

Entertainment Hi End

Musica Digitale e MQA

L'encoding MQA.

MQA si basa sulle più recenti ricerche in neuroscienza uditiva, sulla percezione del suono di alta qualità e sulle nuove tecnologie di codifica digital. Come detto nella newsletter precedente MQA è una filosofia molto più ampia che semplice e sola tecnologia di encoding e decoding per audio ad alta risoluzione. Affronteremo degli argomenti complessi che cercherò di semplificare al massimo, per evitare altre 15 newsletter di spiegazioni.

Le teorie che stanno alla base della codifica MQA: pillole di neuroscienza, elettronica ed acustica

Riassumerò in poche righe alcuni concetti fondamentali multidisciplinari, per capire le teorie che stanno alla base della codifica MQA.

- Quello ascoltiamo non è la forma d'onda acustica analogica o lo spettro che interpretiamo, ma gli impulsi elettrochimici di circa 30.000 neuroni di ognuna delle singole cellule cocleari, le quali sono connesse tra loro attraverso altri neuroni.

I segnali attraversano il tronco cerebrale, il mesencefalo e poi arrivano alla corteccia uditiva (in cui finalmente "udiamo"). In tutti questi passaggi, i neuroni che lavorano suddivisi per gamma di frequenze, inizialmente codificano i segnali per livello, spettro, modulazione, attacco e rilascio per poi passare successivamente quella che è l'informazione a combinazioni neuronali di strutture che scambiano, codificano ed estraggono una varietà di caratteristiche temporali, spettrali ed etologiche.

- Abbiamo già detto che il nostro sistema uditivo è molto più sensibile nel percepire i segnali nel dominio del tempo che in quello della frequenza. Sfruttando metodi di eccitazione neuronale, si è scoperto che la risoluzione temporale dell'apparato uditivo umano può avvicinarsi a 8 μ s. Queste moderne intuizioni dalla psicofisica implicano quindi che dovremmo conservare la struttura temporale dei segnali acustici ad una scala più fine per le nostre orecchie: un sistema di distribuzione che consente una risoluzione dall'inizio alla fine della catena di 8 μ s implica una larghezza di banda minima di circa 44 kHz.

- Ogni filtro interposto lungo il flusso di distribuzione dell'audio introduce sfasamenti temporali.

- Un limite della catena audio saranno i trasduttori a ciascuna estremità, microfoni lato produzione ed altoparlanti lato riproduzione ognuno con i suoi filtri passa alto e passa basso. All'interno di questi due elementi saranno collocati una serie di amplificatori (amplificatore microfonico, amplificatori di segnali nel mixer, ...nel registratore, ... preamplificatore, amplificatore finale). Come mostra la figura qui sotto, per raggiungere l'obiettivo di avere una banda di 44kHz senza sfasamenti temporali attraverso una cascata uniforme di otto stadi (che implicano 8 filtri), ognuno necessiterebbe di una larghezza di banda vicina a 100 kHz !!! Questo lo dice il signor Nyquist ed il signor Bode...

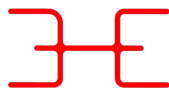
Studio Progettazione

Via G. Mazzini, 4 – 37060 – Mozzecane -VR-

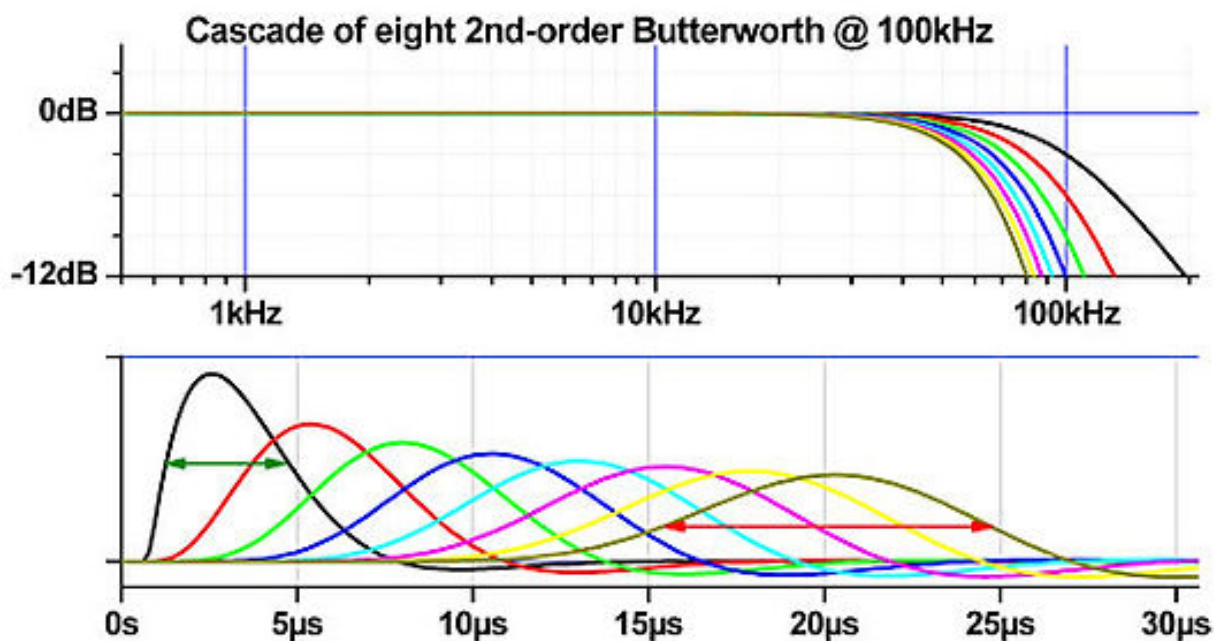
si riceve su appuntamento Mob. +39 347 65 09 541

