INTERNATIONAL STANDARD

ISO/IEC 23003-3

First edition
2012-04-01

AMENDMENT 1
2014-03-15

Information technology — MPEG audio technologies —

Part 3: Unified speech and audio coding

AMENDMENT 1: Conformance

Technologies de l'information — Technologies audio MPEG —
Partie 3: Discours unifié et codage audio
AMENDEMENT 1: Conformité

Reference number
ISO/IEC 23003-3:2012/Amd.1:2014(E)

ISO copyright office
Case postale 56 • CH-1211 Geneva 20
Tel. + 41 22 749 01 11
Fax + 41 22 749 09 47
E-mail copyright@iso.org
Web www.iso.org

Published in Switzerland

Foreword

ISO (the International Organization for Standardization) and IEC (the International Electrotechnical Commission) form the specialized system for worldwide standardization. National bodies that are members of ISO or IEC participate in the development of International Standards through technical committees established by the respective organization to deal with particular fields of technical activity. ISO and IEC technical committees collaborate in fields of mutual interest. Other international organizations, governmental and non-governmental, in liaison with ISO and IEC, also take part in the work. In the field of information technology, ISO and IEC have established a joint technical committee, ISO/IEC JTC 1.

International Standards are drafted in accordance with the rules given in the ISO/IEC Directives, Part 2.

The main task of the joint technical committee is to prepare International Standards. Draft International Standards adopted by the joint technical committee are circulated to national bodies for voting. Publication as an International Standard requires approval by at least 75 % of the national bodies casting a vote.

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO and IEC shall not be held responsible for identifying any or all such patent rights.

Amendment 1 to ISO/IEC 23003-3:2012 was prepared by Joint Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 29, *Coding of audio, picture, multimedia and hypermedia information*.

Information technology — MPEG audio technologies —

Part 3:
Unified speech and audio coding

AMENDMENT 1: Conformance

In Clause 2, "Normative References", add the following entry:

ISO/IEC 14496-26:2010, *Information technology — Coding of audio-visual objects — Part 26: Audio conformance*

In 4.5.4 replace:

Four different hierarchical levels are defined with increasing number of audio channels and increasing complexity. All four levels include Level 2 of the Baseline USAC profile. The definition of the four levels of the Extended HE AAC profile is given in Table 3. All notes in Table 3 and all restrictions listed in the columns 2, 3, 4, and 5 ("Max. channels/object", "Max. AAC sampling rate, SBR not present [kHz]", "Max. AAC sampling rate, SBR present [kHz]", "Max. SBR sampling rate [kHz] (in/out)") of Table 3 apply only when decoding HE AAC v2 profile compliant bit streams.

Table 3 — Levels for the Extended HE AAC profile

Level (NOTE 1)	Max. channels / object	Max. AAC sampling rate, SBR not present [kHz]	Max. AAC sampling rate, SBR present [kHz]	Max. SBR sampling rate [kHz] (in/out)	Max. PCU	Max. RCU	Max. PCU HQ / LP SBR (NOTE 5)	Max. RCU HQ / LP SBR (NOTE 5)
1	NA	NA	NA	NA	NA	NA	NA	NA
2	2	48	24	24/48	12	11	12	11
3	2	48	24/48 (NOTE 3)	48/48 (NOTE 2)	15	11	15	11
4	5	48	24/48 (NOTE 4)	48/48 (NOTE 2)	25	28	20	23
5	5	96	48	48/96	49	28	39	23

NOTE 1: Level 2, 3, and 4 Extended HE AAC profile decoders implement the baseline version of the parametric stereo tool. A level 5 decoder shall not be limited to the baseline version of the parametric stereo tool.

NOTE 2: For level 3 and level 4 decoders, it is mandatory to operate the SBR tool in downsampled mode if the sampling rate of the AAC core is higher than 24kHz. Hence, if the SBR tool operates on a 48kHz signal, the internal sampling rate of the SBR tool will be 96kHz, however, the output signal will be downsampled by the SBR tool to 48kHz.

NOTE 3: If Parametric Stereo data are present the maximum AAC sampling rate is 24kHz, if Parametric Stereo data are not present the maximum AAC sampling rate is 48kHz.

NOTE 4: For one or two channels the maximum AAC sampling rate, with SBR present, is 48kHz. For more than two channels the maximum AAC sampling rate, with SBR present, is 24kHz.

NOTE 5: The PCU/RCU number are given for a decoder operating the LP SBR tool whenever applicable.

with:

A number of hierarchical levels are defined with increasing number of audio channels and increasing complexity. All levels include Level 2 of the Baseline USAC profile. The definition of the levels of the Extended HE AAC profile is given in Table 3. All notes in Table 3 and all restrictions listed in the columns 2, 3, 4, and 5 ("Max. channels/object", "Max. AAC sampling rate, SBR not present [kHz]", "Max. AAC

sampling rate, SBR present [kHz]", "Max. SBR sampling rate [kHz] (in/out)") of Table 3 apply only when decoding HE AAC v2 profile compliant bit streams.

Table 3 — Levels for the Extended HE AAC profile

Level (NOTE 1)	Max. channels / object	Max. AAC sampling rate, SBR not present [kHz]	Max. AAC sampling rate, SBR present [kHz]	Max. SBR sampling rate [kHz] (in/out)	Max. PCU	Max. RCU	Max. PCU HQ / LP SBR (NOTE 5)	Max. RCU HQ / LP SBR (NOTE 5)
1	NA	NA	NA	NA	NA	NA	NA	NA
2	2	48	24	24/48	12	11	12	11
3	2	48	24/48 (NOTE 3)	48/48 (NOTE 2)	15	11	15	11
4	5	48	24/48 (NOTE 4)	48/48 (NOTE 2)	25	28	20	23
5	5	96	48	48/96	49	28	39	23
6	7	48	24/48 (NOTE 4)	48/48	34	37	27	30
7	7	96	48	48/96	67	37	53	30

NOTE 1: Level 2, 3, 4, 6 and 7 Extended HE AAC profile decoders implement the baseline version of the parametric stereo tool. A level 5 decoder shall not be limited to the baseline version of the parametric stereo tool.

NOTE 2: For level 3, 4 and 6 decoders, it is mandatory to operate the SBR tool in downsampled mode if the sampling rate of the AAC core is higher than 24kHz. Hence, if the SBR tool operates on a 48kHz signal, the internal sampling rate of the SBR tool will be 96kHz, however, the output signal will be downsampled by the SBR tool to 48kHz.

NOTE 3: If Parametric Stereo data are present the maximum AAC sampling rate is 24kHz, if Parametric Stereo data are not present the maximum AAC sampling rate is 48kHz.

NOTE 4: For one or two channels the maximum AAC sampling rate, with SBR present, is 48kHz. For more than two channels the maximum AAC sampling rate, with SBR present, is 24kHz.

NOTE 5: The PCU/RCU number are given for a decoder operating the LP SBR tool whenever applicable.

NOTE 6: A Level 6 or 7 decoder is not required to decode a Level 5 stream.

In 5.3.2 amend Table 36 as follows:

Table 36 — Syntax of acelp_coding()

Syntax	No. of bits	Mnemonic
acelp_coding(acelp_core_mode)		
{		
[...]		
switch (acelp_core_mode) {		
case 0		
icb_index[sfr];	**20**	**uimsbf**
break;		
case 1		
icb_index[sfr];	**28**	**uimsbf**
break;		
case 2		
icb_index[sfr];	**36**	**uimsbf**
break;		
case 3		
icb_index[sfr];	**44**	**uimsbf**
break;		
case 4		
icb_index[sfr];	**52**	**uimsbf**
break;		
case 5		
icb_index[sfr];	**64**	**uimsbf**
break;		
case 6		
icb_index[sfr];	**12**	**uimsbf**
break;		
case 7		
icb_index[sfr];	**16**	**uimsbf**
break;		
}		
gains[sfr];	**7**	**uimsbf**
}		
NOTE: coreCoderFrameLength designates the core frame length in samples and is equal to either 1024 or 768. See also 6.1.1.2.		

In 7.14.5.2.1, replace:

Depending on the coding mode, the following codebooks are used:

with:

Depending on the coding mode, the following codebooks are used:

— 12-bit codebook with 2 pulses i_0 and i_1. Pulse i_0 can be selected from either track 0 or 2, pulse i_1 can be selected from either track 1 or 3 (5×2+2)

— 16-bit codebook with 3 pulses on three tracks. One pulse on track 0, one pulse on track 2 and one pulse on either track 1 or 3 (selected track signalled by a 1 bit field), which amounts to (5×3+1) = 16 bits.

Add a new Clause 8, "Conformance testing", as shown below:

8 Conformance testing

8.1 Introduction

The present Clause 8 specifies conformance criteria for both bitstreams and decoders compliant with the USAC standard as defined in this document. This is done to assist implementers and to ensure interoperability.

