INTERNATIONAL STANDARD

ISO/IEC 23003-1

First edition
2007-02-15

AMENDMENT 3
2016-12-15

Information technology — MPEG audio technologies —

Part 1:
MPEG Surround

AMENDMENT 3: MPEG Surround extension for 3D Audio

Technologies de l'information — Technologies audio MPEG —

Partie 1: Ambiance MPEG

AMENDEMENT 3: Extension de l'ambiance MPEG pour audio 3D

Reference number
ISO/IEC 23003-1:2007/Amd.3:2016(E)

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.

The procedures used to develop this document and those intended for its further maintenance are described in the ISO/IEC Directives, Part 1. In particular the different approval criteria needed for the different types of document should be noted. This document was drafted in accordance with the editorial rules of the ISO/IEC Directives, Part 2 (see www.iso.org/directives).

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. Details of any patent rights identified during the development of the document will be in the Introduction and/or on the ISO list of patent declarations received (see www.iso.org/patents).

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

For an explanation on the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the World Trade Organization (WTO) principles in the Technical Barriers to Trade (TBT) see the following URL: www.iso.org/iso/foreword.html

The committee responsible for this document is ISO/IEC JTC 1, *Information technology*, SC 29, *Coding of audio, picture, multimedia and hypermedia information*.

Information technology — MPEG audio technologies —

Part 1:
MPEG Surround

AMENDMENT 3: MPEG Surround extension for 3D Audio

Page 3, 3.1

Add the following after 3.1.12:

3.1.13

N-N/2-N configuration

configuration of MPEG Surround coding system that recreated N channels from half of N downmixed channels with the corresponding spatial parameters

Pages 3 and 4, 3.1

Renumber the terms 3.1.13 to 3.1.26 as 3.1.14 to 3.1.27.

Page 6, 3.5

Add the following variables:

$\mathbf{I}_N$ is unity matrix and subscript index indicate matrix dimension, e.g. N by N unity matrix.

$\mathbf{O}_N$ is null matrix and subscript index indicate matrix dimension, e.g. N by N null matrix.

Add a new Clause 10

10 Outline

10.1 General

The decoding process for N-N/2-N is described in the following clause.

10.2 Syntax

10.2.1 Payloads for N-N/2-N Extension

Table 10.1 — Syntax of SpatialSpecificConfig()

Syntax	No. of bits	Mnemonic
`SpatialSpecificConfig()`		
`{`		
`bsSamplingFrequencyIndex;`	**4**	**uimsbf**
`if (bsSamplingFrequencyIndex == 0xf ) {`		
NOTE 1 SpeakerConfig3d() is defined in ISO/IEC 23008-3:2015, Table 5.		
NOTE 2 numOttBoxes and numTttBoxes are defined by Table 10.2 dependent on bsTreeConfig.		

Table 10.1 *(continued)*

Syntax	No. of bits	Mnemonic
bsSamplingFrequency;	24	uimsbf
}		
bsFrameLength;	7	uimsbf
bsFreqRes;	3	uimsbf
bsTreeConfig;	4	uimsbf
if (bsTreeConfig == '0111') {		
bsNumInCh;	4	uimsbf
bsNumLFE	2	uimsbf
bsHasSpeakerConfig	1	uimsbf
if (bsHasSpeakerConfig == 1) {		
audioChannelLayout = SpeakerConfig3d();		Note 1
}		
}		
bsQuantMode;	2	uimsbf
bsOneIcc;	1	uimsbf
bsArbitraryDownmix;	1	uimsbf
bsFixedGainSur;	3	uimsbf
bsFixedGainLFE;	3	uimsbf
bsFixedGainDMX;	3	uimsbf
bsMatrixMode;	1	uimsbf
bsTempShapeConfig;	2	uimsbf
bsDecorrConfig;	2	uimsbf
bs3DaudioMode;	1	uimsbf
if (bsTreeConfig == '0111') {		
for (i=0; i< NumInCh - NumLfe; i++) {		
defaultCld[i] = 1;		
ottModelfe[i] = 0;		
}		
for (i= NumInCh - NumLfe; i< NumInCh; i++) {		
defaultCld[i] = 1;		
ottModelfe[i] = 1;		
}		
}		
for (i=0; i<numOttBoxes; i++) {		Note 2
OttConfig(i);		
}		
for (i=0; i<numTttBoxes; i++) {		Note 2
TttConfig(i);		
}		

NOTE 1 SpeakerConfig3d() is defined in ISO/IEC 23008-3:2015, Table 5.

NOTE 2 numOttBoxes and numTttBoxes are defined by Table 10.2 dependent on bsTreeConfig.

Table 10.1 *(continued)*

Syntax	No. of bits	Mnemonic
`    if (bsTempShapeConfig == 2) {`		
`        bsEnvQuantMode`	1	uimsbf
`    }`		
`    if (bs3DaudioMode) {`		
`        bs3DaudioHRTFset;`	2	uimsbf
`    if (bs3DaudioHRTFset==0) {`		
`        ParamHRTFset();`		
`      }`		
`    }`		
`    ByteAlign();`		
`    SpatialExtensionConfig();`		
`}`		
NOTE 1 SpeakerConfig3d() is defined in ISO/IEC 23008-3:2015, Table 5.		
NOTE 2 numOttBoxes and numTttBoxes are defined by Table 10.2 dependent on bsTreeConfig.		

