

UIT-T

SECTEUR DE LA NORMALISATION
DES TÉLÉCOMMUNICATIONS
DE L'UIT

G.722

Appendice III
(11/2006)

SÉRIE G: SYSTÈMES ET SUPPORTS DE TRANSMISSION, SYSTÈMES ET RÉSEAUX NUMÉRIQUES

Equipements terminaux numériques – Codage des
signaux analogiques par des méthodes autres que la MIC

Codage audiofréquence à 7 kHz à un débit inférieur
ou égal à 64 kbit/s

Appendice III: Algorithme de haute qualité pour le masquage des pertes de paquets pour le vocodeur G.722

Recommandation UIT-T G.722 – Appendice III

RECOMMANDATIONS UIT-T DE LA SÉRIE G
SYSTÈMES ET SUPPORTS DE TRANSMISSION, SYSTÈMES ET RÉSEAUX NUMÉRIQUES

CONNEXIONS ET CIRCUITS TÉLÉPHONIQUES INTERNATIONAUX	G.100–G.199
CARACTÉRISTIQUES GÉNÉRALES COMMUNES À TOUS LES SYSTÈMES ANALOGIQUES À COURANTS PORTEURS	G.200–G.299
CARACTÉRISTIQUES INDIVIDUELLES DES SYSTÈMES TÉLÉPHONIQUES INTERNATIONAUX À COURANTS PORTEURS SUR LIGNES MÉTALLIQUES	G.300–G.399
CARACTÉRISTIQUES GÉNÉRALES DES SYSTÈMES TÉLÉPHONIQUES INTERNATIONAUX HERTZIENS OU À SATELLITES ET INTERCONNEXION AVEC LES SYSTÈMES SUR LIGNES MÉTALLIQUES	G.400–G.449
COORDINATION DE LA RADIODÉLÉPHONIE ET DE LA TÉLÉPHONIE SUR LIGNES	G.450–G.499
CARACTÉRISTIQUES DES SUPPORTS DE TRANSMISSION ET DES SYSTÈMES OPTIQUES	G.600–G.699
EQUIPEMENTS TERMINAUX NUMÉRIQUES	G.700–G.799
Généralités	G.700–G.709
Codage des signaux analogiques en modulation par impulsions et codage	G.710–G.719
Codage des signaux analogiques par des méthodes autres que la MIC	G.720–G.729
Principales caractéristiques des équipements de multiplexage primaires	G.730–G.739
Principales caractéristiques des équipements de multiplexage de deuxième ordre	G.740–G.749
Caractéristiques principales des équipements de multiplexage d'ordre plus élevé	G.750–G.759
Caractéristiques principales des équipements de transcodage et de multiplication numérique	G.760–G.769
Fonctionnalités de gestion, d'exploitation et de maintenance des équipements de transmission	G.770–G.779
Caractéristiques principales des équipements de multiplexage en hiérarchie numérique synchrone	G.780–G.789
Autres équipements terminaux	G.790–G.799
RÉSEAUX NUMÉRIQUES	G.800–G.899
SECTION NUMÉRIQUE ET SYSTÈMES DE LIGNES NUMÉRIQUES	G.900–G.999
QUALITÉ DE SERVICE ET DE TRANSMISSION – ASPECTS GÉNÉRIQUES ET ASPECTS LIÉS À L'UTILISATEUR	G.1000–G.1999
CARACTÉRISTIQUES DES SUPPORTS DE TRANSMISSION	G.6000–G.6999
DONNÉES SUR COUCHE TRANSPORT – ASPECTS GÉNÉRIQUES	G.7000–G.7999
ASPECTS RELATIFS AUX PROTOCOLES EN MODE PAQUET SUR COUCHE TRANSPORT	G.8000–G.8999
RÉSEAUX D'ACCÈS	G.9000–G.9999

Pour plus de détails, voir la Liste des Recommandations de l'UIT-T.