Mail: entertainmenthiend@gmail.com Sito: <http://www.entertainmenthiend.it>

P.Iva 01713570883 – R.E.A. RG420663



Entertainment Hi End



- Recenti scoperte nelle neuroscienze fanno ipotizzare che l'audio possa essere trasmesso in modo più efficiente se la codifica del segnale è ottimizzata per i suoni ambientali naturali, senza limitazioni (o filtri) piuttosto che con limiti "rettangolari" nel dominio di frequenza ed ampiezza.

- C'è un contenuto significativo sopra i 20 kHz in molti tipi di musica, come rivelato da un'analisi delle registrazioni ad alta frequenza di campionamento (vedi analisi sul grafico qui sotto). Una caratteristica notevole e comune degli spettri degli strumenti musicali è che la potenza diminuisce, spesso in modo significativo, con una frequenza crescente. Anche se alcuni strumenti musicali producono suoni superiori a 20 kHz, non necessariamente ne deriva che un sistema "trasparente" debba riprodurli; ciò che importa è se i mezzi utilizzati per ridurre la larghezza di banda possano essere rilevati dall'ascoltatore umano.

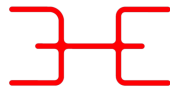
Studio Progettazione

Via G. Mazzini, 4 – 37060 – Mozzecane -VR-

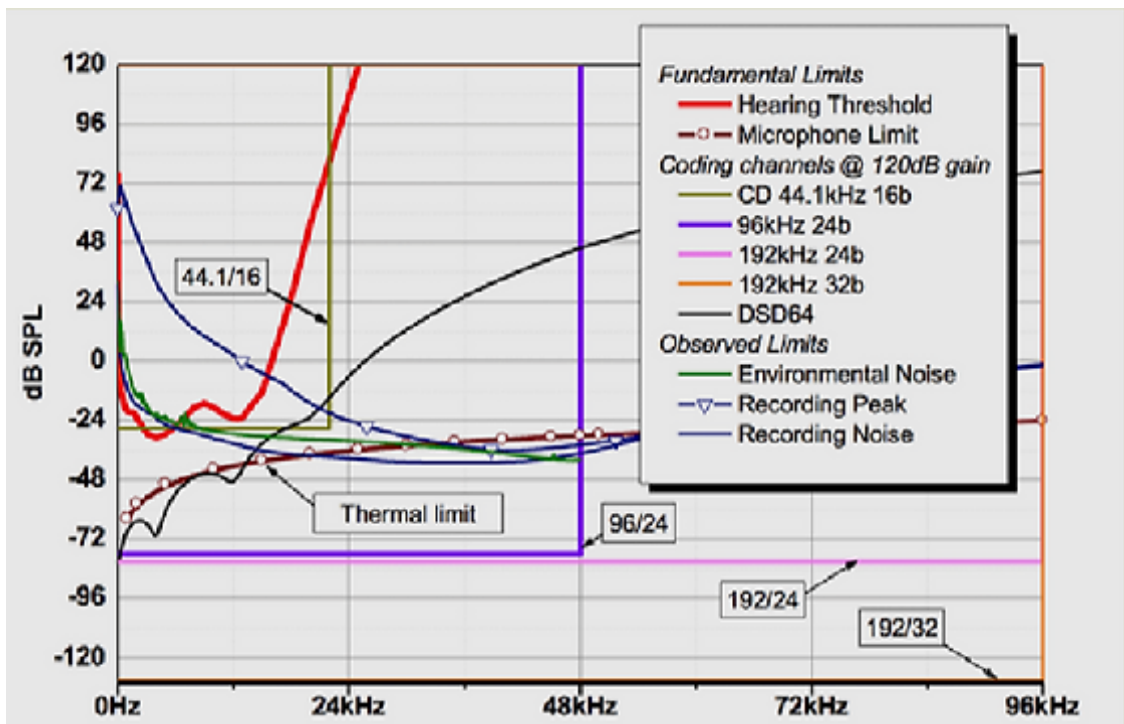
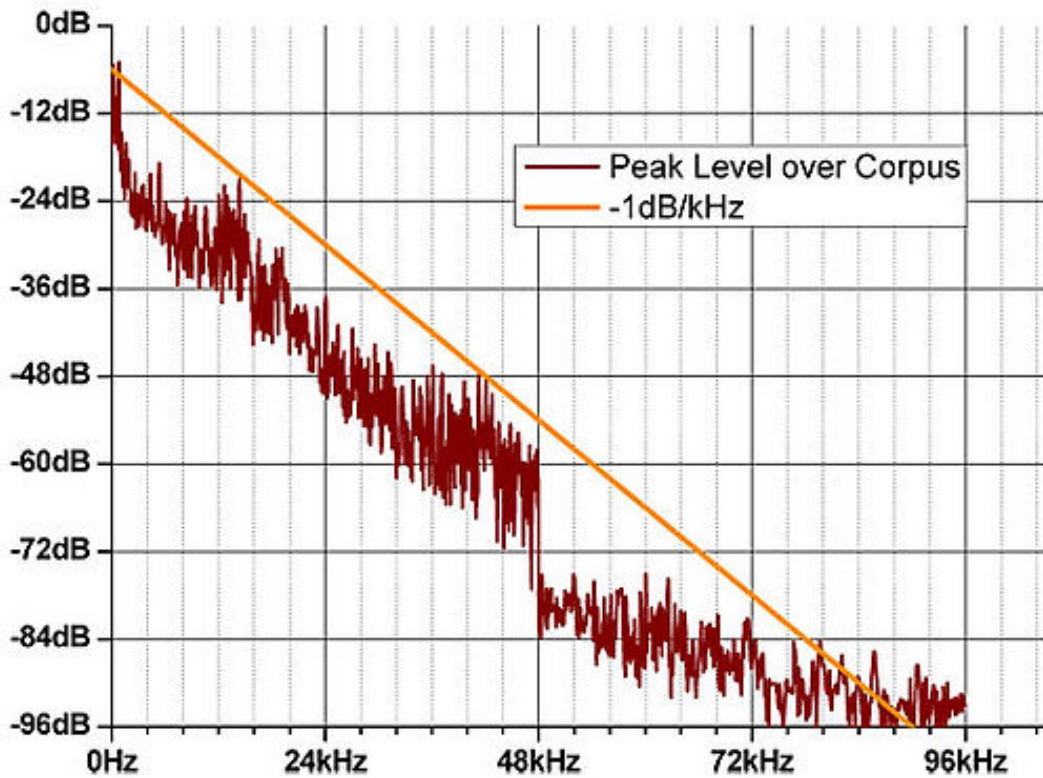
si riceve su appuntamento Mob. +39 347 65 09 541