Table 10.2 — bsTreeConfig

bsTreeConfig	Meaning
0,1,2,3,4,5,6	Identical meaning in ISO/IEC 20003-1:2007, Table 40
7	N-N/2-N configuration numOttBoxes = NumInCh numTttBoxes = 0 numInChan = NumInCh numOutChan = NumOutCh output channel ordering is according to Table 10.5
8…15	Reserved

bsNumInCh Defines number of input DMX channels for N-N/2-N configuration according to:

Table 10.3 — bsNumInCh

bsNumInCh	NumInCh	NumOutCh
0	12	24
1	7	14
2	5	10
3	6	12
4	8	16
5	9	18
6	10	20
7	11	22
8	13	26
9	14	28

Table 10.3 *(continued)*

bsNumInCh	NumInCh	NumOutCh
10	15	30
11	16	32
12,...,15	Reserved	Reserved

bsNumLfe Defines number $\mathbf{N_{LFE}}$ of output Lfe channels for N-N/2-N configuration

Table 10.4 — bsNumLFE

bsNumLFE	NumLfe
0	0
1	1
2	2
3	Reserved

Table 10.5 — Output channel ordering for N-N/2-N configuration

NumOutCh	NumLfe	Output channel ordering
24	2	Rv,Rb,Lv,Lb,Rs,Rvr,Lsr,Lvr,Rss,Rvss,Lss,Lvss,Rc,R, Lc,L,Ts,Cs,Cb,Cvr, C,LFE,Cv,LFE2,
14	0	L,Ls,R,Rs,Lbs,Lvs,Rbs,Rvs,Lv,Rv, Cv,Ts, C,LFE
12	1	L,Lv,R,Rv,Lsr,Lvr,Rsr,Rvr,Lss,Rss,C,LFE
12	2	L,Lv,R,Rv,Ls,Lss,Rs,Rss,C,LFE,Cvr,LFE2
10	1	L,Lv,R,Rv,Lsr,Lvr,Rsr,Rvr,C,LFE

NOTE 1 All of Names and layouts of loudspeaker follows the naming and position in ISO/IEC 23001-8:2013/FDAM1, Table 8.

NOTE 2 Output channel ordering for the case of 16, 20, 22, 26, 30 and 32 is following the arbitrary order from 1 to N without any specific naming of speaker layouts.

NOTE 3 Output channel ordering for the case when bsHasSpeakerConfig == 1 follows the order from 1 to N with associated naming of speaker layouts as specified in ISO/IEC 23008-3:2015, Table 94.

bsHasSpeakerConfig This flag indicates whether the output channels have a different layout than the output channel ordering specified in Table 10.5. If present (bsHasSpeaker-Config == 1), the loudspeaker layout of the output configuration "audioChan-nelLayout" can be used for rendering if the N-N/2-N system is used together with other MPEG standards (e.g. ISO/IEC 23008-3:2015).

audioChannelLayout This structure describes the loudspeaker layout of the output configuration. If the output configuration contains LFE channels, the LFE channels shall be ordered such that each LFE channel is processed together with one non-LFE channel using one OTT box and shall be positioned at the end of the channel list (e.g. L, Lv, R, Rv, Ls, Lss, Rs, Rss, C, LFE, Cvr, LFE2).

10.3 The N-N/2-N configuration

10.3.1 Introduction

In the following subclauses, the general structure for the N-N/2-N system is outlined. For this configuration, N/2 is identical to the number of downmix signals (*NumInCh* = N/2), denoted x_0 to $x_{NumInCh-1}$. Therefore, the number of output signals (i.e. N) should be an even number in order to process N/2 downmix signals, since the number of OTT boxes is equal to N/2.

The input vector to be multiplied by $\mathbf{M}_1^{n,k}$ is a vector containing the N/2 downmix channels. A maximum number of N/2 decorrelators can be used when LFE channels are not included in output channels. However, if the number of output channels exceeds twenty channels, the de-correlation filters are reused according to 10.7. Some of the decorrelator indices are repeated because the number of available decorrelators that ensure orthogonal decorrelated output signals is limited to 10, as defined in ISO/IEC 23003-1:2007. Therefore, the recommended number of output channels for the N-N/2-N configuration is less than 20 (or 24 with two Lfe channels).

The outputs of the decorrelators can be replaced by residual signals for certain frequency regions, depending on the bitstream. No decorrelation is used for the case of OTT based upmix when a LFE channel is one output of the OTT box. No residual signal can be inserted for these OTT boxes.