Recommandation UIT-T G.722

Codage audiofréquence à 7 kHz à un débit inférieur ou égal à 64 kbit/s

Appendice III

Algorithme de haute qualité pour le masquage des pertes de paquets pour le vocodeur G.722

Résumé

Le présent appendice III spécifie un algorithme de haute qualité pour le masquage des pertes de paquets (PLC, *packet loss concealment*) destiné à la Rec. UIT-T G.722. Cet algorithme procède au masquage des pertes de paquets dans le domaine de sortie à 16 kHz du décodeur G.722. On utilise une extrapolation périodique pour remplir la forme d'onde des paquets perdus, mélangés à du bruit filtré suivant les caractéristiques que possédait le signal avant la perte. Le signal extrapolé à 16 kHz est soumis à la banque des filtres d'analyse QMF, et les signaux des sous-bandes sont transmis à des codeurs MICDA dans les sous-bandes partielles pour actualiser les états des décodeurs MICDA desdites sous-bandes. Un traitement additionnel a lieu à chaque perte de paquets pour passer sans heurt de la forme d'onde extrapolée à la forme d'onde décodée à partir des paquets reçus. Entre autres, les états des décodeurs MICDA des sous-bandes sont alignés en phase avec le premier paquet reçu après une perte de paquets, et la forme d'onde décodée est alignée temporellement pour pouvoir s'aligner sur la forme d'onde extrapolée, avant que les deux soient ajoutées par chevauchement (*overlaps-added*) pour lisser la transition. Dans le cas d'une longue perte de paquets, l'algorithme atténue graduellement les signaux de sortie.

L'algorithme opérant selon une longueur de trame intrinsèque de 10 ms peut intervenir sur n'importe quelle longueur de paquet ou de trame multiple de 10 ms. La trame d'entrée la plus longue devient une supertrame pour laquelle le masquage des pertes de paquets est appelé à la longueur de la trame intrinsèque de 10 ms autant de fois que nécessaire, d'où l'absence de décalage supplémentaire par rapport au décodage G.722 normal utilisant la même longueur de trame.

L'algorithme PLC décrit dans le présent appendice satisfait aux mêmes conditions de complexité que l'algorithme PLC décrit dans l'Appendice IV de la Rec. UIT-T G.722. A une complexité additionnelle de 2,8 WMOPS dans le cas le plus défavorable et de 2 WMOPS en moyenne par rapport au décodeur G.722 sans PLC, l'algorithme PLC G.722 décrit dans le présent appendice fournit une qualité vocale nettement meilleure que l'algorithme PLC G.722 décrit dans l'Appendice IV, qui propose lui un autre compromis entre qualité et complexité.

Source

L'Appendice III de la Recommandation UIT-T G.722 a été agréé le 24 novembre 2006 par la Commission d'études 16 (2005-2008) de l'UIT-T.

AVANT-PROPOS

L'UIT (Union internationale des télécommunications) est une institution spécialisée des Nations Unies dans le domaine des télécommunications. L'UIT-T (Secteur de la normalisation des télécommunications) est un organe permanent de l'UIT. Il est chargé de l'étude des questions techniques, d'exploitation et de tarification, et émet à ce sujet des Recommandations en vue de la normalisation des télécommunications à l'échelle mondiale.

L'Assemblée mondiale de normalisation des télécommunications (AMNT), qui se réunit tous les quatre ans, détermine les thèmes d'étude à traiter par les Commissions d'études de l'UIT-T, lesquelles élaborent en retour des Recommandations sur ces thèmes.

L'approbation des Recommandations par les Membres de l'UIT-T s'effectue selon la procédure définie dans la Résolution 1 de l'AMNT.

Dans certains secteurs des technologies de l'information qui correspondent à la sphère de compétence de l'UIT-T, les normes nécessaires se préparent en collaboration avec l'ISO et la CEI.

NOTE

Dans la présente Recommandation, l'expression "Administration" est utilisée pour désigner de façon abrégée aussi bien une administration de télécommunications qu'une exploitation reconnue.

Le respect de cette Recommandation se fait à titre volontaire. Cependant, il se peut que la Recommandation contienne certaines dispositions obligatoires (pour assurer, par exemple, l'interopérabilité et l'applicabilité) et considère que la Recommandation est respectée lorsque toutes ces dispositions sont observées. Le futur d'obligation et les autres moyens d'expression de l'obligation comme le verbe "devoir" ainsi que leurs formes négatives servent à énoncer des prescriptions. L'utilisation de ces formes ne signifie pas qu'il est obligatoire de respecter la Recommandation.

DROITS DE PROPRIÉTÉ INTELLECTUELLE

L'UIT attire l'attention sur la possibilité que l'application ou la mise en œuvre de la présente Recommandation puisse donner lieu à l'utilisation d'un droit de propriété intellectuelle. L'UIT ne prend pas position en ce qui concerne l'existence, la validité ou l'applicabilité des droits de propriété intellectuelle, qu'ils soient revendiqués par un membre de l'UIT ou par une tierce partie étrangère à la procédure d'élaboration des Recommandations.