8.2 Terms and definitions

bitstream
encoded audio data

conformance data
conformance test sequences and conformance tools

conformance tool
tool to check certain conformance criteria

conformance test sequence
generic term for conformance test bitstreams and corresponding reference waveforms

conformance test bitstream
USAC bitstream used for testing the conformance of a USAC decoder

conformance test condition
condition which applies to properties of a conformance test bitstream in order to test a certain functionality of the USAC decoder

conformance test case
combination of one or more conformance test conditions for which a set of conformance test sequences is provided

main audio channel
audio channel conveyed by means of a UsacSingleChannelElement or UsacChannelPairElement

reference waveform
decoded counterpart of a bitstream

USAC bitstream
data encoded according to the USAC standard

UsacCPE
UsacChannelPairElement

UsacEXT
UsacExtElement

UsacLFE
UsacLfeElement

UsacSCE
UsacSingleChannelElement

8.3 USAC conformance testing

8.3.1 Profiles

Profiles are defined in 4.5. Some conformance criteria apply to USAC in general, while others are specific to certain profiles and their respective levels. Conformance shall be tested for the level of the profile with which a given bitstream or decoder claims to comply.

In addition to the conformance requirements described in this clause, a decoder which claims to comply with the Extended HE AAC Profile shall fulfill conformance for the HE AAC v2 profile according to ISO/IEC 14496-26:2010.

8.3.2 Conformance tools and test procedure

To test USAC compliant audio decoders, ISO/IEC JTC 1/SC 29/WG 11 supplies a number of conformance test sequences. Supplied sequences cover all profiles as defined in 4.5. For a supplied test sequence, testing can be done by comparing the output of a decoder under test with a reference waveform also supplied by ISO/IEC JTC 1/SC 29/WG 11. In cases where the decoder under test is followed by additional operations (e.g. quantizing a signal to a 16 bit output signal) the conformance point is prior to such additional operations, i.e. it is permitted to use the actual decoder output (e.g. with more than 16 bit) for conformance testing.

Measurements are carried out relative to full scale where the output signals of the decoders are normalized to be in the range between –1.0 and +1.0.

In ISO/IEC 14496-26:2010 a set of test methods is defined to test the output of the decoder under test against the reference output. RMS/LSB Measurement, Segmental SNR and PNS conformance criteria are used for the comparison. A particular test method for a certain test sequence is specified in 8.5.

For elements producing output that cannot be tested with the methods described in ISO/IEC 14496-26:2010, specific conformance testing procedures are described in 8.5.

8.3.2.1 Conformance data

All test sequences are provided in the shape of a zip archive as an electronic attachment. Furthermore, an MS Excel worksheet (“Usac_Conformance_Tables.xlsx”) is provided as an electronic attachment that lists all test sequences for each module.

For all conformance test sequences, the file names are composed of several parts which convey information about:

— which module of the decoder is tested

— which channelConfigurationIndex is employed

— which test conditions apply to the test sequence

— which coreSbrFrameLengthIndex applies to the test sequence

— which sampling frequency is signalled in the test sequence

The file naming convention given in Table 149 is used. Values in box brackets are optional.

Table 149 — File name conventions

Module	File Name (compressed)	File Name (uncompressed)
Frequency domain coding (FD mode), 8.4.4	Fd_<cCI>_c<cSFLI>_<testCase>_<uSFI>.mp4	FD_<cCI>_c<cSFLI>_<testCase>_<uSFI>.wav
Linear predictive domain coding (LPD mode), 8.4.5	Lpd_<cCI>_c<cSFLI>_<testCase>_[<uSFI>].mp4	Lpd_<cCI>_c<cSFLI>_<testCase>_[<uSFI>].wav
Common core coding tools, 8.4.6	Cct_<cCI>_c<cSFLI>_<testCase>_[<uSFI>].mp4	Cct_<cCI>_c<cSFLI>_<testCase>_[<uSFI>].wav
Enhanced spectral band replication (eSBR), 8.4.7	eSbr_<cSFLI>_<testCase>.mp4	eSbr_<cSFLI>_<testCase>.wav
MPEG Surround 2-1-2, 8.4.10	Mps_<bsFR>_Sc<sCI>_<testCase>.mp4	Mps_<bsFR>_Sc<sCI>_<testCase>.wav

<cCI> channelConfigurationIndex as described in Table 68.

<testCase> Setup string. May consist of a concatenation of one or more abbreviations as listed in Table 150. If no setup string is specified the basic test conditions apply

<cSFLI> coreSbrFrameLengthIndex as described in Table 70.

<uSFI> usacSamplingFrequencyIndex as described in Table 67. If the escape value is specified the used sampling frequency is appended, e.g. “xx_1f_42000.mp4” for a sampling frequency of 42 kHz.

<bsFR> bsFreqRes as described in ISO/IEC 23003-1:2007, Table 39

<sCI> stereoConfigIndex as described in Table 72

Table 150 — Test conditions and abbreviations

FD core mode	
Test Condition	**Abbrev.**
FD window switching test condition	Win
Noise filling test condition	Nf
Tns test condition	Tns
Varying max_sfb test condition	Sfb
Handling of extensions condition	Ex
Arithmetic coder test condition	Ac
Non-meaningful FD window switching test condtion	Nmf
M/S stereo test condition	Ms
Complex prediction stereo test condition	Cp

LPD core mode	
Test Condition	**Abbrev.**
LPC coding test condition	Lpc
ACELP core mode test condition	Ace
TCX and noise filling test condition	Tcx
LPD mode coverage and FAC test condition	Lpd
Bass-post filter test condition	Bpf
AVQ test condition	Avq

Combined core coding	
Test Condition	**Abbrev.**
FD-LPD transition and FAC test condition	Flt
FD/TCX noise filling test condition	Cnf
Bass-post filter test condition	Cbf
synchr. FD-LPD transition and FAC test condition	Flts
asynchr. FD-LPD transition and FAC test condition	Flta
Arithmetic coder test condition	CAc

eSbr	
Test Condition	**Abbrev.**
QMF accuracy test condition	Qma
Envelope adjuster accuracy and SBR pre-processing test condition	Eaa
Header and grid control test condition test condition	Hgt
Inverse filtering test condition	Ift
Additional sine test (missing harmonics) test condition	Ast
Sampling rate test condition	Sr
Channel mode test condition	Cm
interTes test condition	Tes
PVC test condition	Pvc
Harmonic transposition (QMF) test condition	Htq
Harmonic transposition (crossproducts) test condition	Xp
Transposer toggle test condition	Ttt
Envelope shaping toggle (PVC on/off) test condition	Est
Varying crossover frequency test condition	Xo
stereoConfigIndex test condition	Mps

Mpeg surround 212	
Test Condition	**Abbrev.**
TSD test condition	Tsd
Rate mode test condition	Rm
Phase coding test condition	Pc
Decorrelator configuration. test condition	Dc
DMX gain test condition	Dm
Bands phase test condition	Bp<X>
Pseudo lr test condition	Plr
Residual bands test condition	Rb<X>

8.4 USAC Bitstreams

8.4.1 General

8.4.1.1 Characteristics

Characteristics of bitstreams specify the constraints that are applied by the encoder in generating the bitstreams. These syntactic and semantic constraints may, for example, restrict the range or the values of parameters that are encoded directly or indirectly in the bitstreams. The constraints applied to a given bitstreams may or may not be known a priori.

8.4.1.2 Test procedure

Each USAC bitstream shall meet the syntactic and semantic requirements specified in this document. The present subclause defines the conformance criteria that shall be fulfilled by a compliant bitstream. These criteria are specified for the syntactic elements of the bitstream and for some parameters decoded from the USAC bitstream payload.

For each tool a set of semantic tests to be performed on the bitstreams is described. To verify whether the syntax is correct is straightforward and therefore not defined herein after. In the description of the semantic tests it is assumed that the tested bitstreams contains no errors due to transmission or other causes. For each test the condition or conditions that must be satisfied are given, as well as the prerequisites or conditions in which the test can be applied.

8.4.2 USAC Configuration

8.4.2.1 Characteristics

Encoders may apply restrictions to the following parameters of the bitstream:

a) usacSamplingFrequencyIndex

b) usacSamplingFrequency

c) coreSbrFrameLengthIndex

d) channelConfigurationIndex

e) presence of configuration extensions

f) numOutChannels

g) bsOutputChannelPos

h) numElements

i) stereoConfigIndex

j) use of time warped MDCT

k) use of noise filling in FD mode

l) use of the eSBR harmonic transposer

m) use of the eSBR inter-TES tool

n) use of the eSBR PVC tool

o) SBR default header, for details see 8.4.7.

p) MPS config, for details see 8.4.10.