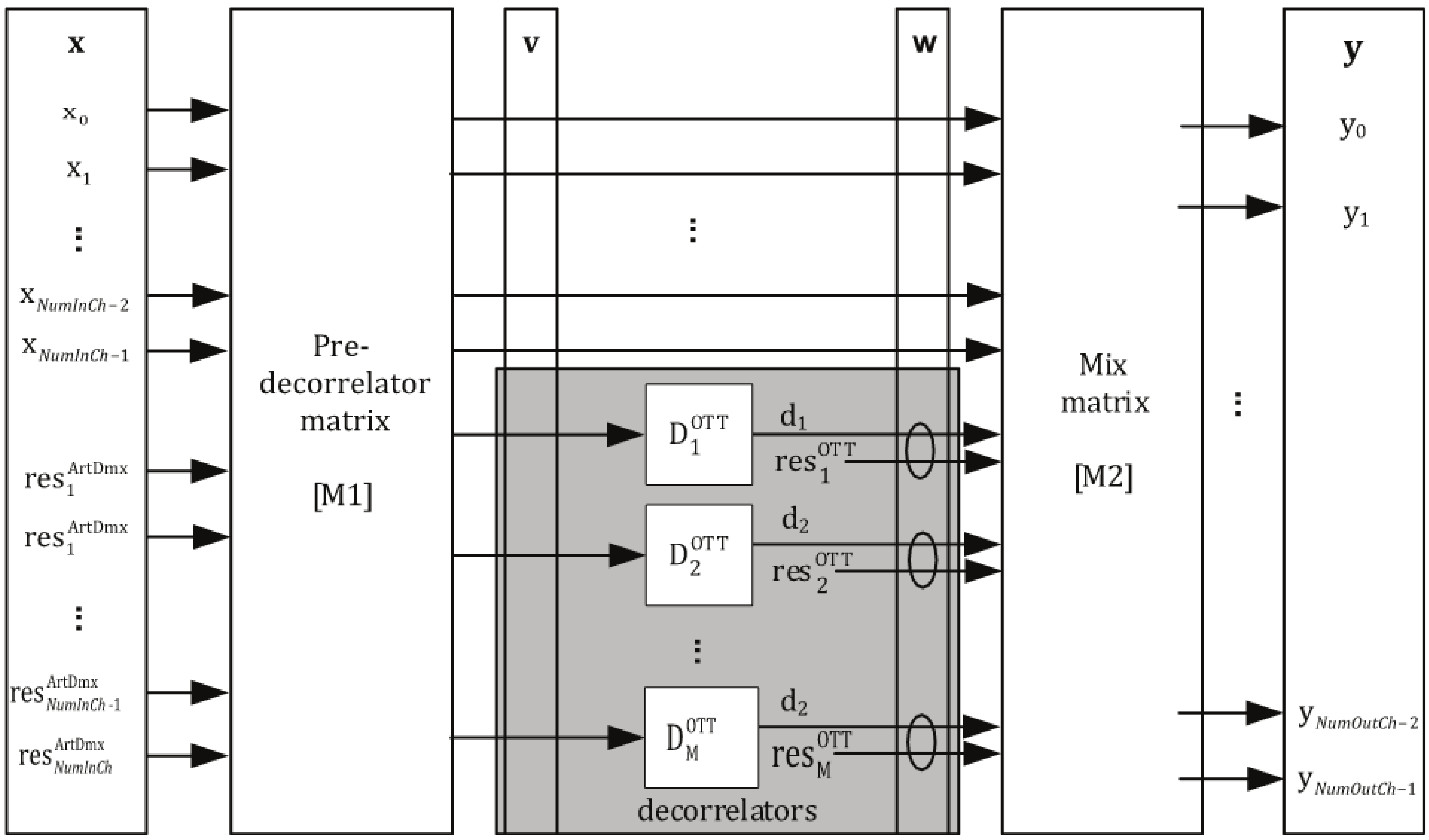

Figure 10.1 — Matrix view of the spatial audio processing for the N-N/2-N configuration

The decorrelators, decorrelated signals and residual signals in Figure 10.1 (labelled "1" to "M (i.e. NumInCh-NumLfe)") correspond to different OTT boxes depending on configuration.

The multi-channel reconstruction for the N-N/2-N configuration can also be visualized by means of a tree-structure. This is outlined in Figure 10.2. In Figure 10.2, every OTT box re-creates two channels based on one input channel, the corresponding CLD and ICC parameters, and residual signal. The OTT boxes and the corresponding data are numbered corresponding to the order they appear in the bitstream.