A la date d'approbation de la présente Recommandation, l'UIT avait été avisée de l'existence d'une propriété intellectuelle protégée par des brevets à acquérir pour mettre en œuvre la présente Recommandation. Toutefois, comme il ne s'agit peut-être pas de renseignements les plus récents, il est vivement recommandé aux développeurs de consulter la base de données des brevets du TSB sous <http://www.itu.int/ITU-T/ipr/>.

© UIT 2008

Tous droits réservés. Aucune partie de cette publication ne peut être reproduite, par quelque procédé que ce soit, sans l'accord écrit préalable de l'UIT.

TABLE DES MATIÈRES

	Page
Appendice III – Algorithme de haute qualité pour le masquage des pertes de paquets pour le vocodeur G.722.....	1
III.1 Domaine d'application.....	1
III.2 Références normatives.....	1
III.3 Abréviations	1
III.4 Conventions.....	2
III.5 Description générale de l'algorithme PLC.....	5
III.6 Extrapolation de la forme d'onde du PLC-MIC-WB des signaux de sortie G.722	8
III.7 Recodage de la sortie PLC	21
III.8 Contrôle des caractéristiques des signaux et de leur utilisation pour le PLC.....	25
III.9 Calcul des décalages temporels	28
III.10 Rephasage.....	31
III.11 Alignement temporel	33
III.12 Description au bit près de l'algorithme PLC G.722.....	36

Recommandation UIT-T G.722

Codage audiofréquence à 7 kHz à un débit inférieur ou égal à 64 kbit/s

Appendice III

Algorithme de haute qualité pour le masquage des pertes de paquets pour le vocodeur G.722

III.1 Domaine d'application

Le présent appendice décrit pour [UIT-T G.722] un algorithme de masquage des pertes de paquets de grande qualité. L'analyse statistique des résultats des essais de sélection des PLC pour [UIT-T G.722] a démontré que l'algorithme PLC du présent appendice était à l'évidence le meilleur des candidats examinés (y compris celui de l'Appendice IV) en termes de qualité vocale pour les applications de [UIT-T G.722] en présence de pertes de paquets. Le présent appendice satisfait aux mêmes prescriptions de complexité que l'Appendice IV, mais sa complexité est plus grande. Grâce à elle, il convient à des applications générales de [UIT-T G.722] pouvant être confrontées à des effacements de trames ou à des pertes de paquets, au nombre desquelles peuvent figurer, par exemple, la téléphonie utilisant le protocole Internet (VoIP, *voice over Internet protocol*), la téléphonie par Wi-Fi, et la prochaine génération des systèmes DECT. L'algorithme du présent appendice ajoute une complexité de 2,8 WMOPS dans le cas le plus défavorable et de 2 WMOPS en moyenne à celle correspondant au décodeur G.722. Il est facile d'implémenter, sauf pour les applications où il n'existe pratiquement pas de marge de complexité après l'implémentation du décodeur G.722 de base sans PLC.

III.2 Références normatives

- [UIT-T G.191 An.A] Recommandation UIT-T G.191 Annexe A (2005), *Outils logiciels pour la normalisation du codage des signaux vocaux et audiofréquence – Liste des outils logiciels disponible (STL)*.
- [UIT-T G.192] Recommandation UIT-T G.192 (1996), *Interface parallèle numérique commune pour la normalisation des signaux vocaux*.
- [UIT-T G.722] Recommandation UIT-T G.722 (1988), *Codage audiofréquence à 7 kHz à un débit inférieur ou égal à 64 kbit/s*.

III.3 Abréviations

Le présent appendice utilise les abréviations suivantes:

DECT	communication numérique sans fil améliorée (<i>digital enhanced cordless telecommunication</i>)
FIR	réponse impulsionnelle finie (<i>finite impulse response</i>)
LPC	codage prédictif linéaire (<i>linear predictive coding</i>)
MIC	modulation par impulsions et codage
MICDA	modulation par impulsions et codage différentiel adaptatif
OLA	superposition/admission (<i>overlap-add</i>)
PLC	masquage de perte de paquets (<i>packet loss concealment</i>)

