1. Scope

This operational practice sets out the requirements for *downmixing* 5.1 and 5.0 channel surround sound audio mixes to 2 channel stereo. This operational practice recommends a number of factors which need to be taken into consideration to establish consistency of reproduction across varying listening environments and programme genres.

2. What is downmixing?

Downmixing is the process of re-mixing a number of audio channels in a multi-channel mix, and combining them into fewer audio channels, for example; 5.1 channel surround into 2 channel stereo.

The application for Australian free-to-air television broadcasting is any program with multichannel audio, either live or pre-recorded, that is to be downmixed to feed an MPEG-1 Layer II stereo or Dolby AC-3 2.0 digital sound part of a TV program service. If the multichannel sound is broadcast on a HD TV service with Dolby AC-3 5.1 channel capability, there should be consideration of a home TV receiver environment automatic downmixing to stereo. With services such as 'catchup TV' or live streaming the important consideration is at all times the maintenance of audibility of dialogue.

3. The Listening Environment

Whilst professional audio engineers should always mix in the best conditions possible, it is important to remember that those resulting mixes may be heard in a number of varying consumer environments such as those with high ambient noise levels or using poor sound reproduction equipment employing small speaker drivers.

4. Quality of a downmix

It should be the objective of a downmix to be high in sound quality; with minimal change to spectral content and dynamic range. Elements of the mix that can be heard by the listener in 5.1 channel surround should also be detectable to the listener in a 2.0 downmix - excluding those elements in the LFE channel.

5. Where does downmixing happen and how does this affect the result?

Downmixing can happen at the originating location and be performed by an experienced audio engineer; for example, in an outside broadcast environment. It can also happen at a broadcast playout facility, via an automated process with little or no capability for operator intervention, or in the home, for example; within the functionality of a DTV consumer set-top box or integrated DTV receiver when tuned to a HD program with 5.1 audio.

6. Live downmixing vs Post downmixing in the DTV receiver

Live downmixing by an experienced audio engineer can be dynamic and adjusted during the course of the program in order to get the best desired result.

Whereas, downmixing undertaken within the functionality of consumer DTV receiver is usually achieved by a static set of downmixing parameters. These downmixing co-efficients can be sent by the broadcaster as metadata embedded in the audio signal. Although the downmix parameters can be changed dynamically, it is normal practise for a broadcaster to select a specified set of parameters as a method for the downmix.

7. Downmixing for simulcast

Some programs are simulcast on HD and SD services. The HD service may have 5.1 channel sound, while the SD service usually has 2.0 channel sound. The 2.0 channel downmix for SD is usually performed by the broadcaster at the studio.

Many HD viewers may have only 2-channel reproduction, and their downmix will take place in the receiver.

It is important therefore for broadcasters to maintain consistency wherever possible between studio downmixes and receiver downmixes by using the same mixing coefficients. Since the broadcaster has relatively little control of downmix coefficients in the receiver, this means that standard (receiver) downmix coefficients should be used at the studio also, unless this is likely to result in quality problems such as disappearing elements in the mix or poor speech intelligibility.

8. Metadata settings; Fixed or variable

Most broadcasters have opted for fixed metadata settings such as *dialnorm*, channel mode, and dynamic range profile, and strive to 'fix' the audio prior to broadcast to match those settings, rather than change metadata on a programme to programme basis.

A specified set of parameters as a method for the downmix avoids potential variations in delivery to the television viewer, whereas varying the parameters may cause wide fluctuations in the delivered sound signal as new metadata settings are transmitted and applied. (e.g. Muting of audio signal during the change from 5.1 channel surround to 2.0 etc.)

9. Downmixing "upmixed" material

Some broadcasters transmit a constant 5.1 channel surround on some services, much of which has been upmixed from stereo two channel sources, in some cases Lt Rt surround encoded.

The upmixing devices, depending on manufacturer, tend to employ a Pro-Logic style of decode, mainly steering in-phase material into the centre channel, and placing more out-of-phase elements into the surround channels. It should be noted this upmixed 5.1 (or 5.0) channel surround may be downmixed by many consumer DTV set-top boxes & integrated DTVs back to 2 channel stereo.

10. Out-of-phase content disappearing on downmix

The surround channels in an original 5.1 mix, when downmixed to stereo, are often given a phase shift to create a Lt Rt 2.0 mix. If heard in mono, these elements of the mix (particularly dialogue), *may* be low level or lost all together due to phase cancellation. It is not recommended for content that is essential to the narrative at any particular moment (like dialog or effects) be placed in the surround channels in the original mix.

11. Binaural summation & relative loudness in different channel formats and different source directions, especially surround channels vs. front channels

In loudness measurement, the surround channels have a higher weighting than the front channels. That is because human psychoacoustic perception hears sounds from behind as louder than sound origination from in-front of the listener.

12. Directional masking / unmasking considerations

This may result in a loud section of a soundtrack which, in the original mix, is carried in the surround channels, tending to override lower level content carried in the front channels. In live sound presentations this effect can also be contributed to by reverberation. When the loud sections of the soundtrack (that are carried in the surround channels) are subsequently downmixed so that they came from the same direction as quieter sounds, this is termed “directional masking”.

As a result, the “danger” of directional masking needs to be taken into consideration when establishing the loud and low level sections of the soundtrack in the 5.1 surround channel assignments.

If not taken into consideration when the mix down is undertaken to 2 channel stereo or mono, the loud level section originally carried in the surround channels may completely mask this low level section of audio carried in the front.

