
Hypothesis Paper to support a deeper Technical Analysis of MQA (Master Quality Authenticated) by MQA Limited

Document Version

Version	Changes	Author
1.0 / 06.03.2017	Initial Version	Stephan Hotto / XiVero GmbH

Document References / Sources

#	Document/Web	Content	Publication Date	Applicant/Author
1	Patent Application: WO2013/186561 A2	Doubly Compatible Lossless Audio Bandwidth Extension	December 2013	Meridian Audio Limited
2	Patent Application: WO2015/189533 A1	Digital Encapsulation of Audio Signals	December 2015	Meridian Audio Limited
3	Patent Application: WO 2016087583 A1	Non-linear filter with group delay at pre-response frequency for high res audio	June 2016	Meridian Audio Limited
4	Patent Application: EP3029674 A1	Mastering improvements to audio signals	December 2014	Meridian Audio Limited
5	A Hierarchical Approach to Archiving and Distribution" paper (AES 137 th Convention, Los Angeles, USA, 2014 October)	Most likely explaining the basic implementation of MQA	October 2014	John Robert Stuart Peter G. Craven
6	Sparse Sampling – Theory and Applications	Sampling signals with Finite Rate of Innovation	November 2009	Pier Luigi Dragotti
7	Quantization and Dither: A Theoretical Survey	Subtractive and Non-Subtractive Dithering	May 1992	Stanley P. Lipshitz
8	JAS Journal 2015 Vol. 55 No. 5	About MQA	2015	J. Robert Stuart Keith Howard

Content

1	INTRODUCTION	3
1.1	About the Authors	3
1.2	Document Structure	3
1.3	Do we need MQA for Audio Compression?	3
2	WHAT IS MQA?	5
2.1	Overview	5
2.2	In more Detail.....	5
3	HYPOTHESES APPROACH	8
3.1	Why do we go for Hypotheses?	8
3.2	The Hypotheses.....	8
4	HYPOTHESES	9
4.1	MQA is a lossy audio codec ... degrading the achievable audio quality	9
4.2	Decoder free Alternative Audio Compression Scheme	21
5	TECHNICAL DETAILS	26
5.1	Signal Analysis	26
5.2	Time domain and Frequency Domain are absolutely Equivalent!.....	26
5.3	Bandwidth Limiting Causes Ringing and Signal Blurring	27
5.4	Apodization.....	32
5.5	Aliasing	34
5.6	Nyquist-Shannon Sampling Theorem and Low Pass Filtering.....	38
5.7	Sparse Sampling of Signals with Finite Rate of Innovation	39
5.8	Analog Signal Reconstruction – Interpolation Filter.....	40
5.9	Signal to Noise Ratio (SNR), Quantization Noise and Dithering	41
5.10	De-Blurring – Inverse Filtering	43
6	CONCLUSION.....	44
6.1	Summary.....	44

1 Introduction

1.1 About the Authors

We are a small startup company (XiVero GmbH - <https://www.xivero.com/>) based at Düsseldorf / Germany, specialized in digital signal processing within the high frequency and audio domain, developing solutions for the HiFi enthusiasts and the audio industry.

One of our goals is to hand the audiophiles the same tools used by audio engineers to verify the quality of music recordings, advertised as high resolution audio.

This paper is on analyzing several technical aspects of MQA as conscientious as possible, but of course we are not omniscient. If our readers identify wrong statements, we would be more than happy to correct them in the next version of the hypothesis paper.

MQA is a highly proprietary solution and as of our knowledge there are no software encoders and decoders openly available that we could use to do a real-world analysis of the MQA audio compression scheme. For that reason, we have to rely on publicly available information (e.g. Meridian Patents, MQA-Technical Paper, etc.).

We position two hypotheses to get a grip on the proprietary technology behind MQA to provide the readers with sufficient information to do informed decisions whether MQA is a product they want to consume as music lovers.

1.2 Document Structure

MQA and all other forms of digital signal processing apply some algorithms making the whole analysis a bit more difficult to read.

A “Technical Details” chapter (**pls. see chapter 5**) explains all aspects in-depth. The hypotheses are referencing to the detailed technical information which are critical to understand the methods involved.

1.3 Do we need MQA for Audio Compression?

We as HiFi enthusiasts, sometimes called “Audiophiles”, want to get the best audio experience possible. This includes access to recordings that contain the most detailed audio information we could get our hands on.

With the advent of High Resolution Audio there is now the possibility to playback audio files with sample rates as high as 192kHz at a resolution of 24Bit. Well, there are even recordings in 352.8kHz / 32Bit, but those aren’t subject of today’s discussion.

Unfortunately, there is at least one disadvantage associated with increased sample rates and that is file size. In comparison to a 44.1kHz / 16Bit stereo file having a data rate of 1,411.2Kbit/s, a 192kHz / 24Bit audio record asks for 9,216Kbit/s.

Well, because hard disk storage isn't a scarce commodity anymore we are able to handle those larger files without any issues. Nevertheless, when it comes to downloads it makes sense to reduce the file size by applying several compression methods that must be in the case of native high resolution audio "lossless".

As audiophiles, we don't want any degradation of the audio signal and therefore there is no way to use "lossy" codecs (e.g. MP3, AAC, etc.) that throw away parts of the spectral information based on psychoacoustic models of our auditory system that need to be changed as soon as new research results appear.

In favor of sound quality, it is definitely necessary to choose "lossless" codecs (e.g. FLAC, ALAC, etc.) to reduce the file size for a more efficient download.

Most recently we are in the age of streaming, where streaming provider use "lossy" formats but some want to stream at least at CD audio quality which still works quite well through a limited internet channel.

High resolution audio is a new challenge for internet streaming that asks for a "lossless" codec and a small footprint in terms of file size.

As mentioned above, we could go for FLAC which is currently one of the most efficient lossless audio codecs available, but those data rates are still quite high, at least for today's internet and especially mobile LTE channels.

That is the point in time where Meridian Audio Limited placed its new MQA (Master Quality Authenticated) workflow including an audio compression scheme creating a smaller file size to support today's streaming infrastructures.

The company MQA Limited is responsible to drive the development and marketing of this new invention.

2 What is MQA?

2.1 Overview

We have the highest respect for J. Robert Stuart who is the inventor of numerous audio technologies driven by Meridian Audio Limited.

One of the newest technology launched by the Meridian engineers is MQA (Master Quality Authenticated), advertised not only as a novel audio compression scheme but also as a completely new mastering process that is said to aim improving the whole recording and playback chain, from making sure that a recording is authenticated to applying special technologies like apodized filters.

Within this paper, we just like to focus on the audio compression part of MQA because the strongest argument to go for MQA is currently its smaller file size in comparison to the original high resolution audio files.

We like to look at the high resolution music catalogs, currently batch processed into MQA and we like to ask whether the audio quality is degraded by applying the MQA audio compression scheme or whether there is really an advantage for consumers of MQA in terms of audio quality.

Especially, we like to ask whether it makes sense to provide MQA as a download option although it would be possible to download the unaltered original real native high resolution recordings.

2.2 In more Detail

There are three main claims of MQA we like to look at:

- 1.) MQA achieves a high temporal resolution because of the filters and sampling techniques used.
- 2.) MQA compresses a 192kHz / 24Bit recording transparently in a lossless manner into a 48kHz / 24Bit baseband.
- 3.) MQA claims to improve the sound quality of the original recording.

On item 1:

We will show that MQA in contrast to item 1 in fact limits the bandwidth of 192kHz / 24Bit native high resolution records and therefore reduces the temporal resolution achievable.

Another important point is the application of linear-phase filters. We're going to prove that those filters do not impact the audio signal at all, as long as they operate outside of the audio spectrum. That is especially true for 192kHz / 24Bit recordings where the anti-aliasing and interpolation filters are effective at around 96kHz far above the audio spectrum reaching to around 60-70kHz (**pls. see chapter 5.3**).

We will explain later that limiting the bandwidth of a signal inevitably blurs its time domain resolution, whereas if we don't do any frequency band limiting then the time

domain resolution remains untouched! In simple words, if we don't touch the original high resolution audio file then we don't lose any time domain resolution!

For MQA it is in fact an issue because the end to end channel frequency response drops quite early (**pls. see [8]**).

On item 2:

A quote from the “A Hierarchical Approach to Archiving and Distribution” paper (AES 137th Convention, Los Angeles, USA, 2014 October) written by J. Robert Stuart and Peter G. Craven inventors of MQA show what is meant with transparent channel and lossless.

The term “lossless” is used rather in the sense that the changes to signal are not noticeable by our auditory system:

“Even though some musical instruments produce sounds above 20 kHz [53] it does not necessarily follow that a transparent system needs to reproduce them; what matters is whether or not the means used to reduce the bandwidth can be detected by the human listener.”

Why do we need MQA to compress an audio file, although there are already lossless compression schemes available (e.g. FLAC) to make sure that the de-compressed audio file is bit by bit equal to the original?

Let's Bob Stuart (inventor of MQA) answer the question himself by taking a quote from his paper **[8]**.

Quote:

“High-performance lossless compression can improve matters by reducing the data rate to 2.9 Mbps per channel, a saving of 37 per cent, but this is still inefficient and too high to be ideal for streaming from online music services.”

Information theory tells us that we can't compress a signal with a certain entropy below a dedicated data rate. In our case, we need to put the information of 75% of the original bandwidth into the lower bits of a data rate which is just 25% of the original sample rate.

The available Shannon information space of the native high resolution record is most likely not fully used but there is still the need to reduce the bit-depth of the baseband and the high frequency sub bands to make space for the compressed sub-bands to be placed in the lower 7 bits of the baseband.

If we define “lossy” as changing the content of the audio file in a way that we throw data away on basis of the auditory system then MQA is by definition “lossy”.

The question is, whether the data removed from the original audio file are impacting the perceived audio quality.

As we learned from the past, any compression scheme that works with assumptions about our auditory system (e.g. MP3, AAC, etc.) has been proven wrong with new research at the horizon.

We will show that there are further methods available to reduce the file size of already FLAC encoded high resolution audio by keeping its sample rate and hence temporal resolution. What's more important, those technologies don't need any kind of decoder and the music can be played back by any available audio player. Furthermore, that solution would be free of any royalties.

On Item 3:

If MQA's claim to improve the sound quality of the original recording should mean that MQA encoding of a native high resolution audio file improves its quality during playback then this is quite a claim.

Our understanding, which can be wrong, is that the apodizing filter of the MQA encoder aims at reducing the pre-ringing caused by brick-wall linear phase filters in the audio chain.

We do explain apodization (**pls. see chapter 5.4 Apodization**) in detail and show that it is just a process to reduce the bandwidth in a way to minimize ringing, with the side effect of impacting the temporal resolution.

Chapter 5.3 explains that bandwidth is proportional to time domain resolution and any reduction in bandwidth blurs/smears a signal in the time domain.

Therefore, we will demonstrate that a brick-wall linear phase filter operating outside of the audio spectrum - as it is the case for 192kHz and even in most cases for 96kHz native high resolution audio - provides a higher temporal resolution and therefore less blurring/smearing in the time domain than any other kind of apodized minimum or non-linear-phase filters applied by MQA.

3 Hypotheses Approach

3.1 Why do we go for Hypotheses?

As mentioned above, MQA is a proprietary solution which makes it difficult to look into its inner workings. Only limited information is available (e.g. patents, white papers, MQA papers, etc.) allowing us to do a technical analysis.

Therefore, our goal is to formulate hypotheses to prove or disprove the technical assumptions behind them.

It would be an honor if Bob Stuart himself would assist to get a better understanding of his invention and supports us in disproving the most critical hypothesis, that states that MQA is a “lossy” audio compression scheme doing harm to already available native high resolution recordings that are MQA batched processed, not only for streaming but also for downloading.

3.2 The Hypotheses

We will have a look at the following two hypotheses set up to formulate methods to either prove or disprove the claim.

1. Hypothesis: MQA is a lossy audio codec, not able to reproduce the original high resolution recording during playback, degrading the achievable audio quality.
2. Hypothesis: There is an alternative to the MQA audio compression scheme for the application of streaming that doesn't need a special decoder and doesn't alter the audio signal in any adverse manner.

4 Hypotheses

4.1 MQA is a lossy audio codec ... degrading the achievable audio quality

Hypothesis:

MQA is a lossy audio codec not able to reproduce the original high resolution recording during playback, degrading the achievable audio quality.

Proof:

The hypothesis is proven if the decoded MQA audio file is different from the original high resolution recording in its frequency response (magnitude & phase) and therefore in time domain appearance.

This proof is complicated by claims that changing the audio signal by applying the “Mastering Quality Authentication” process actually improves the original native high resolution audio file. We will explain that this claim cannot stand for 192 kHz native high resolution records. **(pls. see chapter 5 Technical Details).**

Disproof:

The original and the MQA encoded audio file are similar in a sense that the frequency response in magnitude & phase and therefore the time domain appearance of the audio signal are unaltered in any way.

Technical Analysis to either prove or disprove the hypothesis:

Again, we need to emphasize that MQA is a closed and proprietary solution, therefore all our technical statements are conclusions from information openly available, either provided by Meridian/MQA itself or by secondary sources.

For all further discussions, we consider a native high resolution audio file that has been recorded at 192kHz / 24Bit, wherein any filtering during the recording and playback process has happened above the relevant audio spectrum, as input for the 48kHz / 24Bit MQA compression scheme.

MQA Technical Details:

MQA can fold any sample rate up to 384kHz / 24Bit into a 48kHz / 24Bit MQA file. The scheme is hierarchical, which means that the output could be also a 96kHz / 24Bit MQA file.

MQA always remains in the sample base of the source material. In the case of 352.8kHz recordings the maximum compression would be 44.1kHz / 24Bit.

It is important to know that the MQA encoded baseband can be played by any audio player with a bit-depth and therefore SNR similar to a well dithered compact disk.

MQA Overall Transmission Frequency Response

The overall MQA channel transmission (AD- to DA-Conversion) as described by the MQA paper [8] (**Figure 1 - JAS Journal 2015 Vol. 55 No. 5**) shows a drop in spectral magnitudes of already around 4dB at 40kHz. That in itself already changes the frequency response and therefore the temporal resolution of the original high resolution audio signal (**pls. see chapter 5.3**).

Furthermore, because the impulse response of the MQA channel is a non-linear phase filter, as shown in the same graph, the phase response of the native high resolution audio recording is altered too.