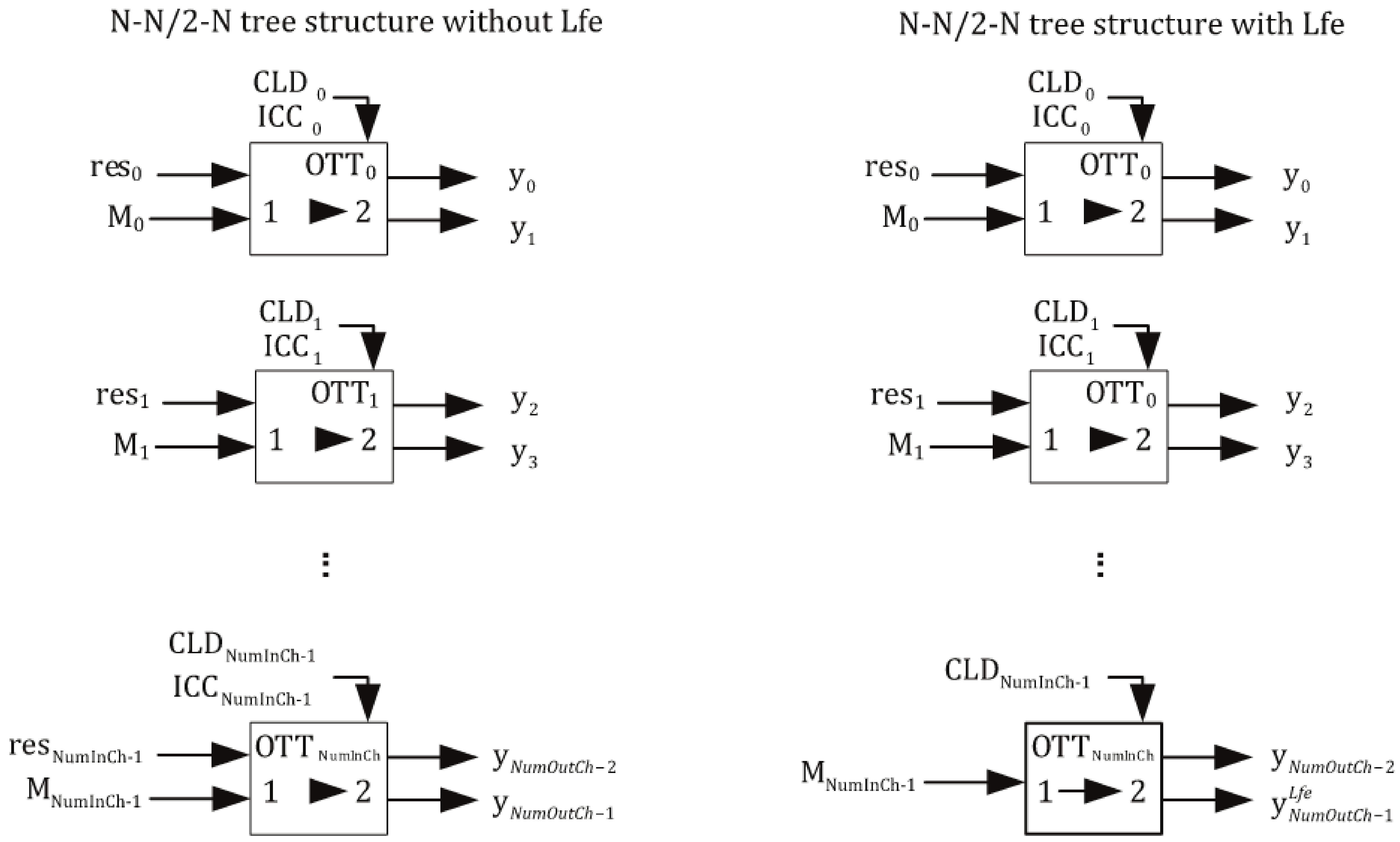

Figure 10.2 — Tree structure view of the spatial audio processing for the N-N/2-N configurations

The definitions of the vectors and matrices for N-N/2-N configuration are used. The matrixes $\mathbf{M}_1^{n,k}$ and $\mathbf{M}_2^{n,k}$ are defined accordingly in 10.5 and 10.6, while the vectors to be multiplied with the matrices in order to form the output are defined in the following subclauses.

10.4 Vector definitions for the N-N/2-N configuration

10.4.1 Operation without temporal shaping tools

For the N-N/2-N configuration, the input signals to the decorrelators are defined by $\mathbf{v}^{n,k}$, which is derived from the input vector $\mathbf{x}^{n,k}$ and the matrix $\mathbf{M}_1^{n,k}$ having N rows and 1 column, according to:

$$\mathbf{v}^{n,k} = \mathbf{M}_1^{n,k}\mathbf{x}^{n,k} = \mathbf{M}_1^{n,k}\begin{bmatrix} x_{\mathrm{M}_0}^{n,k} \\ x_{\mathrm{M}_1}^{n,k} \\ \cdots \\ x_{\mathrm{M}_{\mathrm{NumInCh-1}}}^{n,k} \\ x_{\mathrm{res}_0^{\mathrm{ArtDmx}}}^{n,k} \\ x_{\mathrm{res}_1^{\mathrm{ArtDmx}}}^{n,k} \\ \cdots \\ x_{\mathrm{res}_{\mathrm{NumInCh-1}}^{\mathrm{ArtDmx}}}^{n,k} \end{bmatrix} = \begin{bmatrix} v_{\mathrm{M}_0}^{n,k} \\ v_{\mathrm{M}_1}^{n,k} \\ \cdots \\ v_{\mathrm{M}_{NumInCh-1}}^{n,k} \\ v_0^{n,k} \\ v_1^{n,k} \\ \cdots \\ v_{NumInCh-NumLfe-1}^{n,k} \end{bmatrix}$$

The subscripts for the different elements in the $\mathbf{v}^{n,k}$ vector indicate which OTT box decorrelator the signal is input to, with the exception from $v_{\mathrm{M}_0}^{n,k}$ to $v_{M_{NumInCh-NumLfe-1}}^{n,k}$, which is the direct signal.