8.4.2.2 Test procedure

8.4.2.2.1 UsacConfig()

usacSamplingFrequencyIndex Shall be encoded with a non-reserved value specified in Table 67. For further profile and level dependent restrictions see 8.4.11.

usacSamplingFrequency No restrictions apply. For profile and level dependent restrictions see 8.4.11.

coreSbrFrameLengthIndex no restrictions apply

channelConfigurationIndex Shall be encoded with a non-reserved value specified in Table 68. For further profile and level dependent restrictions see 8.4.11. In the case of channelConfigurationIndex==0 further restrictions apply as described in 8.4.2.2.2.

usacConfigExtensionPresent no restrictions apply

8.4.2.2.2 UsacChannelConfig()

numOutChannels no restrictions apply. For profile and level dependent restrictions see 8.4.11.

bsOutputChannelPos A bsOutputChannelPos of value 3 or 26 (LFE speaker positions) shall be associated with an LFE channel. Any other value shall be associated with a main audio channel.

8.4.2.2.3 UsacDecoderConfig()

numElements the value of this data element shall be such that the accumulated sum of all channels contained in the bitstream complies with the restrictions outlined in 8.4.2.2.1.

usacElementType no restrictions apply. For profile and level dependent restrictions see 8.4.11.

8.4.2.2.4 UsacSingleChannelElementConfig()

No restrictions are applicable to this bitstream element.

8.4.2.2.5 UsacChannelPairElementConfig()

NOTE The UsacChannelPairElementConfig() element and all included elements may only be present when coding more than one output channel (see restrictions applying to UsacConfig() in 8.4.2.2.1).

stereoConfigIndex no restrictions apply

8.4.2.2.6 UsacLfeElementConfig()

No restrictions are applicable to this bitstream element.

8.4.2.2.7 UsacCoreConfig()

tw_mdct no restrictions apply. For profile and level dependent restrictions see 8.4.11.

noiseFilling no restrictions apply

8.4.2.2.8 SbrConfig()

harmonicSBR no restrictions apply

bs_interTes no restrictions apply

bs_pvc no restrictions apply

8.4.2.2.9 SbrDfltHeader()

dflt_start_freq no restrictions apply

dflt_stop_freq no restrictions apply

dflt_header_extra1 no restrictions apply

dflt_header_extra2 no restrictions apply

dflt_freq_scale no restrictions apply

dflt_alter_scale no restrictions apply

dftl_nose_bands no restrictions apply

dflt_limiter_bands no restrictions apply

dflt_limiter_gains no restrictions apply

dflt_interpol_freq no restrictions apply

dflt_smoothing_mode no restrictions apply

8.4.2.2.10 Mps212Config()

bsFreqRes	shall not be encoded with a value of 0
bsFixedGainDMX	no restrictions apply
bsTempShapeConfig	no restrictions apply
bsDecorrConfig	shall not be encoded with a value of 3
bsHighRateMode	no restrictions apply
bsPhaseCoding	no restrictions apply
bsOttBandsPhasePresent	no restrictions apply
bsOttBandsPhase	shall not be encoded with a value larger than the value of numBands as given by ISO/IEC 23003-1:2007, 5.2, Table 39 and depends on bsFreqRes.
bsResidualBands	shall not be encoded with a value larger than the value of numBands as given by ISO/IEC 23003-1:2007, 5.2, Table 39 and depends on bsFreqRes.
bsPseudoLr	no restrictions apply
bsEnvQuantMode	shall be 0

8.4.2.2.11 UsacExtElementConfig()

usacExtElementType	no restrictions apply
usacExtElementConfigLength	no restrictions apply
usacExtElementDefaultLengthPresent	no restrictions apply
usacExtElementDefaultLength	no restrictions apply
usacExtElementPayloadFrag	no restrictions apply

8.4.2.2.12 UsacConfigExtension()

numConfigExtensions	no restrictions apply
usacConfigExtType[]	no restrictions apply
usacConfigExtLength[]	no restrictions apply
fill_byte	should be '10100101'

8.4.3 Framework

8.4.3.1 Characteristics

Encoders may apply restrictions to the following parameters of the bitstream:

a) signalling of independently decodable frames

b) presence of extension elements

c) core_mode

d) presence of TNS

8.4.3.2 Test procedure

8.4.3.2.1 UsacFrame()

usacIndependencyFlag no restrictions apply

8.4.3.2.2 UsacSingleChannelElement

No restrictions are applicable to this bitstream element.

8.4.3.2.3 UsacChannelPairElement

No restrictions are applicable to this bitstream element.

8.4.3.2.4 UsacLfeElement

No restrictions are applicable to this bitstream element.

8.4.3.2.5 UsacExtElement

usacExtElementPresent no restrictions apply

usacExtElementUseDefaultLength no restrictions apply

usacExtElementPayloadLength no restrictions apply

usacExtElementStart no restrictions apply

usacExtElementStop no restrictions apply

usacExtElementSegmentData no restrictions apply

8.4.3.2.6 UsacCoreCoderData

core_mode no restrictions apply.

tns_data_present no restrictions apply

8.4.4 Frequency domain coding (FD mode)

8.4.4.1 Characteristics

Encoders may apply restrictions to the following parameters of the bitstream:

a) use of noise filling

b) window_shape

c) M/S Stereo

d) use of TNS

e) Complex prediction stereo coding

f) max_sfb

g) use of time warped MDCT

h) use of long blocks

i) use of short blocks

8.4.4.2 Test procedure

8.4.4.2.1 fd_channel_stream

global_gain no restrictions apply.

noise_level no restrictions apply

noise_offset no restrictions apply

fac_data_present shall be 0, if the core_mode of the preceding frame of the same channel was 0 or if mod[3] of the preceding frame of the same channel was > 0.

8.4.4.2.2 ics_info

window_sequence A conformant bitstream shall consist of only meaningful window_sequence transitions. However, decoders are required to handle non-meaningful window_sequence transitions as well. The meaningful window_sequence transitions are shown in Table 133.

window_shape no restrictions apply

max_sfb shall be <= num_swb_long or num_swb_short as appropriate for window_sequence and sampling frequency and core coder frame length.

scale_factor_grouping no restrictions apply

8.4.4.2.3 tw_data

tw_data_present no restrictions apply

tw_ratio no restrictions apply

8.4.4.2.4 scale_factor_data

hcod_sf Shall only be encoded with the values listed in the scalefactor Huffman table. Shall be encoded such that the decoded scalefactors sf[g][sfb] are within the range of zero to 255, both inclusive.

8.4.4.2.5 tns_data

n_filt	no restrictions apply
coef_res	no restrictions apply
length	shall be small enough such that the lower bound of the filtered region, does not exceed the start of the array containing the spectral coefficients.
order	shall not exceed the values listed in in Table 130.
direction	no restrictions apply
coef_compress	no restrictions apply
coef	no restrictions apply

8.4.4.2.6 ac_spectral_data

arith_reset_flag	no restrictions apply

8.4.4.2.7 StereoCoreToolInfo

tns_active	no restrictions apply
common_window	no restrictions apply
common_max_sfb	no restrictions apply
max_sfb1	shall be <= num_swb_long or num_swb_short as appropriate for window_sequence and sampling frequency and core coder frame length.
ms_mask_present	no restrictions apply
ms_used	no restrictions apply
common_tw	no restrictions apply
common_tns	no restrictions apply
tns_on_lr	no restrictions apply
tns_present_both	no restrictions apply
tns_data_present	no restrictions apply

8.4.4.2.8 cplx_pred_data

cplx_pred_all	no restrictions apply
cplx_pred_used	no restrictions apply
pred_dir	no restrictions apply
complex_coef	no restrictions apply
use_prev_frame	shall be 0 if the core transform length of previous frame is different from the core transform length of the current frame or if the core_mode of the previous frame is 1.
delta_code_time	no restrictions apply
hcod_sf	no restrictions apply

8.4.5 Linear predictive domain coding (LPD mode)

8.4.5.1 Characteristics

Encoders may apply restrictions to the following parameters of the bitstream:

a) acelp_core_mode

b) lpd_mode (use of ACELP, short TCX, medium TCX, and long TCX)

c) activation of bass-post filter

8.4.5.2 Test procedure

8.4.5.2.1 lpd_channel_stream

acelp_core_mode	shall be encoded with a value in the range of 0 to 5, both inclusive.
lpd_mode	shall be encoded with a non-reserved value listed in Table 89.
bpf_control_info	no restrictions apply
core_mode_last	shall be encoded with the value of data element core_mode of the previous frame
fac_data_present	shall be 0, if the core_mode of the preceding frame of the same channel was 0 and mod[0] of the current frame is > 0, or if mod[0] of the current frame is > 0 and mod[3] of the preceding frame of the same channel was > 0.
short_fac_flag	shall be encoded with a value of 1 if the window_sequence of the previous frame was 2 (EIGHT_SHORT_SEQUENCE). Otherwise short_fac_flag shall be encoded with a value of 0.