Mail: entertainmenthiend@gmail.com Sito: <http://www.entertainmenthiend.it>

P.Iva 01713570883 – R.E.A. RG420663



Entertainment Hi End

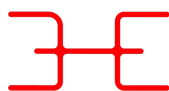


Studio Progettazione

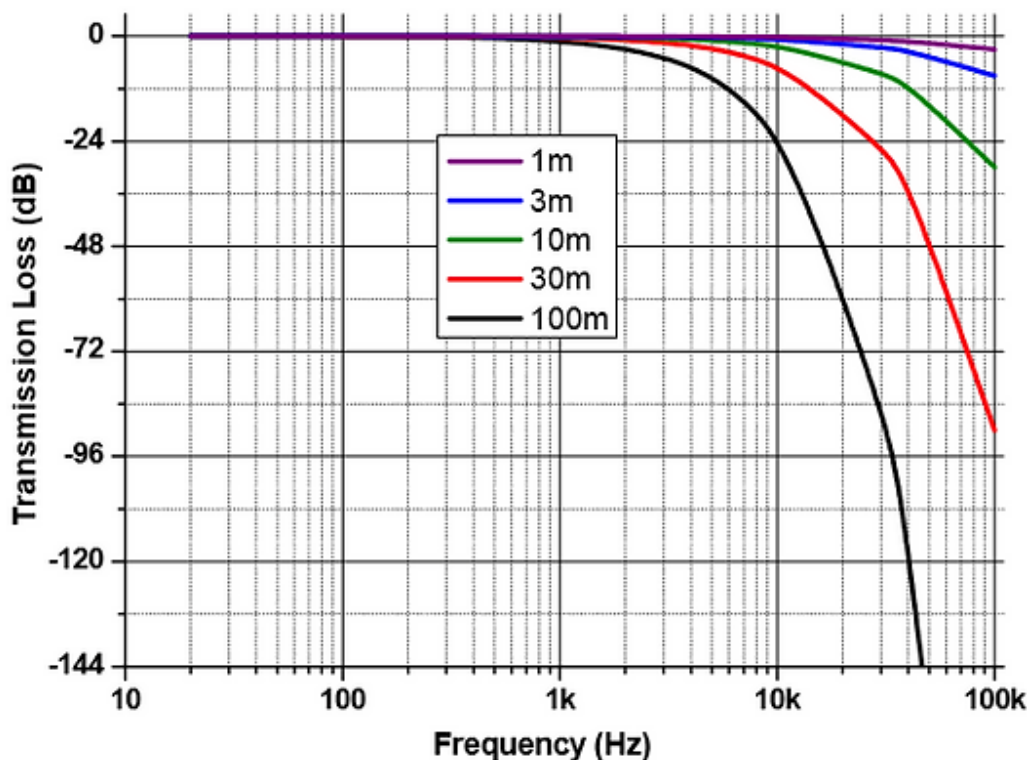
Via G. Mazzini, 4 – 37060 – Mozzecane -VR-
si riceve su appuntamento Mob. +39 347 65 09 541

Mail: entertainmenthiend@gmail.com Sito: <http://www.entertainmenthiend.it>

P.Iva 01713570883 – R.E.A. RG420663



Entertainment Hi End



- La nostra comprensione della tecnologia digitale è stata influenzata dal teorema di Shannon che dimostra che la ricostruzione in modo univoco del segnale analogico può essere possibile se la scomposizione della forma d'onda tramite Fourier è finita ed è uniformemente campionata ad una specifica frequenza detta di (Nyquist) uguale a almeno il doppio della frequenza massima. Il suono non è intrinsecamente limitato dalla banda (i primi sistemi di registrazione analogici mostravano un rolloff ad alta frequenza relativamente delicato), ma la teoria digitale implica precise e chirurgiche limitazioni di banda e la ricostruzione del segnale necessita interpolazioni che possono essere approssimative del segnale reale.