MQA aims at keeping the temporal resolution as high as possible and furthermore they don't want to introduce pre-ringing, making them arrive at the decision to use apodized non-linear-phase filters (**pls. see chapter 5.4 Apodization**) with a shallow slope.

Such approach has unfortunately the draw back that the filter needs to kick in early or that an early frequency droop is traded for aliasing because frequencies above the Nyquist-Frequency aren't sufficiently attenuated.

Quote from the patent WO 2015/189533 A1:

“However, we take the view that filters that would be considered correct in communications engineering are not audibly satisfactory, at least not at sample rates that are currently practical for mass distribution. We accept that aliasing may take place and are proposing to balance aliasing against 'time-smear' of transients due to the lengthening of the system's impulse response caused by filtering.”

We completely agree with MQA that the temporal resolution of an audio recording is the most critical part and we have to do everything to not changing it at any stage in the recording and playback process.

Because temporal resolution is proportional to bandwidth (**pls. see chapter 5.3**), a 192kHz / 24Bit recording provides the highest resolution we could get.

Even if we don't hear a single sinus at 30kHz, those frequency components are very important to achieve the necessary temporal resolution our auditory system seems to be able to recognize (**pls. see [8]**).

Higher sample rates aren't beneficial because the audio spectrum is definitely limited to a frequency range below 96 kHz (**pls. see Figure 1**).

MQA's comparison of their frequency response with that of a 192kHz / 24Bit channel (**pls. see [8]; Fig. 3**) is only half the truth. We are going to show that in contrast to MQA's frequency response those long linear phase brick-wall low-pass filters operate outside of the audio spectrum at 96kHz and therefore don't do any harm (**pls. see chapter 5.3**).

As of [8] - **Figure 11** the area of interest where the audio signal has significant spectral components ends at about 48kHz. Beyond that frequency the audio signal gets submerged in noise.

For further discussions in the scope of this document we define the minimum audio spectrum for high resolution audio recordings as a frequency band between 0Hz – 48kHz with a 1/frequency spectral magnitude envelope. The whole discussion still stands if we take the best recordings examined, which show spectral components above 60 kHz.

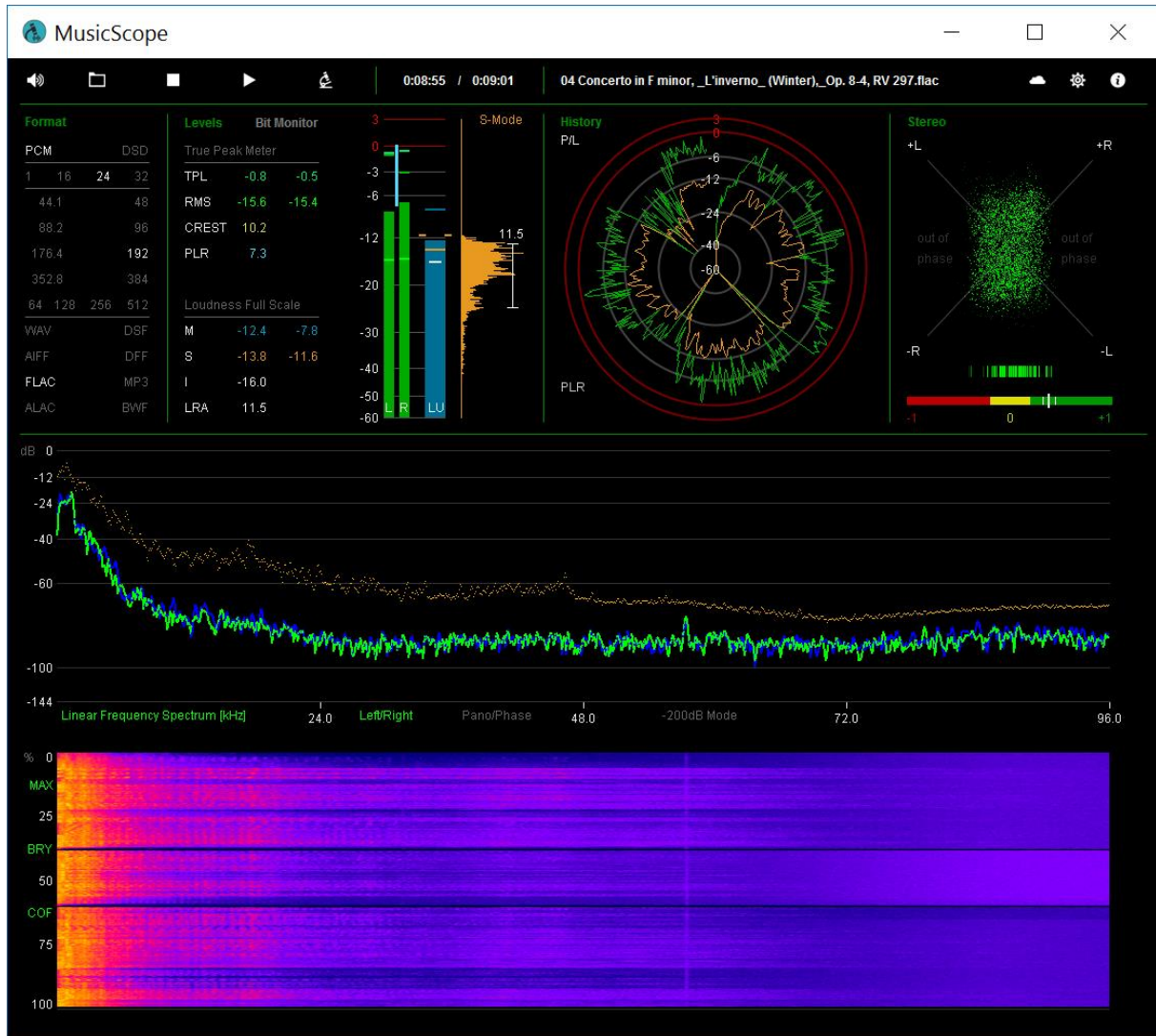


Figure 1: 192kHz/24Bit native high resolution violin recording having still structure within the spectrogram, representing spectral components of music beyond 70kHz!

MQA Origami

A lot has been written about the origami process and it is not that easy to find exact technical details in the patent application papers.

In general, to achieve a baseband of 24kHz (48kHz Sample Rate), the original 96kHz (192kHz Sample Rate) audio baseband needs to be split up into sub bands.

In a second step the content of those sub bands is compressed and placed into the lower bits of the sub sampled baseband represented as a 48kHz / 24Bit WAV or FLAC encoded MQA compressed audio file with 17 Bit audio information and 7 data bits, mask as dithering noise.

We have to emphasize that it is difficult to deduct from the available patent applications how many bits are really used as data bits. The above-mentioned derived numbers are a reasonable conclusion.

Within the proprietary MQA decoder the different sub bands are unfolded and joined together to reconstruct the original high resolution bandwidth and therefore temporal resolution. We will show that this process is in fact not lossless in the defined meaning of the 1st hypothesis.

Subsequently we discuss several techniques implemented in MQA that do change the native high resolution audio content during the encoding process.

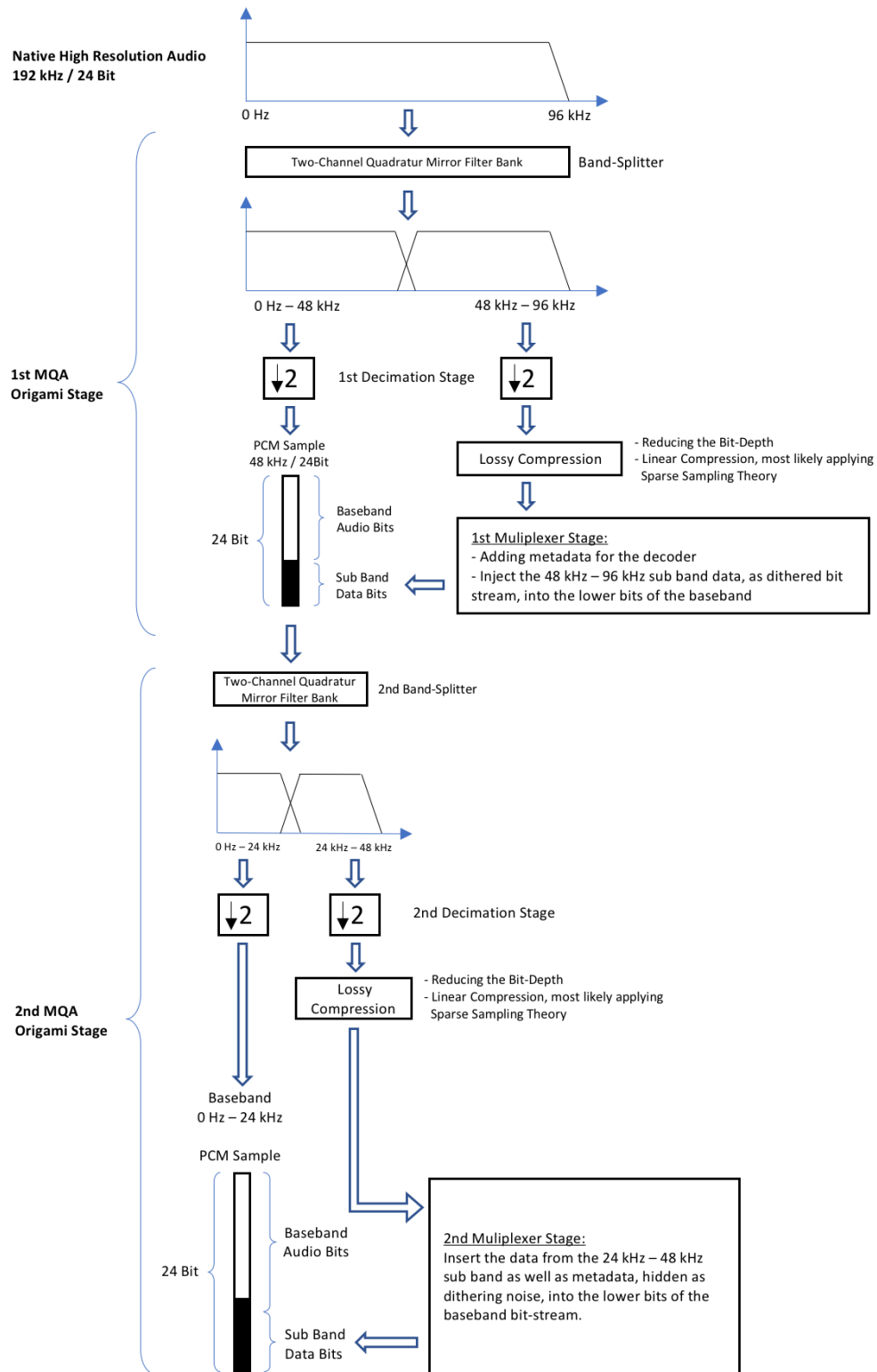


Figure 2: Simplified technical Diagram of the MQA encoder implementation. The real implementation may vary, but the general digital signal processing stages are part of the process.

The MQA Encoder:

To get a better understanding of the digital signal processing technologies involved we like to use the simplified technical diagram (**Figure 2**) outlining the MQA encoder structure, used for a two-fold process that needs to be applied to compress a 192kHz / 24Bit record into a 48kHz / 24 Bit MQA encoded audio file.

The real implementation may vary, but the general digital signal processing stages are part of the process.

1st MQA Encoder Step:

A filter bank separates the 1st baseband (0 – 48kHz) from the sub band (48kHz – 96kHz).

What exact kind of “Two-Channel Quadratur Mirror Filter (QMF) Bank” is used by MQA is not known. Those combined filters have the property to split a signal (Band-Splitter) in an upper and lower frequency range which can be perfectly recombined (Band-Joiner). If the QMFs have been designed in the right way then there is no issue with distortion of the phase or frequency magnitude nor aliasing after the lower and upper frequency bands have been re-joined within the MQA decoder.

The real implementation of MQA allows the playback of the baseband (0Hz – 24kHz) by standard audio players. For the reason that this baseband gets not joined with its upper band before the MQA decoding (**pls. see Figure 5**) it is most likely plagued by phase distortions and aliasing. That reasoning is based on MQA’s claim to designs their filters, and likely the QMFs too, as short non-linear-phase filters to avoid pre-ringing.

Let’s have a look at the baseband of a real MQA record in comparison to a down-sampled version of the original native high resolution record.

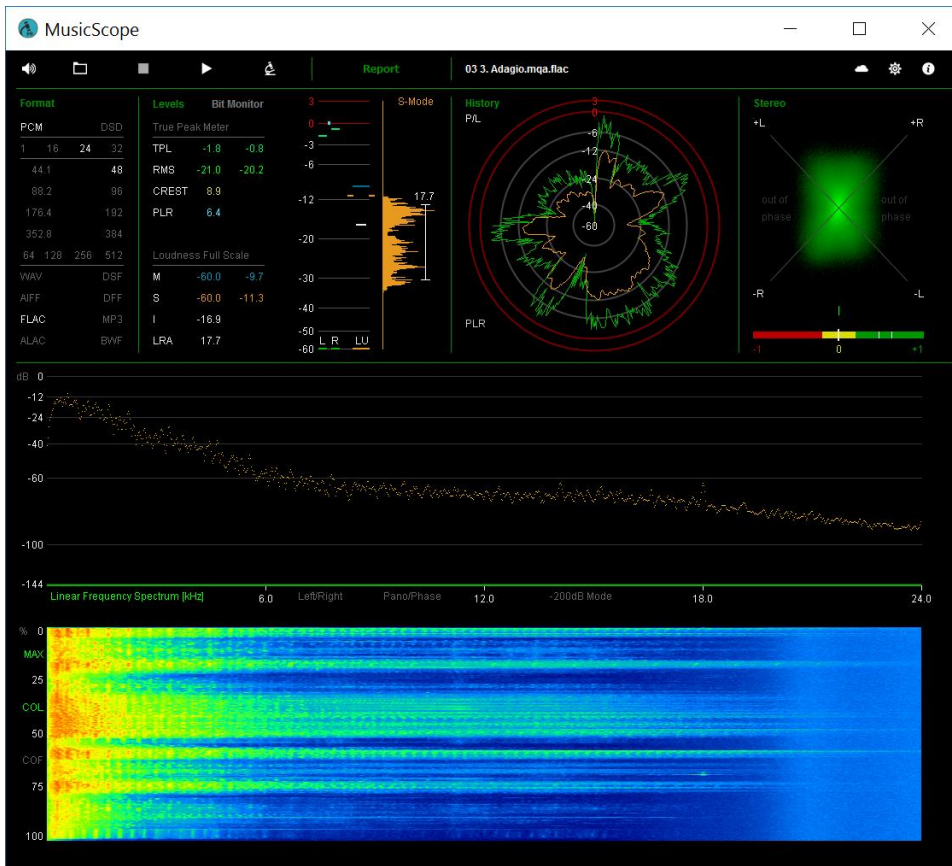


Figure 3: MQA Baseband – The spectrogram exhibits an increased overall noise floor (pls. see in comparison to Figure 4), rising heavily beyond 18kHz.

