1. SCOPE

This Operational Practice sets out guidelines for minimising the various artefacts that may distort audio signals when low bit-rate coding schemes are employed to convey contribution sound in broadcasting for programmes with a quality imperative. The Recommendations in this OP are in no way intended to constrain the applications in event television such as mobile telephone feeds but provide guidance toward maintaining quality in contribution audio.

2. RECOMMENDED BIT RATES FOR AUDIO CODING SCHEMES IN THE BROADCASTING CONTRIBUTION CHAIN

The use of very low bit rate coding schemes should be avoided in television contribution services because they are likely to precipitate unpleasant artefacts when those services are subsequently processed through the high efficiency down-stream emission codecs in the broadcasting chain.

Where lossy compression schemes are utilised to reduce the bit rate of the material to effect lower carriage costs, careful analysis should be undertaken to weigh the cost to the quality of the final on-air product. As a general rule, the use of lossy compression schemes should be confined to single transaction services where there is no concatenation. They are not recommended for use in contribution circuits.

If a lossy compression system must be employed for contribution, its application must take into account that the contribution scheme must be as good as or better than the emission scheme to be utilised plus some overhead to account for the unknown concatenation that may occur between contribution and emission.

The following bit rates are recommended for broadcast contribution employing the most commonly used audio coding schemes:

a) Contribution via MPEG 1 Layer II – 256 kbps stereo for 48 kHz sampling

b) Contribution via MPEG 1 Layer III ("MP3") – preferably 320 kbps stereo for 48 kHz sampling

b) Contribution via AAC – at least 256 kbps stereo for 48 kHz sampling

c) Contribution via HE-AAC – not less than 128 kbps stereo for 48 kHz sampling

Table 1 on pages 2 and 3 provide a convenient reference of recommended bit rates for the codecs commonly used in Australian broadcasting systems.

Annex A to this OP provides further guidance on assessing and managing the application of low bit rate compression schemes.

TABLE 1 - SUGGESTED BIT RATES FOR LOSSY CODECS FOR BROADCAST USE

N.B. base rates are in bold. Other rates are derived from these by a simple formula, to the nearest multiple of 8 kbps. If these rates are not available, use the next highest rate.

Codec	Sample rate kHz	Audio bandwidth kHz	Channel Config.	Usage	Minimum bit rate for emission use kbps	Minimum bit rate for contribution use kbps
AC-3	48	20	5.1	Speech/Music/FX/Full program	**384**	640
Dolby E			8	Speech/Music/FX/Full program		1920 (20 bit)
Dolby E			6	Speech/Music/FX/Full program		1536 (16 bit)
AAC	48	20	Stereo	Speech/Music/FX/Full program		**256**
AAC-LD			Stereo	Speech/Music/FX/Full program		**256**
APT-X			Stereo	Speech/Music/FX/Full program		**384**
HE-AAC			Stereo	Speech/Music/FX/Full program		**128**
MPEG-1 Layer II			Stereo	Speech/Music/FX/Full program	**192**	**256**
MPEG-1 Layer III			Stereo	Speech/Music/FX/Full program		**320**
AAC	44.1	20	Stereo	Speech/Music/FX/Full program		256
AAC-LD			Stereo	Speech/Music/FX/Full program		256
APT-X			Stereo	Speech/Music/FX/Full program		384
HE-AAC			Stereo	Speech/Music/FX/Full program		128
MPEG-1 Layer II			Stereo	Speech/Music/FX/Full program		256
MPEG-1 Layer III			Stereo	Speech/Music/FX/Full program		320