- Qui sotto il grafico riporta i limiti fondamentali del rumore del microfono, la soglia minima dell'udito umano ed un esempio di rumore ambientale. Osservando questo grafico: in alto abbiamo 120 dB SPL che è la soglia del dolore, al centro 0 dB è il silenzio che possiamo percepire mentre 120 dB sotto lo zero ci sta la soglia minima del rumore a 32 bit. Quindi anche se più bit ci danno una scala più fine, la capacità di informazione trasmessa del canale è largamente sovradimensionata.

- A frequenze di campionamento superiori a 48 kHz possiamo utilizzare filtri di ricostruzione con effetti ridotti di pre e post-ringing. Nelle registrazioni di musica a banda larga, i segnali al di sopra dei 48 kHz normalmente cadono al di sotto del rumore di fondo, suggerendo che una frequenza di campionamento di 96 kHz potrebbe preservare la maggior parte del contenuto spettrale. L'ulteriore aumento della

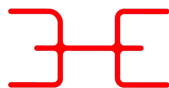
Studio Progettazione

Via G. Mazzini, 4 – 37060 – Mozzecane -VR-

si riceve su appuntamento Mob. +39 347 65 09 541

Mail: entertainmenthiend@gmail.com Sito: <http://www.entertainmenthiend.it>

P.Iva 01713570883 – R.E.A. RG420663



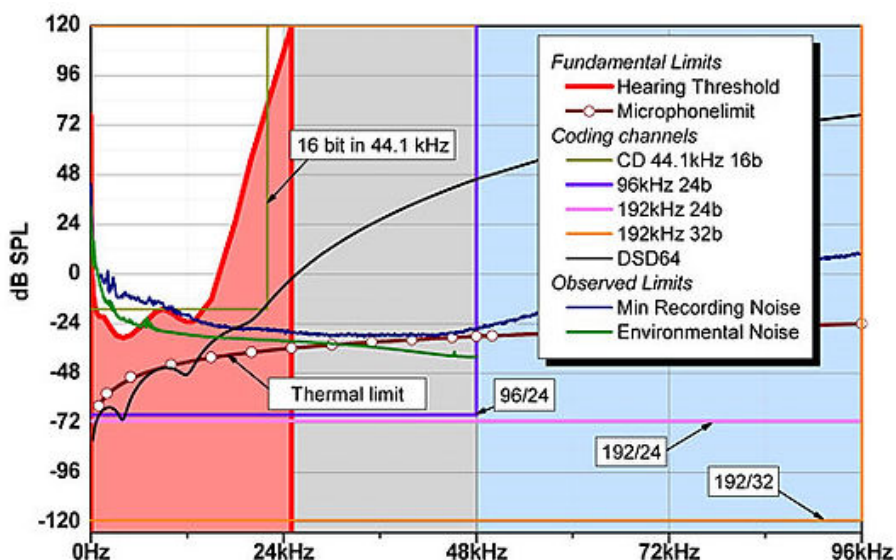
Entertainment Hi End

frequenza di campionamento, produce miglioramenti ridotti: i benefici sono derivanti da un minor numero di decimazioni / passaggi di interpolazione nei convertitori e da minor effetto pre e post ringing.

- L'aumento delle profondità di bit da 16 a 24 o addirittura 32 offre rendimenti decrescenti. Mentre più bit forniscono gradazioni di ampiezza più fine, la profondità di bit determina solo la gamma dinamica, non la risoluzione dell'ampiezza. Analizzando il rumore di fondo di un microfono ed il suo livelli di pressione massima, il suo spettro può essere descritto in un canale LPCM a 192.5 kHz a 17.5 bit. Il rumore di fondo nelle registrazioni non giustifica l'utilizzo più di 18 bit. Fino ad ora, l'alta risoluzione in audio non è stata definita in modo utile, il che è un peccato perché senza un ponte sicuro tra scienza uditiva e ingegneria audio, lo sviluppo può essere casuale e non razionale.

- Se un sistema di riproduzione deve essere impeccabile per l'ascoltatore umano, allora i suoi errori dovrebbero essere sia naturali che plausibili. E' stato proposto che gli errori di sistema audio dovrebbero solo assomigliare a quelli introdotti quando il suono viaggia a breve distanza attraverso l'aria: entro limiti ragionevoli, l'aria non introduce distorsione, ma aggiunge rumore termico e sfocatura temporale attenuando progressivamente le frequenze più alte. Un tale sistema, posto tra l'ascoltatore e l'esecutore, potrebbe non essere notato. Questo è l'obiettivo che una catena di registrazione-riproduzione con MQA si vuole porre. Qui sotto l'effetto dell'aria sulla trasmissione audio:

- Nel diagramma qui sotto: i suoni che percepiamo singolarmente come tali sono racchiusi nella parte bianca del grafico. Con lo stato attuale della tecnologia, i componenti sopra i 48kHz superano la capacità di riproduzione di quasi tutti gli altoparlanti e le cuffie.