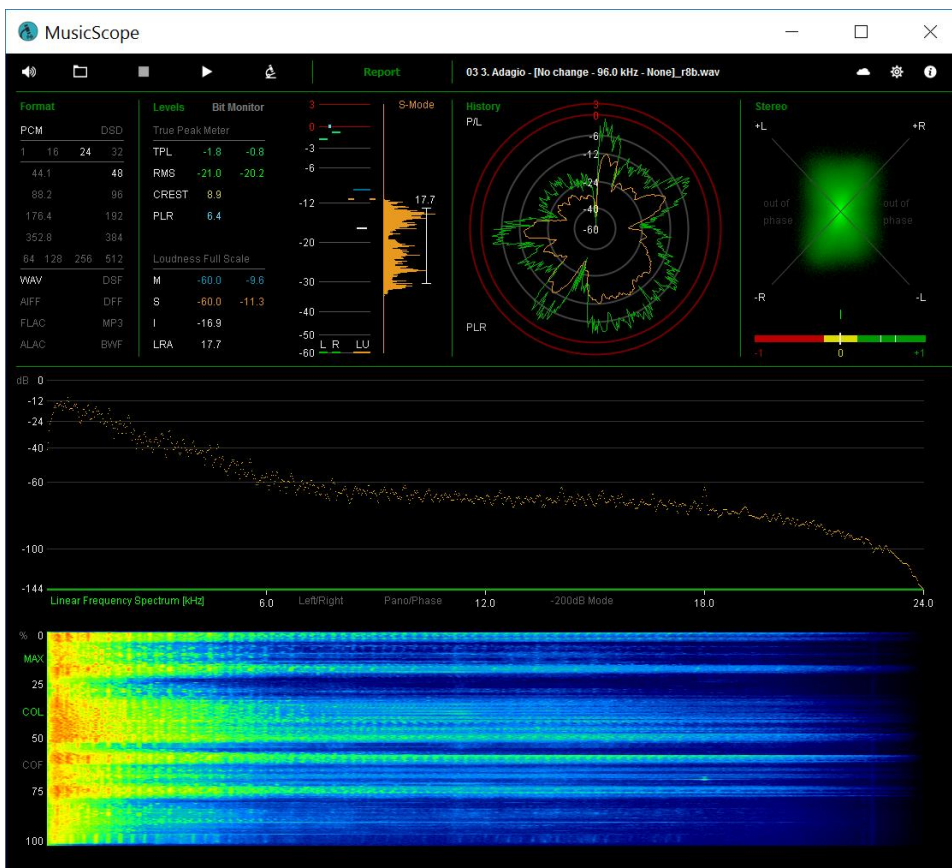


Figure 4: Baseband of the down-sampled native high resolution audio record.

The MQA baseband noise floor is significantly higher than the low noise floor level of the 48kHz / 24Bit down sampled native high resolution record.

MQA states that they reach a Signal to Noise Ratio (SNR) - due to noise shaping - comparable to 20 Bits. We don't know whether that claim can be kept.

The strong rise of noise beyond 18 kHz is either caused by noise shaping or by aliasing.

If it is aliasing then it would disturb standard audio players (**pls. see chapter 5.5**) but the MQA decoder could compensate it by re-joining (QMF Band Joiner) the baseband with the complementary upper frequency band.

IMPORTANT!

The observed reduced SNR does not get better if the MQA decoder adds the high frequency sub-bands because the 7Bits of the 24Bit recording are lost to store the data for the sub bands. Nevertheless, MQA applies nonadaptive noise-shaped dithering to try to compensate at least partly for the loss of bits (**pls. see chapter 5.9**).

2nd MQA Encoder Step:

After separating the frequency bands by applying a "QMF Bank" the two separated signal paths are decimated by a factor of two to reduce the sample rate from 192kHz to 96kHz.

The upper frequency band gets "lossy" compressed by applying methods like sparse sampling compression. The reason why that stage is in fact lossy is the bit reduction to around 17Bits.

Up to now we haven't found information whether the compression algorithm can be put in a state of overload when the upper frequency energy and entropy is that high that the linear prediction algorithm generates residual errors too numerous to be encoded in the lower 7 bits of the baseband. If that is the case then the compressor would be quite lossy.

A multiplexer creates the necessary metadata for the decoder and generates a bit-stream to be written into the lower bits of the baseband.

At this point in time the first MQA Origami step, as described in **[8]** has been done.

The 2nd MQA Origami step is quite similar, so that we end up with the MQA baseband of 48kHz sample rate with a bit depth for the noise shaped dithered audio signal of 17Bits and 7bits of data to be used by the decoder to reconstruct the upper sub bands (24kHz – 96kHz).

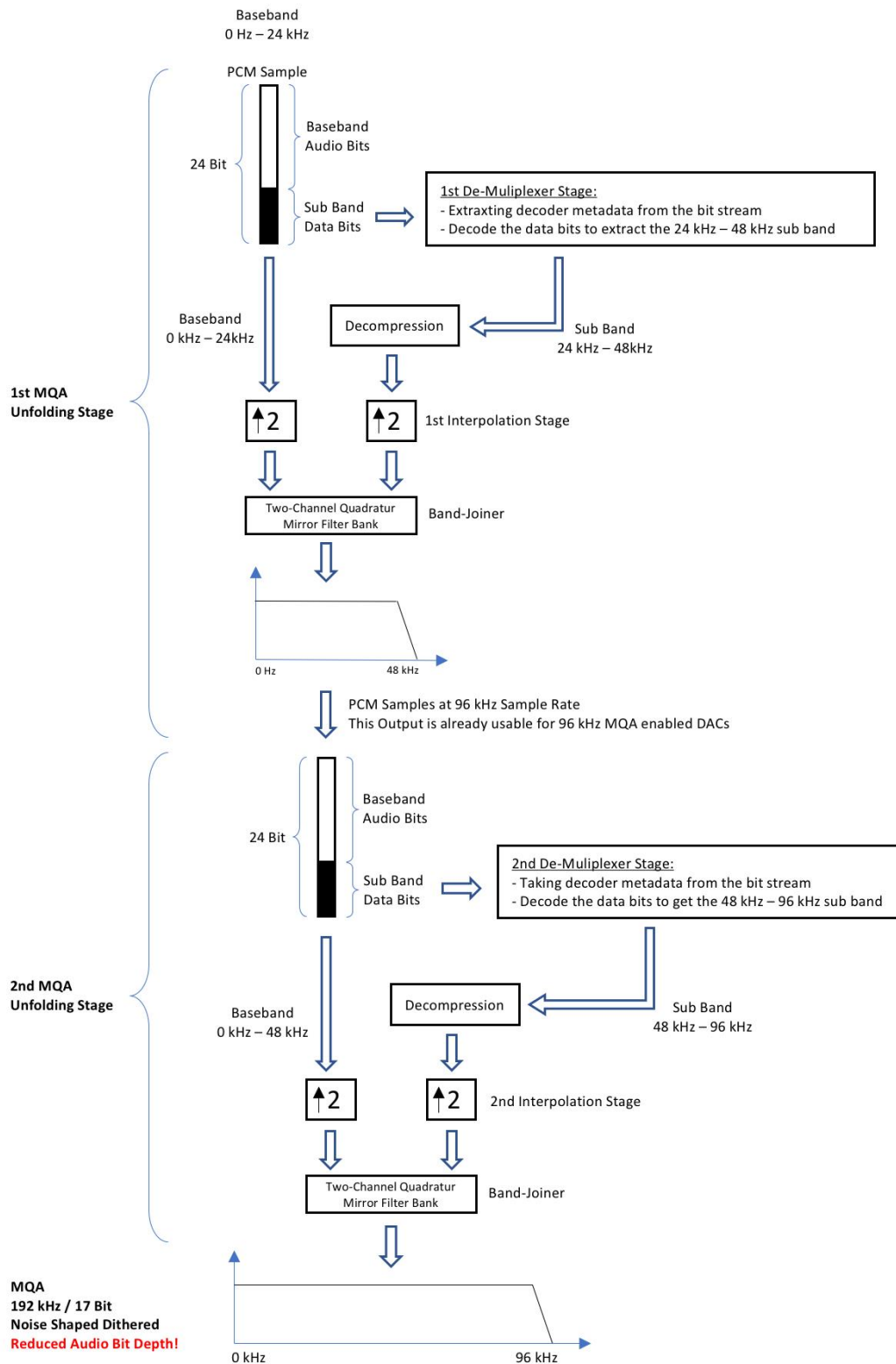


Figure 5: Simplified technical diagram of the MQA decoder implementation. The real implementation may vary, but the general digital signal processing stages are part of the process.

The MQA Decoder:

The simplified technical diagram (**Figure 5**) outlines the MQA decoder structure used to unfold the 48kHz / 24Bit MQA baseband into the 192kHz / 17Bit noise shaped MQA decoded audio file.

The real implementation may vary, but the general digital signal processing stages are part of the process.

MQA Decoder Steps:

- A De-Multiplexer reads the metadata buried in the 7bit data stream to get the information how the data has been encoded.
That is also the stage where the “Authentication Check” happens. Any change to the bit structure makes it impossible for the MQA decoder to do its work.
- A decompressor generates the audio bits for the upper frequency band. We have to keep in mind that this is a bit-depth reduced version of the upper sub band.
- The 1st interpolation stage up-samples both streams to 96kHz.
- A band-joiner (QMF Bank) outputs the first MQA unfold (2. MQA Encoder Origami Step), creating a baseband of 0Hz – 48kHz.

A second decoder segment applies a similar process to unfold the 1st MQA Encoder Origami Step to output the full spectrum baseband 0Hz – 96kHz, with a reduced bit-depth of 17 Bits.

Implications of the Bit-Truncation:

MQA describes in [8] that real-world systems do show thermal noise rendering some bits unusable and that the loss of further bit-depth can be compensated by appropriate dithering (**pls. see chapter 5.9**).

Most uncooled physical systems are able to reach a thermal limited noise floor at around -120dB below full scale. Nevertheless, a theoretical 24Bit system is able to push the quantization noise down to -144dBFS (without dithering) and therefore allows for 24dB = 4Bits of Headroom to place the information of the upper frequency bands.

BUT! We did statistical investigations into existing high resolution recordings and we can confirm that sometimes even the 2nd LSB already holds information, most likely due to applied dithering processes. If we go along with the MQA argument then during playback on today’s HiFi systems that information gets mask in device thermal noise.

Of course, the MQA process uses more than the maximum 4Bits submerged in noise. From the documentation of MQA, it is not fully clear how many bits are occupied with the high frequency sub bands, but we would deduct from the patent applications that we’re talking about 7Bits.

So, that would leave us with effectively around 17Bits in the critical baseband of 0Hz – 24kHz. We have to keep in mind that this limitation remains to be valid, even after MQA-Decoding!

Now comes in the idea of dithering that is most likely implemented as noise shaping (**pls. see chapter 5.9**) assisting to push the higher quantization error and therefore additional noise to frequencies where it is less critical for our auditory system, because the overall noise level cannot be reduced!

The claim that there are still 20 effective bits is therefore most likely only true in the lower frequency range. Furthermore, we're talking about increasing the SNR in relation to the quantization noise, meaning we could not retrieve the real information that has been within the bits before they have been truncated.

Dithering is a statistical process and it is therefore difficult to compare real information within the bits truncated with the dithered version of the signal.

We think that it is safe to say that decoded MQA has not the same bit depth as native high resolution audio, especially in the case of well dithered recordings that show audio information already in the 2nd LSB!

Conclusion of the 1st Hypothesis:

We want to encourage the reader to read first the technical details in **chapter 5** to get a deeper understanding why we came to the following conclusion.

The deduction from the discussed technical arguments proves the hypothesis that MQA is in fact “lossy” is true.

MQA alters the bit-depth as well as the frequency response (magnitude & phase) and therefore the time domain appearance of the original high resolution audio file.

It is debatable whether those alterations are audible, but as we learned from the past, any compression scheme that works with assumptions about our auditory system (e.g. MP3 & AAC) have been proven wrong with new research at the horizon.

As long as streaming is not able to provide larger bandwidth more cost efficient, MQA could be a solution to stream audio better than Compact Disk, MP3 or AAC quality.

As we all know, in a couple of years the bandwidth provided, even in mobile networks, will be large enough to distribute the real native high resolution content, satisfying the audiophiles demands.

For downloads there is no need to go for MQA because the channel allows us to get native high resolution audio files in FLAC format, with the highest temporal resolution achievable that are not altered in any way by applying technologies like MQA.

We will show in our 2nd Hypothesis (**pls. see chapter 4.2**) that a different compression scheme can reduce already compressed FLAC files by an average of 30% - 50% as a decoder and royalty free alternative to MQA for streaming applications.

4.2 Decoder free Alternative Audio Compression Scheme

Hypothesis:

There is an alternative to the MQA audio compression scheme for the application of streaming that doesn't need a special decoder and that doesn't alter the audio signal in any adverse manner.

Proof:

This hypothesis is proven if the new compression scheme archives a significant higher compression rate than the well-known and highly efficient FLAC (Free Lossless Audio Codec).

Furthermore, the compressed audio file needs to be similar to the native high resolution audio input file in a sense that the frequency response in magnitude & phase and therefore the time domain appearance of the audio signal are unaltered in any way, after the signal has been decoded with any available FLAC capable audio player.

Disproof:

The new scheme alters the compressed native high resolution audio file in its frequency response (magnitude & phase) and therefore in time domain appearance. Furthermore, the hypothesis is disproved if the new compression scheme cannot reduce the average file size of already FLAC encoded audio files by around 30%.

Technical Analysis to either prove or disprove the hypothesis:

An alternative to MQA shall avoid impacting the frequency response and therefore should preserve the highest time domain resolution of a well-made 192kHz / 24Bit native high resolution recording.

