Annex 2 – TV receiver[1] performance testing – requirements

1. Purpose

The purpose of this proposed consultancy work is to investigate:

- TV receiver performance in single frequency network (SFN) reception environments, including possible 'wider area' SFN operations using DVB-T2
- the ability of TV receivers to operate with possible shared multiplex configurations.

Testing should be performed using a representative sample of the TV receivers currently available for purchase on the Australian market and where possible older devices that are currently in use by viewers but no longer available to purchase on the market. To the extent possible, the intent is to provide data on a representative sample of the deployed TV receiver fleet in Australia.

The project products must include reports regarding TV receiver performance (SFN performance analysis and shared multiplex performance analysis). The reports must also outline the receiver sampling methodology, the testing methodologies and receiver test results. A correlation analysis of the performance of the receivers with specific chip sets should also be included in the reports where this information is available.

2. Scope

a) TV receiver sample selection

- For the purpose of this consultancy, TV receivers include receivers embedded in TV units as well as stand-alone TV receivers, including those in set-top boxes.
- A sample of TV receivers that is a reasonable representation of the TV receivers currently available on the Australian market and currently owned and used by households should be selected for testing. Attention should be given to ensuring that this also includes a representative sample of chipsets used in TV receivers.
- When available, older TV receivers still in use by Australian households should be included in the sample.
- To determine a representative sample of TV receivers, the successful supplier may use their own records (if existent) or may need to liaise with retailers, TV manufacturers, the broadcasting industry, or any other entity they may consider a reliable source of information.
- The TV receiver sample selection methodology and the potential size of the sample should be detailed in the response to the ATM, including justification why the sample is representative.
- The sample should, to a practical extent, include both Freeview compliant and Freeview non-compliant receivers. The proportion of Freeview non-compliant

[1] TV receivers include receivers embedded in TV units as well as stand-alone TV receivers, including those in set-top boxes.

receivers within the sample should, in general, reflect the proportion in the general population.

> The report should provide a description and justification on the receiver sampling methodology and sample size used.

> The agreed list of 30 receiver models to be tested is as follows:

[Redacted]

> This sample may be adjusted if new information becomes available with regard to improving the sample as a representation of the receivers available in Australian market and/or in use in households. Any adjustment should be agreed by the ACMA.

b) Transport stream generation

> The generation of transport streams, particularly for the multiplex testing components of the project, is critical. The testing scenarios should, as much as practical, emulate real broadcast operation in terms of the program content.

> Transport streams will need to be evaluated by the ACMA to ensure they are representative of typical Australian broadcast programming and typical program scenarios across channels.

c) TV receiver testing

> TV receiver testing should be performed in an appropriate laboratory with appropriate testing equipment.

Part 1: Single Frequency Network (SFN) performance testing

Transmission parameters for SFN performance testing

DVB-T tests: DVB-T tests should be performed with the following technical parameters: 7 MHz TV channel, 8k OFDM, 64QAM, 3/4 forward error correction (FEC) and 1/16 guard interval which are the parameters currently in use for most terrestrial broadcasting services in Australia[2].

Ideally, whenever possible the thresholds for the above measurements for DVB-T should be determined on a target BER of 2×10^{-4} (post-Viterbi), measured between the inner and outer codes, before Reed-Solomon decoding. This should produce a quasi error-free (QEF) reception of BER<1×10^{-11} at the input of the MPEG decoder[3].

However, for domestic grade receivers such measurements may not be possible in which case the Subjective Failure Point (SFP) method should be performed as described in Annex 8 of Recommendation ITU-R BT.1368-13[4] (tests at one input level is sufficient. An input level of -60dBm is proposed).

[2] Free TV Australia Operational Practice OP-71, Page 13.
[3] As described in Recommendation ITU-R BT.1368-13.
[4] Available from the ITU website.

DVB-T2 tests:

DVB-T2 tests should be performed both in DVB-T2 @32Mbit/s Mode (B) and DVB-T2 @36Mbit/s Mode (D) (see Report ITU-R BT.2467[5]).

Ideally, whenever possible the thresholds for the above measurements for DVB-T2 should be determined on a target BER of 1×10^{-7} (post-LDPC), measured between the inner and outer codes, before BCH decoding. This should produce a quasi-error-free (QEF) reception of BER<1×10^{-11} at the input of the MPEG decoder[6].

