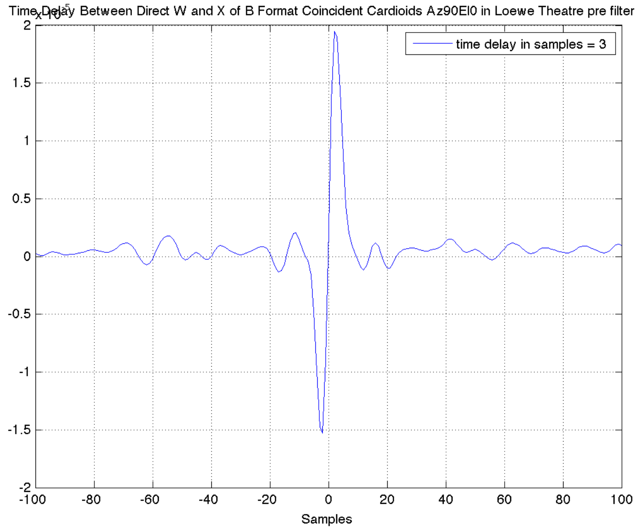


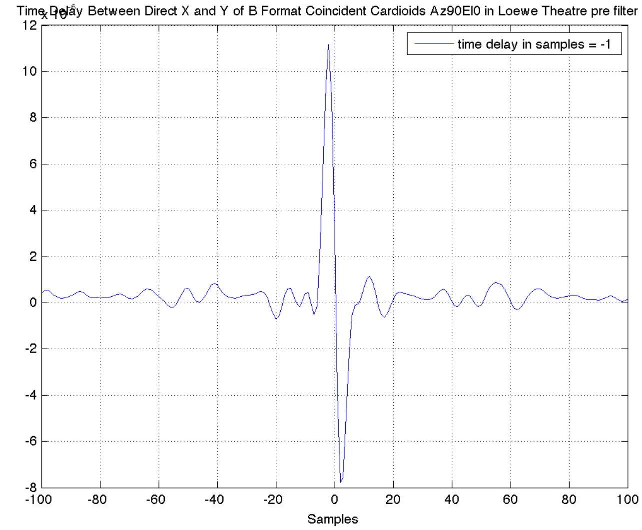


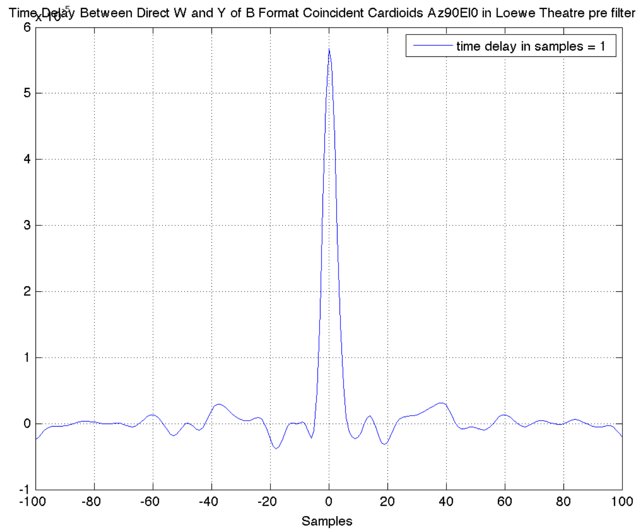


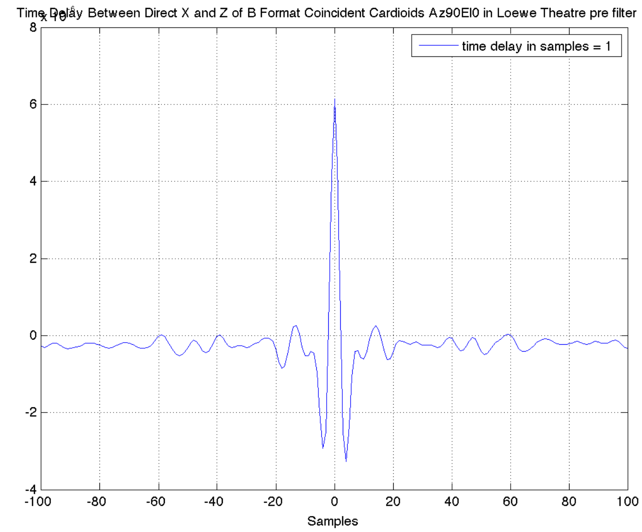


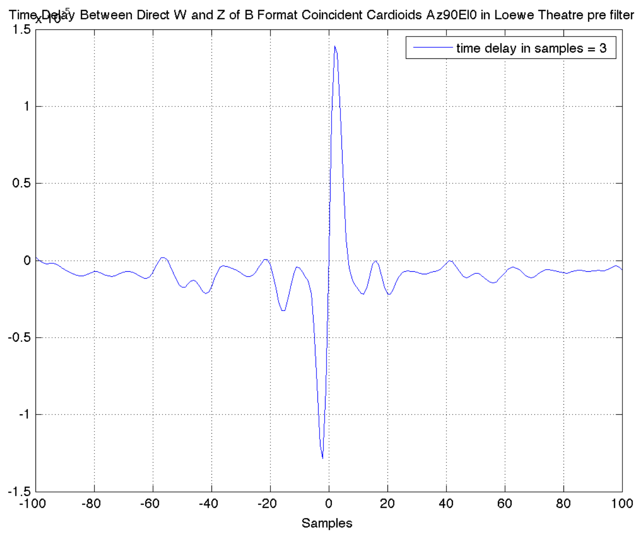
Virtual True Coincident DPA 4011A Tetra Configuration Az90

