

Recommandation Technique

CST - RT – 040 - TV – v1.2 – 2016

Recommandation Technique PAD fichiers Editeurs (CST/FICAM/HDFORUM)

Ce document spécifie les contraintes techniques applicables à la fabrication de fichiers HD pour acceptation PAD (Prêt A Diffuser) des programmes livrés aux chaînes de télévision en France

Version	Date de publication
V 1.1	18 août 2016
V 1.2	13 décembre 2016

SOMMAIRE

1 – OBJET

- 1.1 – Préambule
- 1.2 – Périmètre couvert par le présent document
- 1.3 – Règles typographiques
- 1.4 – Normes et recommandations utilisées dans ce document
 - 1.4.1 – Container et formats des essences audio et vidéo
 - 1.4.2 – Vidéo
 - 1.4.3 – Audio

2 – SPECIFICATIONS TECHNIQUES DES SIGNAUX

2.1 – Spécifications techniques de l'essence vidéo

- 2.1.1 – *Standard de fabrication*
- 2.1.2 – *Suppression verticale*
- 2.1.3 – *Spécifications métrologiques*
- 2.1.4 – *Qualité subjective*

2.2 – Spécifications techniques de l'essence audio

- 2.2.1 – *Conditions de mesure de l'écoute*
 - 2.2.1.1 – Types d'enceintes acoustiques, positionnement
 - 2.2.1.2 – Calibration des écoutes
 - 2.2.1.3 – Niveau d'écoute
 - 2.2.1.4 – Mesure des niveaux
- 2.2.2 – *Caractéristiques des signaux*
 - 2.2.2.1 – Niveau d'alignement
 - 2.2.2.2 – Niveau maximum des crêtes des programmes (PCM ou Dolby E)
 - 2.2.2.3 – Dynamique et niveau d'intensité sonore (sonie, loudness)
 - 2.2.2.4 – Synchronisation audio/vidéo
 - 2.2.2.5 – Format audio numérique
- 2.2.3 – *Qualité subjective du son*
- 2.2.4 – *Fiche technique de mixage*
- 2.2.5 – *Spécificités pour le format PCM*
 - 2.2.5.1 – Mono
 - 2.2.5.2 – Stéréo Lo/Ro et Lt/Rt
- 2.2.6 – *Spécificités pour le format Dolby E*
 - 2.2.6.1 – Mise en phase des trames Dolby E et vidéo HD à 25 images/seconde
 - 2.2.6.2 – Affectation des canaux au sein d'un flux Dolby E
 - 2.2.6.3 – Compatibilité de la réduction stéréo (*downmix*)
 - 2.2.6.4 – Métadonnées (*metadata*) du Dolby E

3 – LIVRABLES

3.1 – Généralités

- 3.1.1 – *Dénomination des fichiers*
- 3.1.2 – *Identification du support physique*
- 3.1.3 – *Livraison dématérialisée*

3.2 – Format de fichier

- 3.2.1 – *Format d'encapsulation MXF*
- 3.2.2 – *Format audio*
- 3.2.4 – *Format des sous-titres*
- 3.2.5 – *Format du fichier de métadonnées*

4 – GLOSSAIRE

4.1 – Termes techniques

4.2 – Organismes

1 - OBJET

1.1 - PREAMBULE

Ce document est issu du travail commun des Editeurs du HDForum, de la CST et de la FICAM. Il définit des recommandations communes aux différents Editeurs, concernant la nature des signaux audio, vidéo, container et métadonnées, à respecter lors de la livraison sous forme « fichier » de contenus HD (Haute Définition) PAD (Prêts à Diffuser). Il constitue de fait un document de référence, mais n'exclut aucunement l'ajout de spécificités propres à l'Editeur.

1.2 - PERIMETRE COUVERT PAR LE PRESENT DOCUMENT, CONDITIONS DE MODIFICATIONS

Le périmètre couvert par le présent document se restreint à la description des signaux audio et vidéo encapsulés dans un container MXF. Il constitue une étape de mutualisation des recommandations PAD Editeurs.

Pour les spécifications techniques audio et vidéo, le périmètre de ce document s'applique à tout élément de programme délivré aux Editeurs : auto-promo, parrainage, téléfilm, documentaire, série, jingle, publicité, films, court-métrage, directs, flux, etc.

Les demandes des Editeurs concernant le contenu et la nature des supports de livraison ne sont pas mentionnées dans ce document, excepté ce qui concerne les signaux d'identification, car il a été décidé que ce domaine devait être défini spécifiquement par chaque Editeur en fonction de ses propres besoins et des cadres contractuels propres à chaque chaîne et à chaque programme.

L'amendement de ces recommandations (par incrément du numéro de version) est effectué selon le type et la nature des modifications apportées au présent document. Si le numéro de version est de la forme A.B :

- A sera incrémenté de 1 dans le cas d'ajouts ou modifications fonctionnels majeurs impactant le produit livré.
- B sera incrémenté de 1 dans le cas de modifications correctives, ajouts de précisions, etc.

Tout amendement du présent document devra faire l'objet d'une validation par l'ensemble des parties rédactrices de la présente version.

1.3 - REGLES TYPOGRAPHIQUES :

Les passages en italique indiquent :

- Des recommandations sur des valeurs subjectives faisant appel au bon sens de chacun.
- Les titres des paragraphes de niveau 3 et 4
- Les dénominations des champs de Meta data

Les passages en **gras** indiquent :

- Les titres des paragraphes
- Les références à d'autres documents normatifs ou recommandations dont le contenu devra impérativement être respecté dans le cadre du présent document.
- Des éléments sur lesquels une attention particulière sera portée par les différentes parties pour différentes raisons (ex : nouveaux éléments de la norme impliquant des changements d'habitudes de travail).

1.4 - REFERENCES DES NORMES ET RECOMMANDATIONS UTILISEES DANS LE DOCUMENT :

1.4.1 – Container et format des essences audio et vidéo

L'ensemble des références concernant le format « *AS-10 MXF for Production Specification* », dans sa version courante, est intégré au document disponible sur le site Web de l'AMWA (<http://www.amwa.tv>).