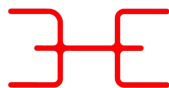
Studio Progettazione

Via G. Mazzini, 4 – 37060 – Mozzecane -VR-

si riceve su appuntamento Mob. +39 347 65 09 541

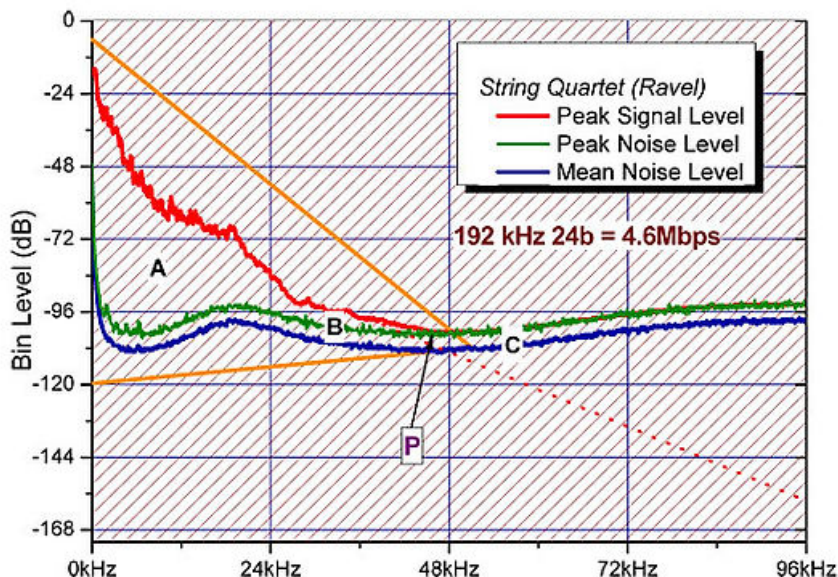
Mail: entertainmenthiend@gmail.com Sito: <http://www.entertainmenthiend.it>

P.Iva 01713570883 – R.E.A. RG420663



Entertainment Hi End

- La pratica dell'encoding MQA.



- Il diagramma qui sopra è un diagramma di Shannon che mostra l'intero spazio di codifica per un segnale sorgente a 24-bit/192kHz, che corrisponde ad un bit rate in LPCM a 4,6Mbps. All'interno di esso, la curva rossa è lo spettro di frequenza di picco dell'audio reale, in questo caso un movimento da un quartetto d'archi da un'opera di Ravel. La curva verde e blu, rispettivamente, mostrano il picco e la media del rumore di sottofondo della registrazione: complessivamente lo spazio racchiuso dalle curve è il suono che ascoltiamo .

Questa registrazione è stata scelta come esempio perché è realizzata con strumenti reali in un'acustica naturale. Le corde pizzicate sono difficili da riprodurre e lo spettro mostra componenti armoniche ben oltre i 20kHz, (anche se sopra la soglia udibile, contribuiscono alla creazione del suono complessivo).

Il livello di picco rilevato dal microfono mostra uno spettro decrescente con frequenza crescente.

Si noti che la musica e le curve di rumore convergono in "P" - probabilmente ci sono componenti del segnale al di sopra di quella frequenza che sono comunque perse, "affogate" nel rumore.

- La regione contrassegnata con "A" si trova nella banda audio convenzionale: noi siamo sensibili alle frequenze fino a 20kHz.
- La regione "B" contiene contenuti musicali, ma nessuno di questi intervalli è udibile se ascoltato separatamente; tuttavia gli elementi in "B" contribuiscono alla risoluzione temporale e all'involuppo sonico, ma non appaiono isolatamente. L'esperienza mostra che la rimozione di questi abbassa la fedeltà audio percepita.