The vector $\mathbf{w}^{n,k}$ holding the direct signal, decorrelated signals, and the residual signals is defined according to:

$$\mathbf{w}^{n,k}=\begin{bmatrix} v_{\mathrm{M}_0}^{n,k} \\ v_{\mathrm{M}_1}^{n,k} \\ \cdots \\ v_{\mathrm{M}_{NumInCh\text{-}1}}^{n,k} \\ \delta_0(k)D_0\left(v_{M_0}^{n,k}\right)+\left(1-\delta_0(k)\right)v_{\mathrm{res}_0}^{n,k} \\ \delta_1(k)D_1\left(v_{M_2}^{n,k}\right)+\left(1-\delta_1(k)\right)v_{\mathrm{res}_1}^{n,k} \\ \cdots \\ \delta_{NumInCh-NumLfe-1}(k)D_{NumInCh-NumLfe-1}\left(v_{M_{NumInCh-NumLfe-1}}^{n,k}\right)+\left(1-\delta_{NumInCh-NumLfe-1}(k)\right)v_{\mathrm{res}_{NumInCh-NumLfe-1}}^{n,k} \end{bmatrix}=\begin{bmatrix} w_{\mathrm{M}_0}^{n,k} \\ w_{\mathrm{M}_1}^{n,k} \\ \cdots \\ w_{\mathrm{M}_{NumInCh\text{-}1}}^{n,k} \\ w_1^{n,k} \\ w_2^{n,k} \\ \cdots \\ w_{NumInCh-NumLfe-1}^{n,k} \end{bmatrix}$$

where $\delta_{\mathrm{x}}(k)=\begin{cases}0 & ,0\le k\le \max\{k_{\mathrm{set}}\}\\ 1 & ,\text{otherwise}\end{cases}$ and where k_{set} is the set for which all values of k fulfil $\kappa(k)<\mathbf{m}_{\mathrm{resProc}}(X)$ given by Table A.31, and where $D_X\left(v_X^{n,k}\right)$ is the output from decorrelator D_X given the input signal $v_X^{n,k}$.

The subscripts for the different elements indicate which OTT box the signal corresponds to the numbering of OTT boxes for the 5-1-5$_1$ configuration as given by Figure 23. Hence, $D_X\left(v_X^{n,k}\right)$ is the decorrelator output from box OTT$_X$ and $v_{\mathrm{res}_X}^{n,k}$ is the corresponding residual signal.

The subband output signals are subsequently defined for every time-slot n, and every hybrid subband k, by $\mathbf{y}^{n,k}$, which is derived from the vector $\mathbf{w}^{n,k}$ and the matrix $\mathbf{M}_2^{n,k}$ having *NumOutCh* rows and *NumInCh-NumLfe* columns, according to

$$\mathbf{y}^{n,k}=\mathbf{M}_2^{n,k}\mathbf{w}^{n,k}=\mathbf{M}_2^{n,k}\begin{bmatrix} w_{\mathrm{M}_0}^{n,k} \\ w_{\mathrm{M}_1}^{n,k} \\ \cdots \\ w_{\mathrm{M}_{NumInCh\text{-}1}}^{n,k} \\ w_1^{n,k} \\ w_2^{n,k} \\ \cdots \\ w_{NumInCh\text{-}NumLfe-1}^{n,k} \end{bmatrix}=\begin{bmatrix} y_0^{n,k} \\ y_1^{n,k} \\ \cdots \\ y_{NumInCh-2}^{n,k} \\ y_{NumInCh-1}^{n,k} \end{bmatrix}.$$

The elements of $\mathbf{M}_2^{n,k}$ are defined in 10.6, and the hybrid subband signals defined in $\mathbf{y}^{n,k}$ are synthesized to the time-domain by the hybrid synthesis filterbank as defined in 6.3.

10.4.2 Operation with temporal shaping tools

If temporal shaping is used, the vector $\mathbf{v}^{n,k}$ is defined identically to the previous subclause, however, two $\mathbf{w}^{n,k}$ vectors are defined. The first, $\mathbf{w}^{n,k}_{\text{direct}}$ holds the direct signal and the residual signals, while the second $\mathbf{w}^{n,k}_{\text{diffuse}}$ holds the decorrelator output signals, according to:

$$\mathbf{w}^{n,k}_{\text{direct}} = \begin{bmatrix} v^{n,k}_{\mathrm{M}_0} \\ v^{n,k}_{\mathrm{M}_1} \\ \cdots \\ v^{n,k}_{\mathrm{M}_{NumInCh-1}} \\ \left(1-\delta_0\left(k\right)\right) v^{n,k}_{\text{res}_0} \\ \left(1-\delta_0\left(k\right)\right) v^{n,k}_{\text{res}_1} \\ \cdots \\ \left(1-\delta_2\left(k\right)\right) v^{n,k}_{\text{res}_{NumInCh-NumLfe-1}} \end{bmatrix} = \begin{bmatrix} w^{n,k}_{\mathrm{M}_0} \\ w^{n,k}_{\mathrm{M}_1} \\ \cdots \\ w^{n,k}_{\mathrm{M}_{NumInCh\text{-}1}} \\ w^{n,k}_0 \\ w^{n,k}_1 \\ \\ w^{n,k}_{NumInCh-NumLfe-1} \end{bmatrix}$$