PWE	extrapolation périodique d'ondes (<i>periodic waveform extrapolation</i>)
QMF	filtre miroir en quadrature (<i>quadratic mirror filter</i>)
STL2005	liste 2005 des outils logiciels (<i>software tool library 2005</i>)
VoIP	téléphonie utilisant le protocole Internet (<i>voice over Internet protocol</i>)
WB	large bande (<i>wideband</i>)
WiFi	fidélité sans fil (<i>wireless fidelity</i>)

III.4 Conventions

Le présent appendice utilise les conventions suivantes:

- l'algorithme PLC opérant avec une longueur intrinsèque de trame de 10 ms est décrit uniquement pour des trames de 10 ms; pour les paquets de plus grande taille (multiple de 10 ms) le paquet reçu est décodé en sections de 10 ms;
- l'indice temporel discret des signaux au taux d'échantillonnage de 16 kHz est en général indiqué soit par "j", soit par "i";
- le temps discret des signaux au niveau d'échantillonnage à 8 kHz est en général indiqué par "n";
- les signaux en bande basse (0-4 kHz) sont identifiés par un symbole "L";
- les signaux en bande haute (4-8 kHz) sont identifiés par un symbole "H";
- le présent appendice s'efforce de réutiliser les conventions de [UIT-T G.722].

La liste suivante est la liste des symboles les plus fréquemment utilisés et de leur correspondance.

$x_{out}(j)$	Sortie du décodeur G.722 à 16 kHz
$x_{PLC}(i)$	Sortie PLC G.722 à 16 kHz
$x_{out, FGF}(j)$	Sortie du décodeur G.722 à 16 kHz pour la "première bonne trame" reçue après la perte de paquets
$w(j)$	Fenêtre LPC
$x_w(j)$	Parole à la fenêtre
$r(i)$	Autocorrélation
$\hat{r}(i)$	Autocorrélation après lissage spectral et correction du bruit blanc
\hat{a}_i	Coefficients intermédiaires du prédicteur LPC
a_i	Coefficients du prédicteur LPC
$d(j)$	Signal d'erreur de prédiction sur le court terme à 16 kHz
avm	Ampleur moyenne du signal résiduel d'erreur de prédiction sur le court terme
a'_i	Coefficients pondérés du filtre de synthèse sur le court terme
$xw(j)$	Parole pondérée à 16 kHz
$xwd(n)$	Parole pondérée télé-échantillonnée (2 kHz)
b_i	Filtre passe-bas du 60 ^e ordre pour télé-échantillonnage
$c(k)$	Corrélation pour analyse approximative du pitch (2 kHz)
$E(k)$	Energie pour analyse approximative du pitch (2 kHz)
$c2(k)$	Corrélation mise au carré signée pour analyser le pitch approximatif (2 kHz)

cpp	Période de pitch approximatif
$cpplast$	Période de pitch approximatif de la dernière trame
$Ei(j)$	$E(k)$ interpolé (à 16 kHz)
$c2i(j)$	$c2(k)$ interpolé (à 16 kHz)
$\tilde{E}(k)$	Energie pour affinement du pitch (16 kHz)
$\tilde{c}(k)$	Corrélation pour affinement du pitch (16 kHz)
$ppfe$	Période de pitch pour effacement de trames
$ptfe$	Dérivation du pitch pour effacement de trames
ppt	Dérivation du prédicteur de pitch
$merit$	Facteur de mérite de périodicité
G_r	Facteur d'échelle pour composante aléatoire
G_p	Facteur d'échelle pour composante périodique
$ltring(j)$	Sonnerie (pitch) de longue durée
$ring(j)$	Sonnerie finale (y compris de brève durée)
$wi(j)$	Fenêtre d'ouverture en fondu
$wo(j)$	Fenêtre de fondu de fermeture
$wn(j)$	Sortie du générateur de bruit
$wgn(j)$	Sortie à l'échelle du générateur de bruit
$fn(j)$	Bruit filtré et bruit à l'échelle
$cfecount$	Compteur d'effacements de trames de 10 ms consécutives
$w_i(j)$	Fenêtre pour superposition/admission
$w_o(j)$	Fenêtre pour superposition/admission
h_i	Coefficients de filtre QMF
$x_L(n)$	Signal en sous-bande, bande basse (8 kHz)
$x_H(n)$	Signal en sous-bande, bande haute (8 kHz)
$I_L(n)$	Indice pour codeur MICDA, bande basse (8 kHz)
$I_H(n)$	Indice pour codeur MICDA, bande haute (8 kHz)
$s_{Lz}(n)$	Signal prédit, bande basse, contribution à la section des zéros
$s_{Lp}(n)$	Signal prédit, bande basse, contribution à la section des pôles
$s_L(n)$	Signal prédit, bande basse
$e_L(n)$	Signal d'erreur de prédiction en bande basse
$r_L(n)$	Signal reconstruit en bande basse
$p_{Ll}(n)$	Signal tronqué, partiel, reconstruit en bande basse
$\nabla_L(n)$	Facteur d'échelle logarithmique en bande basse
$\Delta_L(n)$	Facteur d'échelle en bande basse
$\nabla_{L,m1}(n)$	Facteur d'échelle logarithmique en bande basse, première moyenne