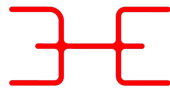
Studio Progettazione

Via G. Mazzini, 4 – 37060 – Mozzecane -VR-

si riceve su appuntamento Mob. +39 347 65 09 541

Mail: entertainmenthiend@gmail.com Sito: <http://www.entertainmenthiend.it>

P.Iva 01713570883 – R.E.A. RG420663

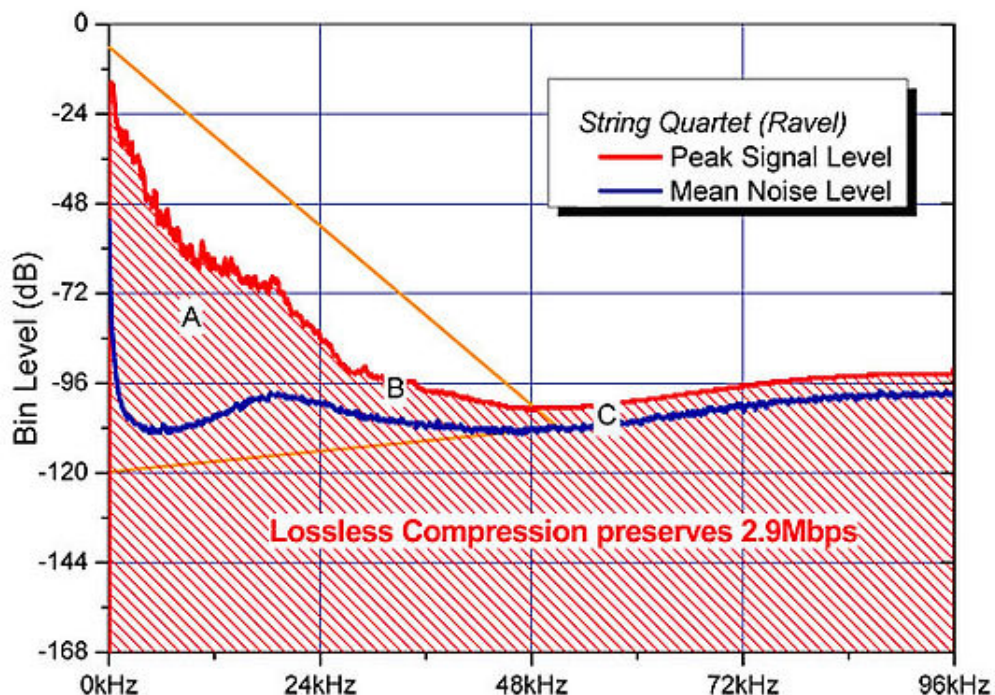


Entertainment Hi End

- La regione "C" è diversa: non ha informazioni salienti sulla musica, è al di sopra della banda passante di quasi tutti i microfoni e gli altoparlanti ed è anche al di sotto e al di là delle soglie umane per i segnali di rumore. Questo rumore costante è impercettibile. Tuttavia se ci teniamo questo spettro senza filtrarlo, con le frequenze di campionamento più elevate (= banda passante maggiore) è possibile avere una migliore risoluzione nella regione "A". Maggiore è la velocità di campionamento minore è la "sfocatura" temporale, molto ben percepita dal sistema uditivo umano.

Come notato in precedenza, nelle registrazioni acustiche, il punto "P" è osservato tra 30 e 60kHz, tipicamente a 40kHz. Quindi, quando le frequenze di campionamento sono superiori a 96kHz, ad esempio a 192 o 352,8kHz, la regione "C" si estende su una larghezza di banda più ampia. In quella regione, tutti i componenti correlati alla musica sono ben al di sotto del rumore. Come già ribadito è possibile migliorare complessivamente la riproduzione del suono perché la maggiore frequenza di campionamento consente una minore sfocatura causata dal filtro ed una migliore convergenza temporale.

Il triangolo arancione nel diagramma sopra, racchiude tutta la parte musicalmente più rilevante del segnale e ha un'area di circa un sesto dell'intero spazio di codifica.



Per tutti i ragionamenti esposti sopra concordiamo di tenerci uno spettro ampio 96KHz, per cui adotteremo 192Khz di frequenza di campionamento.

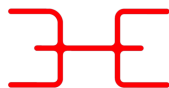
Studio Progettazione

Via G. Mazzini, 4 – 37060 – Mozzecane -VR-

si riceve su appuntamento Mob. +39 347 65 09 541

Mail: entertainmenthiend@gmail.com Sito: <http://www.entertainmenthiend.it>

P.Iva 01713570883 – R.E.A. RG420663



Entertainment Hi End

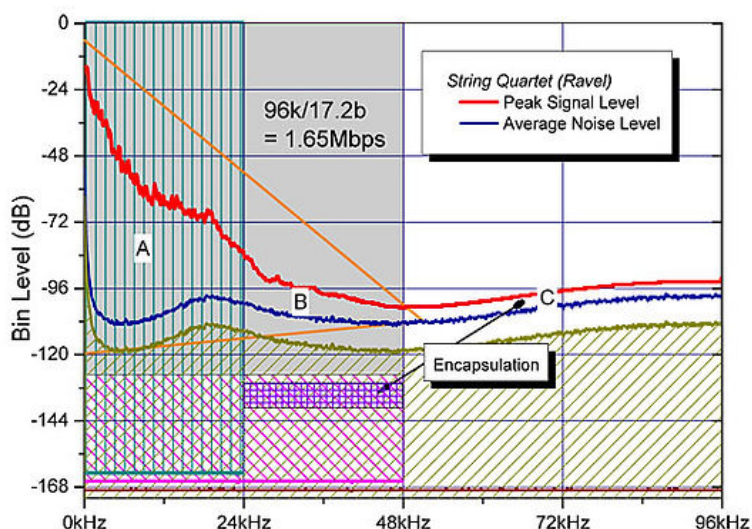
Se comprimessimo mediante una codifica lossless (FLAC O ALAC) un segnale 192KHz/24 bit otterremmo uno streaming a 2,9Mbps, con un risparmio del 37% rispetto al segnale non compresso, ma questo sarebbe ancora inefficiente per essere ideale per lo streaming da servizi di musica online. E' qui che entra in gioco la tecnica di encoding di MQA.