As a second important requirement, it must be possible to playback the compressed audio file without the need for a special proprietary decoder to get access to the full native high resolution bandwidth!

The scheme we are suggesting has been already discussed by individuals as a possible alternative for MQA.

We want to emphasize that this is not our invention!

The “FLAC Entropy Optimizer” implements, besides of the below mentioned idea, additional statistical algorithms to analyze how many bits are really submerged in noise and how high the bandwidth of the music's spectral components reaches.

As a small transcoding tool “XiFEO” is able to reduce the FLAC file size of high resolution audio for mobile players and streaming without impacting the temporal resolution of the native high resolution record, by staying within the same sample rate.

The technology itself is not patented, so that there are no royalties for those who want to implement the idea in the compression process for streaming!

FLAC Entropy Optimization:

A FLAC encoder becomes much more efficient if we reduce the entropy of an audio file before it gets compressed.

The methods FLAC applies for lossless compression are based on the same assumptions as the technology of “Sparse Sampling of Signals with Finite Rate of Innovation” (**pls. see chapter 5.7**).

A piece of an audio signal is approximated either by a simple polynomial or linear predictive coding (LPC).

An audio sample sequence containing noise is not really a special signal of finite rate of innovation (FIR) that could be sparsely sampled without taking care about the residual error and therefore it is not enough to just encode the coefficients of the polynomial but also the residual error!

The residual error increases with the entropy of the data to be compressed. Simply said, if we could reduce the noise then the FLAC encoder would operate much more efficient.

As a side note, this is another disadvantage of MQA which is almost always distributed as FLAC compressed audio to be compatible for legacy audio players. Because of the dithering used by MQA all bits that carry the high frequency bands are highly random and therefore just noise for the FLAC-Encoder which compromises the achievable compression rate due to a high entropy.

XiFEO’s goal is to reduce the entropy of the file in a two-step process:

- 1.) Statistical Analysis of the bit-structure to find out how many LSB bits are just carrying noise.

- 2.) An analysis that identifies the highest frequency components that still contain spectral information of the music to remove out of band noise.

On item 1:

All LSB-Bits that just contain real noise can be truncated by simply setting them to zero. The effect is a strong reduction in entropy.

The approach is comparable to MQA but they don’t care about the real noise level and therefore truncate always the same number of bits.

During our tests, we learned that sometimes well dithered native high resolution audio files still have audio information within the 2nd LSB and therefore we could only throw away one bit to avoid losing any critical bit-depth. That is quite in contradiction to MQA’s approach to replace those bits with an admittedly sophisticated dithering.

Of course, any truncation asks for dithering which we apply.

On Item 2:

As we will learn in **chapter 5.3** filtering outside of the audio spectrum does not do any harm, even if we apply very long linear phase low-pass brick-wall filters.

So, if we have identified up to what frequency the highest spectral components of the music reach, we can just filter everything else above those frequencies because it is simply out of band noise.

Achievable Compression Factors:

In the case of very well recorded and dithered 192kHz / 24Bit native high resolution audio files it is difficult to achieve a high compression rate because then it would be at most possible to truncate a small number of bits.

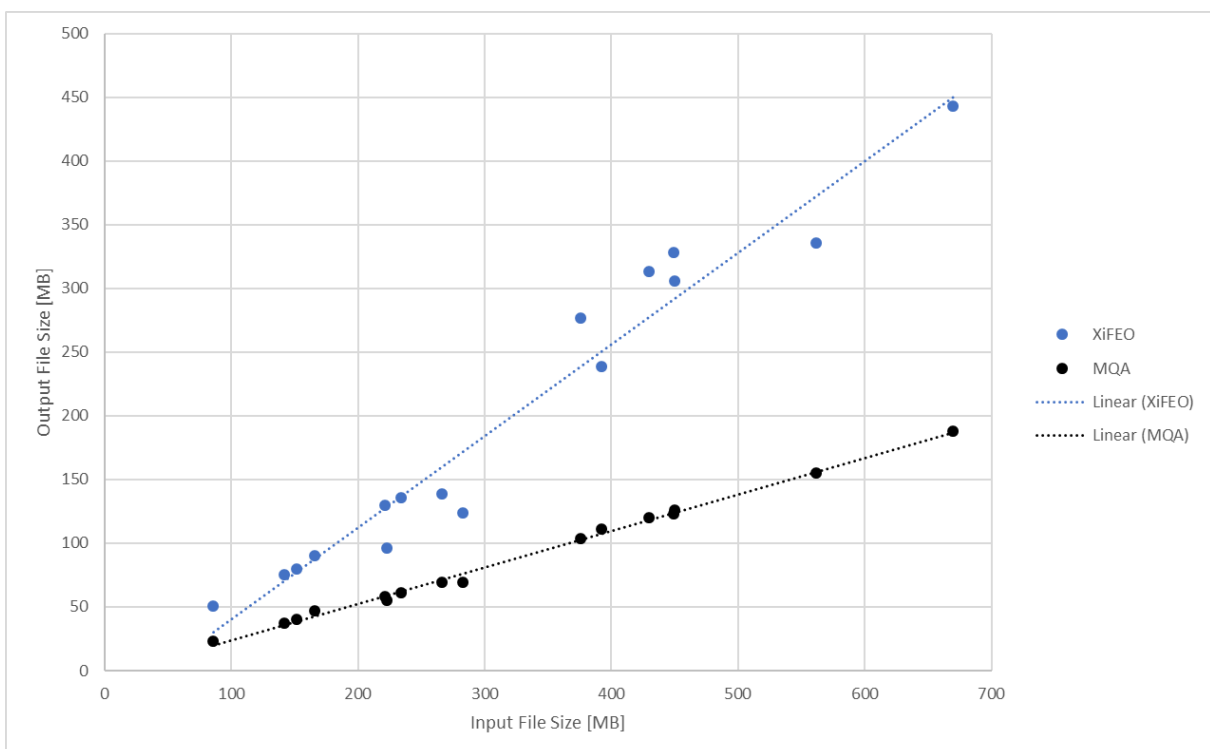


Figure 6: Comparing the compression rate between MQA and XiFEO for 192kHz / 24Bit test files

MQA shows a stable compression rate for the 192kHz / 24Bit input files of around **3.7**. That is an expected value because the output file is of the format 48kHz / 24bit which is 25% of the original data rate.

XiFEO exhibits a variable compression rate depending on the number of bits that are allowed to be truncated. The selected high resolution records are of quite high quality, therefore the average compression rate for the three test albums is at around **1.7**.

In a manual mode, it is possible to set the number of bits to be truncated to increase the compression rate significantly, trading real bit-depth like MQA does.

For more common 96kHz / 24Bit high resolution audio files that additionally do not exhibit extreme fidelity, the difference between the MQA and XiFEO compression rates is not that significant anymore. Nevertheless, it needs to be noted that XiFEO's compression rate depends on the input material and therefore is not static.

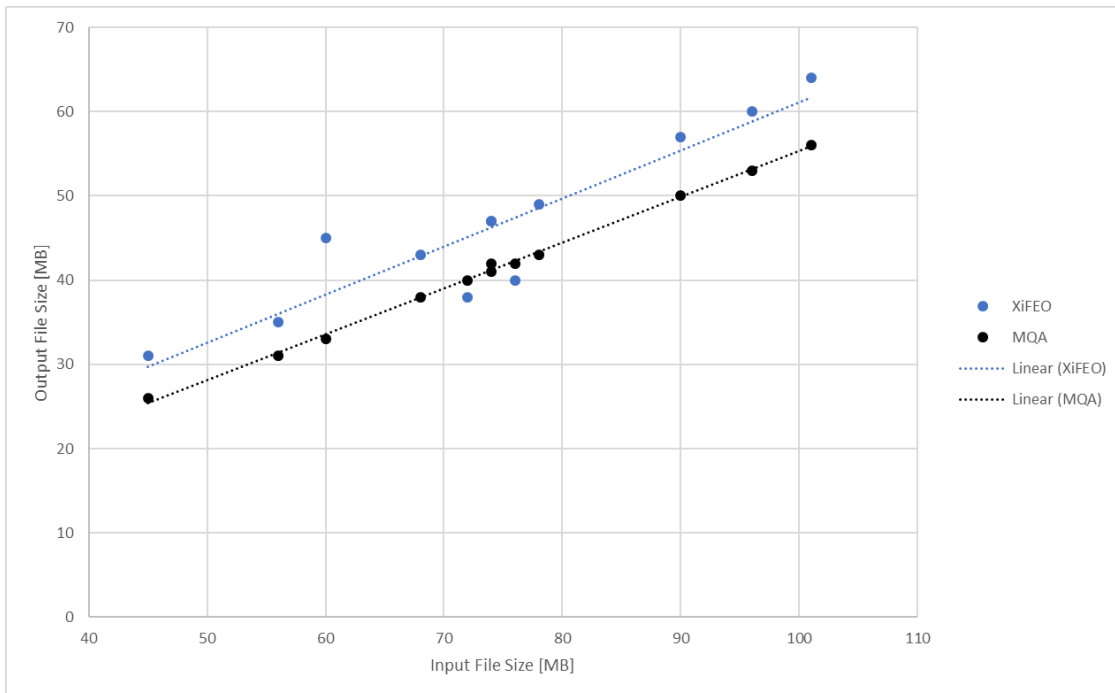


Figure 7: Comparing the compression rate between MQA and XiFEO for 96kHz / 24Bit input audio files of medium quality. XiFEO shows for two files in this test series a better compression.

XiFEO Analyzation & Transcoding:

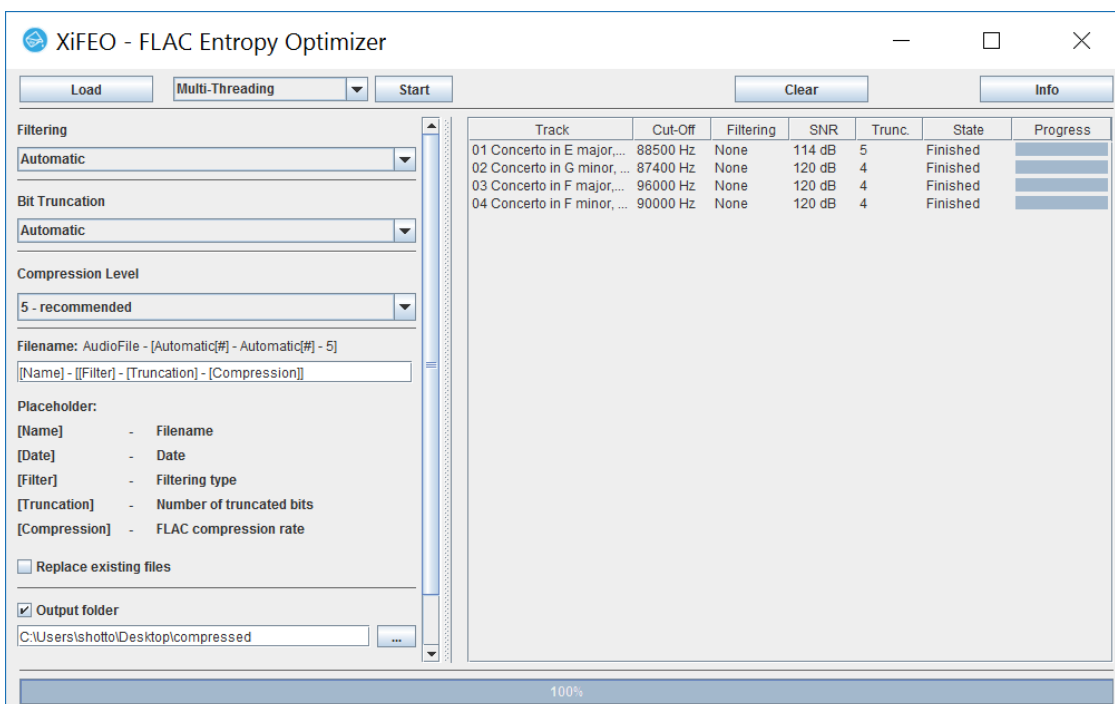


Figure 8: The analyzing output displays the identified upper bandwidth for the music's spectral components as well as the number of bits submerged in noise. The transcoding process takes that information to act accordingly.

Conclusion of the 2nd Hypothesis:

By reducing the entropy of the high resolution audio files, before they get compressed by any standard FLAC encoder, their file size can be reduced by a good amount.

The XiFEO compression rates for 192kHz / 24Bit high resolution audio are admittedly not as high as the results of the MQA audio compression scheme, because only bits that represent noise are truncated. This implies that the compression rate varies for different audio files.

The compression results for the more common 96kHz / 24Bit high resolution audio files are even more promising.

The most important advantages of XiFEO:

- 1.) The sample rate of the original native high resolution audio file is not changed to make sure that the temporal resolution is preserved.

- 2.) There is no need for a special proprietary decoder because any FLAC capable player has access to the full bandwidth of the high resolution audio file.

The transcoding (compression) process takes FLAC, ALAC, WAV and AIFF audio files as input and generates entropy optimized FLAC files.

5 Technical Details

5.1 Signal Analysis

The mathematical models behind the signal analysis enable us to look deeper into the inner workings of linear systems. In the context of audio systems, it is interesting to analyze how they handle the input signal and whether the output signal is a precise representation of the recording.

There are a lot of mathematical techniques available to represent the same signal in different spaces (Time Domain, Fourier Analysis, Laplace- and Hilbert-Transformation, etc.). What we need to know is that all those mathematical models have been invented/discovered to make it easier to do calculations. They are all absolute equivalent.

For example, within the time domain a systems behavior is represented by its impulse response (IR). That is the system's answer to a Dirac impulse. If we want to know what happens with an input signal within a system described by its IR we have to convolve the input signal with the IR of the system. This is quite a calculation intensive process. Therefore, we could transform the input signal and the systems IR into the frequency domain (Fourier Transformation) (**pls. see chapter 5.2**) to just do a complex multiplication. We finally transfer the output back into the time domain to get the result. That sounds tedious, but in fact that is a much faster process than just doing a long convolution within the time domain.

So, let's conclude that there were a couple of smart mathematicians in the 18th & 19th century laying the foundation for our today's digital signal processing algorithms.

5.2 Time domain and Frequency Domain are absolutely Equivalent!

Both domains are interchangeable and if a signal is converted from one domain into the other and back again it remains completely unaltered.