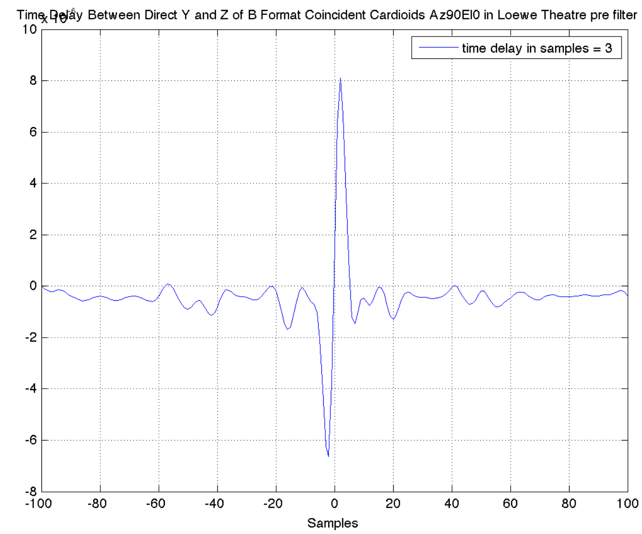




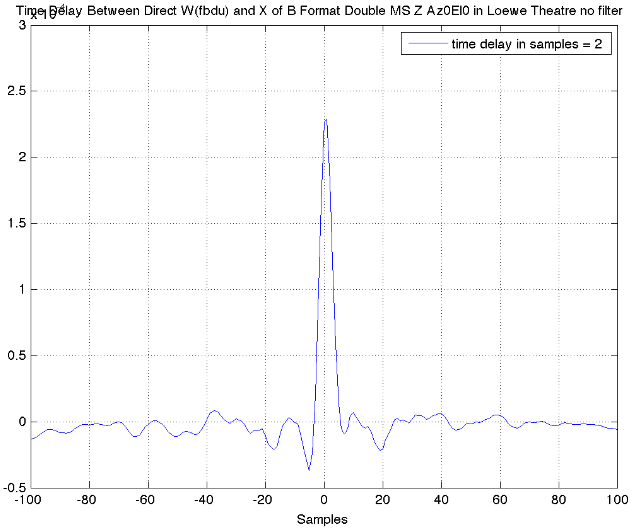


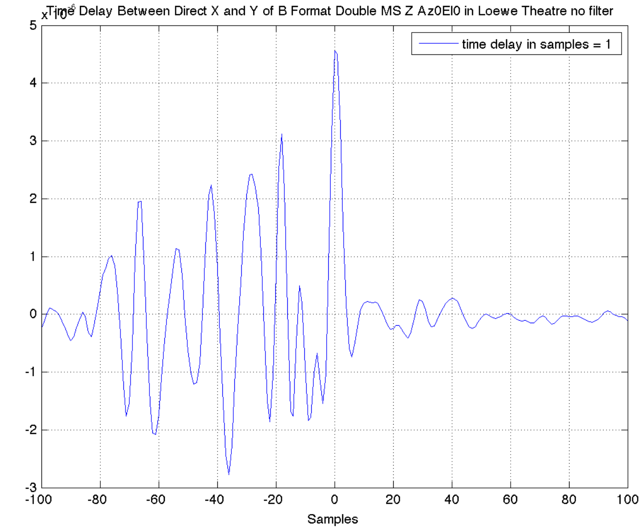


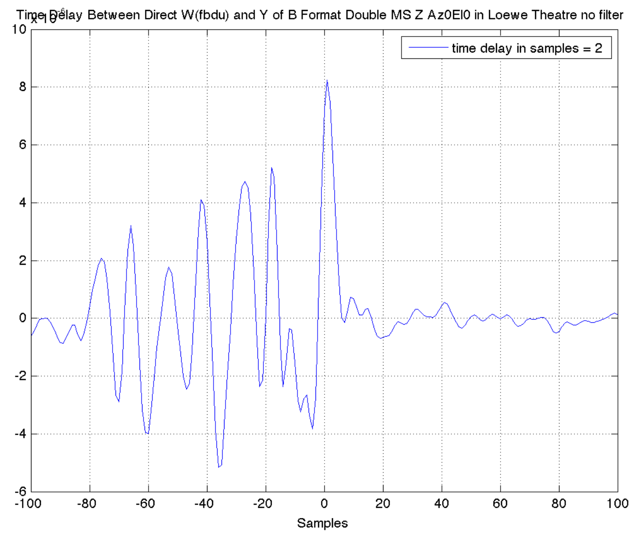


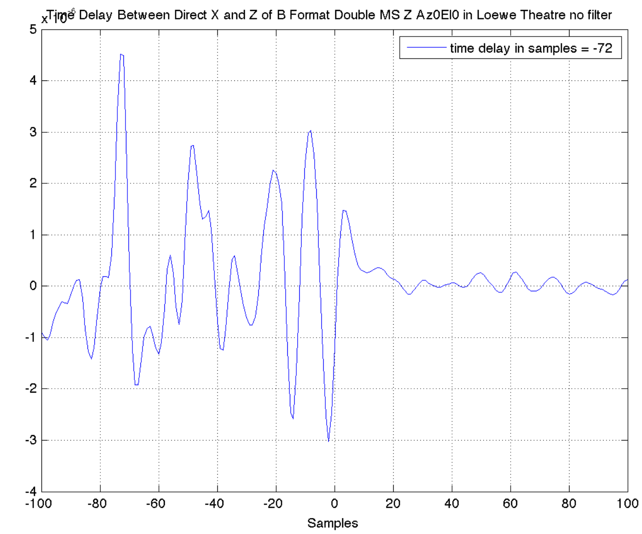


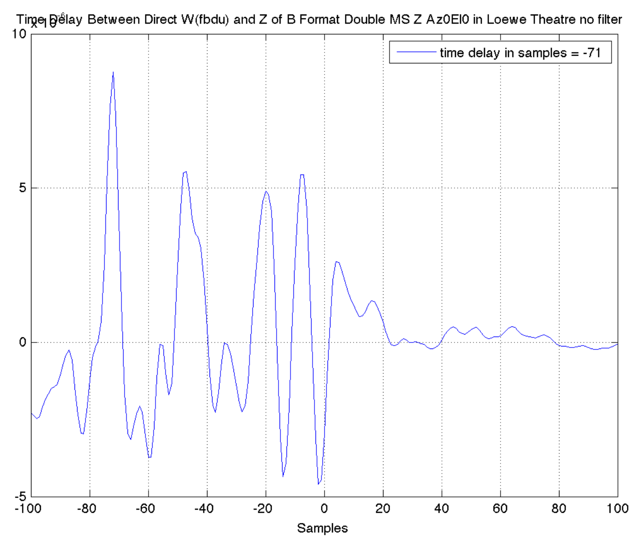
Physical Near Coincident Double MS Z Configuration
Sennheiser Twins Az0