L'obiettivo di MQA è quello di fornire il contenuto del triangolo arancione con estrema precisione, evitando la sfocatura temporale, per cui in qualche modo veicolando lo spettro sopra i 44.1KHz ed evitare la modulazione del rumore nei convertitori

Per raggiungere questi obiettivi, MQA va oltre il concetto di "lossless", nel senso che alle radici della sua filosofia, MQA ha la consapevolezza che il concetto "senza perdita" non è garanzia di qualità del suono finale. Per completare l'idea di riproduzione più vicina possibile al reale MQA prende in considerazione anche aspetti più intimi relativi all'hardware di produzione/riproduzione, fino a tenere in considerazione le caratteristiche, con i loro difetti, dei convertitori A/ D e D / A. Un file digitale a 16-bit / 44,1KHz può essere consegnato senza perdita di dati, ma ciò non garantisce il raggiungimento di una qualità audio end-to-end irrepreensibile, partendo dal mondo analogico, passando per il digitale, tornando all'analogico

Lo stesso vale per il segnale audio digitale Hi-Res convenzionale perché, sebbene potenzialmente superiore a 16-bit / 44.1kHz, non è creato ed ottimizzato per corrispondere alla percezione nel dominio temporale dell'udito umano, come le ultime ricerche hanno evidenziato.

Quindi l'obiettivo di MQA non è solo un'operazione "senza perdita di dati" in senso stretto e tecnico. L'obiettivo è quello di catturare e fornire tutto ciò che possiamo sentire, senza la confusione intrinseca del campionamento convenzionale, senza l'incertezza della quantizzazione del convertitore e per trasmettere il suono analogico ascoltato nello studio di mastering all'utente finale senza modifiche. Questo lo rende 'senza perdita' in un modo molto più profondo e rilevante.



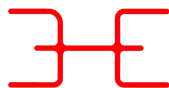
Studio Progettazione

Via G. Mazzini, 4 – 37060 – Mozzecane -VR-

si riceve su appuntamento Mob. +39 347 65 09 541

Mail: entertainmenthiend@gmail.com Sito: <http://www.entertainmenthiend.it>

P.Iva 01713570883 – R.E.A. RG420663



Entertainment Hi End

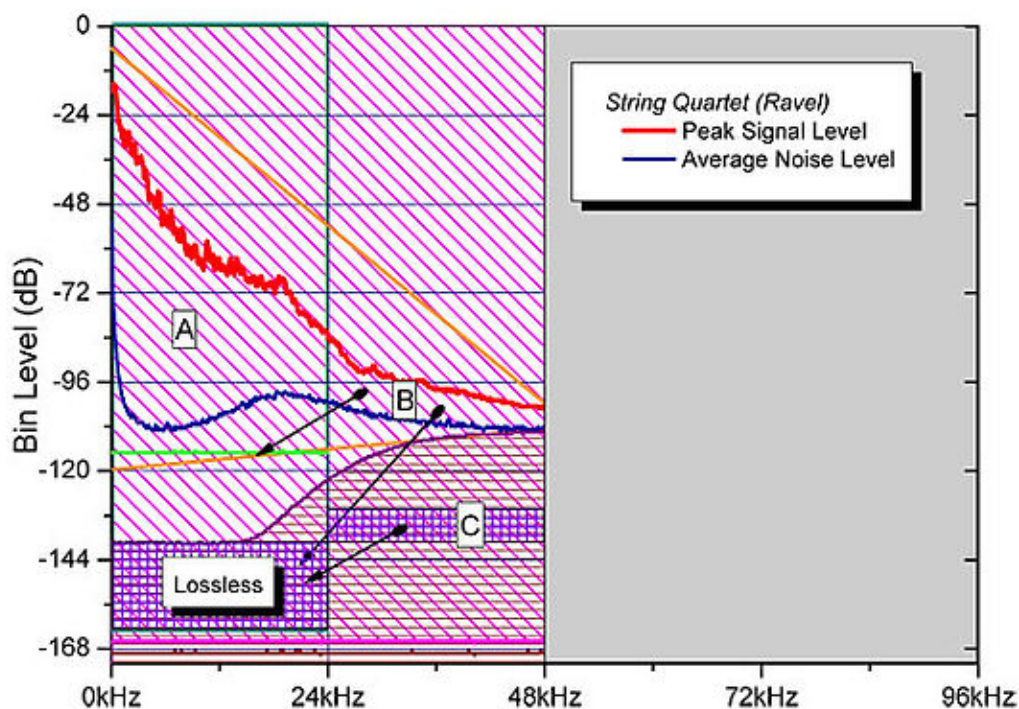
Riprendiamo ad analizzare, per semplicità lo spettro dell'opera di Ravel.

MQA, per ridurre la dimensione del file (e dello streaming) del segnale audio in alta risoluzione lavora "ripiegando" porzioni di spettro all'interno del file stesso, nascondendolo questa porzione sotto il livello di rumore. Cerchiamo di capire meglio. Si ottiene così un file, idealmente assimilabile ad un ORIGAMI.