1.4.2 - Vidéo

- SMPTE 274M : Paramètres signal HD 1920x1080 <http://www.smpte.org>
- ITU-R BT.709 : Paramètres de codage vidéo HD
- ITU-R BT.1886 : Paramètres de calibration des moniteurs
- EBU R 103 : Tolérances sur le signal vidéo dans les systèmes de télévision numérique (v2.0 – Juin 2016)

Le signal vidéo correspond aux caractéristiques 1080i (entrelacé) à 25 images/seconde de ces normes, quel que soit le format du signal vidéo d'origine. Les formats 1080/25p et 1080/25PsF sont **proscrits**.

1.4.3 - Audio

- ITU-R BS. 1770-4 : Algorithme de mesure audio Leq RLB
- ITU-R BS.775 : Disposition des haut-parleurs 5.1
- UER-R91-1998 : Allocation des canaux en audio 5.1
- EBU-R128 et R128s1: normalisation de l'intensité sonore (ou sonie, ou loudness) et niveau maximum permis
- EBU-Tech 3341, 3342, 3343, 3344 : recommandations techniques relatives à l'EBU R128 et R128s1
- EBU-Tech 3276 et 3276s1 : calibration des systèmes d'écoute
- EBU Tech 3304 : Signaux de test audio 5.1
- SMPTE 299M : Audio 24-bit pour HD TV
- IEC 60268-5 : Ecoutes audio

2 - SPECIFICATIONS TECHNIQUES DES SIGNAUX PAD

2.1 - SPECIFICATIONS TECHNIQUES DE L'ESSENCE VIDEO

2.1.1 - Standard de fabrication

Sauf mention contraire de la part de l'éditeur, le standard de fabrication du fichier Prêt à Diffuser est :

1080i – 25 images/seconde - 4:2:2 - ratio image 16/9

Ceci n'interdit aucunement la production et la post-production dans un autre standard.

2.1.2 - Suppression verticale

Le tableau suivant spécifie le positionnement vertical des lignes actives dans le signal vidéo suivant le format de l'image. Une tolérance de +/- 2 lignes est admise :

Format image	Vidéo 16/9		
	Ligne de début	Ligne de fin	Nombre de lignes
1,33	21	560	540
1,37	21	560	540
1,66	21	560	540
1,77	21	560	540
1,85	32*	550*	519*
2,35	87*	495*	409*
2,39	90*	492*	403*

* : Valeur qui peut être modifiée sur demande de recadrage spécifique validée par le ou les ayants-droits

Horizontalement, le SAV (Start of Active Video) doit donc être immédiatement suivi d'un contenu utile. Un écart de 0,1 µs entre le SAV et le début du contenu utile est toléré. Pour un programme donné, cet écart, s'il existe, sera constant sur la durée du programme.

2.1.3 - Spécifications métrologiques

Le tableau suivant spécifie les tolérances admises par les Éditeurs sur les valeurs numériques des composantes vidéo HD.

Conformément à la recommandation EBU R 103 (v2.0 juin 2016), et afin de garantir le respect de l'espace colorimétrique (« Color Gamut »), il est recommandé que les valeurs numériques des composantes RGB, ainsi que la valeur Y de luminance correspondante, se situent impérativement entre les valeurs minimum et maximum, soit 5-246 en 8 bits (voir le tableau ci-dessous).

Système Profondeur d'analyse (en bits)	Amplitude en valeurs (code) d'échantillons numériques		
	Amplitude attendue par composante	Valeurs minimum et maximum acceptées	Amplitude totale du signal vidéo
8 bits	16 – 235	5 - 246	1 - 254
10 bits	64 – 940	20 – 984	4 - 1019
12 bits	256 – 3 760	80 – 3 936	16 – 4 079
16 bits	4 000 – 60 160	1 280 – 62 976	256 – 65 279

La proportion de pixels dont les valeurs sont hors « Color Gamut » (voir tableau ci-dessus) ne doit pas excéder 1% du nombre total de pixels de l'image active.

2.1.4 - Qualité subjective image

- *Pour les images issues d'un télécinéma, une attention particulière sera apportée pour traiter les défauts de copie film et de télécinéma (exemptes de rayures, poussières, éclats de gélatines, défauts de déclenchement de correction colorimétrique...)*
- *De façon générale, l'image ne devra pas comporter un niveau excessif de bruit, de défauts de moiré, d'artefacts de compression, de défauts liés à l'utilisation excessive d'un procédé de réduction de bruit, etc.*
- *Les scènes les plus sombres du programme devront dans tous les cas avoir un rendu acceptable lors de la visualisation sur un téléviseur « grand public ». Les noirs ne devront pas être collés ou écrasés et les détails devront rester perceptibles dans les zones sombres.*
- *Les génériques de fin devront être lisibles en vidéo.*
- *Une attention particulière sera portée aux conversions de quantification, notamment dans le cadre de tournages réalisés à l'aide de moyens typés "cinéma" (codecs 10 à 16 bits), lors de la conformation du signal vidéo 8 bits destiné au fichier Prêt A Diffuser.*

2.2 - SPECIFICATIONS TECHNIQUES DE L'ESSENCE AUDIO

OBJECTIFS

Les contraintes et les objectifs de la diffusion sonore en télévision ayant servi de préalable à la rédaction de ce document sont les suivants :

- Respecter l'œuvre, notamment dans ses choix artistiques
- Garantir le confort d'écoute chez le téléspectateur par la parfaite intelligibilité du message sonore dans les différentes configurations (multicanal tous types dont 5.1, downmix, stéréo, mono)
- Garantir une continuité de la perception du niveau sonore dans l'enchaînement des différents programmes
- Optimiser l'utilisation des métadonnées inhérentes aux codages *Dolby Digital* et *Dolby Digital Plus*.