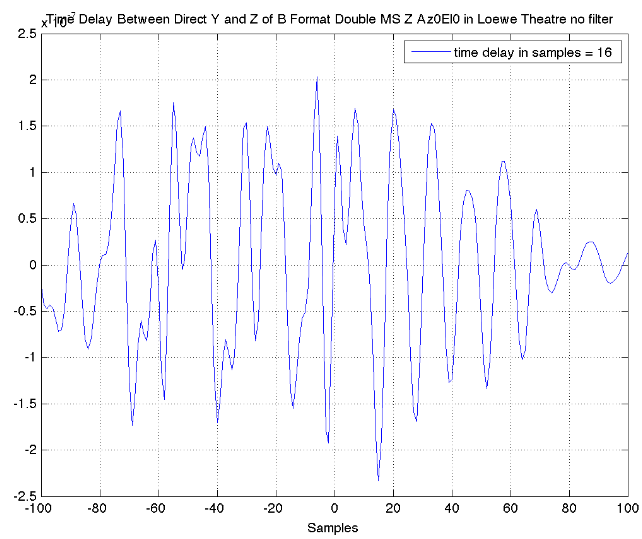




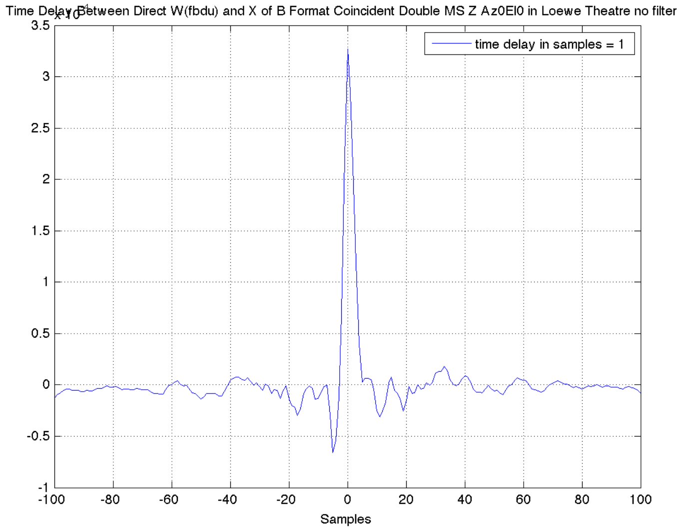


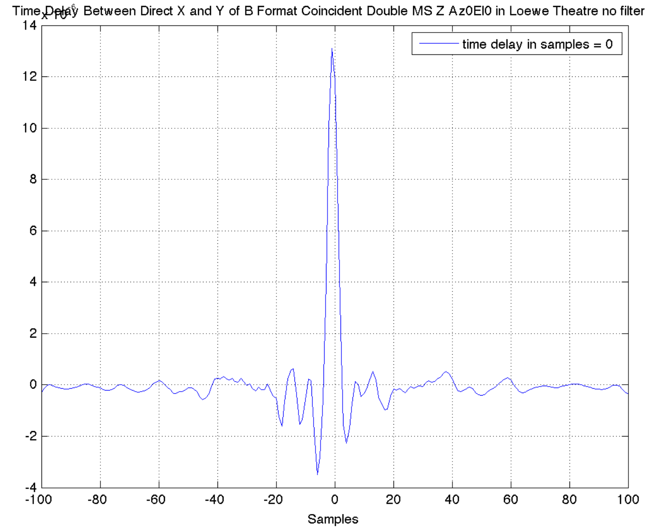


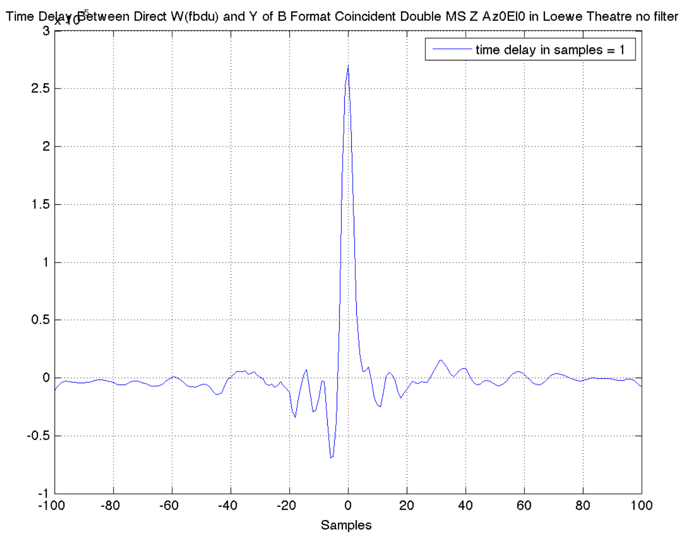


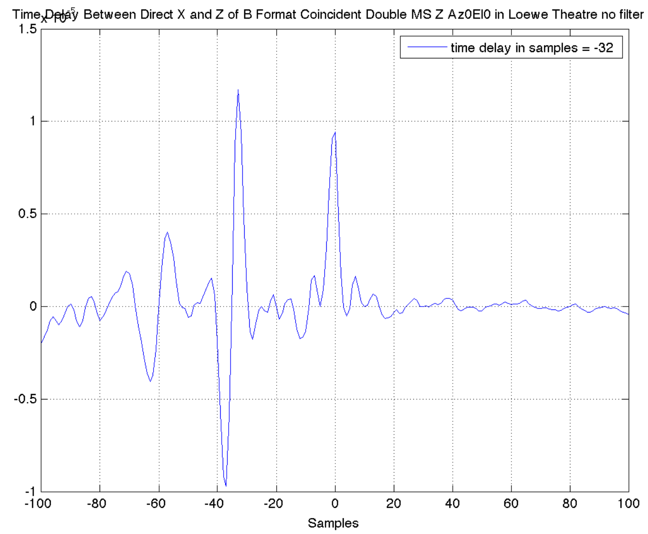


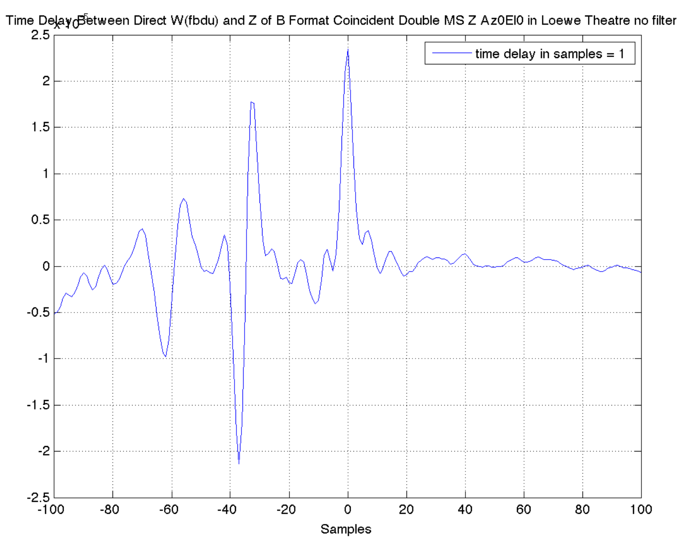
Virtual True Coincident Double MS Z Configuration Sennheiser Twins Az0

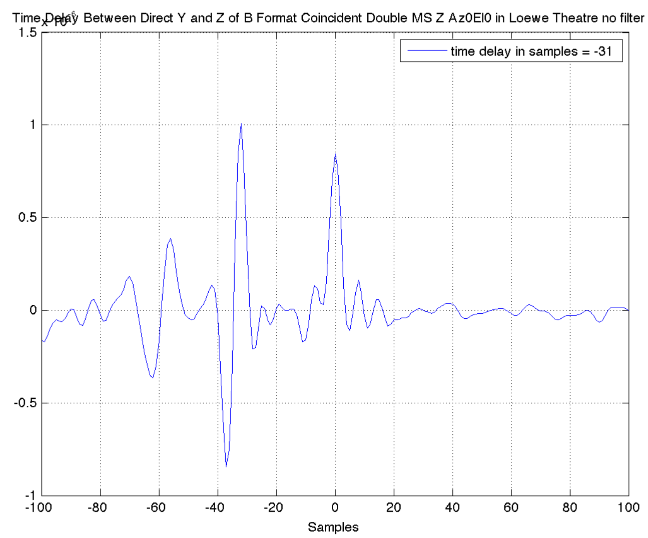

Time Delay Between Direct W(fbdu) and X of B Format Coincident Double MS Z Az0El0 in Loewe Theatre no filter
time delay in samples = 1
Samples


Time Delay Between Direct X and Y of B Format Coincident Double MS Z Az0El0 in Loewe Theatre no filter
time delay in samples = 0
Samples


Time Delay Between Direct W(fbdu) and Y of B Format Coincident Double MS Z Az0El0 in Loewe Theatre no filter
time delay in samples = 1
Samples


Time Delay Between Direct X and Z of B Format Coincident Double MS Z Az0El0 in Loewe Theatre no filter
time delay in samples = -32
Samples


Time Delay Between Direct W(fbdu) and Z of B Format Coincident Double MS Z Az0El0 in Loewe Theatre no filter
time delay in samples = 1
Samples


Time Delay Between Direct Y and Z of B Format Coincident Double MS Z Az0El0 in Loewe Theatre no filter
time delay in samples = -31
Samples

Appendix D – B Format Correlation Data

Research Lab

Physical Near Coincident Core Sound Tetra Mic Az0

corr B format from A no filter			
1	0.76825	-0.036136	-0.045869
0.76825	1	-0.0061325	-0.067887
-0.03613	-0.006132	1	0.14747
corr direct signal B format from A no filter			
1	0.89323	-0.028092	-0.23427
0.89323	1	-0.005134	-0.3294
-0.02809	-0.005134	1	0.015629
corr B format from A w/VVTetra filter			
1	0.82182	-0.16339	-0.18299
0.82182	1	-0.13325	-0.17437
-0.16339	-0.13325	1	0.18409
corr direct signal B format from A w/VVTetra filter			
1	0.95686	-0.34177	-0.47772
0.95686	1	-0.32371	-0.55982
-0.34177	-0.32371	1	-0.000932

Virtual True Coincident Core Sound Tetra Mic Az0

corr B format from A			
1	0.82226	-0.14735	-0.054829
0.82226	1	0.0016901	0.11014
-0.14735	0.0016901	1	0.25386
corr direct signal B format from A			
1	0.94554	-0.14472	-0.094481
0.94554	1	0.037587	0.027754
-0.14472	0.037587	1	0.89909

Physical Near Coincident Core Sound Tetra Mic Az90

corr B format from A pre filter			
1	0.15189	0.75988	0.17711
0.15189	1	0.18098	0.30243
0.75988	0.18098	1	0.13302
corr direct signal B format from A pre filter			
1	0.42172	0.88222	0.33053
0.42172	1	0.37513	0.87009
0.88222	0.37513	1	0.2856
corr B format from A post filter			
1	-0.79973	0.18832	0.16626
-0.79973	1	-0.14053	-0.14985
0.18832	-0.14053	1	0.31942
corr direct signal B format from A post filter			
1	-0.94086	0.39865	0.33822
-0.94086	1	-0.30873	-0.25184
0.39865	-0.30873	1	0.85635

Virtual True Coincident Core Sound Tetra Mic Az90

corr B format from A			
1	0.16496	0.82578	0.17699
0.16496	1	0.15515	0.58724
0.82578	0.15515	1	0.15598
corr direct signal B format from A			
1	0.19114	0.97617	0.20586
0.19114	1	0.28228	0.96341
0.97617	0.28228	1	0.27825

Virtual True Coincident DPA 4011A Tetra Configuration Az0

corr B format from A			
1	0.76809	-0.082492	-0.099525
0.76809	1	0.035273	0.026416
-0.082492	0.035273	1	0.31637
corr direct signal B format from A			
1	0.9442	-0.10264	-0.23684
0.9442	1	0.13443	-0.026518
-0.10264	0.13443	1	0.94118

Virtual True Coincident DPA 4011A Tetra Configuration Az90

corr B format from A			
1	-0.42203	0.6426	-0.59623
-0.42203	1	-0.54383	0.53048
0.6426	-0.54383	1	-0.87599
corr direct signal B format from A			
1	-0.94936	0.9196	-0.96395
-0.94936	1	-0.90192	0.98291
0.9196	-0.90192	1	-0.89109

Physical Near Coincident Double MS Z Configuration Sennheiser Twins Az0

corr B format from Double MS Z (W:fbdu)			
1	0.1317	0.78107	-0.10282
0.1317	1	0.08409	0.011462
0.78107	0.08409	1	-0.11314
corr Direct B format from Double MS Z (W:fbdu)			
1	0.35889	0.90963	-0.30801
0.35889	1	0.28435	-0.14321

0.90963	0.28435	1	-0.34903

Virtual True Coincident Double MS Z Configuration Sennheiser Twins Az0

corr B format from Double MS Z (W:fbdu)			
1	0.78183	-0.24551	-0.24981
0.78183	1	-0.18751	-0.20146
-0.24551	-0.18751	1	0.090165
corr Direct B format from Double MS Z (W:fbdu)			
1	0.91204	-0.94009	-0.66756
0.91204	1	-0.84136	-0.71923
-0.94009	-0.84136	1	0.62012