$\nabla_{L,m2}(n)$	Facteur d'échelle logarithmique en bande basse, seconde moyenne
$\nabla_{L,track}(n)$	Facteur d'échelle logarithmique en bande basse, suiviage (<i>tracking</i>)
$\nabla_{L,chg}(n)$	Facteur d'échelle logarithmique en bande basse, degré de change
$MPTH$	seuil tonal multiple
$MPDTH$	seuil de déviation tonal multiple
$\beta_L(n)$	Marge de stabilité de la section des pôles en bande basse
$\beta_{L,MA}(n)$	Moyenne mouvante de la marge de stabilité de la section des pôles en bande basse
$\beta_{L,min}$	Marge de stabilité minimale de la section des pôles en bande basse
$s_{Hz}(n)$	Signal prédit en bande haute, contribution à la section des zéros
$s_{Hp}(n)$	Signal prédit en bande haute, contribution à la section des pôles
$s_H(n)$	Signal prédit en bande haute
$e_H(n)$	Signal d'erreur de prédiction en bande haute
$r_H(n)$	Signal reconstruit en bande haute
$r_{H,HP}(n)$	Signal reconstruit filtré avec un filtre passe-haut, en bande haute
$p_H(n)$	Signal reconstruit, partiel, en bande haute
$p_{H,HP}(n)$	Signal reconstruit, partiel, filtré avec un filtre passe-haut, en bande haute
$\nabla_H(n)$	Facteur d'échelle logarithmique en bande haute
$\nabla_{H,m}(n)$	Facteur d'échelle logarithmique en bande haute, moyenne
$\nabla_{H,track}(n)$	Facteur d'échelle logarithmique en bande haute, suiviage
$\nabla_{H,chg}(n)$	Facteur d'échelle logarithmique en bande haute, degré de change
$\alpha_{LP}(n)$	Coefficient pour filtrage passe-bas du facteur d'échelle logarithmique en bande haute
$\nabla_{H,LP}(n)$	Facteur d'échelle logarithmique en bande haute, filtré avec filtre passe-bas
$r_{Le}(n)$	Estimation du signal d'erreur reconstruit en bande basse
$es(n)$	Signal extrapolé pour le calcul du décalage du rephasage
$R_{SUB}(k)$	Corrélation croisée normalisée, sous-échantillonnée
$R(k)$	Corrélation croisée normalisée
T_{LSUB}	Décalage sous-échantillonné
T_L	Décalage pour rephasage
$es_{tw}(n)$	Signal extrapolé pour affinement du décalage aux fins d'alignement temporel
T_{Lwarp}	Décalage aux fins d'alignement temporel
$X_{warp}(j)$	Signal aligné en temps (16 kHz)
$es_{ola}(j)$	Signal extrapolé pour superposition/admission (16 kHz)

III.5 Description générale de l'algorithme PLC

Pour faciliter la compréhension, six types de trames ont été définis et référencés dans le texte, à savoir:

- Type 1: trame reçue au-delà de la trame 8 après perte de paquets;
- Type 2: trames 1 et 2 perdues lors de pertes de paquets;
- Type 3: trames 3 à 6 perdues lors de pertes de paquets;
- Type 4: trame perdue au-delà de la trame 6 lors de pertes de paquets;
- Type 5: trame reçue immédiatement après perte de paquets;
- Type 6: trames 2 à 8 reçues après perte de paquets.

Ces types de trames sont illustrés à la Figure III.1 avec un exemple de ligne de distribution temporelle; l'algorithme PLC fonctionne sur la base d'une longueur de trame intrinsèque de 10 ms.