CONSIDERATIONS GENERALES

Tous les travaux de conformation des mixages aux dispositions de la présente recommandation doivent être réalisés sous l'autorité des ayant-droits ou par tout prestataire désigné ou mandaté par eux. Le travail fourni doit être satisfaisant pour tous les intervenants (producteur, post producteur, Editeurs).

L'utilisation de solutions de compression dynamique du signal ne doit pas aboutir à des niveaux d'énergie constants, sans « aération » du signal sonore dans le temps.

UNITES

LUFUS signifie Loudness Unit relative à une échelle Full Scale. Une valeur en LUFUS est l'expression d'une mesure pondérée K (Leq(R2LB)) sur une échelle Full Scale.

L'échelle LUFUS est graduée par pas de 1 dB. Appliquer un gain de + 1 dB au signal mesuré entraînera la lecture d'un gain de + 1 dB sur l'échelle LUFUS.

CONDITIONS D'ECOUTE ET DE MIXAGE

Le mixage de ces bandes sonores doit être réalisé dans un auditorium dont les équipements techniques et les dimensions sont adaptés à l'écoute télévisuelle (écoute de proximité). Par exemple, les auditoriums de type cinéma de grandes dimensions ne sont pas adaptés pour ces mixages, et sont donc déconseillés.

On trouvera la définition du niveau d'écoute pour la « validation » des bandes sonores des PAD dans une salle de vérification dans les dispositions des recommandations EBU – Tech 3276 et 3276 s1.

Le mixeur pourra ajuster son niveau d'écoute (niveau de pression acoustique SPL), validé avec un bruit rose référencé (EBU Technical recommandation R68), en fonction de la salle et du système d'écoute proposé, autour de la valeur 79 dB (C) par canal (hors LFE).

2.2.1 - Conditions de mesure et d'écoute

2.2.1.1 - Types d'enceintes, positionnement

Les installations de mixage, de « mastering », d'encodage et de contrôle qualité audio doivent être équipées d'enceintes conformes à la norme **IEC 60268-5**.

Le positionnement des enceintes doit être conforme à la norme **ITU-R BS.775**.

2.2.1.2 - Calibration des écoutes

La calibration du système d'écoute doit se faire relativement aux stipulations de la recommandation EBU Tech 3276 s1.

2.2.1.3 - Niveau d'écoute

Les programmes sont contrôlés sur des systèmes d'écoute dont le niveau d'alignement, pour un bruit rose diffusé à -18dBFS, est de 74 dB(C) SPL, sur chaque canal, excepté le canal LFE diffusé à 78dB(C) SPL.

Ce niveau est considéré comme étant celui auquel la dynamique globale, l'intelligibilité des dialogues et les timbres sont jugés conformes à une diffusion télévisuelle.

2.2.1.4 – Mesure des niveaux

Niveau crête

Le contrôle de niveau, exprimé en dBTP, est assuré par un crête-mètre instantané «True Peak»

Recommandations en vigueur :

- **SMPTE 299M** : "24-Bit Digital Audio Format for SMPTE 292 M Bit-Serial Interfaces".
- **ITU-R BS.1770-4**

Niveaux d'énergie (Loudness)

Historique

Suite aux travaux menés depuis 2000 dans le cadre des groupes de travail Ficam/CST sur le confort d'écoute et les niveaux sonores, complétés par les travaux du Département Son de la CST et du HD Forum, il est proposé une méthodologie de mesure du niveau d'énergie (« Loudness ») permettant de valider la valeur de ce niveau pour tous types de programmes destinés à la diffusion télévision, quel que soit le format (PCM, Dolby E, etc.) et quel que soit le type de mixage (multi canal, stéréo, mono, etc.).

Cette valeur permettra également de renseigner la ou les métadonnées d'éventuels encodeurs.

Méthodologie de mesure

La méthodologie de mesure retenue consiste à mesurer la totalité du programme utile à l'aide d'un équipement implémentant l'algorithme décrit dans la recommandation R 128 éditée par l'EBU, ainsi que dans ses annexes 3341, 3342, 3343 et 3344 et délivrant une valeur en dB LUFS.

2.2.2 - Caractéristiques des signaux

2.2.2.1 - Niveau d'alignement

Le niveau d'alignement lu sur un crête mètre numérique Full Scale est de -18dB FS pour un signal sinusoïdal à la fréquence de 1000 Hz.

Un signal d'alignement 1000 Hz à -18 dBFS doit afficher un niveau de Loudness de -18 LUFS sur un équipement de mesure en EBU Mode, si le signal est présent sur les canaux gauche et droite d'un programme stéréo.

Si le signal n'est présent que sur un seul canal, le niveau affiché sera de -21 LUFS.

Pour mémoire, en analogique : le niveau d'alignement lu sur PPM quasi-crête 10ms DIN 45406 est de -9 dB et le niveau d'alignement est lu à 0 Vu sur un Vu mètre.

2.2.2.2 - Niveaux maximum des crêtes des programmes (PCM ou Dolby E)

Le niveau des crêtes du signal audio, mesuré en True Peak, ne doit pas dépasser -3,0 dBTP.

2.2.2.3 – Dynamique et niveau d'intensité sonore (sonie, loudness)

Pour chaque version de mixage (française, multilingue, audiodécrite ...), le tableau suivant résume les recommandations sonores à l'attention des fournisseurs.

Programme court : un programme dont la durée est inférieure ou égale à 2 minutes.