However, for domestic grade receivers such measurements may not be possible in which case Subjective Failure Point (SFP) method should be performed as described in Annex 5 of Recommendation ITU-R BT.2033-2[7] (tests at one input level is sufficient. An input level of -60dBm is proposed).

SFN performance testing

The testing methodology should be detailed in the report.

The following tests are to be conducted for each sample receiver:

Mode DVB-T@ 23Mbit/s MPEG2/MPEG4 HD/SD

1. Threshold test (Emin)
2. Co-channel Protection Ratio test
3. SFN. 2 TX test. 0-dB (equal levels). TX 2 varying +/- 1.2x GI.
4. SFN 4 TX test. -6 dB Tx 1; 0 dB Tx 2(ref); -10 dB Tx 3; Tx 4 varying. Tx 1 0 µs (ref); Tx 2 +25% GI; Tx 3 +70% GI; Tx 4 varying +/- 1.2x GI.
    - Commence Tx4 at -1.2 GI and - 6dB relative to Tx2.
    - If SFP occurs reduce Tx4 amplitude until signal returns.
    - Repeat for steps of 0.05 GI to -1 GI
    - If and when no power reduction is required beyond the – 6dB level, use 0.2 GI steps from -1 GI to 1GI
    - From 1 GI to 1.2 GI use 0.05 GI steps reducing Tx4 amplitude when necessary to ensure signal is restored
    - For all timing steps – record the level of Tx4 when the signal is restored / maintained

Mode DVB-T2 @32Mbit/s Mode (B) MPEG4/HEVC HD/SD

1. Threshold test (Emin)
2. Lower adjacent channel Protection ratio test
3. Co-channel Protection Ratio test
4. Upper adjacent channel Protection Ratio test
5. SFN. 2 TX test. 0-dB (equal levels). TX 2 varying +/- 1.2x GI.
6. SFN. 3 TX test 1. 0 dB Tx1; -1dB for Tx2 & Tx3. Tx 1 0µs; Tx 2 1/2 GI; Tx 3 varying +/- 1.2x GI.
7. SFN. 3 TX test 2. 0 dB Tx1; -1dB for Tx2 & Tx3. Tx 1 0µs; Tx 2 90% GI; Tx 3 varying +/- 1.2x GI.
8. SFN. 3 TX test 3. 0 dB Tx 1; -6dB for Tx 2 and Tx3 . Tx 1 0µs; Tx 2 1/2 GI; Tx 3 varying +/- 1.2x GI.

[5] https://www.itu.int/pub/R-REP-BT.2467
[6] See Report ITU-R BT.2341-0 and ETSI TS 102 831 V1.2.1.
[7] Available from the ITU website.

9. SFN. 3 TX test 4. 0 dB Tx 1; -6dB for Tx 2 and Tx3. Tx 1 0µs; Tx 2 90% GI; Tx 3 varying +/- 1.2x GI.
10. SFN. 4 TX test. -6 dB Tx 1; 0 dB Tx 2; -10 dB Tx 3; Tx 4 varying. Tx 1 0 µs; Tx 2 +25% GI; Tx 3 +70% GI
    - Commence Tx4 at -1.2 GI and - 6dB relative to Tx2.
    - If SFP occurs reduce Tx4 amplitude until signal returns.
    - Repeat for steps of 0.05 GI to -1 GI
    - If and when no power reduction is required beyond the – 6dB level, use 0.2 GI steps from -1 GI to 1GI
    - From 1 GI to 1.2 GI use 0.05 GI steps reducing Tx4 amplitude when necessary to ensure signal is restored
    - For all timing steps – record the level of Tx4 when the signal is restored / maintained
11. Receiver synchronisation
    - Receiver power-on from standby mode after varying delay of one Tx
    - Disconnection and reconnection of input RF signal
12. Receiver echo timing v's echo level tolerance. 2 Tx test.
    - Set Tx1 to 0 dB (reference); Tx2 level to be varied to SFP (where possible) at each GI step setting
    - Commence Tx2 at -1.2 GI and 0 dB relative to Tx1.
    - If SFP occurs reduce Tx2 amplitude until signal returns.
    - Repeat for steps of 0.05 GI to -1 GI
    - If and when no power reduction is required beyond the 0 dB level, use 0.2 GI steps from -1 GI to 1GI
    - From 1 GI to 1.2 GI use 0.05 GI steps reducing Tx2 amplitude when necessary to ensure signal is restored
    - For all timing steps – record the level of Tx2 when the signal is restored / maintained