8.4.5.2.2 lpc_data

lpc_first_approximation_index no restrictions apply

8.4.5.2.3 qn_data

qn	the codebook number shall be encoded as described in 7.13.7.2.
qn_base	no restrictions apply
qn_ext	no restrictions apply

8.4.5.2.4 get_mode_lpc

binary_code	shall be encoded with the values listed in Table 143 in the column Binary Code

8.4.5.2.5 code_book_indices

code_book_index	no restrictions apply
kv	no restrictions apply

8.4.5.2.6 acelp_coding

mean_energy	no restrictions apply
acb_index	the adaptive codebook index shall be encoded as described in 7.14.5.1.
ltp_filtering_flag	no restrictions apply
icb_index	the innovation codebook excitation shall be encoded as described in 7.14.5.2.
gains	no restrictions apply

8.4.5.2.7 tcx_coding

noise_factor	no restrictions apply
global_gain	no restrictions apply
arith_reset_flag	no restrictions apply

8.4.6 Common core coding tools

8.4.6.1 Characteristics

Encoders may apply restrictions to the following parameters of the bitstream:

a) use of arithmetic coder reset

8.4.6.2 Test procedure

8.4.6.2.1 arith_data

acod_m	shall be encoded as described in 7.4.3
acod_r	shall be encoded as described in 7.4.3
s	no restrictions apply

8.4.6.2.2 fac_data

fac_gain	no restrictions apply

8.4.7 Enhanced spectral band replication (eSBR)

8.4.7.1 Characteristics

Encoders may apply restrictions to the following parameters of the bitstream:

a) use of the eSBR harmonic transposer

b) use of Crossproducts in eSBR harmonic transposer

c) use of the eSBR inter-TES tool

d) choice of SBR ratio

e) choice of amplitude resolution

f) choice of SBR crossover band

g) use of SBR preprocessing (prewhitening)

h) use of the eSBR PVC tool

8.4.7.2 Test procedure

The present subclause defines the conformance criteria that shall be fulfilled by a compliant bitstream that utilize the Enhanced SBR tool.

8.4.7.2.1 UsacSbrData

sbrInfoPresent	no restrictions apply
sbrHeaderPresent	no restrictions apply
sbrUseDfltHeader	no restrictions apply

8.4.7.2.2 SbrInfo

bs_amp_res	no restrictions apply
bs_xover_band	shall define a value that does not exceed the limits defined in ISO/IEC 14496-3:2009, 4.6.18.3.6.
bs_sbr_preprocessing	no restrictions apply
bs_pvc_mode	shall be encoded with a non-reserved value specified in Table 96

8.4.7.2.3 SbrHeader

bs_start_freq	shall define a frequency band that does not exceed the limits defined in 7.5.5 and ISO/IEC 14496-3:2009, 4.6.18.3.6.
bs_stop_freq	shall define a frequency band that does not exceed the limits defined in 7.5.5 and ISO/IEC 14496-3:2009, 4.6.18.3.6.
bs_header_extra1	no restrictions apply
bs_header_extra2	no restrictions apply
bs_freq_scale	no restrictions apply
bs_alter_scale	no restrictions apply
bs_noise_bands	shall define a value that does not exceed the limits defined in ISO/IEC 14496-3:2009, 4.6.18.3.6.
bs_limiter_bands	no restrictions apply
bs_limiter_gains	no restrictions apply
bs_interpol_freq	no restrictions apply
bs_smoothing_mode	no restrictions apply

8.4.7.2.4 sbr_single_channel_element

sbrPatchingMode	no restrictions apply
sbrOversamplingFlag	no restrictions apply
sbrPitchInBinsFlag	no restrictions apply
sbrPitchInBins	no restrictions apply
bs_add_harmonic_flag	no restrictions apply

8.4.7.2.5 sbr_channel_pair_element

bs_coupling	no restrictions apply
sbrPatchingMode	no restrictions apply
sbrOversamplingFlag	no restrictions apply
sbrPitchInBinsFlag	no restrictions apply
sbrPitchInBins	no restrictions apply
bs_add_harmonic_flag	no restrictions apply

8.4.7.2.6 sbr_grid

bs_frame_class	shall define a value that does not exceed the limits defined in 7.5.1.3 and ISO/IEC 14496-3:2009, 4.6.18.3.6.
tmp	(determines bs_num_env), no restrictions apply
bs_freq_res	no restrictions apply
bs_pointer	shall be encoded with a value listed in ISO/IEC 14496-3:2009, Table 4.174.

The restrictions defined in ISO/IEC 14496-26:2010, 7.17.1.2.1.3 sbr_grid() shall be applied to the following corresponding bitstream elements:

bs_var_bord_0
bs_var_bord_1
bs_num_rel_0
bs_num_rel_1

bs_noise_position	shall be chosen so that the time slot borders for noise floors fall within the leading and trailing SBR frame borders (i.e. the SBR frame boundaries)
bs_var_len_hf	shall be encoded with a non-reserved value specified in Table 97

8.4.7.2.7 sbr_envelope

bs_env_start_value_balance	no restrictions apply
bs_env_start_value_level	no restrictions apply
bs_codeword	shall be encoded as defined in sbr_huff_dec() in ISO/IEC 14496-3:2009, 4.A.6.1.

Additionally, the restrictions defined in ISO/IEC 14496-26:2010, 7.17.1.2.1.5 sbr_envelope() apply.

8.4.7.2.8 dtdf

bs_df_env no restrictions apply

bs_df_noise no restrictions apply

8.4.7.2.9 sbr_sinusoidal_coding

bs_add_harmonic no restrictions apply

bs_sinusoidal_position_flag no restrictions apply

bs_sinusoidal_position shall be chosen so that the position of the starting time slot for sinusoidals fall within the SBR frame boundaries

8.4.7.2.10 sbr_invf

No restrictions are applicable to this bitstream element.

8.4.7.2.11 sbr_noise

The restrictions defined in ISO/IEC 14496-26:2010, 7.17.1.2.1.6 sbr_noise() apply.

8.4.8 eSBR – Predictive vector coding (PVC)

8.4.8.1 Characteristics

Encoders may apply restrictions to the following parameters of the bitstream:

a) activation of PVC

b) use of IDs from the previous frame

c) length

8.4.8.2 Test procedures

8.4.8.2.1 pvc_envelope

divMode no restrictions apply

nsMode no restrictions apply

reuse_pvcID shall be 0 if the bs_pvc_mode of the preceding SBR frame was 0

pvcID no restrictions apply

length shall be chosen so that the time slot borders for pvcID fall within the SBR frame boundaries

grid_info the first grid_info (grid_info[0]) shall be 1 if the bs_pvc_mode of the preceding SBR frame was 0

8.4.9 eSBR – Inter temporal envelope shaping (inter-TES)

8.4.9.1 Characteristics

Encoders may apply restrictions to the following parameters of the bitstream:

a) activation of inter-TES

8.4.9.2 Test procedure

8.4.9.2.1 sbr_envelope

bs_temp_shape no restrictions apply

bs_inter_temp_shape_mode no restrictions apply

8.4.10 MPEG Surround 2-1-2

8.4.10.1 Characteristics

Encoders may apply restrictions to the following parameters of the bitstream:

a) use of phase coding

b) use of residual coding

c) use of pseudo LR

d) use of Transient Steering Decorrelator

8.4.10.2 Test procedure

8.4.10.2.1 Mps212Data

bsIndependencyFlag no restrictions apply

8.4.10.2.2 FramingInfo

bsFramingType no restrictions apply

bsNumParamSets shall have a value not larger than (numSlots-1)/4, where the division shall be interpreted as an ANSI C integer division

bsParamSlot shall be in the range 0..numSlots-1

8.4.10.2.3 OttData

bsPhaseMode no restrictions apply

bsOPDSmoothingMode no restrictions apply

8.4.10.2.4 SmgData

bsSmoothMode no restrictions apply

bsSmoothTime no restrictions apply

bsFreqResStrideSmg no restrictions apply

bsSmgData no restrictions apply

8.4.10.2.5 TempShapeData

bsTsdEnable no restrictions apply

bsTempShapeEnable no restrictions apply

bsTempShapeEnableChannel no restrictions apply

8.4.10.2.6 TsdData

bsTsdNumTrSlots shall be encoded with 4 or 5 bits depending on numSlots

bsTsdCodedPos no restrictions apply

bsTsdTrPhaseData no restrictions apply

8.4.10.2.7 EcData

bsXXXdataMode shall fulfill the requirements outlined in ISO/IEC 23003-1:2007, 6.1.13. Shall not be encoded with a value of 2 if residual coding is applied. Shall have the value 0 or 3 if ps==0 and **bsIndependencyFlag** is set to 1

bsDataPairXXX shall have the value 0 if setIdx == datasets-1. No further restrictions apply

bsQuantCoarseXXX no restrictions apply

bsFreqResStrideXXX no restrictions apply

8.4.10.2.8 EcDataPair

bsPcmCodingXXX no restrictions apply

8.4.10.2.9 GroupedPcmData

bsPcmWord no restrictions apply

8.4.10.2.10 DiffHuffData

bsDiffType	no restrictions apply
bsCodingScheme	no restrictions apply
bsPairing	no restrictions apply
bsDiffTimeDirection	no restrictions apply