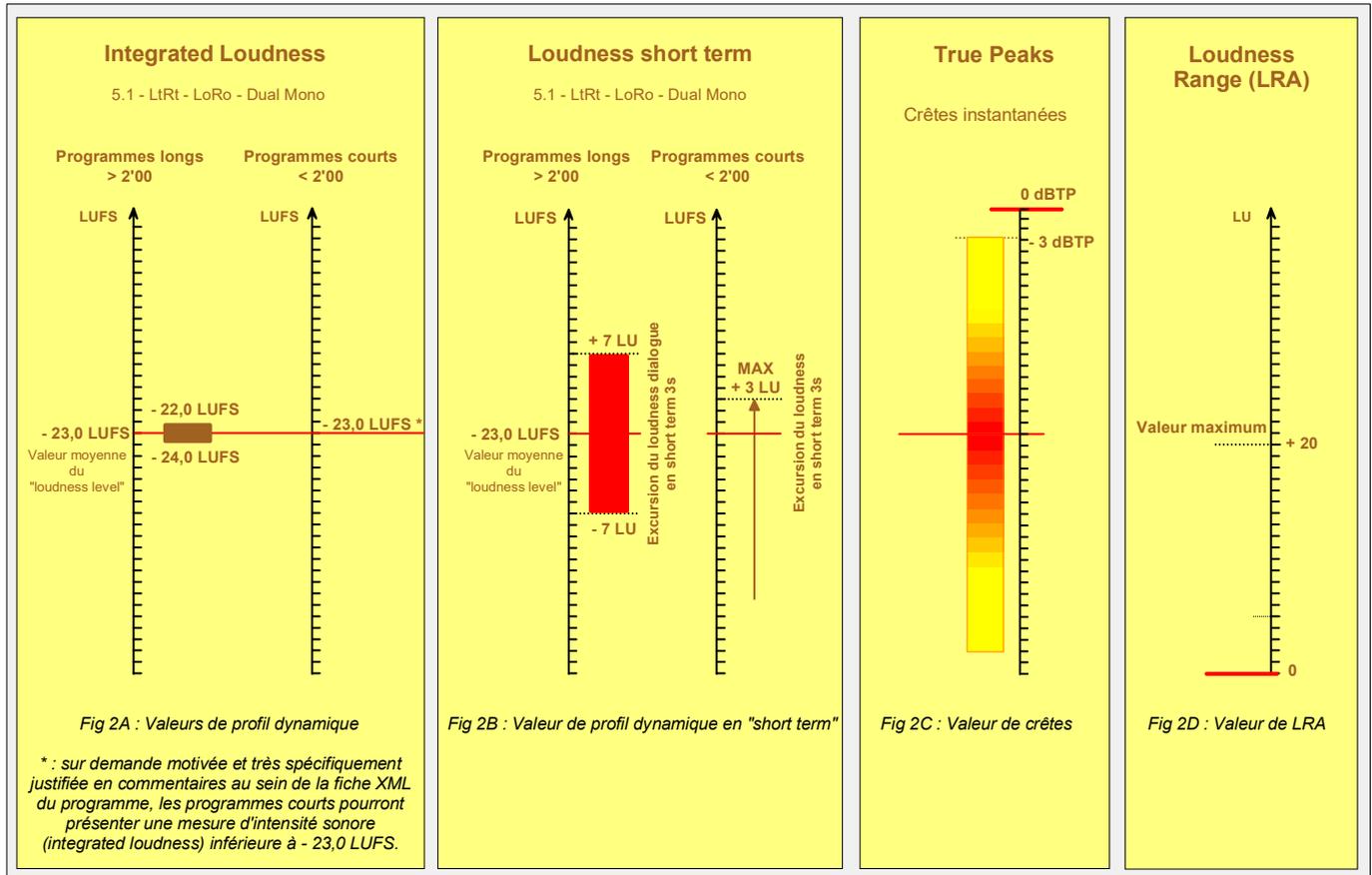
Programme long : un programme dont la durée est supérieure à 2 minutes.

	<i>Programmes courts</i>	<i>Programmes longs</i>
<i>Mesure du niveau d'intensité sonore sur le programme complet (Integrated loudness)</i>	-23,0 LUFS	-23,0 LUFS ± 1,0 LU
<i>Excursion dynamique (Loudness Range ou LRA)</i>	Non mesurée	≤ 20,0 LU Il est conseillé de dépasser une valeur de 5 LU pour conserver une bonne aération de la dynamique
<i>Niveau maximum de la mesure d'intensité sonore court terme (Max Short Term Loudness)</i>	≤ -20,0 LUFS	Non mesuré
<i>Plage d'excursion de l'intensité sonore en Short Term des dialogues autour de la valeur cible du Loudness intégré</i>	Non mesurée	-30,0 ≤ Short term des dialogues ≤ -16,0 en LUFS
<i>Valeur maximum des crêtes instantanées (True peaks)</i>	-3,0 dB TP	-3,0 dB TP

Afin de maintenir une intelligibilité suffisante à destination des téléspectateurs et prendre en compte les variations importantes de leurs conditions d'écoute, la mesure d'intensité sonore court terme, sur les passages du programme comportant des dialogues, devra cibler le niveau moyen de -23,0 LUFS et s'autoriser une excursion maximale de 7 LU autour de ce niveau cible : -30,0 LU correspondant à une voix murmurée, et -16,0 LU à une voix criée. Le respect du niveau cible des dialogues doit permettre d'alimenter indistinctement les diffusions linéaires et non linéaires (services de rattrapage) des programmes, un contexte multi-écrans entraînant un nombre croissant de situations d'écoute au casque, ou au moyen de haut-parleurs miniatures tels que ceux équipant ordinateurs personnels, tablettes et téléphones mobiles.

Le schéma suivant fixe les principaux repères sur les indicateurs d'outils de mesure de l'intensité sonore (à gauche) et du niveau maximum des crêtes (à droite).

Sur demande motivée et très spécifiquement justifiée en commentaire au sein de la fiche XML du programme, les programmes courts pourront présenter une mesure d'intensité sonore en Integrated Loudness inférieure à -23,0 LUFS.



2.2.2.4 - Synchronisme audio/vidéo

Les signaux vidéo et audio sont synchrones entre eux dans la tolérance suivante:

- Soit une avance maximale du son sur l'image de 20ms
- Soit un retard maximum du son sur l'image de 40ms.