Mode DVB-T2 @36Mbit/s Mode (D) MPEG4/HEVC HD/SD

1. Threshold test (Emin)
2. Lower adjacent channel Protection ratio test
3. Co-channel Protection Ratio test
4. Upper adjacent channel Protection Ratio test
5. SFN. 2 TX test. 0-dB (equal levels). TX 2 varying +/- 1.2x GI.
6. SFN. 3 TX test 1. 0 dB Tx1; -1dB for Tx2 & Tx3. Tx 1 0µs; Tx 2 1/2 GI; Tx 3 varying +/- 1.2x GI.
7. SFN. 3 TX test 2. 0 dB Tx1; -1dB for Tx2 & Tx3. Tx 1 0µs; Tx 2 90% GI; Tx 3 varying +/- 1.2x GI.
8. SFN. 3 TX test 3. 0 dB Tx 1; -6dB for Tx 2 and Tx3 . Tx 1 0µs; Tx 2 1/2 GI; Tx 3 varying +/- 1.2x GI.
9. SFN. 3 TX test 4. 0 dB Tx 1; -6dB for Tx 2 and Tx3. Tx 1 0µs; Tx 2 90% GI; Tx 3 varying +/- 1.2x GI.
10. SFN. 4 TX test. -6 dB Tx 1; 0 dB Tx 2; -10 dB Tx 3; Tx 4 varying. Tx 1 0 µs; Tx 2 +25% GI; Tx 3 +70% GI.
    - Commence Tx4 at -1.2 GI and - 6dB relative to Tx2.
    - If SFP occurs reduce Tx4 amplitude until signal returns.
    - Repeat for steps of 0.05 GI to -1 GI
    - If and when no power reduction is required beyond the – 6dB level, use 0.2 GI steps from -1 GI to 1GI
    - From 1 GI to 1.2 GI use 0.05 GI steps reducing Tx4 amplitude when necessary to ensure signal is restored
    - For all timing steps – record the level of Tx4 when the signal is restored / maintained
11. Receiver synchronisation
    - Receiver power-on from standby mode after varying delay of one Tx

- Disconnection and reconnection of input RF signal

12. Receiver echo timing v's echo level tolerance. 2 Tx test.
    - Set Tx1 to 0 dB (reference); Tx2 level to be varied to SFP (where possible) at each GI step setting
    - Commence Tx2 at -1.2 GI and 0 dB relative to Tx1.
    - If SFP occurs reduce Tx2 amplitude until signal returns.
    - Repeat for steps of 0.05 GI to -1 GI
    - If and when no power reduction is required beyond the 0 dB level, use 0.2 GI steps from -1 GI to 1GI
    - From 1 GI to 1.2 GI use 0.05 GI steps reducing Tx2 amplitude when necessary to ensure signal is restored
    - For all timing steps – record the level of Tx2 when the signal is restored / maintained

Part 2: Multiplex sharing

For all tests required to be performed as per this scope, the testing scenarios should, as much as practical, emulate real broadcast operation in terms of the program content (particularly in prime time). Tests should be performed assuming that broadcasters would share multiplexes under the following scenarios:

a) 3 broadcasters share the entire capacity of one multiplex on equal basis (1/3, 1/3, 1/3)

b) 2 broadcasters share one multiplex on (2/3, 1/3) basis

c) 2 broadcasters share the entire capacity of one multiplex on equal basis (1/2, 1/2)

d) 1 broadcaster using the entire capacity of one multiplex (current arrangement) – test control

All receivers should first be tested using the non-shared content (d) above (i.e. the 'control') and the results should be compared with the shared scenarios, as identified above (a) to c)). The purpose is to test whether there are any inherent issues with the receivers which are not related to the multiplex sharing. All potentially identified issues should be reported in detail and under what circumstances they occur (e.g., any specific receivers, chipsets, Freeview non-compliant, etc).

For all tests, the content should be generated to maximise the number of streams that can be accommodated, at an acceptable picture quality, within a multiplex for the particular testing scenario. Acceptable picture quality would be determined by reference to full-reference quality metrics and quality of current TV services, as currently being implemented.