8.4.10.2.11 HuffData1D

hcodFirstband_XXX	**bsCodeW** shall have a value out of a set of values as defined by column 'codeword' in ISO/IEC 23003-1:2007, Tables A.2 and A.3, for CLD and ICC respectively. For IPD, in Table A.2. Shall have a length as defined by the corresponding entry in column 'length'
hcod1D_XXX_YY	**bsCodeW** shall have a value out of a set of values as defined by column 'codeword' in ISO/IEC 23003-1:2007, Tables A.5 and A.6, for CLD and ICC respectively. For IPD, in Table A.3. Shall have a length as defined by the corresponding entry in column 'length'
bsSign	do not apply to the encoding of IPD parameters. No further restrictions apply

8.4.10.2.12 HuffData2DFreqPair, HuffData2DTimePair

hcodLavIdx	**bsCodeW** shall have a value out of a set of values as defined by column 'codeword' in ISO/IEC 23003-1:2007, Tables A.24, and shall have a length as defined by the corresponding entry in column 'length'
hcod2D_XXX_YY_ZZ_LL_escape	**bsCodeW** shall have a value out of a set of values as defined by column 'codeword' in ISO/IEC 23003-1:2007, Tables A.8 and A.9, for CLD and ICC respectively. For IPD, in Table A.4. Shall have a length as defined by the corresponding entry in column 'length'
hcod2D_XXX_YY_ZZ_LL	**bsCodeW** shall have a value out of a set of values as defined by column 'codeword' of the applicable table in ISO/IEC 23003-1:2007, Tables A.11 to A.18, for CLD and ICC. For IPD, in Tables A.5 to A.8. Shall have a length as defined by the corresponding entry in column 'length'

8.4.10.2.13 SymmetryData

bsSymBit	no restrictions apply

8.4.10.2.14 LsbData

bsLsb	no restrictions apply

8.4.11 Restrictions depending on profiles and levels

8.4.11.1 Introduction

Depending on the profile and level associated with the USAC bitstream, further restrictions may apply.

8.4.11.2 Baseline USAC profile

8.4.11.2.1 usacSamplingFrequencyIndex

For Baseline USAC Profile usacSamplingFrequencyIndex shall be encoded with a value specified in Table 151.

Table 151 — Specification of usacSamplingFrequencyIndex and usacSamplingFrequency in Baseline USAC Profile

	Level				
	1	2	3	4	5
usacSamplingFrequencyIndex/ usacSamplingFrequency	0x03...0x0c, 0x11...0x1b 0x1f / ≤ 48000	0x03...0x0c, 0x11...0x1b 0x1f / ≤ 48000	0x03...0x0c, 0x11...0x1b 0x1f / ≤ 48000	0x03...0x0c, 0x0f...0x1b 0x1f / ≤ 96000	N / A

Furthermore, for the Baseline USAC Profile the employed sampling rates shall be one out of those listed in Table 3.

8.4.11.2.2 channelConfigurationIndex

For Baseline USAC Profile channelConfigurationIndex shall be encoded with a value specified in Table 152.

Table 152 — Specification of channelConfigurationIndex in Baseline USAC Profile

	Level				
	1	2	3	4	5
channelConfigurationIndex	0, 1	0, 1, 2, 8	0..6, 8..10	0..6, 8..10	N / A

8.4.11.2.3 numOutChannels

For Baseline USAC Profile numOutChannels shall be encoded with a value specified in Table 153. Further restrictions apply to the number of main audio channels (channels conveyed in UsacSCEs and UsacCPEs) and LFE channels (conveyed in UsacLFEs) as shown in Table 153.

Table 153 — Specification of numOutChannels for Baseline USAC Profile

	Level				
	1	2	3	4	5
numOutChannels	≤ 1	≤ 2	≤ 6	≤ 6	N / A
number of main audio channels	≤ 1	≤ 2	≤ 5	≤ 5	N / A
number of LFE channels	0	0	≤ 1	≤ 1	N / A

8.4.11.2.4 usacElementType

For the Baseline USAC Profile usacElementType shall take values such that the number of main audio channels and LFE channels comply with the restrictions outlined in 8.4.11.2.3.

8.4.11.2.5 tw_mdct

For Baseline USAC Profile tw_mdct shall be encoded with 0.

8.4.11.2.6 tw_data

tw_data should not be present in Baseline USAC Profile complying bitstreams, due to restrictions of bitstream element tw_mdct.

8.4.11.3 Extended HE AAC profile

8.4.11.3.1 usacSamplingFrequencyIndex

For Extended HE AAC Profile usacSamplingFrequencyIndex shall be encoded with a value specified in Table 154.

Table 154 — Specification of usacSamplingFrequencyIndex and usacSamplingFrequency in Extended HE AAC Profile

	Level				
	1	2	3	4	5
usacSamplingFrequencyIndex/ usacSamplingFrequency	N / A	0x03...0x0c, 0x11...0x1b 0x1f / ≤ 48000	0x03...0x0c, 0x11...0x1b 0x1f / ≤ 48000	0x03...0x0c, 0x11...0x1b 0x1f / ≤ 48000	0x03...0x0c, 0x11...0x1b 0x1f / ≤ 48000

8.4.11.3.2 channelConfigurationIndex

For Extended HE AAC Profile channelConfigurationIndex shall be encoded with a value specified in Table 155.

Table 155 — Specification of channelConfigurationIndex in Extended HE AAC Profile

	Level				
	1	2	3	4	5
channelConfigurationIndex	N / A	0, 1, 2, 8	0, 1, 2, 8	0, 1, 2, 8	0, 1, 2, 8

8.4.11.3.3 numOutChannels

For Extended HE AAC Profile numOutChannels shall be encoded with a value specified in Table 156.

Table 156 — Specification of numOutChannels for Extended HE AAC Profile

	Level				
	1	2	3	4	5
numOutChannels	N / A	≤ 2	≤ 2	≤ 2	≤ 2

8.4.11.3.4 tw_mdct

For Extended HE AAC Profile tw_mdct shall be encoded with 0.

8.4.11.3.5 tw_data

The bitstream element tw_data should not be present in Extended HE AAC Profile complying bitstreams, due to restrictions of bitstream element tw_mdct.

8.5 USAC Decoders

8.5.1 General

This document describes a set of test conditions that shall be applied to verify that a given USAC decoder implementation complies with this standard. Test conditions are designed such that each tool can be tested isolated, thus setting the constraints for the corresponding conformance test sequences.

However, some tools show interactions and dependencies. To cover that fact, test cases are defined that can be composed of one or more test conditions.

Every line in the electronic attachment "Usac_Conformance_Tables.xlsx" represents a test case. For each test case in the worksheet a set of conformance test sequences are provided as an electronic attachment to this document. Which tool or tool combination is tested by a given test sequence can be deduced from its filename, as it follows the nomenclature defined in Table 149. In most cases a conformance test sequence consists of an USAC encoded bitstream wrapped in the MP4 file format and the corresponding decoded wave file. Decoded wave files are always supplied with 24 bit resolution (RIFF (little-endian) data, WAVE audio, Microsoft PCM, 24 bit).

To claim conformance, every test sequence mandatory for a certain profile / level combination has to meet the conformance criteria specified for the given test. Bitstream restrictions depending on profile and level are described in 8.4.11.

For each test case varying conformance criteria may apply. The output of the implementation under test has to be tested against the reference by applying the appropriate test procedure. Test procedures as well as constraints for each test case are listed in the electronic attachment "Usac_Conformance_Tables.xlsx". All test procedures are defined in 8.3.2.

8.5.2 FD core mode tests

This Subclause describes test conditions to test the transform based (FD: frequency domain) part of the decoder.

A full list of all FD core related test cases is shown in the attachment "Usac_Conformance_Tables.xlsx": "FD core UsacSCE" and "FD core UsacCPE".

If not stated otherwise, the RMS test method shall be applied to all mandatory test cases. The RMS test method always includes the LSB test (RMS/LSB). The RMS/LSB measurement is defined in ISO/IEC 14496-26:2010. The decoder under test shall satisfy the conformance criteria for at least 16 bit.

If no test method is specified, a check of conformance using appropriate measurements, e.g. the LSB criterion or objective perceptual measurement systems, is not mandatory but highly recommended. This also applies to bitstreams with non-meaningful window sequences.

NOTE The MPEG-4 conformance tool ssnrcd can be used to apply the RMS/LSB test procedure. The tool is part of the MPEG-4 reference software.