$$\mathbf{w}^{n,k}_{\text{diffuse}} = \begin{bmatrix} v^{n,k}_{\mathrm{M}_0} \\ v^{n,k}_{\mathrm{M}_1} \\ \cdots \\ v^{n,k}_{\mathrm{M}_{NumInCh\text{-}1}} \\ \delta_0\left(k\right) D_0\left(v^{n,k}_0\right) \\ \delta_1\left(k\right) D_1\left(v^{n,k}_1\right) \\ \cdots \\ \delta_{NumInCh-NumLfe-1}\left(k\right) D_{NumInCh-NumLfe-1}\left(v^{n,k}_{NumInCh-NumLfe-1}\right) \end{bmatrix} = \begin{bmatrix} w^{n,k}_{\mathrm{M}_0} \\ w^{n,k}_{\mathrm{M}_1} \\ \cdots \\ w^{n,k}_{\mathrm{M}_{NumInCh\text{-}1}} \\ w^{n,k}_0 \\ w^{n,k}_1 \\ \cdots \\ w^{n,k}_{NumInCh-NumLfe-1} \end{bmatrix}$$

where $\delta_X\left(k\right) = \begin{cases} 0 & ,0 \le k \le \max\left\{k_{\text{set}}\right\} \\ 1 & ,\text{otherwise} \end{cases}$ and where k_{set} is the set for which all values of k fulfil $\kappa\left(k\right) < \mathbf{m}_{\text{resProc}}\left(X\right)$ given by Table A.31, and where $D_X\left(v^{n,k}_X\right)$ is the output from decorrelator D_X given the input signal $v^{n,k}_X$. The subscripts are used as outlined in the previous subclause.

Two temporary output vectors are derived, $\mathbf{y}^{n,k}_{\text{direct}}$ holding the direct signal, and $\mathbf{y}^{n,k}_{\text{diffuse}}$ holding the diffuse signal. They are calculated from $\mathbf{w}^{n,k}_{\text{direct}}$ and $\mathbf{w}^{n,k}_{\text{diffuse}}$, using $\mathbf{M}^{n,k}_2$ which is identical to that used if no temporal shaping is applied. The output is derived from these as outlined in 10.4.2.1, if the STP tool is used, and 10.4.2.2 if the GES tool is used, as indicated by data stream element *bsTempShapeConfig*.

10.4.2.1 Subband Domain Temporal Processing (STP) for N-N/2-N configuration

The subband domain temporal processing tool is applied as described in 6.7 with the following modifications for the N-N/2-N configuration.

The downmix of the spatial upmix is computed as described in 6.7.3, using the following definition for the direct downmix signals.

For N-N/2-N configuration, (*NumInCh-NumLfe*) direct downmix signals are obtained as follows:

$$\hat{z}_{\text{direct, d}}^{n,sb} = \sum_{ch \in ch_d} \tilde{z}_{\text{direct},ch}^{n,sb} \quad ,0 \le d < \left(NumInCh - NumLfe\right)$$

where ch_d comprises the pair-wise output channels depending on the *d* value according to the output channel ordering in Table 10.5 for the N-N/2-N configuration except for the pair with Lfe. It can be defined as:

Table 10.6 — Defining ch_d for N-N/2-N configuration

Configuration	ch_d
N-N/2-N	$\left\{ch_0, ch_1\right\}_{d=0}, \left\{ch_2, ch_3\right\}_{d=1}, \ldots, \left\{ch_{2d}, ch_{2d+1},\right\}_{d=NumInCh-NumLfe}$

The broadband envelopes of the downmix and the envelopes of the diffuse signal portion of each upmix channel are estimated as described in ISO/IEC 23003-1:2007, 6.7.4 using the following definition for the normalized direct energy.

For N-N/2-N configuration, since there are (*NumInCh-NumLfe*) direct signals, $E_{\text{direct\_norm, d}}$ with $0 \le d < \left(NumInCh - NumLfe\right)$ can be obtained in a similar manner as obtained for 5-1-5 configuration.

The scale factors for the final envelope processing are obtained as described in 6.7.5, using the following definition:

For N-N/2-N configuration

$$scale_{ch}^{n} = \sqrt{\frac{E_{\text{direct\_norm, d}}^{n}}{E_{\text{diffuse\_norm},ch}^{n} + \varepsilon}} \quad ,ch \in \left\{ch_{2d}, ch_{2d+1}\right\}_d$$

with $0 \le d < \left(NumInCh - NumLfe\right)$.

10.4.2.2 Guided Envelope Shaping (GES) for N-N/2-N configuration

The Guided Envelope Shaping tool is applied as described in 6.8 with the following modifications for the N-N/2-N configuration.