AAC	32	15	Stereo	Speech/Music/FX/Full program		176
AAC-LD			Stereo	Speech/Music/FX/Full program		176
APT-X			Stereo	Speech/Music/FX/Full program		256
HE-AAC			Stereo	Speech/Music/FX/Full program		88
MPEG-1 Layer II			Stereo	Speech/Music/FX/Full program		176
MPEG-1 Layer III			Stereo	Speech/Music/FX/Full program		216
AAC	16	7	Stereo	Speech only		96
AAC-LD			Stereo	Speech only		96
APT-X			Stereo	Speech only		144
HE-AAC			Stereo	Speech only		48
MPEG-1 Layer II			Stereo	Speech only		96
MPEG-1 Layer III			Stereo	Speech only		120
AAC	48	20	Mono	Speech/Music/FX/Full program		128
AAC-LD			Mono	Speech/Music/FX/Full program		128
APT-X			Mono	Speech/Music/FX/Full program		192
HE-AAC			Mono	Speech/Music/FX/Full program		64
MPEG-1 Layer II			Mono	Speech/Music/FX/Full program	**96**	128
MPEG-1 Layer III			Mono	Speech/Music/FX/Full program		160

AAC	44.1	20	Mono	Speech/Music/FX/Full program		128
AAC-LD			Mono	Speech/Music/FX/Full program		128
APT-X			Mono	Speech/Music/FX/Full program		192
HE-AAC			Mono	Speech/Music/FX/Full program		64
MPEG-1 Layer II			Mono	Speech/Music/FX/Full program		128
MPEG-1 Layer III			Mono	Speech/Music/FX/Full program		160
AAC	32	15	Mono	Speech/Music/FX/Full program		88
AAC-LD			Mono	Speech/Music/FX/Full program		88
APT-X			Mono	Speech/Music/FX/Full program		128
HE-AAC			Mono	Speech/Music/FX/Full program		48
MPEG-1 Layer II			Mono	Speech/Music/FX/Full program		88
MPEG-1 Layer III			Mono	Speech/Music/FX/Full program		112
AAC	16	7	Mono	Speech only		48
AAC-LD			Mono	Speech only		48
APT-X			Mono	Speech only		72
HE-AAC			Mono	Speech only		24
MPEG-1 Layer II			Mono	Speech only		48
MPEG-1 Layer III			Mono	Speech only		64

ANNEX A - ASSESSING AND MANAGING THE APPLICATION OF LOW BIT RATE COMPRESSION SCHEMES

1. LOSSY COMPRESSION AND CONCATENATION

A number of coding schemes enable digital audio to be stored and reticulated at conservative bit rates i.e. they are data reduction schemes. “Lossy” coding schemes compress the digital information by discarding those portions of the audio image that can be omitted without substantially affecting the listener’s perception of the original material. In a single transcription of a compact-disc music track to a personal *MP3* player, for example, substantial bit rate reductions are possible without objectionable alteration to an average listener’s perception of the original sound. However, the resulting compressed audio is only intended for reproduction from the end user’s player. Any further compression of the low bit rate image may discard essential parts of the audio and may precipitate unpleasant artefacts that were not part of the original sound. This cascade process, otherwise known as concatenation, is one of the Achilles’ heels of all lossy audio (and video) coding schemes.

2. MANAGING COMPRESSION IN THE BROADCASTING SYSTEM

MPEG 1 layer II and proprietary audio coding schemes are used in digital television to conserve bandwidth in recording systems, distribution circuits and within the broadcasting channels. These are also “lossy” coding schemes but rely on high bit rates and careful management of consecutive stages of compression to minimise distortions.

Some codecs commonly used by Australian broadcasters \ provide a convenient method of multiplexing 8 audio channels onto a single AES pair while maintaining the relationship between digital sound and picture frames to aid in editing and switching. Since these schemes are able to sustain around eight encode and decode cycles without appreciable quality loss, they are considered appropriate for services where operations may impose concatenation.

Other multiplexed multi-channel audio schemes such as MPEG Surround and AAC Multi-channel have been designed as “one shot” emission schemes and are not suitable for broadcast contribution or distribution (just as AC3 is not suitable for these purposes).