There are often discussions that the frequency domain cannot represent time domain phenomena like ringing. This is simply not true! If a signal is low pass filtered, that means the frequency domain shows nearly zero magnitudes for higher frequencies then it is clear that the impulse response of the system exhibits ringing, because we cut off terms from the Fourier-Series resulting in ringing.

Please refer to the Fourier Analysis (https://en.wikipedia.org/wiki/Fourier_analysis) to learn more about the process to transfer a signal from the time into the frequency domain.

If we take an audio signal as input for the Fourier-Transformation we get complex spectral components, which can be represented by its magnitude and phase.

The spectral magnitudes and phase describe a signal and system entirely!

What does time domain accuracy mean if we know that it all comes down to spectral magnitude and phase information?

As long as we are able to convey the full complex spectrum then there is no harm done to the time domain audio signal!

5.3 Bandwidth Limiting Causes Ringing and Signal Blurring

It is important to understand that a high audio bandwidth is equal to a high temporal resolution!

The temporal resolution is impacted as soon as the bandwidth is reduced by low-pass filtering!

Unfortunately, misleading marketing slides (**pls. see [8] – Fig. 3**) comparing impulse responses of linear phase filters and short minimum phase filters suggesting that the filter impulse response itself is a representation of the ringing in the time domain and therefore always an indicator for time domain blurring/smearing.

That is not the full truth, because we will show that as long as a filter works outside of the audio spectrum it does not do any harm to the time domain resolution!

This is effectively one of the biggest advantages of native high resolution audio (192kHz / 24Bit) to allow for anti-aliasing and interpolation filtering far beyond the highest spectral components of audio signals.

As a pre-condition for the further technical analysis we have to look at the MQA supported statement that the spectral components of music are mostly limited to a bandwidth of around 48kHz. Checking numerous high resolution audio recordings, we can confirm that above that frequency the spectral energy is in fact extremely low and that beyond 75kHz no music signal can be expected.

So, why are marketing departments using a mathematical signal called Dirac impulse, that has an infinite bandwidth and does not exist in reality, to show the ringing effect of long linear phase filters instead of applying real world transients that really occur within a 48kHz bandwidth limited system?

The following setup creates a real-world transient covering the frequency range of 0Hz - 48kHz with a high temporal resolution (maximum at around 30 μ s and a length of approximately 60 μ s).

Such transient that would be in approximation the shortest real world signal possible within a bandwidth limited audio channel, with an 1/frequency spectral magnitude envelope, is filtered by an extreme long brick-wall low-pass filter:

Sample Rate:	192kHz
Cut-Off Frequency:	48kHz
Filter-Length:	1349 Taps = approximately 7ms
Filter Type:	Linear Phase FIR

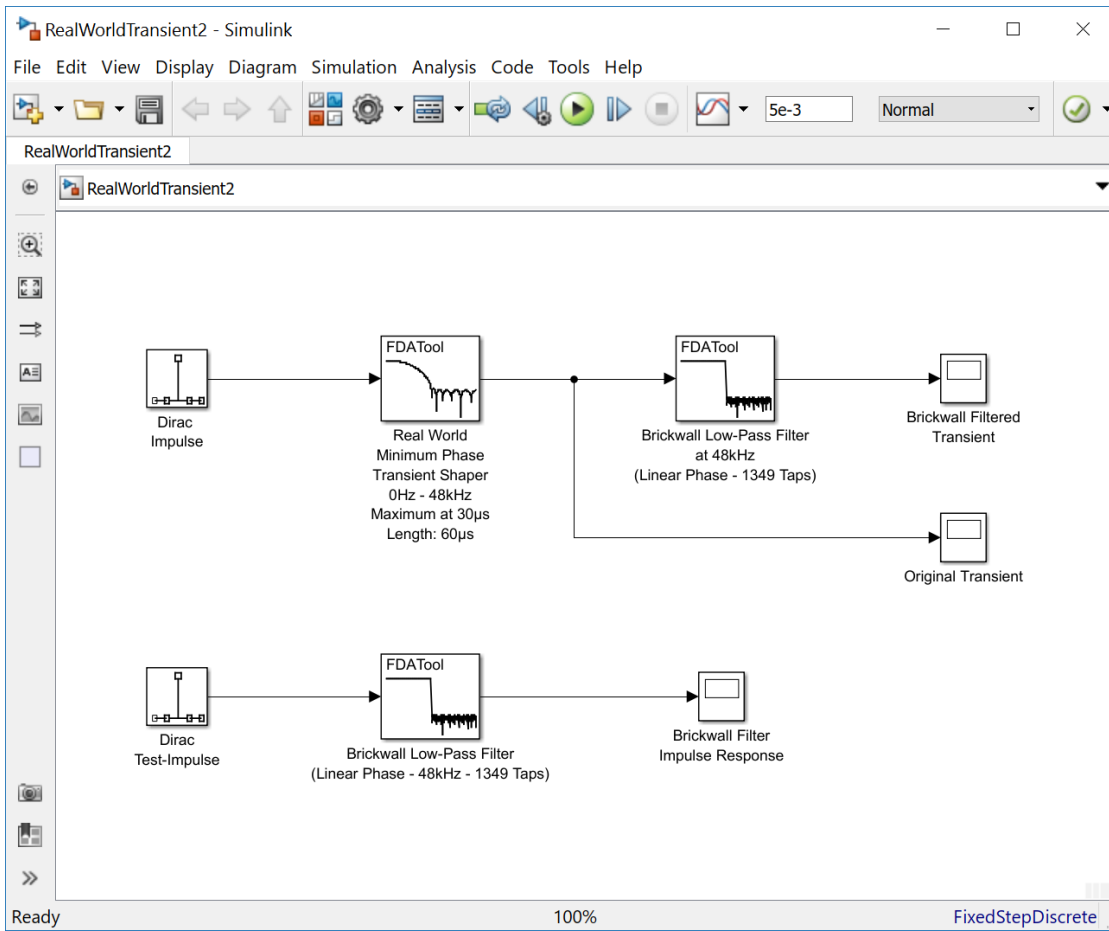


Figure 9: Creating a real world 48kHz bandwidth limited transient and filtering it through an extremely long linear phase brick-wall filter.

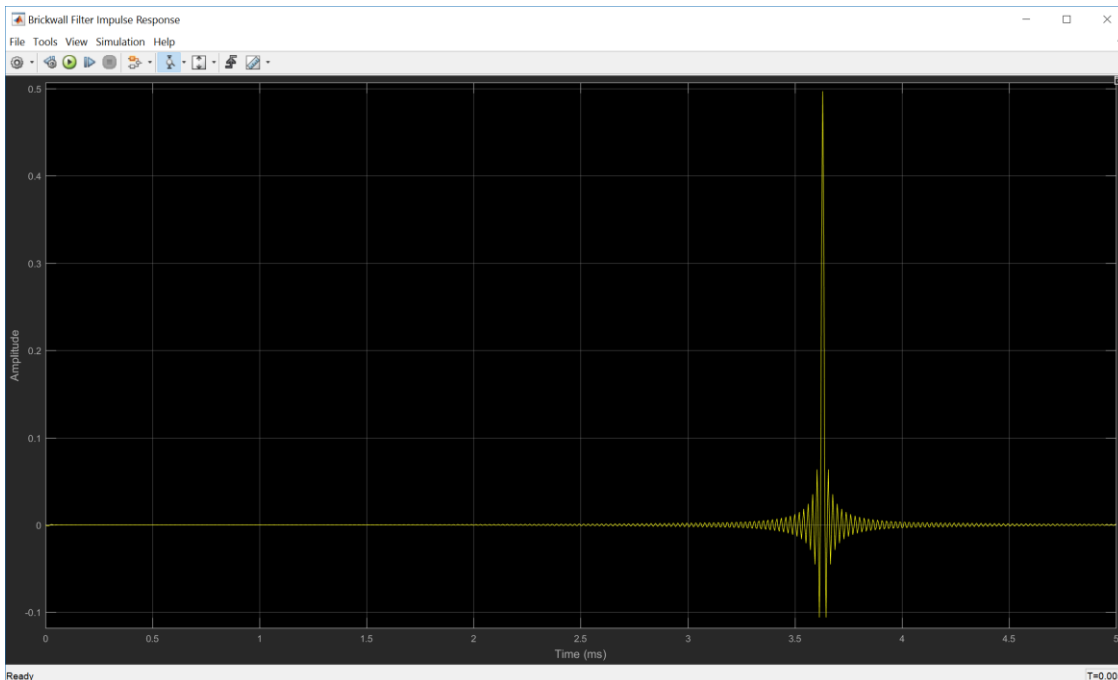


Figure 10: 1349 Taps (approx. 7ms) long linear phase low-pass filter impulse response.

The following transient is fed into the above mentioned linear phase filter.

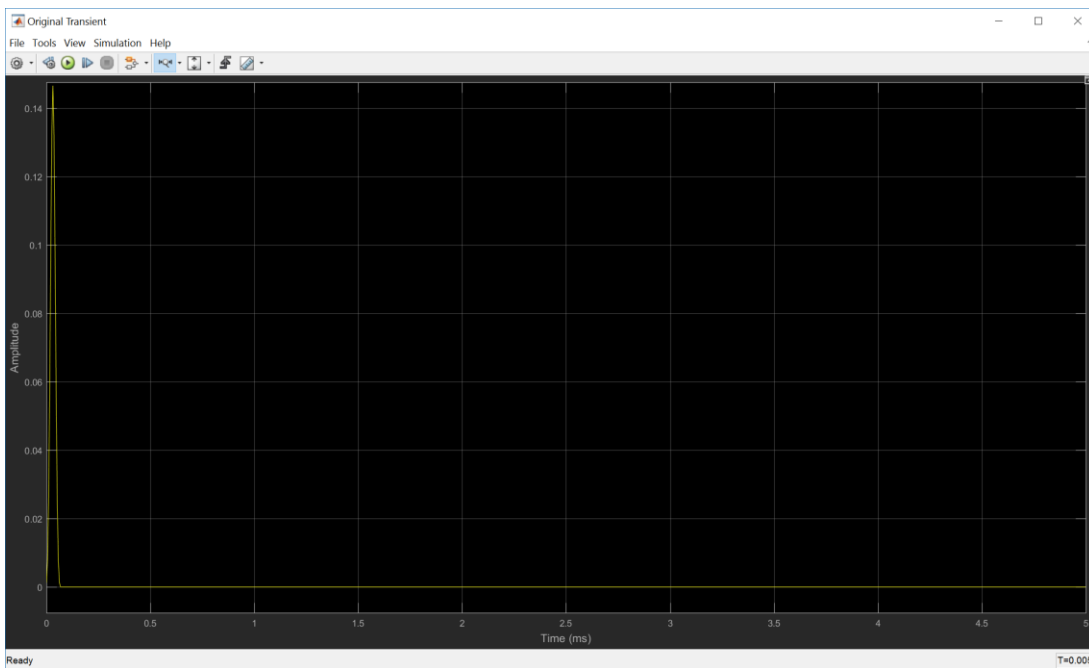


Figure 11: Real world 48kHz bandwidth limited transient.

The following measurement shows the output signal of the low-pass filter which is just delayed by half of the filter length.

There is no ringing or any kind of distortion at all!

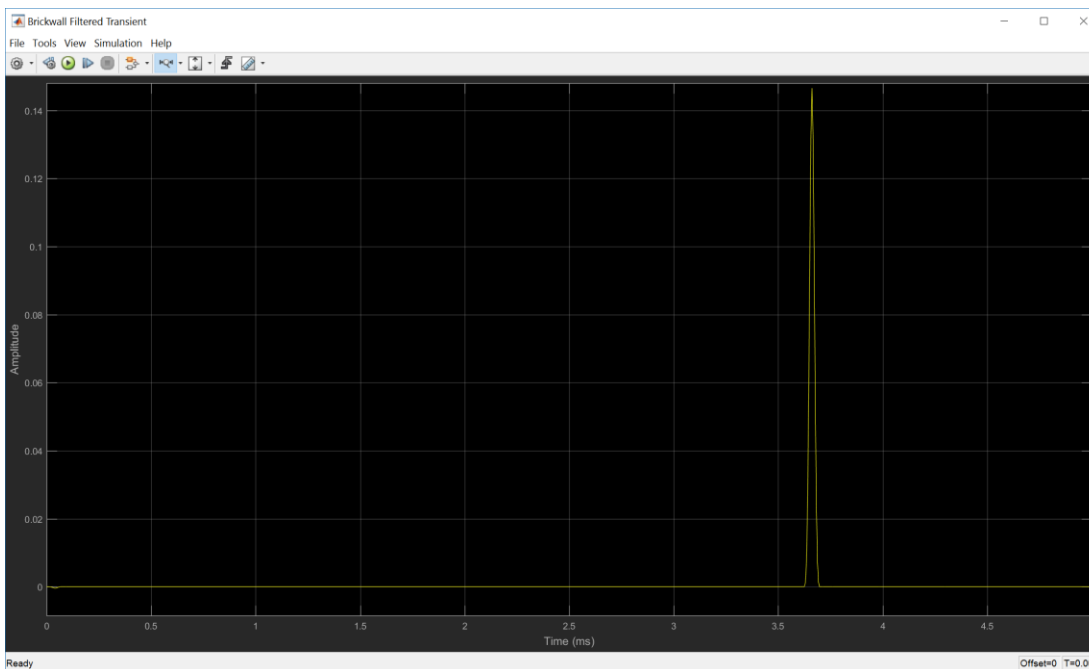


Figure 12: Filtered transient – No distortions!

Why does the signal not show any ringing?

Just because the filter operates outside of the spectrum occupied by the transient (music)!

It is a significant advantage of native high resolution audio to allow filtering out of the audio signal bandwidth!

For MQA this is really an issue because the overall MQA channel transmission (AD-to DA-Conversion) as described by the MQA paper [8] shows a drop in spectral magnitudes of already around 4dB at 40kHz.

To simulate a sub-sampling from 96kHz to 48kHz, leaving the domain of high resolution audio, we use a similar brick-wall filter as above at 24kHz. The input signal is still the real-world 48kHz bandlimited transient.

The following measurement presents the transient after its filtering with a long linear phase filter at 24kHz in the middle of the critical audio spectrum.