Similar to Table 10.7, the output channel order for the N-N/2-N configuration is defined as:

Table 10.7 — Output channels ch_{output} for N-N/2-N configuration

Configuration	ch_{output}
N-N/2-N	$0 \le ch_{out}$ <2(NumInCh-NumLfe)

Similar to Table 10.8, the input channel order for the N-N/2-N configuration is defined as:

Table 10.8 — Input channels ch_{input} for N-N/2-N configuration

Configuration	ch_{input}
N-N/2-N	$0 \le ch_{input} <$ (NumInCh-NumLfe)

Similar to Table 10.9 — Downmix $Dch\left(ch_{\text{ouput}}\right)$ for various configurations the downmix channel mapping function for the N-N/2-N configuration is defined as:

Table 10.9 — Downmix for N-N/2-N configuration

Configuration	bsTreeConfig	$Dch\left(ch_{\text{ouput}}\right)$
N-N/2-N	7	$Dch\left(ch_{\text{ouput}}\right) = d \quad ,if \quad ch_{\text{ouput}} \in \left\{ch_{2d}, ch_{2d+1}\right\}_d$ with: $0 \le d <$ (NumInCh-NumLfe)

10.5 Definition of pre-matrix M1

10.5.1 Introduction

The definition of pre-matrix $\mathbf{M}_1^{n,k}$ for N-N/2-N tree configuration is identical to ISO/IEC 23003-1:2007, 6.5.2.1. The following subclauses additionally define the matrices $\mathbf{R}_1^{l,m}$, $\mathbf{G}_1^{l,m}$, and $\mathbf{H}^{l,m}$ for N-N/2-N configuration.

10.5.2 Calculation of R1 for N-N/2-N

10.5.2.1 Introduction

The $\mathbf{R}_1^{l,m}$ matrix controls the amount of input to the decorrelators. For the case of N-N/2-N configuration, all of input channels are two channel based coupled in order to feed OTT modules so that no OTT module is cascaded. Therefore the number of OTT modules is equal with N/2 for this configuration. In this case, $\mathbf{R}_1^{l,m}$ is only depending on the number of OTT modules equal with the column size of the input vector $\mathbf{x}^{n,k}$. But, Lfe upmix based OTT is not considered since it does not need decorrelators. All elements in $\mathbf{R}_1^{l,m}$ are either 1 or 0.

10.5.2.2 N-N/2-N configuration

For the N-N/2-N configuration $\mathbf{R}_1^{l,m}$ is defined according to:

$$\mathbf{R}_1^{l,m} = \begin{bmatrix} \mathbf{I}_{NumInCh} \\ \hline \mathbf{I}_{NumInCh-NumLfe} \end{bmatrix}, \quad 0 \le m < M_{\text{proc}}, 0 \le l < L$$

In this configuration, all the OTT boxes represent parallel processing stages and no OTT box can be connected with any other OTT boxes. Thus $\mathbf{R}_1^{l,m}$ consists of two unity matrices $\mathbf{I}_{NumInCh}$ and $\mathbf{I}_{NumInCh-NumLfe}$. For instance, the unity matrix $\mathbf{I}_N$ is unity matrix of size N by N.

10.5.3 Calculation G1 for N-N/2-N with no external downmix compensation $bsArbitraryDownmix = 0$

If no external downmix compensation is applied, the $\mathbf{G}_1^{l,m}$ matrix is defined according to the following.

For the N-N/2-N configuration it is defined according to:

$$\mathbf{G}_1^{l,m} = \left[\mathbf{I}_{NumInCh} \mid \mathbf{O}_{NumInCh} \right]$$

where, $\mathbf{O}_{NumInCh}$ is null matrix of size $NumInCh$ by $NumInCh$.

10.5.4 Calculation G1 for N-N/2-N with external downmix compensation $bsArbitraryDownmix = 1$

If external downmix compensation is applied, the $\mathbf{G}_1^{l,m}$ matrix is defined according to the following.

For the N-N/2-N configuration it is defined according to:

$$\mathbf{G}_1^{l,m} = \left[\underbrace{\begin{matrix} g_0^{l,m} & 0 & \cdots & 0 & 0 \\ 0 & g_1^{l,m} & 0 & \cdots & 0 \\ \vdots & 0 & \ddots & 0 & \vdots \\ 0 & \cdots & 0 & g_{NumInCh-2}^{l,m} & 0 \\ 0 & 0 & \cdots & 0 & g_{NumInCh-1}^{l,m} \end{matrix}}_{NumInCh \times NumInCh} \quad \mathbf{O}_{NumInCh} \right]$$

where

$$g_X^{l,m} = \mathbf{G}\left(X, l, m\right), \quad 0 \le X < NumInCh, \; 0 \le m < M_{\text{proc}}, \; 0 \le l < L$$

10.5.5 Calculation G1 for N-N/2-N with residual coding based external downmix compensation *bsArbitraryDownmix* = 2

For the N-N/2-N configuration it is defined according to:

$$\mathbf{G}_1^{l,m} = \begin{cases} \left[\underbrace{\begin{matrix} \alpha \cdot g_0^{l,m} & 0 & \cdots & 0 & 0 \\ 0 & \alpha \cdot g_1^{l,m} & 0 & \cdots & 0 \\ \vdots & 0 & \ddots & 0 & \vdots \\ 0 & \cdots & 0 & \alpha \cdot g_{NumInCh-2}^{l,m} & 0 \\ 0 & 0 & \cdots & 0 & \alpha \cdot g_{NumInCh-1}^{l,m} \end{matrix}}_{NumInCh \times NumInCh} \quad \mathbf{I}_{NumInCh} \right], & m \leq m_{\mathrm{ArtDmxRes}}(i) \\ \left[\underbrace{\begin{matrix} g_0^{l,m} & 0 & \cdots & 0 & 0 \\ 0 & g_1^{l,m} & 0 & \cdots & 0 \\ \vdots & 0 & \ddots & 0 & \vdots \\ 0 & \cdots & 0 & g_{NumInCh-2}^{l,m} & 0 \\ 0 & 0 & \cdots & 0 & g_{NumInCh-1}^{l,m} \end{matrix}}_{NumInCh \times NumInCh} \quad \mathbf{O}_{NumInCh} \right], & \text{otherwise} \end{cases}$$

where

$$g_X^{l,m} = \mathbf{G}\left(X, l, m\right), \quad 0 \leq X < NumInCh, \; 0 \leq m < M_{\mathrm{proc}}, \; 0 \leq l < L$$

and the value of α is updated as for the previous configurations.