For all tests, Service Information and other overheads should be implemented in a way which would emulate current implementation by broadcasters. There have been suggestions that if a single broadcaster transmits their content over two multiplexes, some parts of SI need to be duplicated and transmitted over both multiplexes for the SI to be displayed correctly. The service information (or other overhead) duplication should only occur if necessary for normal display/operation.

The following tests should be performed regarding multiplex sharing:

1. *General operation under shared multiplexing scenarios*

   Tests should determine whether TV receivers would work properly under the shared multiplexing scenario, i.e., the entire content (all the streams) should display correctly with no degradation in quality. All potentially identified issues

should be reported in detail and under what circumstances they occur (e.g., any specific receivers, chipsets, Freeview non-compliant, etc).

2. *Service Information (SI) display testing*

   Testing should determine whether TV receivers have any issues with the Service Information (SI) and correct display of its components including Event Information Table (EIT), i.e., whether under any circumstances the tested receivers display incorrect SI (including not displaying at all). There have been suggestions that if a single broadcaster transmits their content over two multiplexes, some parts of SI need to be duplicated and transmitted over both multiplexes for the SI to be displayed correctly. The service information (or any other overhead) duplication should only occur if necessary for normal display/operation, and the observations of such duplications should be included in the corresponding report.

   The tests should make observations about part of the multiplex capacity required for providing the SI. The studies should also determine whether overall capacity requirement for SI would increase per 7 MHz RF channel due to sharing provisions. The overall SI capacity requirement should be compared with the capacity utilised as per the existing arrangements. All potentially identified issues should be reported in detail and under what circumstances they occur (e.g., any specific receivers, chipsets, Freeview non-compliant, etc).

3. *Logical Channel Number (LCN) display testing*

   Testing should determine whether TV receivers have any issues with the LCN display, i.e., whether under any circumstances the tested receivers display incorrect LCN (including not displaying at all). All potentially identified issues should be reported in detail and under what circumstances they occur (e.g., any specific receivers, chipsets, Freeview non-compliant, etc).

4. *Statistical multiplexing testing*

   The performance of the TV receivers should be tested under statistical multiplexing conditions. For this study, the multiplexing content should be similar to usual broadcasting content from different broadcasters, including a variety of contents such as high-capacity content (e.g., HD sporting events) and low-capacity content.

   The studies should test and compare statistical multiplexing performance for various video coding scenarios, i.e.:

   a) mix of MPEG-2 and MPEG-4 with SD and HD (similar to the existing arrangements),
   b) all MPEG-4 with SD and HD
   c) all MPEG-4 with HD only.
   d) mix of MPEG4 and HEVC with SD and HD (and HD+ or UHD) content (using DVB-T2 only).

   The above studies should take into account types of content normally provided by broadcasters, such as sports.

   The studies should be performed with both DVB-T and DVB-T2. Studies should make observations if there is any dependence of the statistical multiplexing gain on the number of streams within the multiplex, and any dependence of statistical multiplexing gain on formats (SD and HD) and/or video coding algorithms (MPEG-2, MPEG-4 and HEVC).

Since the number of HEVC capable receivers may be limited in the Australian market, a subset of the HEVC capable receivers within the receiver sample should be proportional to the general population.

Statistical multiplexing gain should also be observed in terms of so-called “sequestered” and “holistic” multiplexing. This gain should be observed relative to the case where no statistical multiplexing has been implemented.

Sequestered multiplex is where the multiplex is fragmented into segments each of fixed capacity and where the statistical multiplexing is separately applied to the video services within each of the segments. For instance, if two providers share one multiplex (1/2, 1/2), then one segment with one half of the capacity is allocated to each provider and the statistical multiplexing is then applied within each segment separately.

Holistic multiplexing is without any segmentation, i.e., all the streams within the multiplex are statistically managed in a holistic manner.

In the case of holistic multiplexing, testing should determine whether, and under what circumstances, some sharing entities would be disadvantaged in terms of being allocated a smaller than equal portion of the entire multiplex capacity. The testing should be conducted over a period of time to determine the effect of the holistic statistical multiplexing on both instantaneous and average over time individual shares of capacity allocated to each broadcaster, combining a variety of different content streams that should be similar to the real content provided by terrestrial broadcasting transmissions in Australia.