La prima "piega", chiamata E, dell'origami che vediamo qui sopra riduce la velocità di trasmissione da 192kHz (96KHz di banda passante) a 96kHz (48KHz di banda passante), Il processo di piegatura non sta filtrando il segnale: la frequenza di campionamento intrinseca e la profondità di bit rimangono inalterati !

Il processo è gerarchicamente scalabile, quindi se il file sorgente alla fonte fosse ad esempio, un file a 352.8kHz, useremo tre volte la "piega" per raggiungere la velocità di trasmissione finale di 44.1kHz.

La prima 'piega' E è chiamata tecnicamente incapsulamento: la velocità di trasmissione è dimezzata e il contenuto del segnale di C è memorizzato sotto il livello di rumore, nell'area B. Tecniche già note, chiamate noise shaping.



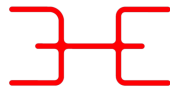
Studio Progettazione

Via G. Mazzini, 4 – 37060 – Mozzecane -VR-

si riceve su appuntamento Mob. +39 347 65 09 541

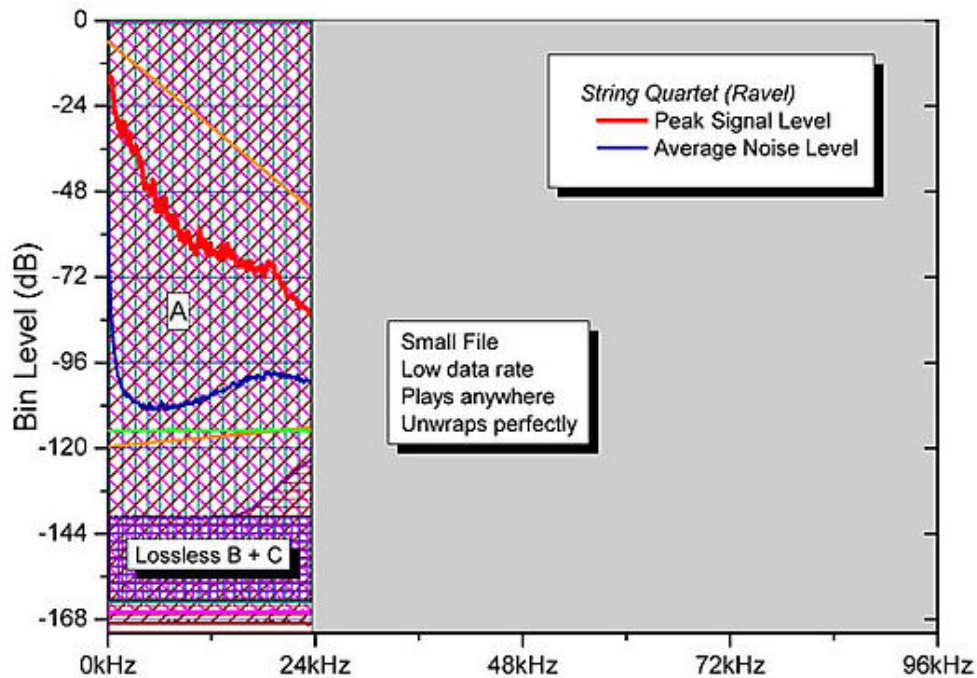
Mail: entertainmenthiend@gmail.com Sito: <http://www.entertainmenthiend.it>

P.Iva 01713570883 – R.E.A. RG420663



Entertainment Hi End

La seconda piega è chiamata "L": la velocità di trasmissione viene nuovamente dimezzata. Il segnale nell'area B viene seppellito senza perdite e nascosto al di sotto del rumore dello spettro della banda base, insieme a C.



Nella figura qui sopra vediamo il file MQA completato. Le pieghe E ed L sono contenute all'interno del rumore. Se è presente un decodificatore in riproduzione MQA, i due processi di riapertura dell'origami vengono attuati con precisione. Se non è presente alcun decoder, viene riprodotto il segnale della banda base compatibile.

Ciò che questi diagrammi e schemi di codifica non mostrano apertamente, sono gli algoritmi dei processi che MQA usa "sotto il cofano" per ottenere questi piegamenti in quanto alcuni di questi processi sono coperti da brevetti importanti. Ad esempio, vengono utilizzati nuovi kernel di campionamento che si adattano alle statistiche del segnale ambientale e musicale, che sono adattati per gli ascoltatori umani e forniscono una stretta risoluzione temporale. Questi si trovano nelle workstation che codificano il file MQA. MQA sfrutta anche la sua architettura end-to-end per abilitare le tecniche di dither sottrattive nel decoder per garantire che, a differenza di altri formati, non vi sia alcuna modulazione del rumore. Questa è una caratteristica unica rispetto a qualsiasi altro sistema analogo presente sul mercato. I singoli DAC vengono analizzati dai laboratori MQA.

Racchiuso nel firmware di decodifica MQA presente nell'hardware DAC di playback ci sono parametri che servono a correggere ed ottimizzare riproduzione del file contenete la codifica MQA.

Studio Progettazione

Via G. Mazzini, 4 – 37060 – Mozzecane -VR-

si riceve su appuntamento Mob. +39 347 65 09 541

Mail: entertainmenthiend@gmail.com Sito: <http://www.entertainmenthiend.it>

P.Iva 01713570883 – R.E.A. RG420663