If not stated otherwise the following constraints apply to all USAC FD core mode test cases:

— Tests are carried out with coreSbrFrameLengthIndex 0 (768) and 1 (1024), respectively

— The value of max_sfb is set to the maximum allowed value depending on the given sampling rate

— Sampling frequencies as defined in Table 157 are included in the tests

— All test conditions apply to both UsacSingleChannelElement() and UsacChannelPairElement()

Table 157 — Subset of sampling rates under test ("SET")

sampling rate / Hz	samplingFrequencyIndex
7350	0x0c
14400	0x19
22050	0x07
28800	0x14
44100	0x04
88200	0x01

The sampling frequencies in Table 157 are composed of a subset of values in Table 67 and were chosen to cover all available scale factor tables. This subset of sampling frequencies is also referred to as "SET" in this document and in the electronic attachment.

8.5.2.1 Basic FD test condition

8.5.2.1.1 Scope

The "Basic FD test condition" represents a minimum setup of the FD core coder for both single channel and channel pair element.

8.5.2.1.2 Conformance test sequences

The test sequences cover the test of the basic functionalities of the USAC FD core coder. All compressed bitstreams are solely composed of long transform blocks (ONLY_LONG_SEQUENCE).

The tests are carried out at both coreSbrFrameLengthIndex 0 (768) and 1 (1024). For 1024 core coder frame length (coreSbrFrameLengthIndex == 1), additional sampling frequencies are included in the basic FD test case, as there are:

— All allowed values for the usacSamplingFrequencyIndex in Table 67 (ALL)

— The sampling frequencies 55425 Hz and 46008 Hz (arbitrary: ARB)

The sampling frequencies have to be mapped according to Table 79 to properly deduce all sampling frequency dependent tables.

For ARB sampling frequencies no usacSamplingFrequencyIndex is available. The sampling rate has to be transmitted by means of usacSamplingFrequency (24 bit, UsacConfig()).

The corresponding files can be identified by the names Fd_[1|2]_c1__<uSFI>*, where uSFI denotes the usacSamplingFrequencyIndex. If no index is available, uSFI is replaced by the given sampling frequency.

8.5.2.2 FD window switching test condition [Win]

8.5.2.2.1 Scope

This test condition shall be applied to verify the proper decoder behaviour in case a meaningful FD window sequence transition is triggered by a bitstream. Meaningful window sequence transitions are listed in Table 131. Furthermore, the test condition focuses on correct processing of all allowed short block groupings and window shapes.

8.5.2.2.2 Test sequences

Test sequences trigger window transitions as described in Table 158.

Table 158 — Window transitions

Frame	Window Sequence
1	ONLY_LONG_SEQUENCE
2	ONLY_LONG_SEQUENCE
3	LONG_START_SEQUENCE
4	EIGHT_SHORT_SEQUENCE
5	EIGHT_SHORT_SEQUENCE
6	LONG_STOP_SEQUENCE
7	ONLY_LONG_SEQUENCE
8	LONG_START_SEQUENCE
9	LONG_STOP_SEQUENCE
10	LONG_START_SEQUENCE
11	STOP_START_SEQUENCE
12	EIGHT_SHORT_SEQUENCE
13	STOP_START_SEQUENCE
14	STOP_START_SEQUENCE
15	LONG_STOP_SEQUENCE

For the FD window switching test condition [Win], the window sequences listed in Table 158 are run through twice using sine (window_shape 0) and KBD (window_shape 1). The next two frames are window_sequence ONLY_LONG_SEQUENCE and LONG_START_SEQUENCE, respectively. The next 128 frames have window_sequence of EIGHT_SHORT_SEQUENCE only and all possible combinations of scale_factor_grouping are transmitted. The values of scale_factor_grouping vary in the range from 0 to 127. The next frame has window_sequence LONG_STOP_SEQUENCE, after which the cycle repeats.

For test cases that combine the FD window switching test condition [Win] with other test conditions (e.g. WinNf), the window sequences listed in Table 158 are run through a first time using sine (window_shape 0) and a second time using KBD (window_shape 1). This set of window sequences and window_shapes is then repeated for the remainder of the bitstream.

8.5.2.3 Noise filling test condition [Nf]

8.5.2.3.1 Scope

This test condition shall be applied to verify the proper behaviour of the noise filling tool of USAC and the correct signalling of its parameters.

8.5.2.3.2 Test sequences

All bitstreams activate the noise filling tool in the UsacCoreConfig. The values of noise_level and noise_offset vary from frame to frame. All possible combinations of noise_filling and noise_offset are triggered at least once by the bitstream.

8.5.2.4 TNS test condition [Tns]

8.5.2.4.1 Scope

This test condition shall be applied to verify the proper behaviour of the temporal noise shaping (TNS) tool of USAC and the correct signalling of its parameters.

8.5.2.4.2 Test sequences

All bitstreams contain TNS data indicated by the bit tns_data_present. TNS parameters are applied as summarized in Table 159.

NOTE TNS short block combination is covered by the test case labelled “WinTns”.

For both mono and stereo test sequences (channelConfigIndex 1 and 2) supplied bitstreams contain at least TNS values as indicated in Table 159.

Table 159 — Tns bitstream values

Bitstream element	Value
n_filt	1..3 (0, 1)
coef_res	0, 1
Length	1, maxSfb
Order	15 (7), 7 (3), 1
Direction	0, 1
coef_compress	0, 1
Coef	0, 15
NOTE: The values in parenthesis are applied to short blocks.	

Table 160 shows TNS values only present in stereo test cases (channelConfigIndex 2).

Table 160 — Tns stereo bitstream values

Bitstream element	Value
tns_data_present[1]	0, 1
tns_on_lr	1
tns_present_both	0, 1
common_tns	0, 1

8.5.2.5 Varying max_sfb test condition [Sfb]

8.5.2.5.1 Scope

This test condition shall be applied to ensure the correct decoder behaviour in case varying values of max_sfb are signalled by the bitstream.

8.5.2.5.2 Test sequences

The value of max_sfb transmitted in ics_info() varies in the range from 0 to maximum. The upper bound is determined by the given sampling rate.

NOTE Varying max_sfb short block combinations is covered by the combined test case labelled “WinSfb”

Additional constraints apply to USAC channel pair element. Different values of max_sfb are transmitted for each channel in the channel pair element.

8.5.2.6 Handling of extensions test condition [Ex]

8.5.2.6.1 Scope

This test condition shall be applied to ensure the proper behaviour of the extension payload mechanism of the USAC decoder.

A USAC decoder shall at least be able to skip over all extensions – both configuration and payload – and decode the embedded USAC single channel element properly.

8.5.2.6.2 Test sequences

Bitstreams contain extensions to both configuration and payload. Extensions to the configuration are summarized in Table 161.

Table 161 — Values of UsacConfigExtension

Bitstream Element	Value			
numConfigExtensions	4			
usacConfigExtType	0	15	255	65805
usacConfExtLength	1	1	1	1
tmp / fill_byte	165	49	50	51

Extensions to the payload are transmitted by means of an USAC extension element. For each extension element one configuration is embedded in the USAC decoder configuration. Table 162 shows the decoder configuration of the bitstream. The audio data are carried in element 2 (UsacSCE). The extension payload is transmitted via element 0, 1, 3 and 4 (UsacEXT). The test is only carried out for USAC single channel element.

Table 162 — USAC decoder configuration

Element Index	0	1	2	3	4
Element Type	UsacEXT	UsacEXT	UsacSCE	UsacEXT	UsacEXT
usacExtElementType	15	255	-	65805	0 (FILL)
usacExtElementConfigLength	4	4	-	4	0
usacExtElementDefaultLengthPresent	1	1	-	0	0
usacExtElementDefaultLength	8	65790	-	0	0
usacExtElementPayloadFrag	0	1	-	0	0
Tmp	“Ex_1”	“Ex_2”	-	“Ex_3”	-

The extension payload transmitted by means of and USAC extension element can vary from frame to frame. Table 163 shows the affected bitstream values.

Table 163 — USAC extension payload

Element Index	0	1	3	4
usacExtElementPresent	0, 1	0, 1	0, 1	0, 1
usacExtElementUseDefaultLength	0, 1	0	0	0, 1
usacExtElementPayloadLength	1..16	1..16	1..16	arbitrary
usacExtElementStart	-	0, 1	-	-
usacExtElementStop	-	0, 1	-	-

In case of fragmented extension payload (element 1), the payload is divided into 9 frames (distance between usacExtElementStart and corresponding usacExtElementStop flag). The payload transmitted for elements 0, 1 and 3 consists of the string “+++ USAC Conformance Test Extension Element [0,1,2] +++”.

Element 4 is used to write fill bytes (10100101) to into the bitstream if needed. The payload may only be present in a few frames at startup.

8.5.2.7 Arithmetic coder test condition [Ac]

8.5.2.7.1 Scope

This test condition shall be applied to ensure the proper behaviour of the arithmetic decoder of USAC.