Observations should be made about the overall capacity, i.e., if there are any potential loss in terms of how much content can be accommodated and its quality. The results should be compared with the existing arrangements. If losses are identified, it should be specified whether they are in terms of the number of streams that can be accommodated, the picture quality (i.e., the available bit rate of individual streams is reduced), or both. Various options (combinations) should be explored in terms of number of streams/bit rate to gain a better understanding about how potential overall reduction in capacity would affect the losses and to what extent.

The following 13 tests are to be conducted for the scenarios and combinations listed in Table 1:

1. Receiver rescan/service discoverability
2. Service navigation
3. SI response including codec and service identification accuracy etc
4. EIT behaviour incl. EIT (other)
5. EIT display
6. LCN behaviour
7. Channel/mux change response
8. QEF video decoding
9. QEF audio decoding
10. Notable anomalies
11. PSNR
12. SSIM
13. VMAF

Table 1: Multiplex Sharing Tests

Sharing Scenario		RF Modulation / Bitrate	Codec/s scenario	Video Resolution/s	Statistical mux design
Sharing scenario (D)	No sharing	DVB-T @23Mbit/s	MPEG2/MPEG4	HD/SD	Holistic
		DVB-T2 @36Mbit/s	MPEG4/HEVC	HD/SD	Holistic
Sharing scenario (A)	3x equal sharers	DVB-T @23Mbit/s	Codec scenario (B) MPEG4	HD/SD	Sequestered
					Holistic
		DVB-T2 @36Mbit/s	Codec scenario (B) MPEG4	HD/SD	Sequestered
					Holistic
		DVB-T2 @36Mbit/s	Codec scenario (C) MPEG4	HD only	Holistic
		DVB-T2 @36Mbit/s	Codec scenario (D) HEVC/MPEG4	UHD/HD/SD or HD+/HD/SD	Holistic
Sharing scenario (B)	2x broadcasters 2/3, 1/3	DVB-T @23Mbit/s	Codec scenario (B) MPEG4	HD/SD	Sequestered
					Holistic
		DVB-T2 @36Mbit/s	Codec scenario (B) MPEG4	HD/SD	Sequestered
					Holistic
		DVB-T2 @36Mbit/s	Codec scenario (C) MPEG4	HD only	Holistic
		DVB-T2 @36Mbit/s	Codec scenario (D) HEVC/MPEG4	UHD/HD/SD or HD+/HD/SD	Holistic
Sharing scenario (C)	2x broadcasters 1/2, 1/2	DVB-T @23Mbit/s	Codec scenario (B) MPEG4	HD/SD	Holistic
		DVB-T2 @36Mbit/s	Codec scenario (B) MPEG4	HD/SD	Sequestered
					Holistic
		DVB-T2 @36Mbit/s	Codec scenario (C) MPEG4	HD only	Holistic
		DVB-T2 @36Mbit/s	Codec scenario (D) HEVC/MPEG4	UHD/HD/SD or HD+/HD/SD	Holistic

3. Testing results presentation and observations/statistics

Detailed results of all the tests specified in this document should be presented in Test Reports (including raw data), one for SFN performance testing and one for multiplex sharing testing or a combined final report. The reports should provide detailed receiver sampling methodology, receiver sample size methodology, detailed methodology for each of the tests performed, the equipment used, and the results of the tests. The reports should also provide other relevant observations not covered in the above requirements. Detailed description of the equipment used, and the test setup should also be included in the reports.

For all testing scenarios, detailed information should be provided about the transmission technology used (DVB-T or DVB-T2), data/frame structure, information about SI and other overhead (payload), zero padding, number of streams used in a

test, bit rate per stream, type of content (sports, movie, etc), format for each stream (SD, HD, UHD or HD+), video coding for each stream (MPEG-2, MPEG-4, HEVC), type of statistical multiplexing used, the duration of each individual test.

Reports should also include observations about any other limitations in the receivers that may affect the overall service capacity, i.e., any potential loss in terms of how much content can be accommodated and its quality. The results should be compared with the existing arrangements. If losses are identified, it should be specified whether they are in terms of the number of streams that can be accommodated, the picture quality (i.e. the available bit rate of individual streams is reduced), or both. Various options (combinations) should be explored in terms of number of streams/bit rate, to gain a better understanding about how potential overall reduction in capacity would affect the losses and to what extent.