13. Redirecting dialogue from the centre channel

Preparation of a suitably mixed stereo sound track preserves dialogue intelligibility in the downmix. It has been widely accepted that the centre channel be re-directed equally and in-phase to the left and right channels of the 2.0 downmix (in addition to a -3dB offset, because it is now present in two channels instead of one).

14. Bass management systems and the LFE channel

Bass management systems are used to redirect bass between channels for optimal reproduction. There are two standard configurations and decoder manufacturers may offer more options as well. Configuration One redirects bass from all channels into a subwoofer. Configuration Two is primarily for redirecting bass into L and R channels, with optional augmentation from a subwoofer.

The LFE signal in a digitised 5.1 mix and the amplifier signal that is sent to a subwoofer in a monitoring system is not the same signal. This is important to note, as the

subwoofer signal may contain LFE (with a +10dB offset) ***plus*** bass that is re-directed from the other main channels (i.e. if the speakers are small, and have little deep bass output). Also, some systems allow the user to direct LFE to the main front Left and Right speakers, if they have large woofers.

It should be noted that most metadata based downmix systems, as included in DTV set-top boxes and DTV integrated receivers do not include the LFE channel content as part of the 2 channel down-mix. This is consistent with most professional down-mix equipment. For this reason, downmixes from 5.1 to 5.0 surround or to stereo should not include any content from the LFE channel of the original mix.

15. Receiver downmix parameters

Mix coefficients at the receiver can be controlled by the broadcaster using metadata.

Feature	Two-Channel Decoder	Multichannel Decoder	Multichannel Adapter	Multichannel DVD Player	Comments
Line Mode	√	√	√	√	
Dialog Normalization	√	√	√	√	
Lt/Rt Downmix	√	√	√	√	
Lo/Ro Downmix	optional	optional	optional	optional	
Bass Management		√	√	√*	*Simplified design option
Dolby Surround Pro Logic		√	optional	optional	

Table 1: Consumer Decoder Product Features (from [11])

As Table 1 indicates, all receivers can perform Lt/Rt downmix. Not all receivers can perform LoRo downmix.

An important element of a down mix will be to set the metadata flags in an appropriate manner at the time of the original mix. Refer Free TV OP62

In addition, the “Pro-Logic” or “Not Indicated” status of each sound track should be considered. Pro-Logic flagging in a AC-3 bitstream will engage Pro-Logic decoding inside a consumer decoder, if it is left in its default or automatic setting.

Another consideration is that this flag should be set to ‘on’ when dealing with Dolby Surround Lt Rt programme material. However, it should also be noted a broadcaster may not have the ability or resources to assess all incoming material, and continually turn the flag to the appropriate mode. As a result, the broadcaster must make a decision to set the flag to one mode only, based on the overwhelming majority of surround Lt Rt material, or lack thereof. A broadcaster may also want to limit the consumers’ decoders

distractingly switching into and out of surround decoding modes, and stay with one mode continually.

16. Dynamic range management and the effect of downmixing on compression and limiting systems

Establishing the dynamic range of a sound track will be largely determined by the listening environment(s) the mix is intended to serve. As components of the original soundtrack mix the amount of compression and limiting applied will have an impact on the final mix.

In the original mix of a soundtrack the parameters for dynamic range, compression and limiting should take account of the sound quality of the 5.1 surround channel down mix to 2 channel stereo and mono.

Free TV OP48 Clause 5 provides advice on compression, audio limiting and spectral manipulation / equalisation of soundtracks for television advertising content broadcast in Australia. Free TV OP 36 Annex D also provides advice in relation to soundtracks for advertising content.

17. Headroom considerations

An often specified “headroom” for analogue audio sound was that the transient audio peaks should not be +8dB above 0VU. In a DTV broadcasting environment many broadcasters consider, where the reference is -20dBFS, true audio peaks[1] should be not above -2dBFS.

These parameters should be taken into consideration with respect to the multichannel headroom required when a future down-mix to stereo may occur. Summing of coherent channels may cause overload in the downmix.

18. The effect of downmixing on loudness measurement

Australian sound engineering tests have initially shown that the difference in loudness measurements of 5.1 and 2.0 on the same content is max ± 2LU, and in most cases much less.

19. References

1. BBC Summary of delivery formats
http://www.bbc.co.uk/guidelines/dq/pdf/tv/hd_summary_delivery_formats_v1_09.pdf
2. Recommendation ITU-R BS.775 Multichannel stereophonic sound system with and without accompanying picture
3. Recommendation ITU-R BS.1196-1 Audio coding for digital terrestrial television broadcasting

[1] True peak meters are oversampling peak meters as defined in [5].

4. Recommendation ITU-R BS.1423 Guidelines for producing multichannel soundtracks using surround matrix techniques.
5. Recommendation ITU-R BS.1770-1 Algorithms to measure audio programme loudness and true-peak audio level
6. Recommendation ITU-R BS.1864 Operational practices for loudness in the international exchange of digital television programmes
7. ATSC Recommended Practice A-85: Techniques for Establishing and Maintaining Audio Loudness for Digital Television
8. Dolby Surround Mixing Manual Part No. 91536 Issue 2
9. Dolby 5.1 Channel production Guidelines Issue 1 S00/12957
10. Dolby 5.1-Channel Music Production Guidelines Issue 3
11. Dolby Digital Professional Encoding Guidelines Issue 1 S00/12972