8.5.2.7.2 Test sequences

Bitstreams are designed such that:

— The window sequence repeatedly cycles through the following values: ONLY_LONG_SEQUENCE, LONG_START_SEQUENCE, EIGHT_SHORT_SEQUENCE, LONG_STOP_SEQUENCE.

— Window shape is always set to 0, i. e. sine window.

— The reset of the arithmetic decoder is triggered at least every 3 frames.

— The bitstream is divided into at least 4 sections, each 100 frames long. The first 4 sections repeat if the bitstream consists of more than 400 frames.

— In section 1 quantized MDCT values are set to zero. The value of max_sfb is increased frame by frame up to the maximum allowed value.

— In section 2 the amplitude of quantized MDCT values is limited to 3, only positive values are transmitted.

— In section 3 the value of the quantized coefficients is increased frame by frame. Spectral coefficients are coded both with and without STOP symbol.

— In section 4 the amplitude of quantized MDCT values is limited to 3 while the sign is altered.

Test sequences are provided for both 768 and 1024 transform length. The sampling rate is always 48 kHz.

8.5.2.8 Non-meaningful FD window switching test condition [Nmf]

8.5.2.8.1 Scope

This test condition should be applied to monitor the decoder behaviour in case FD window sequence transitions not specified in Table 131 occur in a given bitstream.

8.5.2.8.2 Test sequences

All non-meaningful FD window transitions are triggered at least once by the bitstream. It should be ensured that the decoder does not crash during decoding. This test is not mandatory but highly recommended.

The decoder behaviour at non-meaningful FD window transitions is not covered by the standard, hence no decoded waveforms are provided.

8.5.2.9 M/S stereo test condition [Ms]

8.5.2.9.1 Scope

This test condition shall be applied to verify the proper behaviour of the M/S stereo tool of the USAC decoder.

8.5.2.9.2 Test sequences

Bitstreams make use of the M/S stereo tool. An overview of affected bitstream parameters is shown in Table 164.

Table 164 — M/S stereo parameters

Bitstream element	Value	Description
ms_mask_present	0	M/S not active
	1	M/S active on some scale factor bands
	2	M/S active on all scale factor bands
ms_used	0, 1	Indicates the use of M/S stereo per scale factor band

All bitstreams activating the M/S stereo tool shall cover the values as described above.

8.5.2.10 Complex prediction stereo test condition [Cp]

8.5.2.10.1 Scope

This test condition shall be applied to ensure the functionality of the complex prediction stereo tool of the USAC decoder.

8.5.2.10.2 Test sequences

Bitstreams activate the Complex Prediction stereo tool of USAC. The affected bitstream values are listed in Table 165.

Table 165 — Complex prediction stereo parameters

Bitstream element	Value	Description
ms_mask_present	0	Complex prediction not active
	3	Complex prediction active
cplx_pred_used	0, 1	Indicates the use of complex prediction per prediction band
cplx_pred_all	0, 1	Complex prediction on all prediction bands
complex_coef	0, 1	Transmit complex coefficients (1) or real only coefficients(0)
delta_code_time	0, 1	Time differential coding (1) or frequency differential coding (0)
use_prev_frame	0, 1	Use only current frame (0) or use both current and previous frame (1) for MDST estimation
pred_dir	0, 1	Prediction from mid to side (0) or from side to mid (1)

All bitstreams activating the Complex Prediction stereo tool shall cover all values as described above

8.5.3 LPD core mode tests

8.5.3.1 General

This Subclause describes test cases that have to be applied to verify the behaviour of the USAC decoder when operated in LPD coding mode. A full list of all LPD core coding mode related test cases is shown in the attachment “Usac_Conformance_Tables.xlsx”: “LPD core UsacSCE”, and “LPD core UsacCPE”.

The decoded signals (reference and decoder-under-test) are always time-aligned, low-pass filtered and downsampled to twice the audio bandwidth of the LPD core before computing the conformance measure. The free resampling tool "ResampAudio" from the AFsp package, which is also required by the USAC Reference software, can be used for this purpose. Unless specified otherwise, the audio bandwidth of the LPD core is equal to 6400 Hz when coreSbrFrameLengthIndex=1 (frame length equal to 1024 samples) and 4800 Hz when coreSbrFrameLengthIndex=0 (frame length equal to 768 samples).

The conformance measure depends on the test case. For the LPC coding test, the RMS log LPC spectral distance between the reference signal and the output of the decoder-under-test and the segmental SNR of the output of the decoder-under-test compared to the reference signal are used. For the other tests, the segmental SNR of the output of the decoder-under-test compared to the reference signal is used.

The computation of these measures is described in ISO/IEC 14496-26:2010. Alternatively, an implementation of the RMS log LPC spectral distance can be found in the free "libtsp" TSP Signal Processing Library (function called "SPlpcLSdist"), and the segmental SNR can be computed using the "CompAudio" tool from the AFsp package.

The tests are carried out for both 768 and 1024 core coder frame length (coreSbrFrameLengthIndex equal to 0 and 1).

For coreSbrFrameLengthIndex=1 (frame length equal to 1024 samples) three distinct test vectors are used to validate the operation of the USAC decoder under test at different internal sampling frequencies, namely 6000, 12800 and 24000 Hz. These are identified by the file names Lpd_c1_Lpd_<uSFI>*, where uSFI denotes the usacSamplingFrequencyIndex. The audio bandwidth of the LPD core is equal to half the internal sampling frequency.

8.5.3.2 LPC coding test condition [Lpc]

8.5.3.2.1 Scope

The test condition shall be applied to verify the functionality of the linear predictive coding (LPC) filter and the proper decoding of LPC parameters in the bitstream.

8.5.3.2.2 Test sequences

The test bitstream is designed such that:

— All frames are encoded using MDCT-based TCX

— For each of the 4 LPC filters LPC1, LPC2, LPC3 and LPC4, every possible absolute and relative quantization mode from Table 143 is used at least once

— Each of the 256 entries in the first stage approximation codebook (see 7.13.6) is used at least once

Furthermore, the test bitstream is designed to test the decoder on "extreme" LPC filters (in particular, exhibiting high resonances that cover well the entire audio spectrum).

8.5.3.2.3 Conformance criteria

The conformance criteria for the LPC coding test condition is based on the RMS log LPC spectral distance between the reference signal and the output of the decoder-under-test and on the segmental SNR of the output of the decoder under test compared to the reference signal.

The RMS log LPC spectral distance between the reference signal and the output of the decoder under test shall not exceed 0.6 dB. Also, the segmental SNR of the output of the decoder under test compared to the reference signal shall not be less than 40 dB.

8.5.3.3 ACELP core mode test condition [Ace]

8.5.3.3.1 Scope

This test condition shall be applied to verify the correct decoding of frames encoded with the ACELP coding scheme.

8.5.3.3.2 Test sequences

The test bitstream is designed such that:

— All frames are encoded using ACELP (no MDCT-based TCX)

— A complete and balanced coverage of the algebraic codebooks listed in 7.14.5.2.1 is ensured. Specifically, the usage of the algebraic codebooks is as follow:

100 frames encoded using the 20-bit codebook, followed by

100 frames encoded using the 28-bit codebook, followed by

100 frames encoded using the 36-bit codebook, followed by

100 frames encoded using the 44-bit codebook, followed by

100 frames encoded using the 52-bit codebook, followed by

100 frames encoded using the 64-bit codebook, followed by

100 frames encoded using the 12–bit codebook, followed by

100 frames encoded using the 16–bit codebook

— Every possible value of the bitfields mean_energy (4 possibilities, see Table 145), acb_index[·] (512 or 64 possibilities, depending on the subframe position), ltp_filtering_flag[·] (two possibilities) and gains[·] (128 possibilities) is used at least once

— The LPC filters exhibit weak resonances

— The bass-post filter is always disabled (bpf_**control_info**=0)

8.5.3.3.3 Conformance criteria

The conformance criteria for the ACELP core mode test condition is based on the segmental SNR of the output of the decoder under test compared to the reference signal.

The length of the segments is equal to 256 samples.

The segmental SNR of the output of the decoder under test compared to the reference signal shall not be less than 50 dB.

8.5.3.4 TCX and noise filling test condition [Tcx]

8.5.3.4.1 Scope

This test condition shall be applied to verify the correct decoding of frames encoded with the TCX coding scheme. Furthermore, the TCX noise filling is covered.