10.5.6 Calculation H matrix for N-N/2-N for Matrix compatibility

For N-N/2-N configurations, the number of downmix channels is normally larger than five. Therefore, according to description of ISO/IEC 23003-1:2007, 6.5.2.4, the inversion matrix **H** defaults to the unity matrix with a size that is equal to the number of columns in input signal vector $\mathbf{x}^{n,k}$ for all parameter sets and processing bands.

10.6 Definition of mix-matrix M2 for N-N/2-N

10.6.1 Introduction

The definition of post-matrix $\mathbf{M}_2^{n,k}$ for N-N/2-N tree configuration is identical to ISO/IEC 23003-1:2007, 6.5.3.1. The following subclause additionally defines the matrices $\mathbf{R}_2^{l,m}$ for the N-N/2-N configuration.

10.6.2 N-N/2-N configuration

The $\mathbf{R}_2^{l,m}$ matrix for the N-N/2-N configuration is defined according to:

$$\mathbf{R}_2^{l,m} = \begin{bmatrix} \begin{bmatrix} H11_{OTT_0}^{l,m}(n) & H12_{OTT_0}^{l,m}(n) \\ H21_{OTT_0}^{l,m}(n) & H22_{OTT_0}^{l,m}(n) \end{bmatrix} & \mathbf{O}_2 & \cdots & \mathbf{O}_2 \\ \mathbf{O}_2 & \ddots & & \\ \vdots & & \begin{bmatrix} H11_{OTT_i}^{l,m}(n) & H12_{OTT_i}^{l,m}(n) \\ H21_{OTT_i}^{l,m}(n) & H22_{OTT_i}^{l,m}(n) \end{bmatrix} & \vdots \\ & & \ddots & \mathbf{O}_2 \\ \mathbf{O}_2 & \cdots & \mathbf{O}_2 & \begin{bmatrix} H11_{OTT_{numOttBoxes-1}}^{l,m}(n) & H12_{OTT_{numOttBoxes-1}}^{l,m}(n) \\ H21_{OTT_{numOttBoxes-1}}^{l,m}(n) & H22_{OTT_{numOttBoxes-1}}^{l,m}(n) \end{bmatrix} \end{bmatrix}$$

where, the matrix elements are specified using the definition of arbitrary matrix elements $\mathrm{H}11_{OTT_X}^{l,m} \cdots H22_{OTT_X}^{l,m}$ as outlined in ISO/IEC 23003-1:2007, 6.5.3.2, for which,

$$CLD_X^{l,m} = \mathbf{D}_{\mathrm{CLD}}(X,l,m)$$
$$ICC_X^{l,m} = \mathbf{D}_{\mathrm{ICC}}(X,l,m),$$

for $0 \le X < NumInCh$, $0 \le m < M_{\mathrm{proc}}$, $0 \le l < L$. The OTT_X indexing corresponds to the labels of the OTT boxes as given in Figure 23.

10.7 Decorrelators for N-N/2-N configuration

The decorrelators as depicted in Figure 10.1 for N-N/2-N configuration are implemented in the same manner in ISO/IEC 23003-1:2007, 6.6.

In order to ensure orthogonal decorrelated signals, different decorrelators are derived from different filter coefficients, as given in Table A.26 to Table A.29 for the decorrelators $X = 0,\ldots,9$ in ISO/IEC 23003-1:2007. For the N-N/2-N configuration, N/2 decorrelators are needed but given number of decorrelators are limited to 10. It means if the N/2 which is identical to the number of OTT modules is exceed ten without Lfe mode the decorrelators are reused corresponding to the exceeded number of OTT modules by operation of 10 basis modulo.

Table 10.10 — Decorrelator-index as a function of decoder configuration

	Decorrelator $X = 0,\ldots,rem(N/2-1,10)$								
configuration	0	1	2	...	9	10	11	...	N/2-1
N-N/2-N	$D_0^{\mathrm{OTT}}(\ )$	$D_1^{\mathrm{OTT}}(\ )$	$D_2^{\mathrm{OTT}}(\ )$	...	$D_9^{\mathrm{OTT}}(\ )$	$D_0^{\mathrm{OTT}}(\ )$	$D_1^{\mathrm{OTT}}(\ )$	...	$D_{\mathrm{mod}(N/2-1,10)}^{\mathrm{OTT}}(\ )$

10.8 Support for lower and higher sampling frequencies

The downsampled and upsampled operation modes of MPEG Surround shall not be used in combination with an MPEG-H 3D Audio coded downmix signal.

ICS 35.040.40

Price based on 13 pages