Par défaut, les pistes encodées en Dolby E sont enregistrées synchrones avec la vidéo (In-sync). Il en résultera 1 (une) image de retard de l'audio par rapport à la vidéo après décodage Dolby E sans compensation de délai vidéo.

2.2.2.5 - Format Audio Numérique

Les canaux audio doivent être livrés au format numérique suivant :

- Fréquence d'échantillonnage : 48 kHz
- Quantification linéaire : 24 Bits pour le format PCM et 20 Bits avant encodage Dolby E
- Les formats Dolby E doivent être exempts d'erreur CRC (Cyclic Redundance Code)

2.2.3 - Qualité subjective

La bande son doit être naturellement équilibrée et réalisée conformément aux règles de l'art actuellement en vigueur. A l'écoute en environnement calme et sur un système de reproduction ayant une bande passante plate, on ne doit percevoir aucun artefact : clicks de désynchronisation des horloges, dialogues avec sifflantes, pleurage ou ronflements aux basses fréquences (buzz).

Les conversions de format à variation de vitesse (24 à 25 images/seconde) nécessitent généralement une harmonisation afin de rester dans la tonalité musicale d'origine. Les procédés d'harmonisation sur des signaux 5.1

sont actuellement aux limites de la technologie et s'avèrent dans certains cas terriblement destructeurs. L'harmonisation est donc une opération délicate qui doit être réécoutée et validée par une personne de l'art.

Lors de la vérification, l'écoute se fait en stéréo et en multicanal (au besoin via un dématricage Dolby Surround ou un décodage Dolby E). La compatibilité du downmix des programmes multicanaux 5.1 vers stéréo et mono doit être obligatoirement validée en simulant l'action du RF mode Film Standard, a minima avec une vérification par sondage pour les programmes longs et en totalité pour les programmes courts.

INFORMATION IMPORTANTE : en diffusion il arrive fréquemment que la version française, lorsqu'elle s'avère disponible en son multicanal, soit la composante audio principale de la chaîne. Pour les téléspectateurs non équipés d'installation 5.1, c'est cette composante qui sera « downmixée » par le récepteur TNT pour être écoutée en stéréo ou mono.

2.2.4 - FICHE TECHNIQUE DE MIXAGE

Pour les besoins de fabrication du fichier Prêt à Diffuser, chaque livraison d'un mixage doit être accompagnée d'une « Fiche d'informations de mixage » standardisée et fournie en annexe en fin de ce document. Ces éléments permettront à l'opérateur d'inscrire les bonnes métadonnées dans la fiche XML.

2.2.5 - SPECIFICITES POUR LE FORMAT PCM

2.2.5.1 - Mono

En monophonie, le signal enregistré en PCM devra être livré sur deux pistes audio contiguës sans compression de débit numérique. Par souci de compatibilité avec les signaux stéréo, ces deux pistes contenant le signal monophonique seront strictement identiques et en phase.

2.2.5.2 - Stéréo Lo/Ro et Lt/Rt

Affectation des pistes

En stéréophonie, le signal enregistré en PCM devra être livré sur deux pistes audio contiguës sans compression de débit numérique. La piste impaire correspondra au canal gauche et la piste paire au canal droit.

Phase audio

C'est le résultat moyen entre canal gauche et canal droit d'une stéréophonie d'intensité et de phase. Pour assurer la compatibilité du downmix mono, la phase doit être très majoritairement positive.

Spécificités pour les signaux stéréo Lt/Rt Surround

Le matricage Surround en stéréo Lt/Rt devra conserver la cohérence spatiale de l'image sonore et ne pas altérer l'intelligibilité ni le timbre du message sonore. Les programmes dont le mixage d'origine est de type Lt/Rt ne devront en aucun cas être dématricés en 5.0 pour être livrés. De même, ils ne devront pas avoir subi une génération de dématricage puis rematricage Lt/Rt.

2.2.6 - SPECIFICITES POUR LE FORMAT DOLBY E

2.2.6.1 - Mise en phase des trames Dolby E et vidéo HD à 25 images/seconde

Pour permettre le montage vidéo et les commutations des signaux sans artefact, le Dolby E Header sera obligatoirement positionné au delà d'une section exempte de données appelée Guard Band, soit :

- entre le 30ème et le 40ème des 1920 échantillons audio associés à chaque image vidéo,
- en phase avec la ligne 21 (+/- 2) du signal vidéo HD à 25 images/seconde.

(Source : <http://www.dolby.com/us/en/technologies/dolby-e-preferred-alignment.html>)

2.2.6.2 - Affectation des canaux au sein d'un flux Dolby E

Le codage DolbyE prend en compte 4 ou 8 canaux audio selon les conditions suivantes :

- programme disponible uniquement en monophonie ou en stéréophonie : le codage Dolby E est réalisé en configuration 2/0 dans laquelle le son monophonique de la version PCM est dupliqué à l'identique respectivement sur les canaux 1 et 2 du Dolby E, et le son stéréophonique de la version PCM est copié à l'identique canal par canal sur les canaux 1 et 2 du Dolby E,
- programme disponible sur 5 ou 6 canaux mixés respectivement aux formats 5.0 ou 5.1 : le codage Dolby E est réalisé en configuration 3/2 selon l'affectation des pistes de la recommandation **EBU-R91-1998**,
- les spécifications du contenu des canaux 7 et 8 du Dolby E sont laissées à la discrétion des Éditeurs.

Ceci implique d'office la configuration **obligatoire** des métadonnées Dolby E suivantes :

- Program Config = 5.1+2 **ou** 2+2
- Channel Mode = 3/2 **ou** 2/0 pour le programme 1
- Channel Mode = 2/0 pour le programme 2
- Se reporter au paragraphe A.2.6.4 pour les autres paramètres.