Dolan Studios

Physical Near Coincident Core Sound Tetra Mic Az0

corr B format from A pre filter			
1	0.67344	0.2315	0.065944
0.67344	1	0.098508	0.19837
0.2315	0.098508	1	0.43678
corr direct signal B format from A pre filter			
1	0.92095	0.36762	0.42124
0.92095	1	0.33585	0.37249
0.36762	0.33585	1	-0.066843
corr B format from A post filter			
1	0.74162	0.15719	-0.006692
0.74162	1	0.050239	0.098264
0.15719	0.050239	1	0.40994
corr direct signal B format from A post filter			
1	0.9458	0.13371	0.24106
0.9458	1	0.1434	0.088558
0.13371	0.1434	1	-0.42021

Virtual True Coincident Core Sound Tetra Mic Az0

corr B format from A			
1	0.77659	0.39974	-0.006538
0.77659	1	0.25495	0.11768
0.39974	0.25495	1	0.39805
corr direct signal B format from A			
1	0.95884	0.3813	0.087984
0.95884	1	0.39607	0.14513
0.3813	0.39607	1	0.89743

Physical Near Coincident Core Sound Tetra Mic Az90

corr B format from A pre filter			
1	0.063412	0.74585	0.0018116
0.063412	1	-0.018516	0.55605
0.74585	-0.018516	1	0.043953
corr direct signal B format from A pre filter			
1	0.43017	0.91209	0.32213
0.43017	1	0.37409	0.61581
0.91209	0.37409	1	0.42711
corr B format from A post filter			
1	-0.76183	0.095166	-0.012513
-0.76183	1	0.031897	-0.039991
0.095166	0.031897	1	0.51594
corr direct signal B format from A post filter			
1	-0.92681	0.34307	0.27825
-0.92681	1	-0.22272	-0.24159
0.34307	-0.22272	1	0.51776

Virtual True Coincident Core Sound Tetra Mic Az90

corr B format from A			
1	0.098055	0.8382	-0.008795
0.098055	1	0.1421	0.35612
0.8382	0.1421	1	0.083342
corr direct signal B format from A			
1	0.53505	0.96898	0.15473
0.53505	1	0.6594	0.83344
0.96898	0.6594	1	0.34508

Virtual True Coincident DPA 4011A Tetra Configuration Az0

corr B format from A			
1	0.74313	-0.056074	0.064632
0.74313	1	-0.027522	0.10062
-0.056074	-0.027522	1	0.23987
corr direct signal B format from A			
1	0.96309	0.14973	0.018656
0.96309	1	0.15537	0.12853
0.14973	0.15537	1	0.83904

Virtual True Coincident DPA 4011A Tetra Configuration Az90

corr B format from A			
1	0.30559	0.58011	0.11905
0.30559	1	0.4577	0.44723
0.58011	0.4577	1	0.22961
corr direct signal B format from A			
1	0.57211	0.90255	0.33733
0.57211	1	0.75721	0.85858
0.90255	0.75721	1	0.49748

Physical Near Coincident Double MS Z Configuration Sennheiser Twins Az0

corr B format from Double MS Z (W:fbdu)			
1	-0.085426	-0.051251	0.62108
-0.085426	1	-0.061402	0.029272
-0.051251	-0.061402	1	-0.01378
corr Direct B format from Double MS Z (W:fbdu)			
1	0.079461	-0.35759	0.8407
0.079461	1	0.067929	0.21906
-0.35759	0.067929	1	-0.16012

Virtual True Coincident Double MS Z Configuration Sennheiser Twins Az0

corr B format from Double MS Z (W:fbdu)			
1	0.73836	-0.1314	0.066325
0.73836	1	-0.11985	0.077903
-0.1314	-0.11985	1	-0.0021042
corr Direct B format from Double MS Z (W:fbdu)			
1	0.91878	-0.51151	-0.11658
0.91878	1	-0.52832	-0.14806
-0.51151	-0.52832	1	-0.024281

Loewe Theatre

Physical Near Coincident Core Sound Tetra Mic Az0

corr B format from A pre filter			
1	0.5553	-0.5733	-0.29999
0.5553	1	-0.32745	-0.57088
-0.5733	-0.32745	1	0.2638
corr direct signal B format from A pre filter			
1	0.91318	-0.8096	-0.79432
0.91318	1	-0.72646	-0.87547
-0.8096	-0.72646	1	0.69106
corr B format from A post filter			
1	0.67008	-0.59528	-0.41322
0.67008	1	-0.43503	-0.54483
-0.59528	-0.43503	1	0.36241
corr direct signal B format from A post filter			
1	0.96528	-0.78601	-0.89852
0.96528	1	-0.70715	-0.89045
-0.78601	-0.70715	1	0.70676

Virtual True Coincident Core Sound Tetra Mic Az0

corr B format from TA'ed A pre filter			
1	0.57005	-0.60696	-0.19854
0.57005	1	-0.33673	-0.42163
-0.60696	-0.33673	1	0.068554

corr direct signal B format from TA'ed A pre filter			
1	0.95529	-0.87901	-0.7088
0.95529	1	-0.78875	-0.66167
-0.87901	-0.78875	1	0.88566

Physical Near Coincident Core Sound Tetra Mic Az90

corr B format from A pre filter			
1	0.5553	-0.5733	-0.29999
0.5553	1	-0.32745	-0.57088
-0.5733	-0.32745	1	0.2638
corr direct signal B format from A pre filter			
1	0.91318	-0.8096	-0.79432
0.91318	1	-0.72646	-0.87547
-0.8096	-0.72646	1	0.69106
corr B format from A post filter			
1	0.59528	0.67008	-0.41322
0.59528	1	0.43503	-0.36241
0.67008	0.43503	1	-0.54483
corr direct signal B format from A post filter			
1	0.78601	0.96528	-0.89852
0.78601	1	0.70715	-0.70676
0.96528	0.70715	1	-0.89045

Virtual True Coincident Core Sound Tetra Mic Az90

corr B format from TA'ed A pre filter			
1	0.06079	0.31157	-0.015565
0.06079	1	-0.16898	-0.24775
0.31157	-0.16898	1	0.23383
corr direct signal B format from TA'ed A pre filter			
1	0.1117	0.91121	-0.038329
0.1117	1	0.37003	0.7626
0.91121	0.37003	1	0.26503

Virtual True Coincident DPA 4011A Tetra Configuration Az0

corr B format from TA'ed A pre filter			
1	0.35492	-0.020493	-0.05069
0.35492	1	-0.071128	-0.03807
-0.020493	-0.071128	1	0.36405
corr direct signal B format from TA'ed A pre filter			
1	0.97418	-0.21186	0.023694
0.97418	1	-0.12136	0.12368
-0.21186	-0.12136	1	0.90639

Virtual True Coincident DPA 4011A Tetra Configuration Az90

corr B format from TA'ed A pre filter			

1	-0.078439	-0.19807	0.21249
-0.078439	1	0.41809	-0.39805
-0.19807	0.41809	1	-0.088494
corr direct signal B format from TA'ed A pre filter			
1	0.058871	0.97558	0.050496
0.058871	1	0.17335	0.97369
0.97558	0.17335	1	0.16355

Physical Near Coincident Double MS Z Configuration Sennheiser Twins Az0

corr B format from Double MS Z (W:fbdu)			
1	0.63918	-0.013452	-0.1761
0.63918	1	0.10601	4.26E-06
-0.013452	0.10601	1	0.033039
corr Direct B format from Double MS Z (W:fbdu)			
1	0.92258	0.36788	-0.12594
0.92258	1	0.45137	-0.098126
0.36788	0.45137	1	0.079223

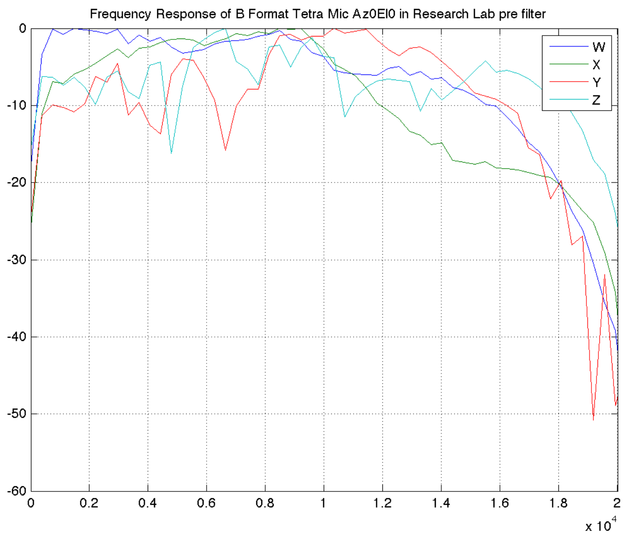
Virtual True Coincident Double MS Z Configuration Sennheiser Twins Az0