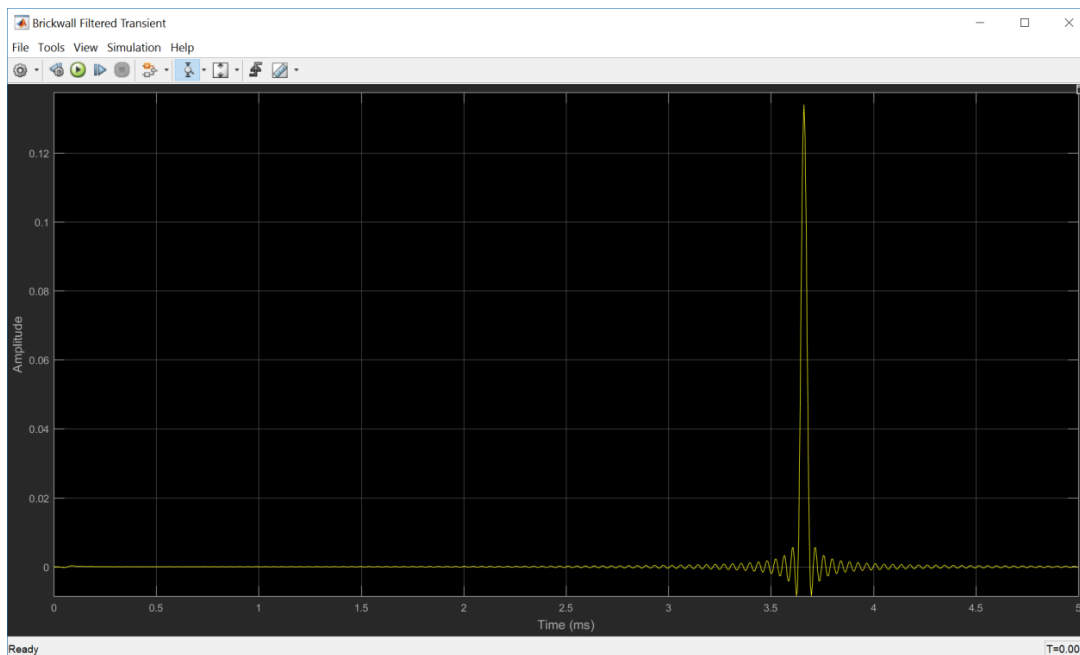


Figure 13: Filtered transient of a 48kHz bandwidth limited signal, filtered at 24kHz. Strong Pre- and Post-Ringing is visible.

As discussed above, if the filtering happens within the audio spectrum, the signal gets distorted, just because we removed frequency components that are important to describe the signal!

Let's keep in mind that we have removed 24kHz of the 48kHz transient which shortened the Fourier series (https://en.wikipedia.org/wiki/Fourier_series) a good amount causing ringing in the time domain.

It is important to know that the ringing energy is proportional to the energy of the spectral components we removed during low-pass filtering.

If we only need to filter very few spectral components with extreme low energy, the ringing would be so small that it does not have such a strong effect.

Figure 14 presents the effect of a minimum phase low-pass filter at 24kHz that avoids the unnatural pre-ringing, using the same real-world transient as input.

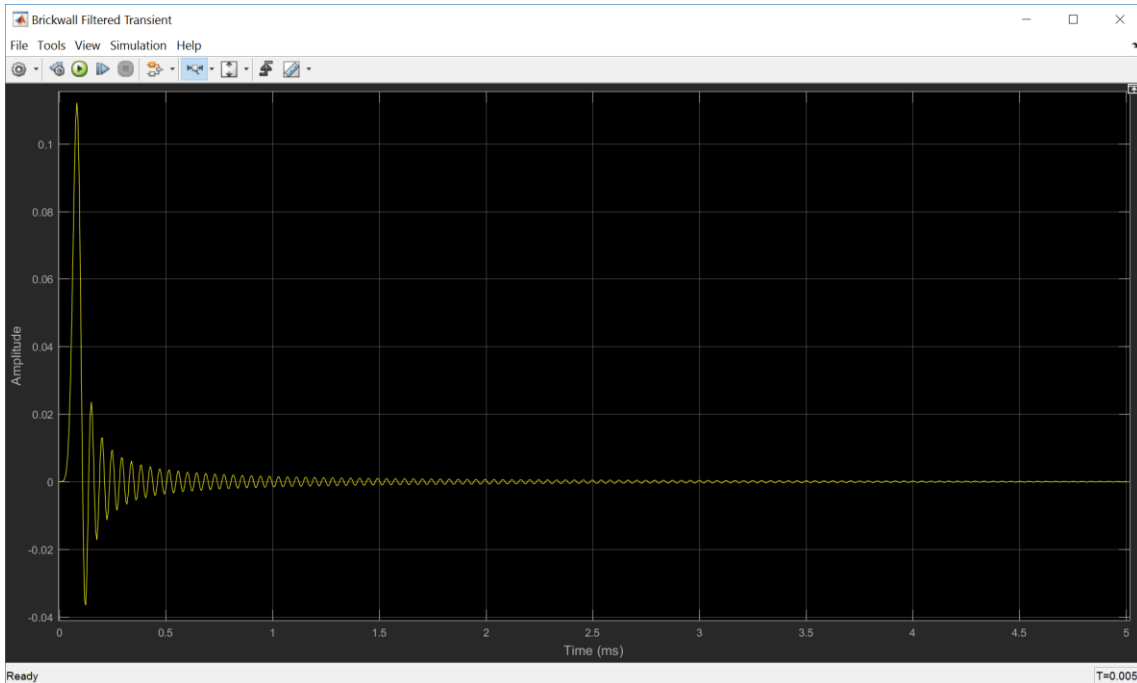


Figure 14: 24kHz filtered transient of a 48kHz bandwidth – Strong post-ringing caused by the minimum-phase filter.

The energy of the pre- and post-ringing of a linear phase filter is concentrated in the post-ringing of a minimum phase filter. Furthermore, a minimum phase filter distorts the phase of the audio signal.

Remember, a signal is described by its spectral magnitudes and phase. As soon as we change one of those components the time domain resolution is impacted!

There are several ways to change the impact of the filter type on the audio signal, but all of them are going to modify it, either by introducing aliasing or by having an early frequency droop losing even more high frequency information and therefore time domain resolution (**pls. see chapter 5.4**).

We want to emphasize that native high resolution audio of a 192kHz sample rate does not exhibit any of the discussed issues, because all filters operate outside of the audio spectrum!

5.4 Apodization

In optics apodization is used to reduce the effect of diffraction, whereas the term “apodizing” during filter calculation means that a special windowing function (https://en.wikipedia.org/wiki/Window_function) is applied to reduce ringing.

Optical apodization is in fact a very good example of avoiding ringing in the spatial domain (<https://en.wikipedia.org/wiki/Apodization>).

The simple trick is to band limit the spatial frequency of the lens to reduce the impact of the aperture that is an optical spatial brick wall filter.

Furthermore, because the CMOS-Sensor samples the projected picture with a dedicated spatial frequency, a special optical low-pass filter is used to avoid moiré patterns.

**There is no free lunch in signal processing.
To reduce the diffraction ringing we have to sacrifice resolution!
The same is true if we apply apodization to audio files!**

What does that mean for audio signals?

Applying a non-linear phase low-pass filter with a flat slope (minimum ringing / no pre-ringing) reduces pre-ringing effects introduced at earlier or later filter stages in the audio chain as long as their cut-off frequencies are above the cut-off frequency of the apodizing filter.

BUT, we thereby blur the audio signal in the time domain by altering its phase and spectral magnitudes!

What are the disadvantages?

As mentioned, the apodizing filter has a very shallow frequency response which makes it unusable for sample rates below 96kHz. Such kind of filters are usually not applicable for 48kHz sample rates because they would affect the magnitude of the critical audio spectral components (**pls. see chapter 5.3**).

As we understand the patent applications, MQA uses this kind of filters, but trades early frequency droop for aliasing (**pls. see chapter 5.5**).

Let's demonstrate the effect of apodization to reduce the ringing effect of brick-wall filters that have been already applied within the critical audio spectrum.

We just take the signal from Fig. 15 and apply an apodization filter with a shallow slope but no aliasing because, as we will see later, that is an irreversible damaging process.

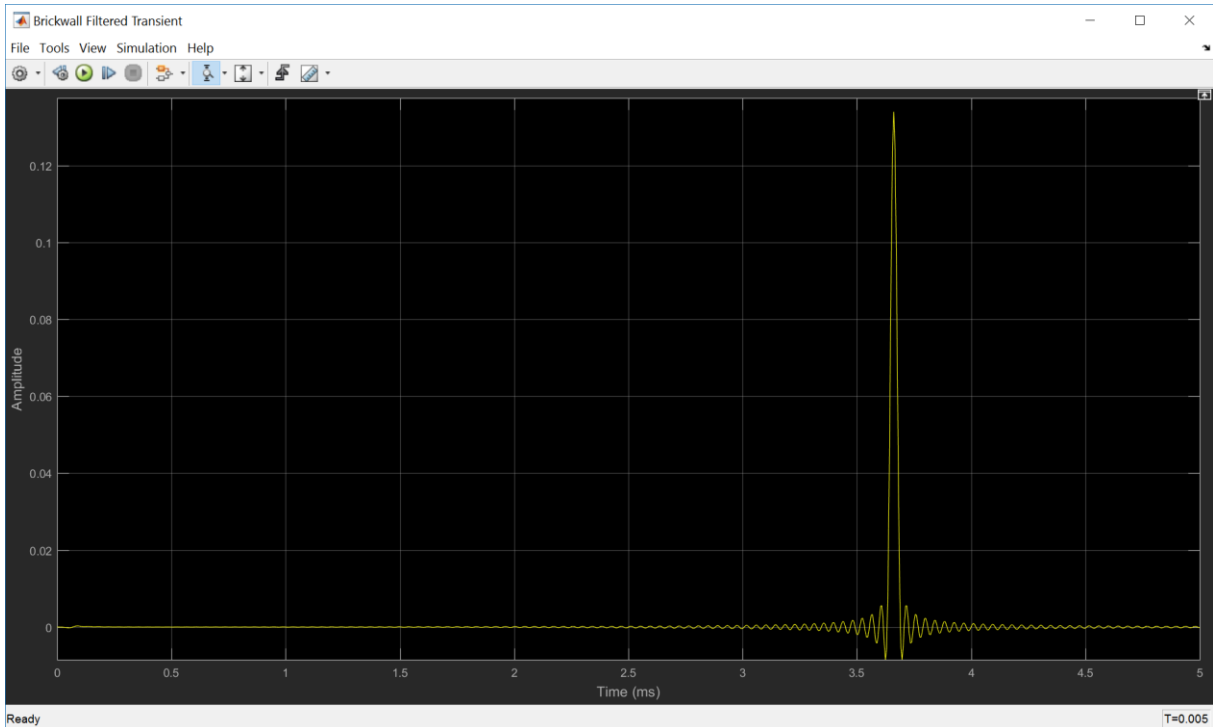


Figure 15: 48kHz Real World Transient plagued with pre- and post-Ringing introduced during low-pass filtering at 24kHz.

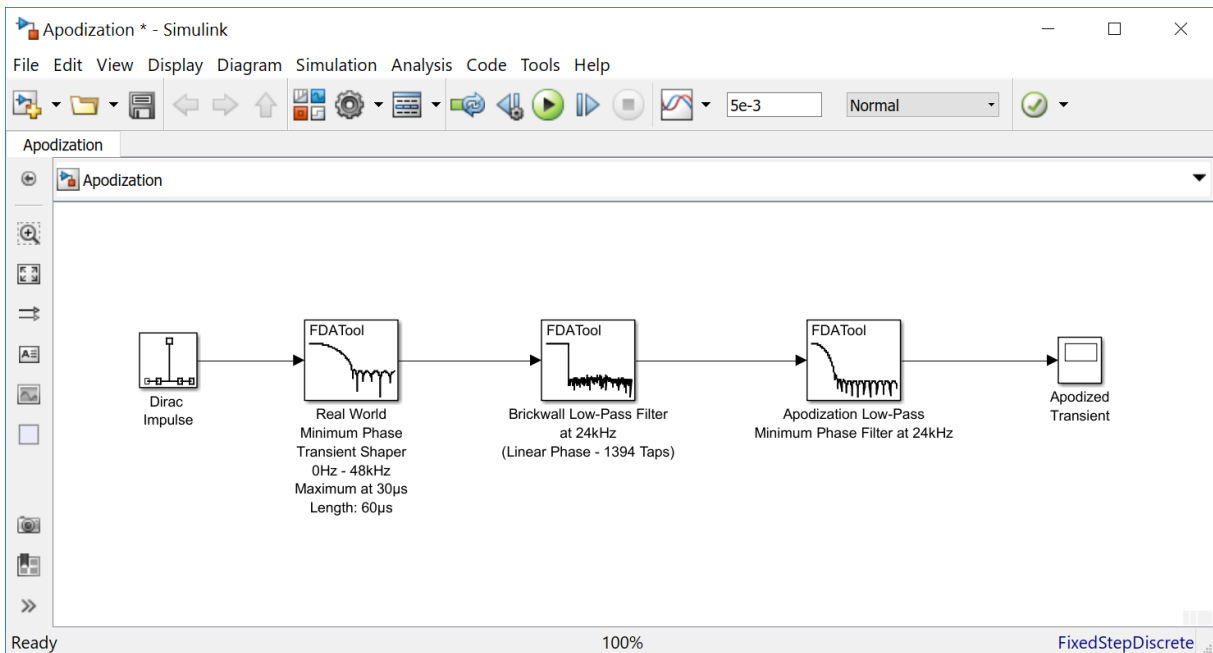


Figure 16: Apodization of the brick-wall filtered real world transient.