Numéro de paire AES (entrées d'encodeur, sorties de décodeur)	Canal du flux Dolby E	configuration 2+2 mixage principal en dual- mono ou stéréo	configuration 5.1+2 mixage principal en 5.1 / 5.0
AES 1 (canaux 1 et 2)	canal 1	canal gauche	canal avant gauche
	canal 2	canal droit	canal avant droit
AES 2 (canaux 3 et 4)	canal 3		canal centre
	canal 4		canal subwoofer
AES 3 (canaux 5 et 6)	canal 5		canal arrière gauche
	canal 6		canal arrière droit
AES 4 (canaux 7 et 8)	canal 7	canal optionnel (à préciser par l'éditeur)	canal optionnel (à préciser par l'éditeur)
	canal 8	canal optionnel (à préciser par l'éditeur)	canal optionnel (à préciser par l'éditeur)

2.2.6.3 - Compatibilité de la réduction stéréo (downmix)

La réduction " 5.1 vers stéréo " (ou downmix) après application des valeurs de gain et de matricage Lt/Rt ou Lo/Ro, paramétrées dans les métadonnées (voir paragraphe suivant), devra :

- conserver la cohérence spatiale de l'image sonore,
- conserver le niveau et l'intelligibilité de la voix,
- conserver l'équilibre entre les éléments du mixage

Les caractéristiques de la phase du signal stéréo issu d'un downmix devront répondre aux mêmes exigences que pour un signal stéréo classique (Cf. §2.2.5).

2.2.6.4 - Métadonnées (metadata) du Dolby E

Typologies des programmes multicanaux (Musique et Autres)

On distinguera deux grandes familles typologiques de programmes multicanaux :

1. les programmes musicaux susceptibles de poser des problèmes en cas de downmix stéréo,
2. tous les programmes autres que musicaux.

Renseignement de la valeur de Dialog Level

Pour tous les programmes mixés conformément aux principes édictés dans cette recommandation, la valeur de *Dialog Level* à inscrire dans les métadonnées sera identique à la mesure d'intensité sonore *Integrated Loudness* effectuée selon la méthodologie décrite dans la recommandation EBU R128.

Cas particulier

Pour certains programmes plus anciens et codés en Dolby E après simple ajustement de leur valeur d'*Integrated Loudness* à -23 LUFS, il est fréquent que le niveau moyen des dialogues diverge sérieusement de cette valeur.

Dans ce cas, afin de maintenir la qualité sonore des phases dialoguées du programme, on renseignera comme métadonnée de Dialog Level la mesure d'intensité sonore rapportée par un équipement utilisant l'algorithme Dialog Intelligence. Dans pareil cas, cette décision sera motivée en commentaire dans la fiche XML du programme.

Métadonnées correspondantes aux programmes 5.1 / 5.0 / 2.0

On distingue trois familles de métadonnées Dolby E à renseigner lors d'un encodage :

1. les métadonnées figées qui dépendent uniquement du paramètre « Channel Mode » (3/2 avec LFE actif pour les programmes en 5.1, 3/2 sans LFE actif pour les programmes en 5.0, 2/0 pour les programmes Stéréo),
2. les métadonnées par défaut pouvant éventuellement être modifiées lors de l'encodage sur demande justifiée à l'aide des fiches de mixage et d'encodage données en annexe,
3. les métadonnées dépendantes uniquement du mixage et pouvant être validées par une mesure.

Toute modification des métadonnées spécifiées dans le tableau ci-après est un motif de refus pour la diffusion si elle n'est pas argumentée dans le cadre d'un commentaire inséré dans la fiche XML du programme et soumise à l'acceptation de l'éditeur.

Note : Le mode Extended BSI doit impérativement être enclenché pour les programmes 5.1 et 5.0. En revanche, il ne doit pas être enclenché pour les formats Stéréo.

Paramètres	Programme 5.1	Programme 5.0	Programme Stéréo (2.0)	Programme mono/dual	Signal de référence
"Dialog level"	- 23,0 LUFS ± 1 (A mesurer)	- 31,0 LUFS			
Program Name	Choix Editeur**	Choix Editeur**	Choix Editeur**	Choix Editeur**	Test_Check
Audio Processing					
Channel Mode	3/2	3/2	2/0	2/0	3/2 ou 2/0
LFE Channel	Enable	Disable	N/A	N/A	5.1: Enable 5.0: Disable 2.0: N/A
DC Filter	Enable	Enable	Enable	Enable	Enable
Low pass Filter	Enable	Enable	Enable	Enable	Enable
LFE Low pass Filter	Enable	Disable	N/A	N/A	5.1: Enable 5.0: Disable 2.0: N/A
Srnd Phase Shift*	Enable	Enable	N/A	N/A	3/2: Enable 2.0: N/A
Srnd 3 dB Attenuation	Disable	Disable	N/A	N/A	3/2: Disable 2.0: N/A
Dyname					
Line Mode Pro	Film LIGHT	Film LIGHT	Film LIGHT	Film LIGHT	None
RF Mode Pro	Film Standard	Film Standard	Film Standard	Film Standard	None
RF Overmod Protect	Disable	Disable	Disable	Disable	Disable
Bitstream info					
Bitstream Mode	Main Complete				
Center Downmix Level	0.707 (-3 dB)				
Srnd Downmix Level	0.707 (-3 dB)				
Dolby Srnd Mode	N/A	N/A	Dolby Surround enabled	Dolby Surround enabled	N/A
Copyright bit	Yes	Yes	Yes	Yes	Yes
Original Bitstream	Yes	Yes	Yes	Yes	Yes
Audio Production Info	Disable	Disable	Disable	Disable	Disable
Mix Level	N/A	N/A	N/A	N/A	N/A
Room type	N/A	N/A	N/A	N/A	N/A
Extended BSI					
Preferred Stereo Downmix*	Lt/Rt	Lt/Rt	N/A	N/A	3/2: Lt/Rt 2.0: N/AN/A
Lt/Rt Center Downmix Level	0.707 (-3 dB)	0.707 (-3 dB)	N/A	N/A	N/A
Lt/Rt Surround Downmix Level	0.707 (-3 dB)	0.707 (-3 dB)	N/A	N/A	N/A
Lo/Ro Center Downmix Level	0.707 (-3 dB)	0.707 (-3 dB)	N/A	N/A	N/A
Lo/Ro Surround Downmix Level	0.707 (-3 dB)	0.707 (-3 dB)	N/A	N/A	N/A
Dolby Surround EX Mode	Not Surround EX	Not Surround EX	N/A	N/A	N/A
A/D Converter Type	Standard	Standard	Standard	Standard	Standard

* Les paramètres Surround Phase Shift et Preferred Downmix sont interdépendants. Se référer au § A.2.6.4.4 pour leur paramétrage.