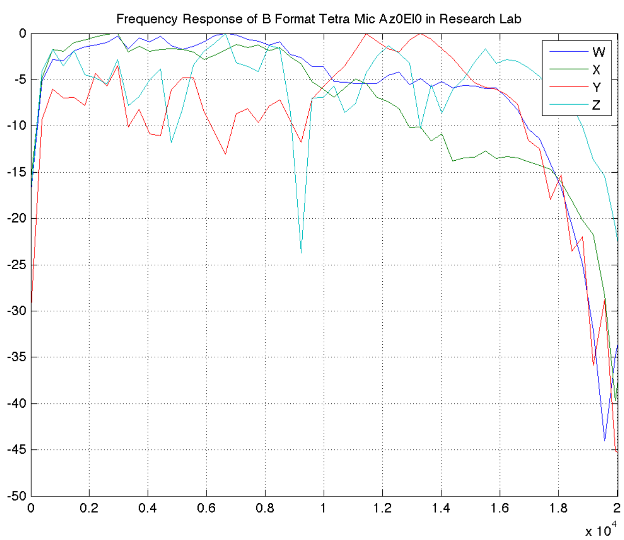
corr B format from Double MS Z (W:fbdu)			
1	0.63781	0.11504	-0.066412
0.63781	1	0.11069	0.02532
0.11504	0.11069	1	-0.041518
corr Direct B format from Double MS Z (W:fbdu)			
1	0.92941	0.88876	0.50519
0.92941	1	0.81338	0.42268
0.88876	0.81338	1	0.44234

APPENDIX E – B FORMAT SPECTRAL PLOTS

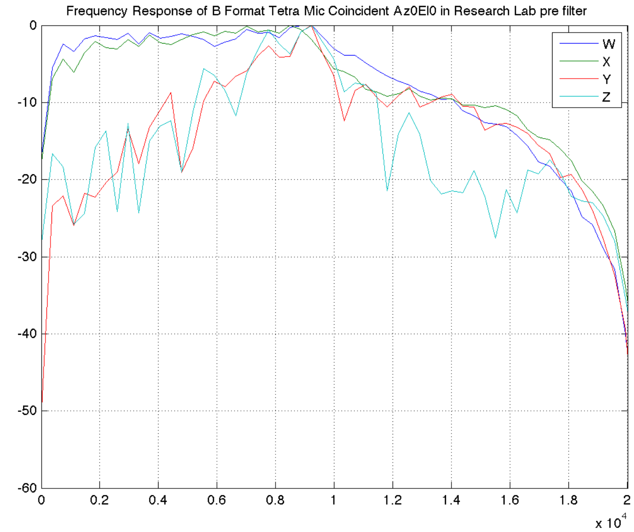
Research Lab

Physical Near Coincident Core Sound Tetra Mic Az0

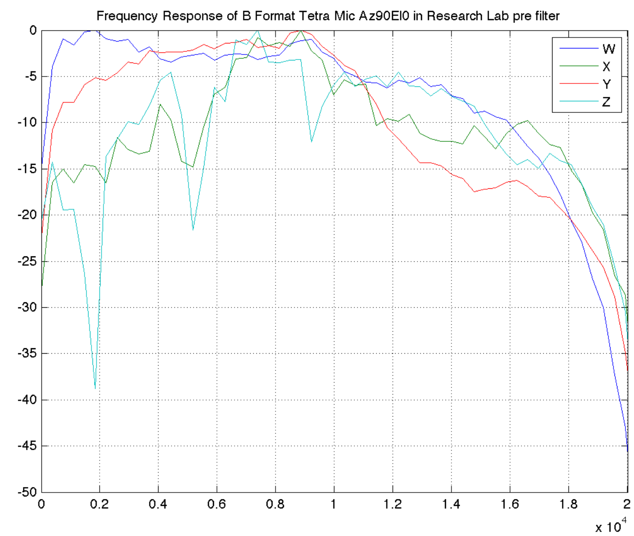


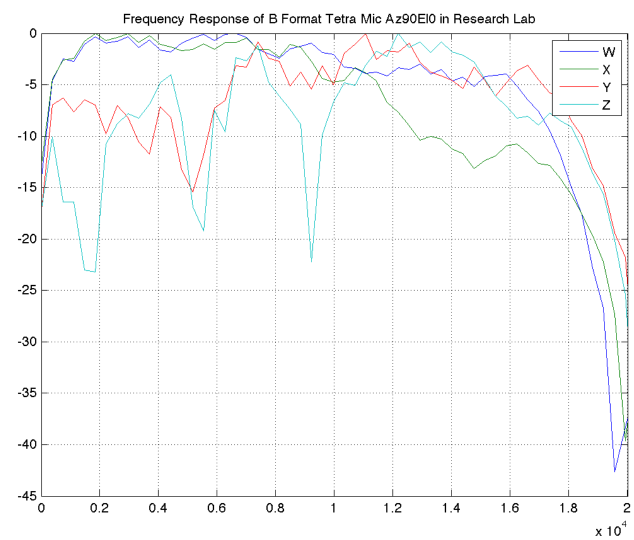


Virtual True Coincident Core Sound Tetra Mic Az0

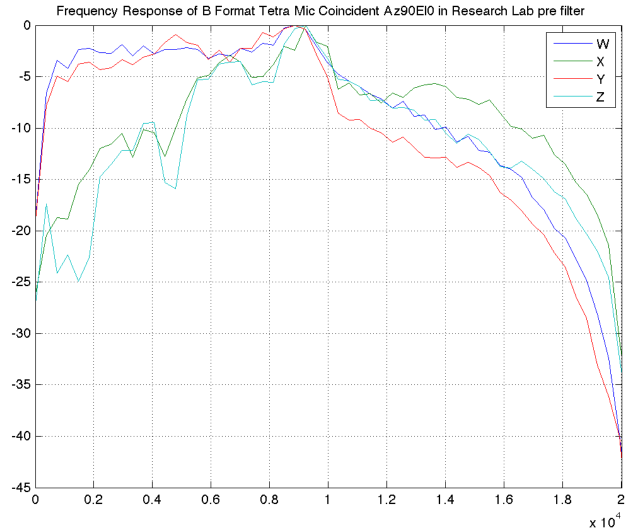


Physical Near Coincident Core Sound Tetra Mic Az90

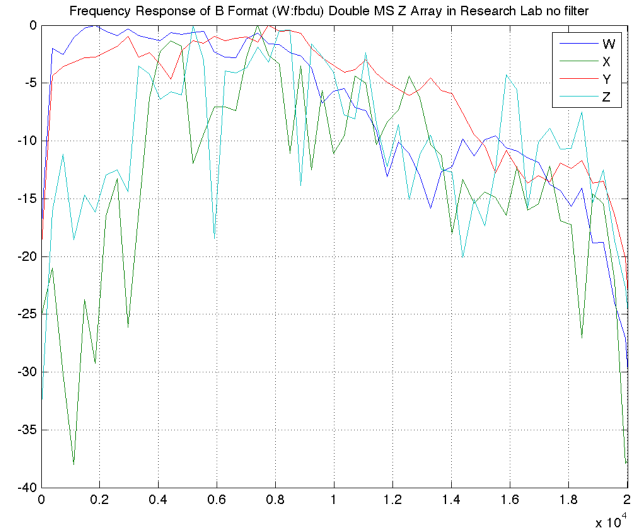




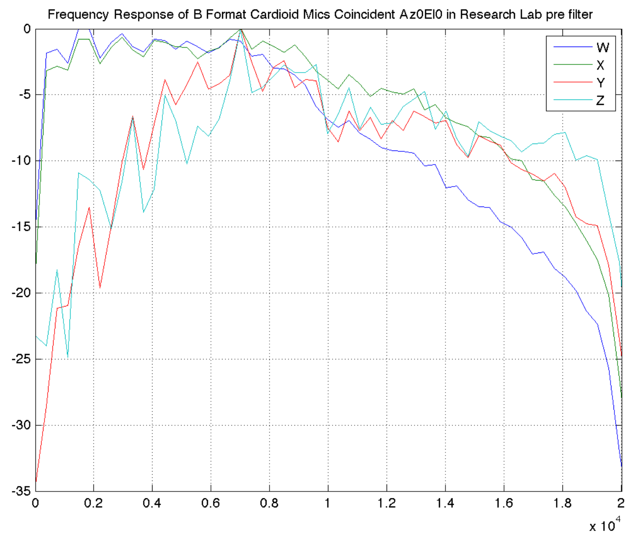
Virtual True Coincident Core Sound Tetra Mic Az90