After apodization the signal looks like the following measurement:

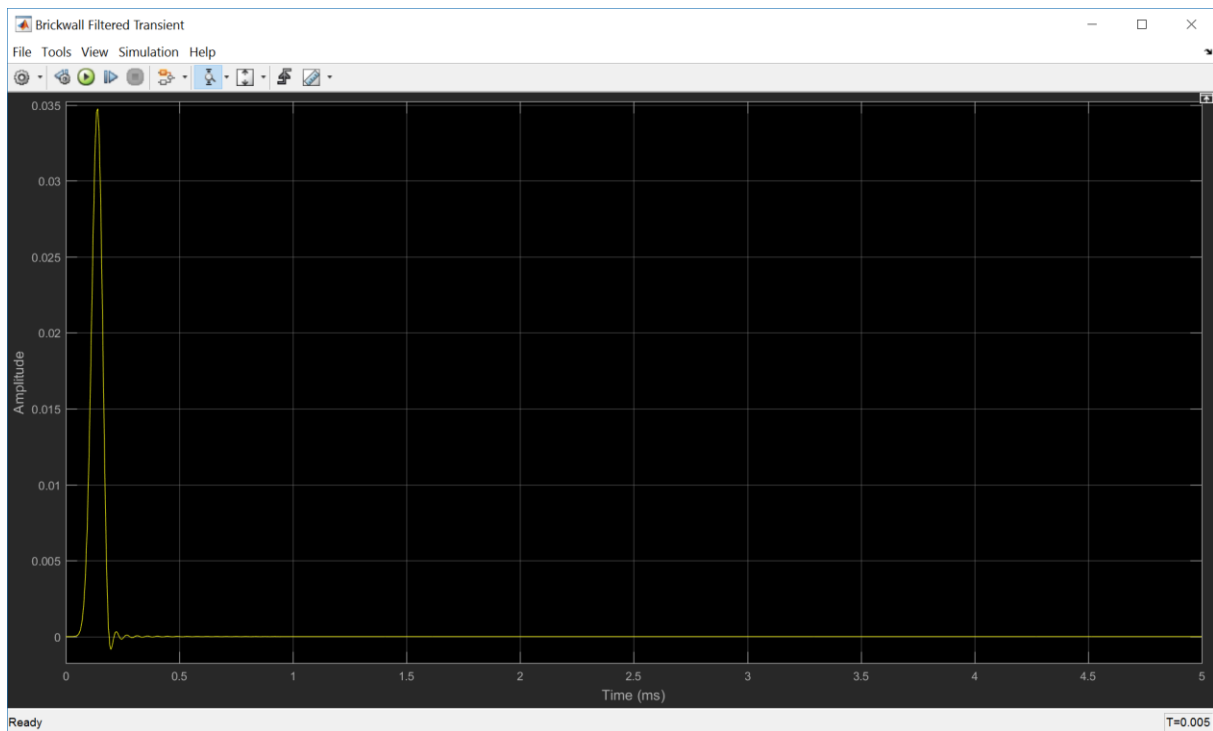


Figure 17: After apodization the pre-ringing is gone and there isn't much post-ringing. BUT, as we can see at the broader transient, we reduced the time domain resolution (sharpness) to get rid of ringing.

**The math behind digital signal processing can be really mean!
If we get rid of ringing then we lose time domain resolution, causing blurring/smearing of the transient. Additionally, we distort the signal phase due to the non-linear-phase filter used!**

As a general rule we can say that a shallower filter slope results in a shorter filter impulse response but wider filter transition which needs space in the frequency domain. Simply said, the filter starts early to roll-off and needs its time to reach the maximum attenuation.

5.5 Aliasing

To reduce the early frequency droop of the MQA apodized filter, that exhibits a shallow slope (**pls. see chapter 5.4**) it is necessary to compromise the attenuation beyond the Nyquist frequency ($1/2 \times$ sample rate).

The Patent Application WO2015/189533 A1 explains that the proposed anti-aliasing filters, which are apodized filters (**pls. see chapter 5.4**), allow a dedicated amount of aliasing.

Quotes from the Patent Application WO2015/189533 A1:

“There is no established criterion for how much aliased components should be reduced relative to original components, but a criterion may be derived based on balancing phase

distortion in the audio band against total noise. We assume that the total response should be minimum-phase in order to avoid pre-responses.”

“Preferably, the downsampler comprises a decimation filter specified at the first sample rate, wherein the alias rejection of the decimation filter is at least 32dB at frequencies that would alias to the range 0-7 kHz on decimation.

If MQA should have implemented its filters as described in their patent applications - and why should they not have done that -, then they obviously made a couple of assumptions how critical aliasing is.

Furthermore, MQA uses minimum-phase or more general non-linear-phase apodized filters to avoid pre-ringing with the adverse effect of early frequency response droop, phase distortion and aliasing.

Let’s have a look at aliasing:

A simple 48kHz sampling of a signal that creates a sinus with an increasing frequency (Chirp) up to 48kHz creates the strongest aliasing we get because if we apply the Nyquist-Shannon-Sampling Theorem (**pls. see chapter 5.6**) we would need to sample with at least 96kHz.

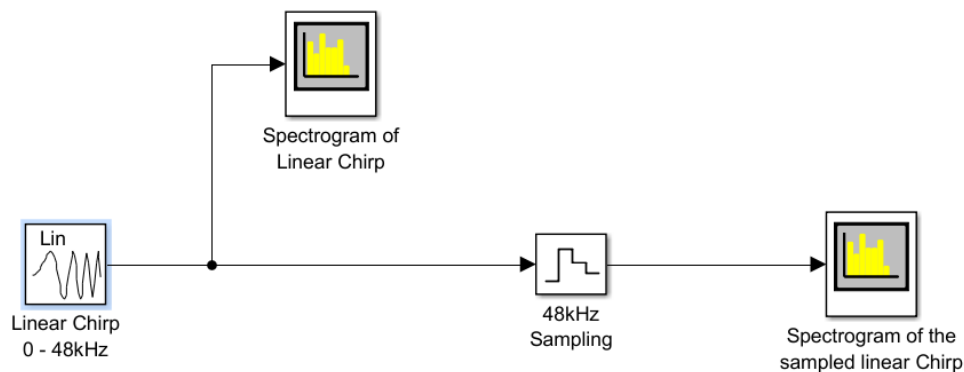


Figure 18: Linear Chirp 0Hz - 48kHz sampled with a sample rate of 48kHz

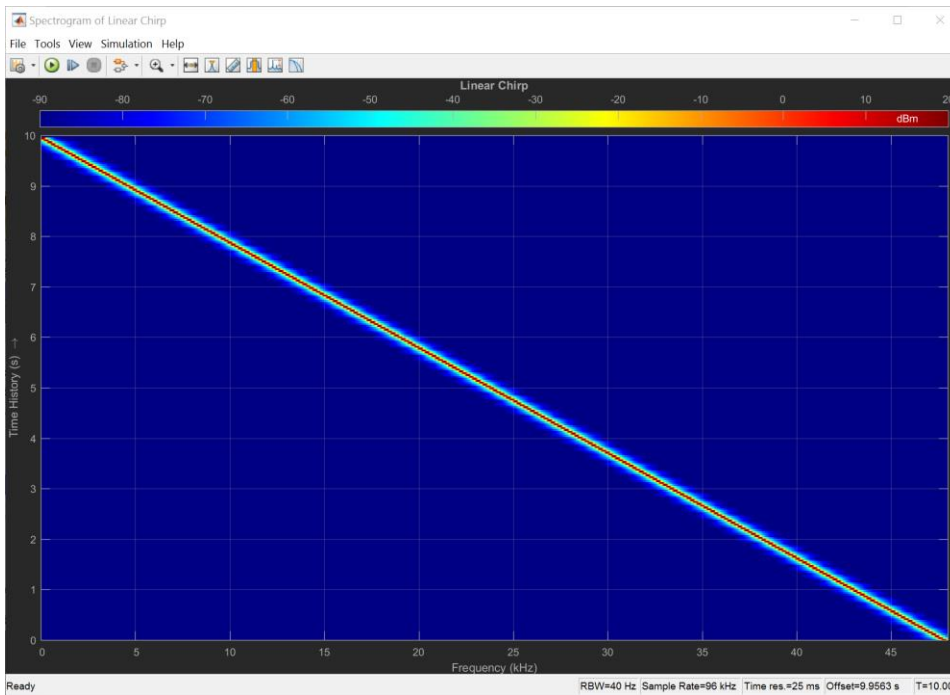


Figure 19: Spectrogram of the 10s linear chirp from 0Hz - 48kHz

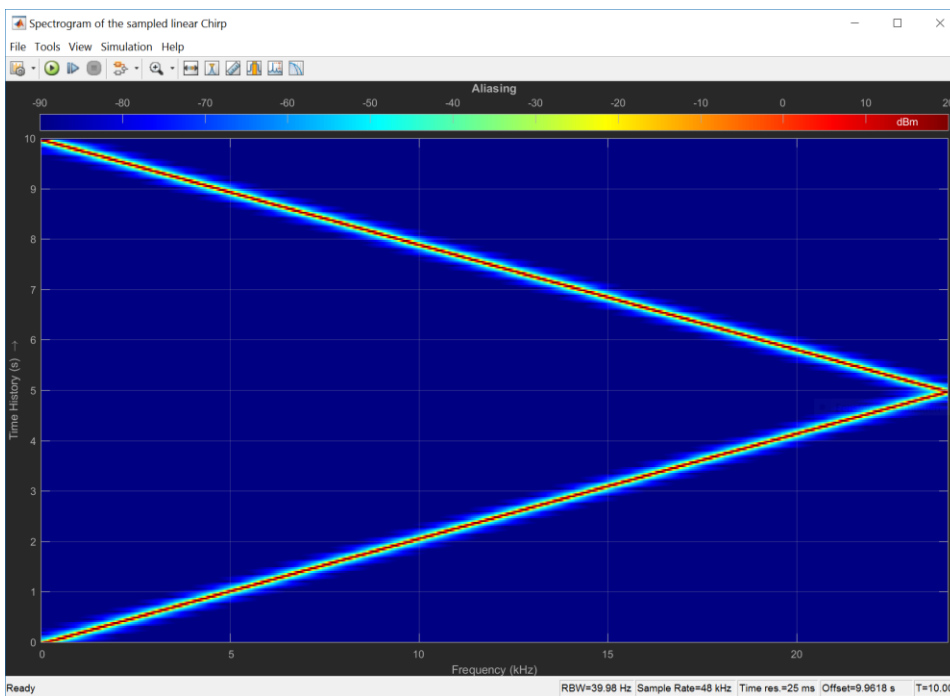


Figure 20: Strong aliasing because of under-sampling

Fig. 19 represents the spectrogram of the 0Hz - 48kHz chirp that takes 10s for the whole frequency range.

The strong aliasing can be seen in Fig. 20 where all frequencies beyond $\frac{1}{2}$ x sample rate are folded back into the spectral range of 0Hz - 24kHz. This means that a frequency component at 47kHz appears at 1kHz.

This is really a very critical issue because our auditory system is very sensitive in the range from 0Hz - 7kHz as confirmed by MQA.

This is even a serious issue if we take real world music signals into consideration where the spectral magnitudes decrease in amplitude proportional to $1/\text{frequency}$.

It is an important learning that aliasing is a non-reversible process if it is introduced during sampling (AD-Conversion) or interpolation (DA-Conversion)!

Now, let's apply a linear-phase brick-wall anti-aliasing filter that suppresses all frequencies beyond 24kHz to avoid any aliasing.

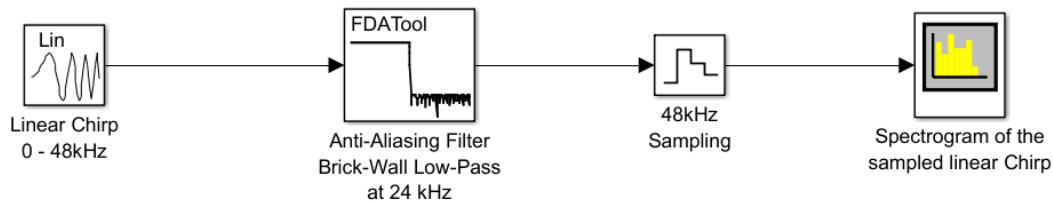


Figure 21: Linear-Phase Brick-Wall Anti-Aliasing Filter

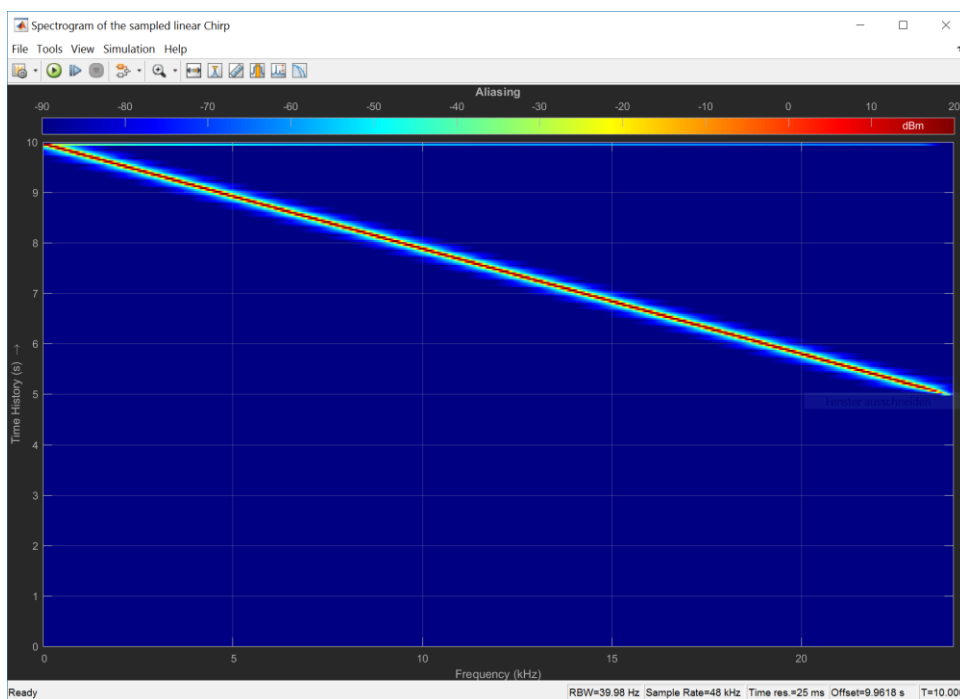


Figure 22: All Aliasing-Components are suppressed

Applying a brick-wall anti-aliasing filter suppresses all aliasing components.

As we know from **chapter 5.3** operating the strongest brick-wall-filter outside of the audio spectrum does not do any harm at all.

That is a strong reason to go for native high resolution audio of 96kHz or even 192kHz to make sure that the anti-aliasing and interpolation filters work at 48kHz or 96kHz, outside of the critical audio spectrum!

Now let's have a look at an apodized anti-aliasing filter that compromises aliasing for early frequency droop.