Similarly, caution must be exercised in the use of compressed audio from consumer based optical discs within a facility or in a programme mix (for example AC-3 on DVD), as their choice of bit-rates and the codecs within the disc-players are intended for single cycle transactions and may result in audible concatenation artefacts if their outputs are subjected to further encoding. For all practical purposes, the use of such material (in-house notwithstanding), can be regarded as *contribution-feeds* from the perspective of this OP and will be better managed by observing the general principles set out in this document.

In the final transmission path of a broadcasting system, compression ratios can be relatively high to meet the demands of limited spectrum availability since the emitted broadcast signal is only intended for a single transaction between the emission coder and the consumers’ receivers and its potential to generate further unwanted artefacts will therefore be limited.

A convenient reference chart of recommended minimum bit-rates for the commonly used emission and contribution coding systems is set out in Table 1 to this OP.

3. IDENTIFYING LOW BIT RATE COMPRESSION ARTEFACTS

It should be noted that, unlike analog distortions that arise from overloads, or non-linear interaction of audio sources, digital compression artefacts are often subtle and difficult to quantify by conventional electronic measurement techniques.

Ongoing research aims to establish a repeatable objective means of measuring digital audio distortions that arise from data reduction coding schemes. The method involves the comparison of samples before and after the coding and decoding processes to measure their departure from a defined mask. The errors might then be quantified.

For the time being, in the absence of a definitive measuring technique, practitioners can learn to recognise the more prevalent effects and avoid those by observing some simple management principles.

There are interactive tutorial resources[1] available that provide audible examples of the common compression artefacts and explain their origins in detail. With the assistance of such a resource, the practised sound professional can learn to recognise common anomalies and anticipate the circumstances that are likely to precipitate them.

4. LOW BIT- RATE CODING SYSTEM MANAGEMENT PRINCIPLES

4.1 Choice of Sampling Rates

Where possible, standardising on a single sampling rate within an organisation is recommended to minimise the need for sample rate conversions. For broadcast television, this sample rate should be 48 kHz. Higher sampling rates mean more bandwidth in transmission and the requirement for more storage space. Sample rates higher than 48 kHz, are not recommended. Similarly, sampling rates can be as low as 16 kHz for some content without appreciable degradation but a sample rate more than double the audio bandwidth of the material is necessary[2].

4.2 Use of MPEG -1 Layer II Codecs

MPEG -1 Layer II (*MP2*) is a widely-used general-purpose coding scheme for interchange of material between and within broadcasting organisations. It is generally more resistant to concatenation effects than higher-compression codecs. The use of Layer II coding for interchange is recommended as a preferred choice wherever possible.

4.3 Use of MPEG – 1 Layer III (*MP3*) and High Efficiency Coding Schemes.

The use of codecs with higher data compression than Layer II (e.g. AC-3) should be limited to the final transmission stage of the broadcasting process. If they are used at earlier stages of production, a significant bit rate margin over the usual single-stage bit rate should be added to minimise concatenation effects (e.g. use MP3 at not less than 192 - 320 kbps rather than 128 kbps for stereo with 48 kHz sampling).

A common problem which can arise is the import into the broadcast system of audio production material which has been coded at a low data rate by a system such as *MP3* to enable it to be conveniently distributed as an E-mail attachment, or transported on inexpensive media, at reduced cost. This practice will almost inevitably result in downstream artefacts being generated due to concatenation of the depleted material with other lossy coding schemes and should be avoided. In principle, contributing material from low audio quality devices via lossy coding systems is not recommended because subsequent downstream coding in the broadcasting system is likely to result in the appearance of unpleasant artefacts

---

[1] One example is *PERCEPTUAL AUDIO CODERS – What to Listen for* – Audio Engineering Society, Inc. 2002 - published as an interactive CD.
[2] In accordance with the *Nyquist* theorem.

4.4 Using AAC and HE-AAC (AAC+) Coding

The Advanced Audio Coding scheme commonly uses the “MP4” file suffix. It has recently gained popularity in production studios because of its high-efficiency and because there is a ready source of music and effects available from on-line vendors who use AAC for distribution.