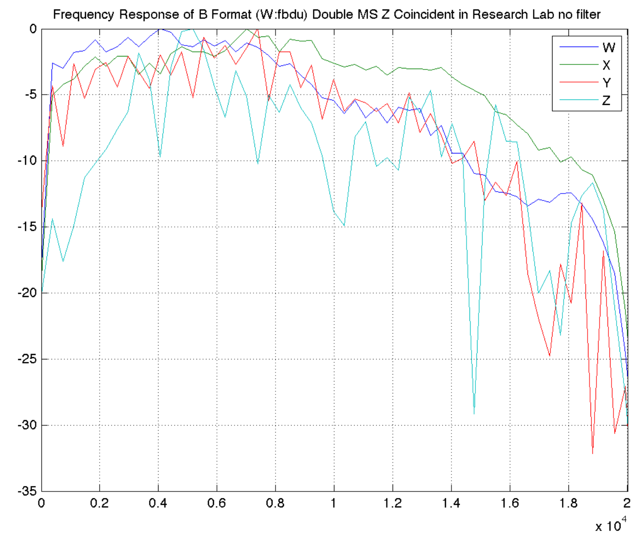
Virtual True Coincident DPA 4011A Tetra Configuration Az0



Virtual True Coincident Double MS Z Configuration Sennheiser Twins Az0

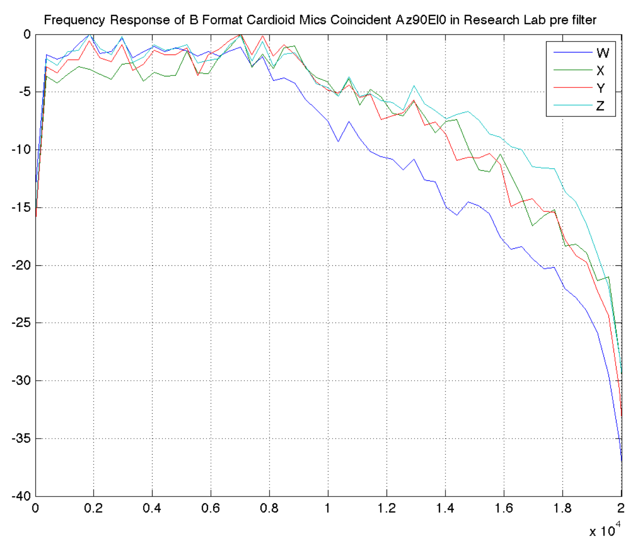


Virtual True Coincident DPA 4011A Tetra Configuration Az90

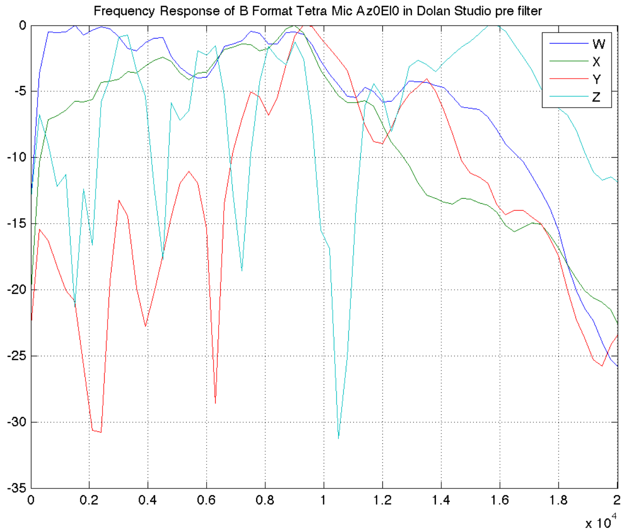


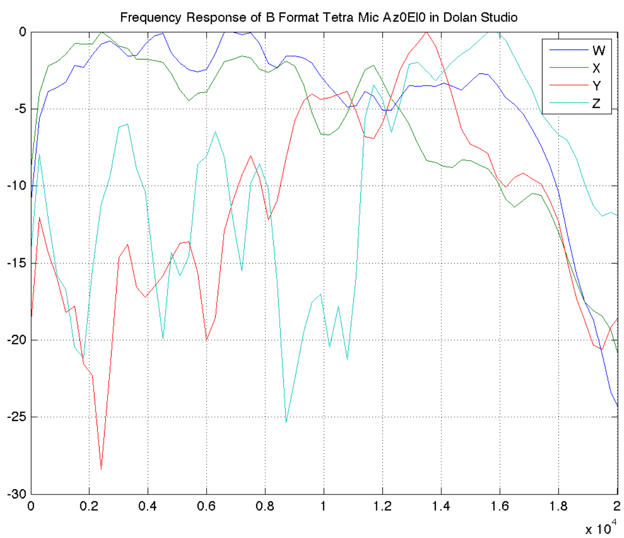
Dolan Studios

Physical Near Coincident Core Sound Tetra Mic Az0

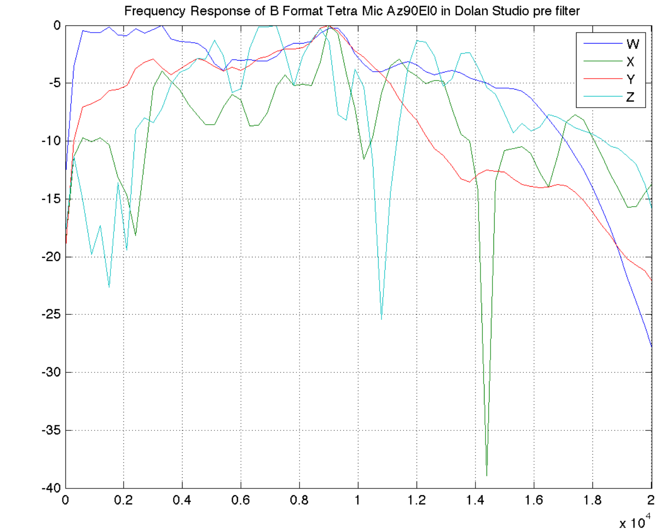


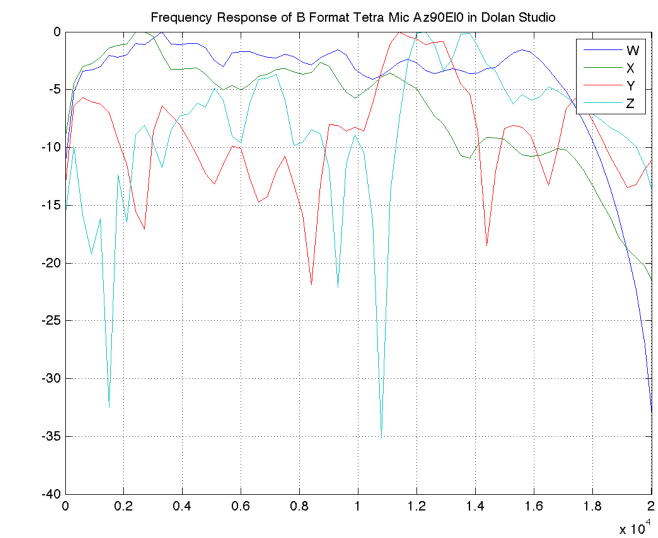
Physical Near Coincident Double MS Z Configuration Sennheiser Twins Az0