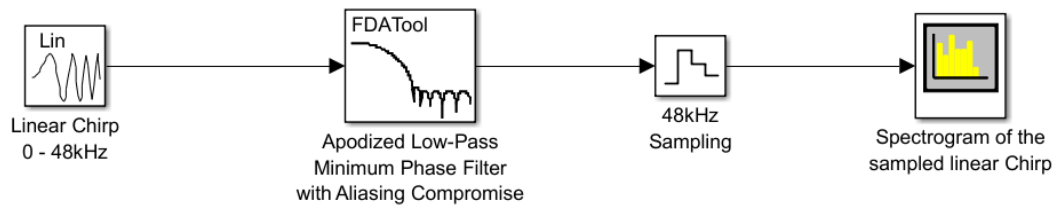


Figure 23: Minimum-Phase Apodized Anti-Aliasing Filter

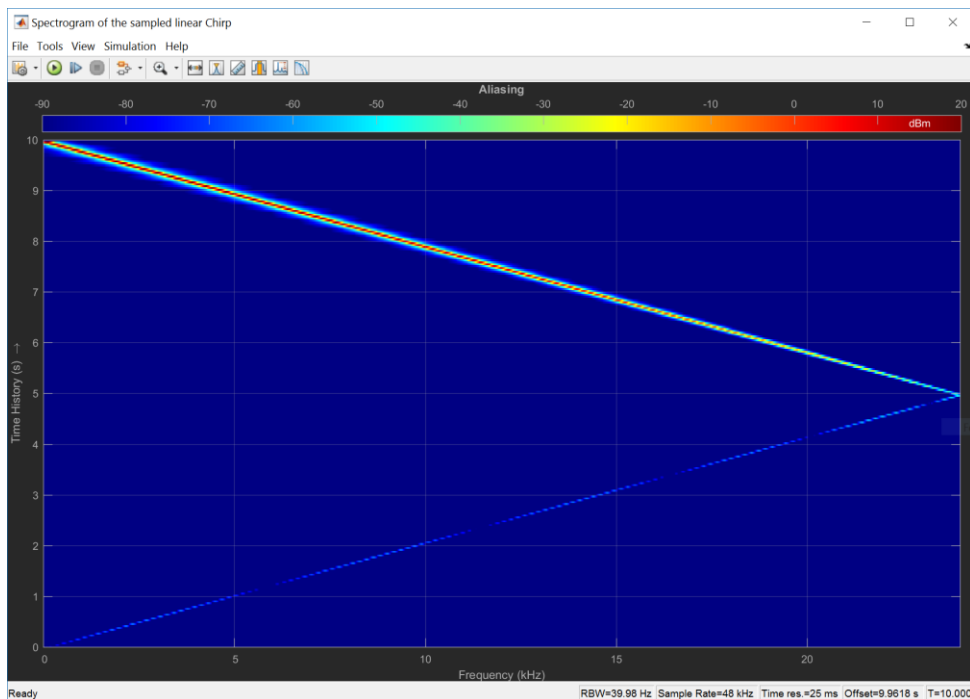


Figure 24: Minimum-Phase Apodized Anti-Aliasing Filter

As we can see easily the filter does not suppress the aliasing components sufficiently.

It is debatable to which degree such amount of aliasing is really impacting the perceived audio quality, but we would always go for the best aliasing suppression which is only achievable by staying within the sample rates of native high resolution audio.

5.6 Nyquist-Shannon Sampling Theorem and Low Pass Filtering

Within the analog audio world there isn't the problem to limit a signal in frequency before it gets further processed. Well, that's probably only partly true but let's pretend that this is a fact because a tape recorder inherently limits the frequency range simply by the inner workings of magnetic tape and the recording head.

In the digital world we have to apply the Nyquist-Shannon Sampling Theorem (https://en.wikipedia.org/wiki/Nyquist%E2%80%93Shannon_sampling_theorem) that dictates that the sample rate must be at least two times higher than the highest frequency component contained in the input signal. That implies that we have to do low pass filtering of the input signal before we transfer the analog signal into the digital domain.

Actually, the signal only needs to be bandlimited to be successfully sampled, but that is more of interest for high frequency analog to digital conversion where that technique is used.

It is known since long that dedicated signals can be sparsely sampled if they adhere to the hard pre-condition of being of limited rate of innovation (pls. see research by Pier Luigi Dragotti). That doesn't mean that the Nyquist-Shannon-Sampling Theorem isn't valid anymore, it just states that for a special kind of signals we could apply an alternative sampling technique using different sampling kernels (**pls. see chapter 5.7**).

To adhere to the Nyquist-Shannon-Sampling-Theorem we have two choices to implement the necessary anti-aliasing filter:

1.) Within Digital Domain:

If we want to implement such filter in the digital domain, the real world input signal needs to be band limited by itself and we have to choose a sample rate that is at least two times higher than the highest relevant frequency component within the input signal. This means that our AD-Converter must be able to apply high enough sample rates.

If we talk about real-world undistorted audio signals, we get away with a sample rate of 192kHz / 24 Bit, not applying any anti-aliasing filter at all!

2.) Within Analog Domain:

That isn't a desirable approach because analog filters are expensive minimum phase implementations and because of tolerances in their filtering parameters are not favorable for audio applications.

5.7 Sparse Sampling of Signals with Finite Rate of Innovation

We don't need to use standard sampling kernels (Sinc-Functions) to sample and reconstruct a signal.

Bandwidth limited signals, which are a pre-condition for correct sampling (**pls. see Chapter 5.6**) are of finite rate of innovation.

If a signal has a special property with an even lower degree of innovation then it would be possible to use sampling kernels of a special type like B-Splines (e.g. 0.5; 1; 0.5 for a linear interpolation) instead of E-Splines (exponential kernels).

For audio compression, we could simplify the case further because segments of music can be considered as piecewise super positioned sinusoidal signals plus noise.

In general, it would be possible to approximate these segments with polynomials or indirect as wavelets (<https://en.wikipedia.org/wiki/Wavelet>) which would allow us to just transmit the polynomial or the low-pass wavelet coefficients to implement a linear compression.

BUT, there is noise which is part of the music!

Even robust sparse sampling algorithms are only able to reduce the residual error in the presence of noise and therefore we lose valuable information to reconstruct the time domain signal.

It is highly likely that MQA uses those methods within the compression and decompression stages of the encoder and decoder.

In sum we have to emphasize that sparse sampling creates a residual error that needs to be encoded into the compressed audio file. The FLAC (Free Lossless Audio Codec) is doing exactly that quite efficiently.

If the noise (entropy) of the signal is high, the residual error increases and therefore the compressor needs more bits to represent the signal perfectly.

We don't know whether there are conditions where the MQA compressor gets overloaded by high frequency content, occupying numerous bits, and therefore has to fall back in a lossy mode.

5.8 Analog Signal Reconstruction – Interpolation Filter

The conversion from the digital into the analog domain is actually not that difficult.

We need again a low pass filter, suppressing all frequency components above $\frac{1}{2}$ x sample rate, to reconstruct the original analog signal.

Such filter is an analog filter. To keep it as simple as possible modern digital to analog converters use the same trick as analog to digital converters which is oversampling the output signal to use simple analog low pass filters.

That oversampling process is indeed an issue because it involves again digital low pass filters build into the DA converters, but if those filters work at 96kHz for a 192kHz sampling rate then they don't do any harm to the audio signal (**pls. see chapter 5.3**).

There are discussions underway that an analog signal between samples cannot be reconstructed. That is absolutely not true!

If we take the Nyquist-Shannon-Sampling Theorem then it would be only necessary to sample a sinus signal two times within one full cycle and we're still able to reconstruct the complete signal.

Furthermore, what we discussed about sparse sampling (**pls. see chapter 5.7**) is also applicable for the reconstruction of the digital signal to transfer it back into the analog domain.

5.9 Signal to Noise Ratio (SNR), Quantization Noise and Dithering

Signal to Noise Ratio:

A huge advantage of native high resolution audio recordings is their high Signal to Noise Ratio (SNR).

It is important to know that this SNR describes the relation between the signal and the quantization noise floor. Each bit we add doubles the SNR and reduces the quantization error and therefore the noise floor by approximately 6dB.

A standard CD recording of 44.1kHz / 16Bit reaches an SNR beyond 96dB by applying techniques like dithering. There are even recordings that claim to reach 120dB SNR, but that is only true for noise shaped recordings. Those show an increased noise floor for the upper frequencies where allegedly our auditory system is not that sensitive for noise. As usual, that is only true until new research proves this otherwise.

Native high resolution audio provides a bit-depth of 24Bit reaching a theoretical SNR of 144dB. As mentioned above, by applying the neat mathematical trick of dithering it would even go beyond this.

MQA engineers explain us that quantization noise below -120dBFS isn't an issue anymore, because the thermal noise of the equipment, we're using during recording and playback, limits the SNR effectively to 120dB.

Now we are discussing the relation between the signal and the real noise floor of physical systems.

Nevertheless, during our statistical analysis we have learned that sometimes well dithered native high resolution audio files still have audio information within the 2nd LSB and it would be only viable to throw away one bit to avoid losing any critical bit-depth.

If we go along with the MQA approach there would be a headroom of 24dB (4Bits of Noise).

MQA needs to use the lower bits to place the information of the upper frequency bands beyond 24kHz. The patent applications describe numerous methods to use different numbers of bits, therefore we can only assume that most likely around 7Bits are used by the current MQA audio compression scheme implemented.

Let's assume that we lose 7Bits then the SNR decreases tremendously, going down to 102dB, not really much better than the SNR of a compact disk.

In the following you will find more information on dithering and what MQA does to increase the SNR after reducing the bit-depth.

Dithering to reduce Quantization Noise

Quantization noise is caused by the non-linear process of quantization and therefore appears as additional harmonics within the spectrum.

If we add noise before quantization, we destroy the correlation between the signal and the quantization itself, achieving a higher SNR than the number of available bits would suggest.

A good understandable dithering process is applied to black and white pictures (<https://en.wikipedia.org/wiki/Dither>).

Simply adding noise before quantization is the simplest form of dithering. Noise shaping is a more sophisticated algorithm to reduce the quantization noise in the lower frequencies but increasing it within the higher frequency range. The total amount of noise always stays the same!

MQA is aiming at the whole recording chain which would allow them to apply subtractive dithering which has some interesting advantages.

MQA needs to hide their high frequency bands in the LSB-Bits of the baseband by making them appear as random noise to avoid any audible artifacts. Just rendering those bits as simple noise would be a waste therefore they use them to apply plain nonadaptive noise-shaped dithered requantization to a constant bit depth (Patent Application: WO2013/186561; Paragraph 15), increasing the SNR to partly compensate the bits truncated.

We are talking here about a statistical process that does not convey the original signal but the difference is allegedly not perceivable.

From our point of view that is a lossy process!

Please refer to the paper of Stanley P. Lipshitz (Quantization and Dither: A Theoretical Survey) to get deeper insights in the different methods of dithering.

If we add the information, that very well made and dithered high resolution audio files contain real information even in the 2nd LSB then we have to conclude that the MQA approach throws viable information away.

Finally, we should keep in mind that there is just a limited Shannon information space available, which forces the MQA algorithm to make room for the probably more important high frequency sub bands.

Let's conclude that dithering is able to increase the SNR in relation to the quantization noise but it cannot fully make up for information loss during bit truncation!

5.10 De-Blurring – Inverse Filtering

There is some talk that MQA compensates for filtering issues introduced during the recording process (e.g. usage of brick-wall linear-phase filters).

We don't know to what degree such a technology, besides of the already mentioned apodization filters (**pls. see chapter 5.4**), is implemented into the MQA-recording chain, but we just like to highlight the limitations of de-blurring.

The Patent Application WO 2016087583 A1 (Non linear filter with group delay at pre-response frequency for high res audio) and the Patent Application EP3029674 A1 (Mastering improvements to audio signals) seem to provide more information about the approach.

But again, this is just another implementation of an apodization filter with the already known effect of limiting the bandwidth and introducing phase distortions, impacting the time domain resolution.

If MQA implements real inverse filtering, then we have to know about its limitations.

Inverse filtering to enhance the temporal resolution of audio signals needs to amplify or re-construct the higher frequencies, lost during the low-pass filtering process. As soon as those frequencies are buried in noise there is no hope to reconstruct the real thing.

The NASA did learn exactly this with their Hubble-Space-Telescope. They were able to recover sharpness (spatial resolution) to a degree by mathematical algorithms because they knew the **exact error** of the mirror. At the end of the day they had to invest billions to achieve the full capacity of Hubble by integrating lenses that corrected the error caused by the wrongly shaped mirror.

If the idea of compensation is limited to the apodization filters, as described in the above Patent Applications and in **chapter 5.4** then this is an innovation that has probably at least equal disadvantages as advantages (e.g. time domain resolution blurring/smearing).

6 Conclusion

6.1 Summary

The discussed technical arguments prove that the 1st and 2nd hypothesis are true!

- 1.) MQA is in fact “lossy” because it alters the bit-depth and frequency response (magnitude & phase) and therefore time domain appearance of the original high resolution audio file by applying none-linear-phase filters impacting the critical audio spectrum (e.g. 4dB attenuation at 40kHz).
- 2.) An alternative compression scheme that does not show those adverse effects has been described above and can be used for streaming applications or mobile high resolution audio players, with limited memory, to store at least 40% more audio files.

For MQA it is debatable whether the alterations introduced by their algorithms are audible and therefore they claim that the whole audio processing chain is “transparent”.

We learned from the past that any compression schemes that work with assumptions about our auditory system (e.g. MP3, AAC, etc.) have been proven wrong with new research at the horizon.

After learning all those technical details about MQA we would like to ask those responsible in the audio industry that this technology should only be used for streaming where it is of advantage to reduce the data rate. This will of course change as the ever-increasing bandwidth of the internet soon renders MQA obsolete, because then we will get the best native high resolution audio experience without altering its spectral components and available bit-depth.

To archive and record music in its best shape we would like to ask the recording industry to go for high quality 192kHz / 24Bit analog to digital converters where any anti-aliasing filters work outside the audio spectrum and therefore don't doing any harm.

Dear responsible operators of download platforms please provide us audiophiles with FLAC encoded native high resolution audio file downloads that are not altered in any way by applying technologies like MQA.

We would like to invite J. Robert Stuart (Meridian Audio Limited & MQA Limited) to hand us a software MQA encoder and decoder to do an in-depth analysis of the MQA audio chain.

As of our knowledge, up to now the whole encoding process is in the hands of MQA Limited, whereas the proprietary hardware decoders only output the already digital to analog converted signal.