8.5.3.4.2 Test sequences

The test bitstream is designed such that:

— All frames are encoded using MDCT-based TCX (no ACELP)

— A complete and balanced coverage of all possible MDCT window lengths is ensured

— Moreover, a complete and balanced coverage of all possible intra-frame and inter-frame transitions between MDCT window lengths is ensured

— Every possible value of the bitfields **noise_factor** (8 possibilities) and **global_gain** (128 possibilities) is used at least once

— The test bitstream contains LPC filters exhibiting weak resonances

In order to guarantee a complete and balanced coverage of all MDCT window lengths and all transitions between these, the usage of the various MDCT window lengths is as follows:

[1 1 1 1] for 150 frames

[2 2 2 2] for 150 frames

[3 3 3 3] for 150 frames

Then a repetition of the following pattern for a total of at least 150 frames:

[1 1 1 1][1 1 2 2][1 1 2 2][2 2 2 2][2 2 1 1][2 2 1 1][3 3 3 3][2 2 2 2][3 3 3 3][3 3 3 3]

where [· · · ·] represents the four LPD coding modes **mod[0..3]** for one frame and 1, 2 and 3 are the mode values that determine the MDCT window length as described in Table 92 (specifically, 1 for short TCX, 2 for medium TCX and 3 for long TCX).

8.5.3.4.3 Conformance criteria

The conformance criteria for the TCX and noise filling test condition is based on the segmental SNR of the output of the decoder under test compared to the reference signal.

The length of the segments is equal to 256 samples.

The segmental SNR of the output of the decoder under test compared to the reference signal shall not be less than 50 dB.

8.5.3.5 LPD mode coverage and FAC test condition [Lpd]

8.5.3.5.1 Scope

This test condition shall be applied to ensure the proper decoding of frames encoded in LPD mode. It also covers all allowed transitions between LPD coding schemes (ACELP / TCX).

8.5.3.5.2 Test sequences

The test bitstream is designed such that:

— Every possible combination of MDCT-based TCX and/or ACELP within a frame is used at least once

— Moreover, a complete and balanced coverage of all possible intra-frame and inter-frame transitions between ACELP and the different MDCT window lengths is ensured

— The test bitstream contains LPC filters exhibiting weak resonances

— The bass-post filter is always disabled (**bpf_control_info**=0)

The first two conditions are guaranteed by using a repetition of the following mode pattern:

A sequence comprising the LPD coding modes corresponding to each of the 26 unreserved values of the bitfield **lpd_mode** from Table 89

followed by:

[0 0 1 1][0 0 1 1][1 1 0 0][1 1 0 0][0 0 2 2][0 0 2 2][2 2 0 0][2 2 0 0][3 3 3 3]

where [· · · ·] represents the four LPD coding modes **mod[0..3]** for one frame and 0, 1, 2 and 3 are the mode values as described in Table 92 (specifically, 0 for ACELP, 1 for short TCX, 2 for medium TCX and 3 for long TCX).

8.5.3.6 Conformance criteria

The conformance criteria for the LPD mode coverage and FAC test condition is based on the segmental SNR of the output of the decoder under test compared to the reference signal.

The length of the segments is equal to 256 samples.

The segmental SNR of the output of the decoder under test compared to the reference signal shall not be less than 50 dB.

8.5.3.7 Bass-post filter test condition [Bpf]

8.5.3.7.1 Scope

This test condition shall be applied to verify the behaviour of the bass-post filter of the USAC decoder in LPD coding mode.

8.5.3.7.2 Test sequences

The test bitstream is designed such that:

— The frames are encoded using alternately the MDCT-based TCX coding mode (5 consecutive frames) and the ACELP coding mode (25 consecutive frames)

— The bass-post filter is switched on (**bpf_control_info**=1) and off (**bpf_control_info**=0) every 5 ACELP frames

— Every possible value of the **acb_index** parameter (512 or 64 possibilities, depending on the subframe position) is used at least once for the ACELP frames where the bass-post filter is enabled

— The test bitstream contains LPC filters exhibiting weak resonances

— For a USAC channel pair element both synchronous (Bpfs) and asynchronous (Bpfa) core coding modes are tested in combination with bass-post filter activity. The Bpfa case occurs when the two channels either use in a different core coding mode (ACELP / TCX) or, when both channels use the ACELP core coding mode but make a reversed use of the bass-post filter (active / inactive).

8.5.3.7.3 Conformance criteria

The conformance criteria for the Bass-post filter test condition is based on the segmental SNR of the output of the decoder under test compared to the reference signal.

The length of the segments is equal to 256 samples.

The segmental SNR of the output of the decoder under test compared to the reference signal shall not be less than 50 dB.

8.5.3.8 AVQ test condition [Avq]

8.5.3.8.1 Scope

This test condition shall be applied to test the AVQ quantization tool of the USAC decoder.

8.5.3.8.2 Test sequences

The test bitstream is designed such that:

— All frames are encoded using alternately ACELP and short MDCT-based TCX (i.e. all frames are encoded using the LPD mode sequence [0 1 0 1])

— As regards the quantization of the FAC information, every absolute leader from Table 141 is used at least once

— The test bitstream contains LPC filters exhibiting weak resonances

— The bass-post filter is always disabled (**bpf_control_info**=0)

8.5.3.8.3 Conformance criteria

The conformance criteria for the AVQ test condition is based on the segmental SNR of the output of the decoder under test compared to the reference signal.

The length of the segments is equal to 256 samples.

The segmental SNR of the output of the decoder under test compared to the reference signal shall not be less than 50 dB.

8.5.4 Combined core coding tests

8.5.4.1 General

This subclause describes test conditions to be applied to the USAC decoder in the case both FD and LPD coding mode are present in a bitstream.

If not stated otherwise, the conformance measure is calculated using the segmental SNR of the output of the decoder under test compared to the reference signal. The length of the segments is equal to 256 samples. The segmental SNR of the output of the decoder under test compared to the reference signal shall not be less than 40 dB.

8.5.4.2 FD-LPD transition and FAC test condition (synchronous / asynchronous) [Flt<a|s>]

8.5.4.2.1 Scope

This test condition shall be applied to ensure the proper decoder behaviour when a given bitstream activates both USAC core coding modes (FD / LPD).

8.5.4.2.2 Test sequences

Bitstreams trigger all allowed transitions between FD and LPD coding modes.

— Bitstreams shall trigger every allowed transition between FD and LPD coding modes as shown in Table 133 at least once

— All allowed combinations of TCX modes and ACELP are triggered at least once

— For USAC channel pair element both synchronous (Flts) and asynchronous (Flta) transitions are triggered. Asynchronous transitions occur when the two channels of the channel pair element use different coding modes (FD / LPD).

— No bass-post filter is used (bpf_contol_info == 0)

8.5.4.3 FD/TCX noise filling test condition [Cnf]

8.5.4.3.1 Scope

This test condition shall be applied to verify the interaction between the FD noise filling and the TCX noise filling functionality.

8.5.4.3.2 Test sequences

Bitstreams activate the noise filling tool in both FD and LPD path. The bitstreams are designed that:

— All allowed values of noise_level and noise_offset are transmitted at least once

— All allowed values of noise_factor are transmitted at least once

— All TCX modes are used at least once

— No ACELP is used

— All valid transitions between FD core mode and LPD core mode as shown in Table 133 are triggered at least once

8.5.4.4 Bass-post filter test condition [Cbf]

8.5.4.4.1 Scope

This test condition shall be applied to ensure the correct behaviour of the bass-post filter at transitions between FD and LPD core mode.

8.5.4.4.2 Test sequences

Bitstreams are designed that:

— The bass-post filter is activated in every frame encoded using LPD coding mode

— All valid transitions between FD core mode and LPD core mode as shown in Table 133 are triggered at least once

— All allowed combinations of TCX modes and ACELP are triggered at least once

8.5.4.5 Arithmetic coder test condition [CAc]

8.5.4.5.1 Scope

This test condition shall be applied to test the arithmetic decoder of USAC when both FD and LPD coding modes are employed.

8.5.4.5.2 Test sequences

Bitstreams are designed such that:

— All valid transitions between FD core mode and LPD core mode as shown in Table 133 are triggered at least once.

— A reset of the arithmetic decoder is triggered in a frame consisting of only ACELP at least once.

Test sequences are provided for both 768 and 1024 transform length. Sampling rate is always 16 kHz.

8.5.5 eSBR Tests

8.5.5.1 General

A full list of all eSBR related test cases is shown in the attachment “Usac_Conformance_Tables.xlsx”:

8.5.5.1.1 eSBR Test procedure

If not stated otherwise, the RMS test method shall be applied to all eSBR test cases. The decoder under test shall satisfy the conformance criteria for at least 16 bit.

8.5.5.2 QMF accuracy test condition [Qma]

8.5.5.2.1 Scope

This test condition shall be applied to verify the implementation of the QMF filter bank.

8.5.5.2.2 Test sequences

The sequence consists of a linear sine sweep from 0 to 8000 Hz (eSbr cross over frequency).

8.5.5.3 Envelope adjuster accuracy and SBR preprocessing test condition [Eaa]

8.5.5.3.1 Scope

This test condition shall be applied to cover the test of the eSbr envelope adjuster as well as the eSbr preprocessing (pre-whitening) functionality.

8.5.5.3.2 Test sequences

Table 166 describes the variables in scope of this test condition.