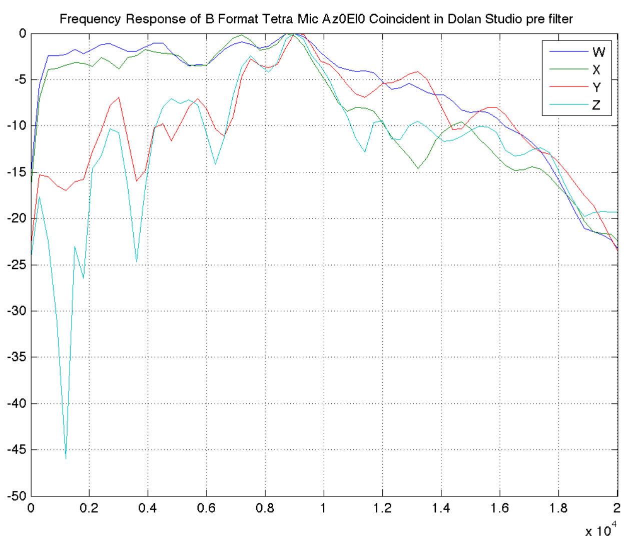


Virtual True Coincident Core Sound Tetra Mic Az0

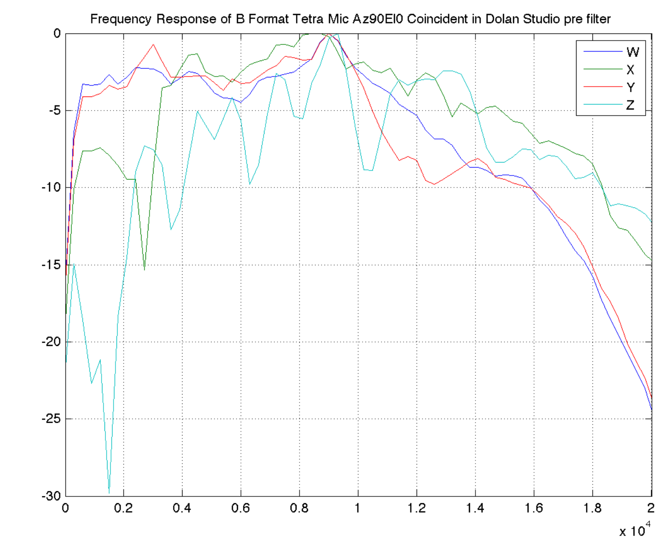




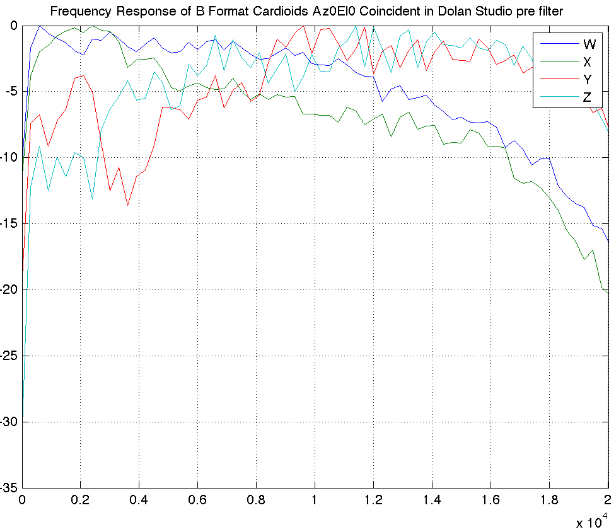
Virtual True Coincident Core Sound Tetra Mic Az90



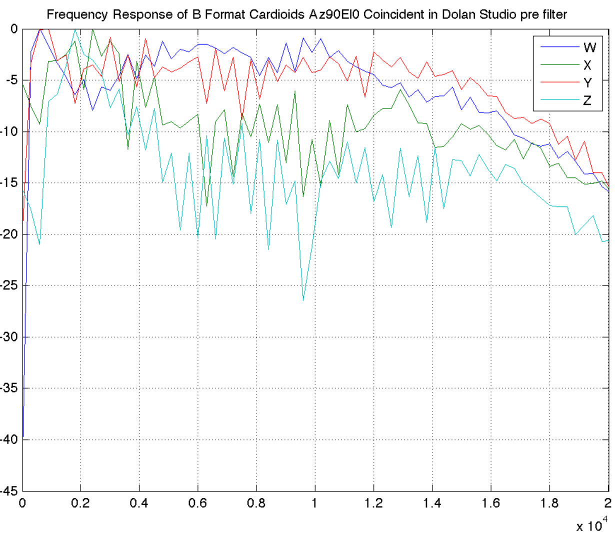
Physical Near Coincident Core Sound Tetra Mic Az90



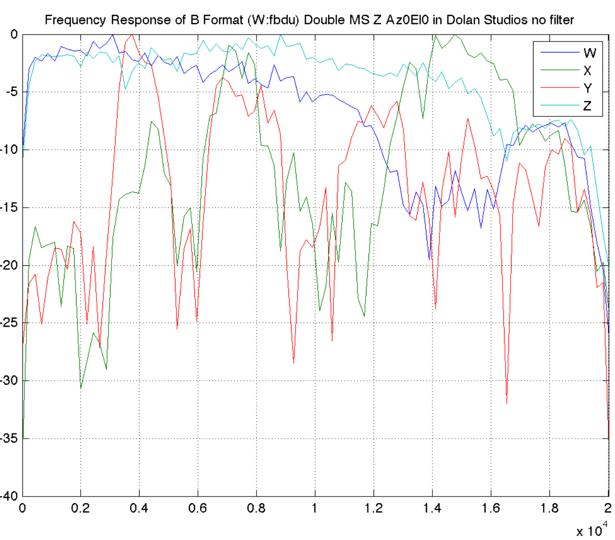
Virtual True Coincident DPA 4011A Tetra Configuration Az0



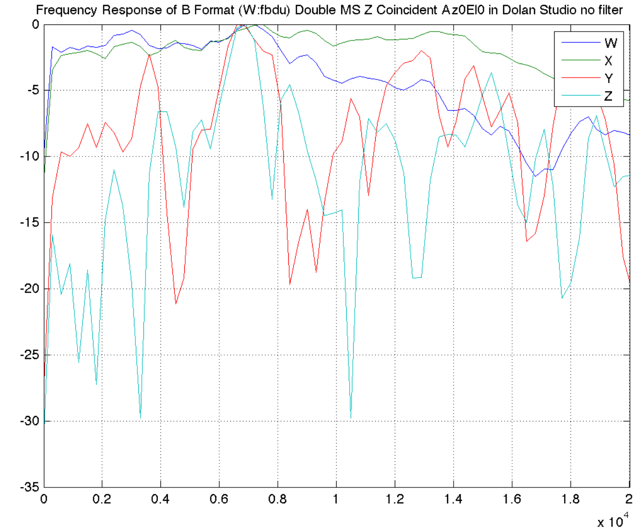
Virtual True Coincident DPA 4011A Tetra Configuration Az90



Physical Near Coincident Double MS Z Configuration Sennheiser Twins Az0

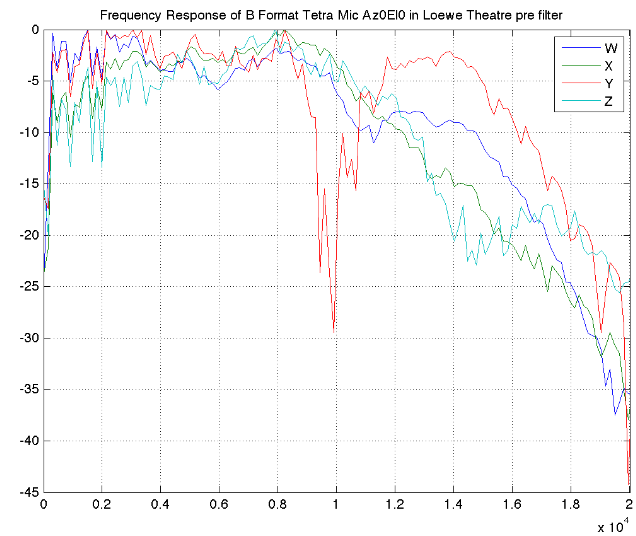


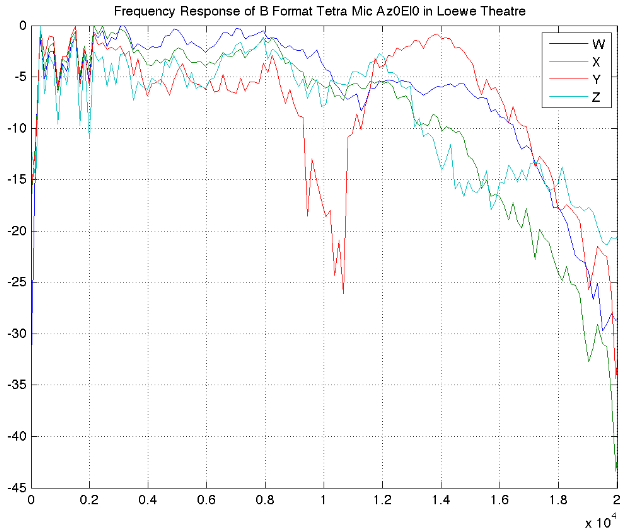
Virtual True Coincident Double MS Z Configuration Sennheiser Twins Az0