** Le Program Name renseigné ne doit contenir aucun caractère accentué ni espace.

Figure 2 : Tableau de paramétrage par défaut des métadonnées du Dolby E pour les formats 5.1, 5.0, 2.0, dual et mono

ABCDE	Métadonnées figées qui dépendent uniquement du paramètre « Channel mode »
ABCDE	Métadonnées par défaut pouvant être modifiées moyennant justification (cf. Notes)
ABCDE	Métadonnées dépendantes du mixage

3 - LIVRABLES

3.1 - GÉNÉRALITÉS

Un programme livré sous forme de fichier est constitué, quelle que soit sa durée :

- d'un seul fichier MXF Op1a contenant l'essence vidéo et les essences audio associées,
- d'une fiche de métadonnées au format XML,
- éventuellement d'un ou plusieurs fichiers STL contenant les sous-titres associés,
- éventuellement d'un listing de sous-titrage, du texte original et de la transcription du générique de fin.

Le début de programme (première image et premier son utile) doit commencer obligatoirement à un Time Code horaire entier (par ex. 00:00:00:00, 01:00:00:00, 10:00:00:00). La présence d'éléments d'amorce avant la première image, de même qu'après la dernière image utile du programme (noir, fonds neutres) restent à la discrétion de l'éditeur.

3.1.1 - DENOMINATION DES FICHIERS

Cet identifiant est à la discrétion de l'Éditeur.

3.1.2 - IDENTIFICATION DU SUPPORT PHYSIQUE

Dans le cas où le programme est livré sur un support physique (disque dur, cartouche informatique...), ce support devra comporter une étiquette avec les informations suivantes :

- l'identifiant de livraison
- le titre du programme (et le cas échéant, le sous-titre ou le numéro d'épisode)
- la durée (HH:MM:SS:II)
- le numéro d'Enregistrement Magnétique, ou Numéro d'Affaire, ou Numéro Identifiant

Ces informations seront également reproduites sur une fiche papier accompagnant le support, accompagnées des suppléments d'information suivants :

- l'identification des canaux audio présents
- les commentaires pouvant s'avérer utiles lors de la vérification : éventuels défauts image ou son constatés, problèmes passagers de synchronisation image/son, etc.

3.1.3 - LIVRAISON DEMATERIALISEE

Chaque Éditeur pourra proposer des solutions de livraison dématérialisée de l'ensemble des fichiers constituant le livrable Prêt à Diffuser.

3.2 - FORMATS DE FICHIER

3.2.1 - FORMAT D'ENCAPSULATION MXF

Le fichier multimédia MXF contenant le programme sera conforme au standard AMWA AS10. Le profil de génération suivra le shim "High_HD_2014" (<http://www.amwa.tv/projects/AS-10.shtml>).

Précautions concernant le Time Code

Les différentes informations de Time Code que peut comporter un fichier MXF devront s'avérer cohérentes entre elles et seront obligatoirement présentes dans le "Material Package", "Source Package" et le "System Item" du fichier MXF.

3.2.3 FORMAT AUDIO

Type et nombre de canaux

Les fichiers MXF Prêts à Diffuser comporteront 8 canaux audio, chaque canal distinct étant transporté par une piste AES3 distincte.

Chaque piste audio du fichier MXF transportera un flux de type AES3, de résolution audio 24 bits, à 48 kHz de fréquence d'échantillonnage. Deux pistes/flux AES3 pourront au besoin transporter un flux encodé en Dolby E. L'allocation des 8 pistes du fichier, ainsi que des canaux éventuellement encodés en Dolby E devra être explicitée dans la fiche XML associée suivant les recommandations de chaque Éditeur.

3.2.4 FORMAT DES SOUS-TITRES

Les programmes en langue étrangère sont diffusés en version originale avec sous-titres français. La plupart des programmes comportent également un sous-titrage spécifique à destination des malentendants.

Format de fichier de sous-titrage

Tout programme devant être diffusé avec un sous-titrage sera accompagné d'un ou de plusieurs fichiers de sous-titrage conformes à la norme N19-2002 de l'UER (fichier STL), respectant les consignes suivantes :

- Le nom des fichiers STL sera indiqué par le diffuseur
- Les sous-titres seront synchrones par rapport au Time Code du fichier MXF livré.
- Le Time Code de sortie du dernier sous-titre du programme sera positionné au plus tard 3 secondes avant la dernière image utile du programme pour éviter que ce dernier sous-titre ne reste figé en diffusion sur le programme suivant.
- Les informations générales du "blocGSI" dans le fichier STL (nom du programme, Time Code de début de programme, etc.) seront dûment renseignées. Le tableau suivant fournit des informations complémentaires, l'Éditeur se réservant le droit d'en modifier les paramètres à sa discrétion :

	Mode télétexte	
<i>Titre du programme (OPT)</i>	A remplir	
<i>Version</i>	Français	Sourds et Malentendants
<i>Code de page (CPN)</i>	850	
<i>Format du Time Code (TC)</i>	25 i/s (STL25.01)	
<i>Norme d'affichage (DSC)</i>	Teletext – Présentation de niveau 1 (1)	
<i>Code de caractères (CCT)</i>	Latin (00)	
<i>Langue (LC)</i>	Français (0F)	
<i>Time Code de début de programme (TCP)</i>	Conforme au time code du début de fichier MXF livré	
<i>Taille des caractères</i>	Double hauteur – Simple largeur	
<i>Nombre maximal de caractères par ligne</i>	37	
<i>Couleur des sous-titres</i>	Blanc	Suivant la charte CSA*
<i>Nombre de rangées (MNR)</i>	23	23
<i>Position des sous-titres</i>	2 lignes de texte maximum Centrage horizontal	Suivant la charte CSA*

* <http://www.csa.fr/content/download/20043/334122/file/Chartesoustitrage122011.pdf>

3.2.5 - FORMAT DU FICHER DE METADONNEES

Une fiche de programme au format XML sera obligatoirement livrée avec chaque fichier MXF.