Standard AAC should be run at 256 kbps or better wherever possible. It should be noted that a major online music vendor commonly uses AAC at 256 kbps. AAC coded material should not be confused with HE-AAC products.

HE-AAC coding has several refinements and is more efficient than standard AAC. Although HE-AAC was deemed suitable for broadcast emission systems in 2005 and added to the DVB specification (ETSI TS 102 005 V1.2.0, 2005), it is only recommended for single transaction services and with bit rates not less than 128 kbps. In Australian broadcasting, HE-AAC is not recommended for contribution services where there may be a number of concatenated processes.

4.5 The Number of Concatenations and Quality

If a production chain has many encode-decode stages and/or if the quality of the material is critical, bit rates of 256kbps stereo and 128 kbps mono are recommended for 48 kHz sample rate in Layer II coding.

If audio quality is less critical, or there are few stages in the production chain, or budget or bandwidth are issues, lower bit rates may be acceptable but the minimum should be 128kbps joint stereo or 64 kbps mono for 48 kHz sample rate in Layer II coding. If a lower sampling rate than 48 kHz is used, the bit rates suggested above should be scaled downwards proportionally.

4.6 Codec Latency

Where sound is synchronised with picture, codec delay should be taken into account and delay equalisation may be necessary to maintain synchronisation. In live two-way links, low-delay codecs should be used wherever possible. Delays under 30ms are preferred in this situation. Low-delay picture codecs should also be used if there is a picture monitor feed and latency in the picture monitor electronics may also be a factor requiring delay equalization.

4.7 Consistency in Monitoring

The assessment of digital artefacts is usually done by critical listening. Organisations therefore should endeavour to maintain consistent monitoring facilities and listening room acoustics between work areas so that work will sound similar when replayed in any site. Any artefacts such as changes in timbre caused by compression anomalies can then be confidently attributed to the electronic process and not confused with colourations due to loudspeakers and room acoustics. Alternatively, a single, known quality assessment area and a known assessor might be recommended.

5 GLOSSARY

Codec	Abbreviation of enCoder-deCoder
MP2	MPEG-1 Layer II as defined in ISO/IEC 11172-3
MP3	MPEG-1 Layer III as defined in ISO/IEC 11172-3
MP4	Generic term for MPEG-4 Audio as defined in ISO/IEC 14496-3. Usually refers to standard AAC coding.
AAC	Advanced Audio Coding – defined in MPEG-2 (ISO/IEC 13818-3) and MPEG-4 (ISO/IEC 14496-3)
HE-AAC	(AAC+) A more efficient version of AAC with refinements, defined in MPEG-4 (ISO/IEC 14496-3).

6 REFERENCES

ISO/IEC 11172-3 Information technology -- Coding of moving pictures and associated audio for digital storage media at up to about 1,5 Mbit/s -- Part 3: Audio (MPEG-1)

ISO/IEC 13818-3 Information technology -- Generic coding of moving pictures and associated audio information -- Part 3: Audio (MPEG-2)

ISO/IEC 14496-3 Information technology -- Coding of audio-visual objects - Part 3: Audio (MPEG-4)

ETSI TS 102 005 V1.4.1, 2010 Digital Video Broadcasting (DVB); Specification for the use of Video and Audio Coding in DVB services delivered directly over IP protocols

EBU Tech 3296
Listening Tests - low-bitrate audio codecs June 2003

EBU - TECH 3309
Evaluations of Cascaded Audio Codecs
June 2005

EBU - TECH 3329
A Tutorial on Audio Contribution over IP May 2008

EBU - TECH 3324
EBU Evaluations of Multichannel Audio Codecs
September 2007

EBU - TECHNICAL REVIEW – January 2006
Audio Coding for Today’s Digital Media World
S. Meltzer and G. Moser