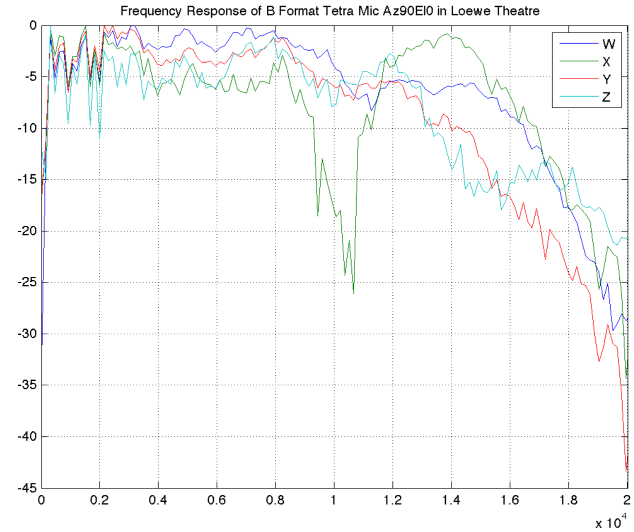


Loewe Theatre

Physical Near Coincident Core Sound Tetra Mic Az0

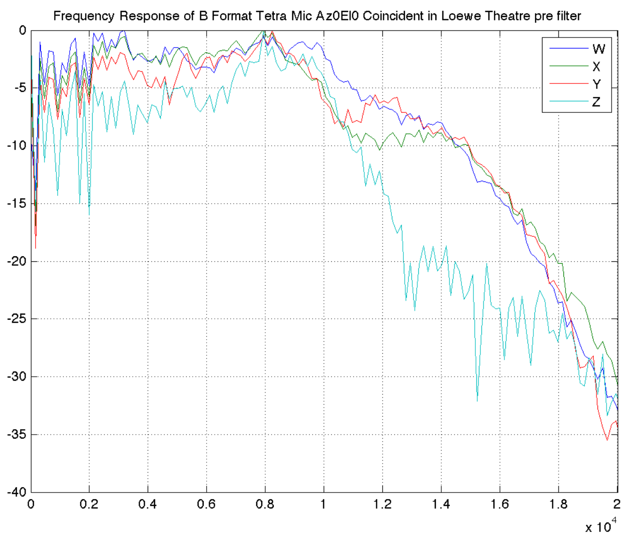


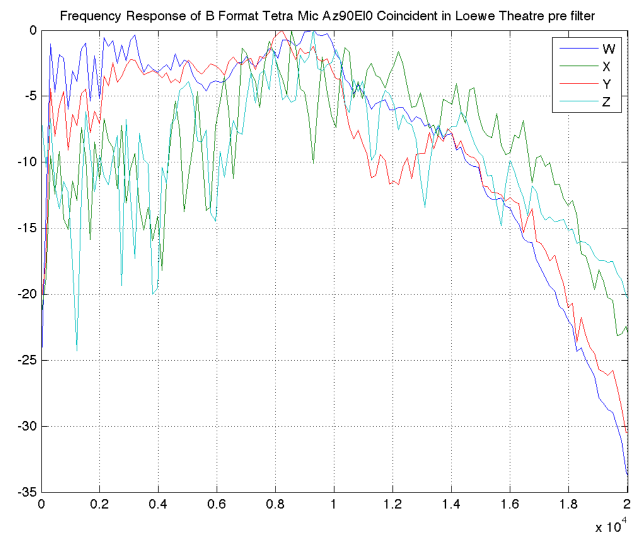




Virtual True Coincident Core Sound Tetra Mic Az0

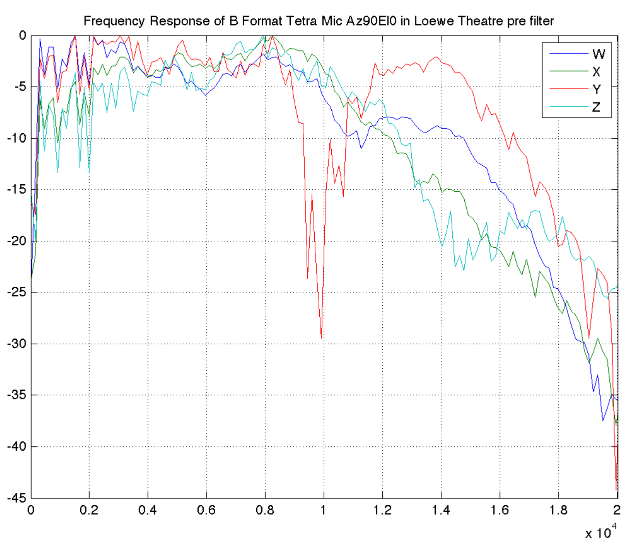
Virtual True Coincident Core Sound Tetra Mic Az90

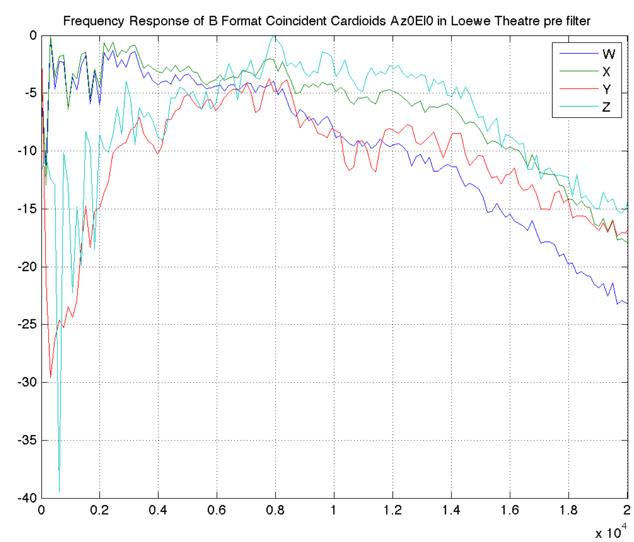




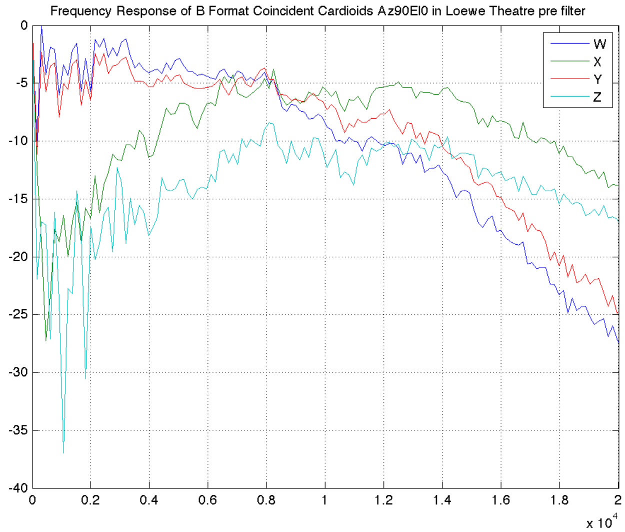
Physical Near Coincident Core Sound Tetra Mic Az90

Virtual True Coincident DPA 4011A Tetra Configuration Az0

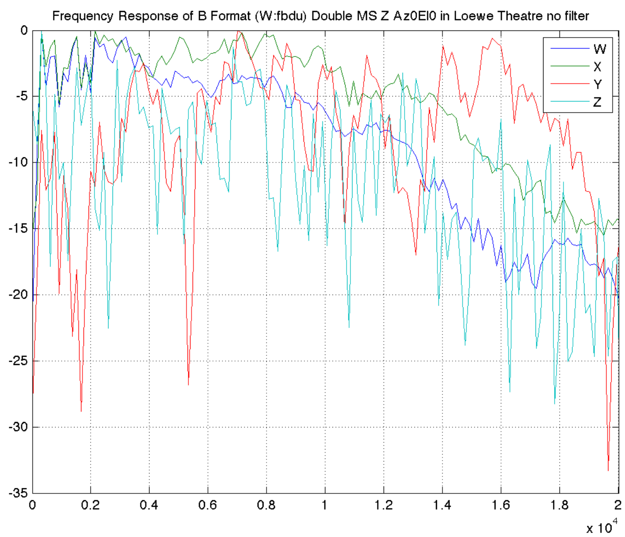




Virtual True Coincident DPA 4011A Tetra Configuration Az90



Physical Near Coincident Double MS Z Configuration Sennheiser Twins Az0



Virtual True Coincident Double MS Z Configuration Sennheiser Twins Az0