Les Éditeurs pourront fournir un exemple de fichier XML, ainsi qu'un fichier Microsoft Excel « MetaPADTV V1-x » comportant des zones de saisies préétablies pour renseigner les différents champs éditoriaux et techniques qui décrivent le programme.

Ce fichier Excel « MetaPADTV V1-x » comporte des macros permettant de générer le fichier XML.

	A	B	C	L
1				
2		MetaPADTV	Fabriquer le XML / Generate XML	
3				
4				
5		Général / General		Obligatoire / Compulsory
6		Nom du MXF		OUI / YES
10				
11		Description Fichier		
14		Diffuseur / Broadcaster		OUI / YES
15		Fournisseur du fichier / File provider		OUI / YES
16		Nature de programme / Program nature	Programme	OUI / YES
18		HouseID		OUI / YES
21		Titre Programme Local / Local Program Title		OUI / YES
22		Titre Episode Local / Local Episode Title		NON / NO
23		Titre Programme Original / Original Program Title		NON / NO
24		Titre Episode Original / Original Episode Title		NON / NO
25		RevisionID / RevisionID		OUI / YES
32				
33		Informations techniques / Technical informations		
34		TC in Fichier (amorçe) / File TC in (amorçe)	10:00:00:00	OUI / YES
35		TC out Fichier (postroll) / File TC out (postroll)	10:14:59:24	OUI / YES
36		TC in Programme / Program TC in	10:00:00:00	OUI / YES
37		TC out Programme / TC out Program	10:14:59:24	OUI / YES
38		Spécification fichier	AS10	OUI / YES

Les macros du fichier Excel génèrent également automatiquement le nom du fichier MXF à partir des données saisies lors de la génération du XML.

Pour un programme Multi-parties, les Time Codes d'entrée/sortie de chaque partie devront être obligatoirement renseignés dans le fichier XML.

49		Description de la segmentation / Segmentation description		
50		Nombre de parties / Number of parts	2	OUI / YES
51				
52		Description de la partie 1 / Part 1 description		
53		Numéro de partie / Part number	1	OUI / YES
54		TC in (Programme ou message) / TC in (Program or message)	10:00:00:00	OUI / YES
55		TC out (Programme ou message) / TC out (Program or message)	10:04:59:24	OUI / YES
56				
57		Description de la partie 2 / Part 2 Description		
58		Numéro de partie / Part number	2	OUI / YES
59		TC in (Programme ou message) / TC in (Program or message)	10:05:00:00	OUI / YES
60		TC out (Programme ou message) / TC out (Program or message)	10:14:59:24	OUI / YES

4 - GLOSSAIRE DES ACRONYMES

4.1 TERMES TECHNIQUES

ACRONYMES	DÉFINITIONS
AS	Application Specification : rédigée par l'AMWA, une AS correspond à une spécification précise des paramètres techniques d'un format complexe, pour simplifier son interopérabilité en correspondant à des cas d'usages.
CRC	Cycling Redundance Code : code de redondance pour améliorer la robustesse d'une information numérique
EAV	End of Active Video
K ou R2LB	Courbe de pondération dérivée de la courbe B appliquée à une source audio dans le cadre d'une mesure d'intensité sonore
LTC	Longitudinal Time Code
LU ou LUFS	Loudness Unit ou Loudness Unit relative à l'échelle Full Scale : unité de mesure d'intensité sonore définie par la recommandation EBU R128
MXF	Material eXchange Format : format de fichier élaboré dédié à l'échange de médias professionnels
PCM	Pulse Code Modulation : signal audio numérique non compressé, représentation d'un signal analogique dans les limites de sa fréquence d'échantillonnage et de son niveau de quantification numérique.
PPM	Peak Program Meter : type de crête-mètre utilisé pour la mesure d'un signal audio
SAV	Start of Active Video
SDI	Serial Digital Interface : standard pour la transmission d'un signal numérique non compressé de télévision (audio, vidéo, ancillary data)
VANC	Vertical ANCillary data space : information numérique non intégrée dans l'image numérique
VITC	Vertical Internal Time Code

4.2 ORGANISMES

ACRONYMES	DÉFINITIONS
AES	Audio Engineering Society : organisation internationale des ingénieurs du son, émettant des standards dans le domaine du son
AMWA	Advanced Media Workflow Association : organisation spécifiant des usages et recommandations pour la production dématérialisée
CSA	Conseil Supérieur de l'Audiovisuel : organe de régulation des chaînes de télévision en France
CST	Commission Supérieure Technique de l'Image et du Son : association des techniciens de l'image et du son en cinéma et télévision
EBU	European Broadcast Union (Union Européenne de Radio Télévision), organisation européenne regroupant les chaînes de télévision et de radio de service public
FICAM	Fédération des Industries du Cinéma, de l'Audiovisuel et du Multimédia
HDFORUM	Association visant à promouvoir en France l'amélioration de la qualité des services de télévisions. Elle comporte 3 collèges : d'éditeurs, d'industriels et de distributeurs
IEC	International Electrotechnical Commission : organisation émettant des standards pour l'industrie électrique et électronique
ITU	International Telecommunication Union (ou Union Internationale des Télécommunications), en charge des recommandations et de la planification internationale des télécommunications
SMPTE	Society of Motion Picture and Television Engineers : association américaine des ingénieurs de la télévision et du cinéma