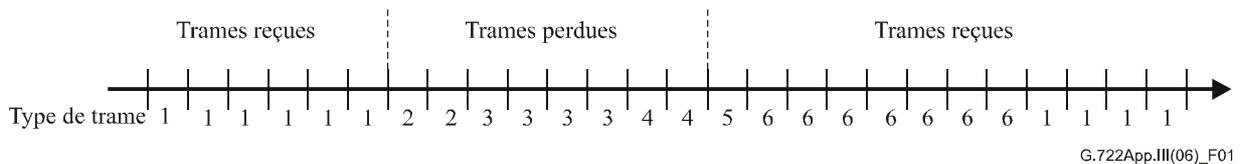


Figure III.1 – Ligne temporelle des types de trame

Type 1

Les trames du type 1 sont décodées selon [UIT-T G.722] avec en ajout le maintien d'un peu de mémoire d'état et un traitement pour faciliter le masquage PLC et le traitement associé; la Figure III.2 en est l'illustration.

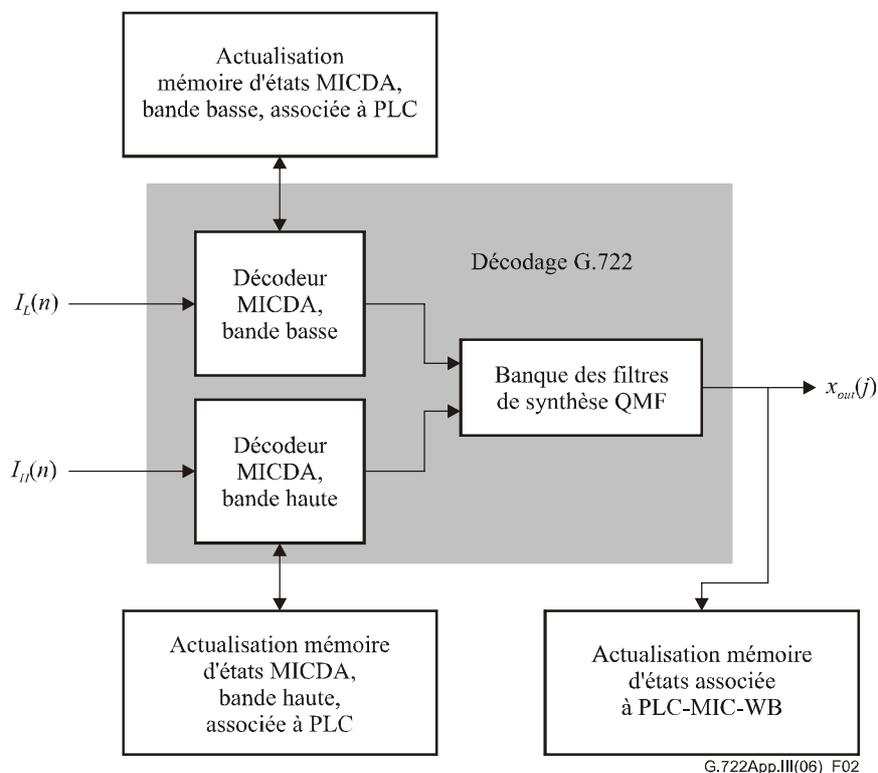


Figure III.2 – Décodage régulier selon la Rec. G.722 (type 1)

Types 2, 3 et 4

L'algorithme procède au masquage PLC-MIC-WB (*wideband*, large bande) dans le domaine vocal en sortie à 16 kHz pour les trames des types 2, 3 et 4. Un schéma du masquage PLC-MIC-WB est reproduit à la Figure III.3. La parole en sortie passée de la Rec. G.722 est mémorisée et transférée au PLC-MIC-WB, lequel est fondé sur une extrapolation périodique d'ondes (PWE, *periodic waveform extrapolation*), l'estimation du pitch étant une composante importante du masquage PLC-MIC-WB. Au début, on estime un pitch approximatif à partir d'un signal échantillonné (à 2 kHz) dans le domaine vocal pondéré. Ensuite, on affine cette estimation à pleine résolution à l'aide de l'échantillonnage original à 16 kHz. La sortie PLC-MIC-WB est une combinaison linéaire de la forme d'onde périodiquement extrapolée et du bruit mis en forme par le codage LPC. Dans le cas d'effacements de longue durée, la forme d'onde de sortie est graduellement atténuée; l'atténuation débute après 20 ms de perte de trames et s'achève après 60 ms.

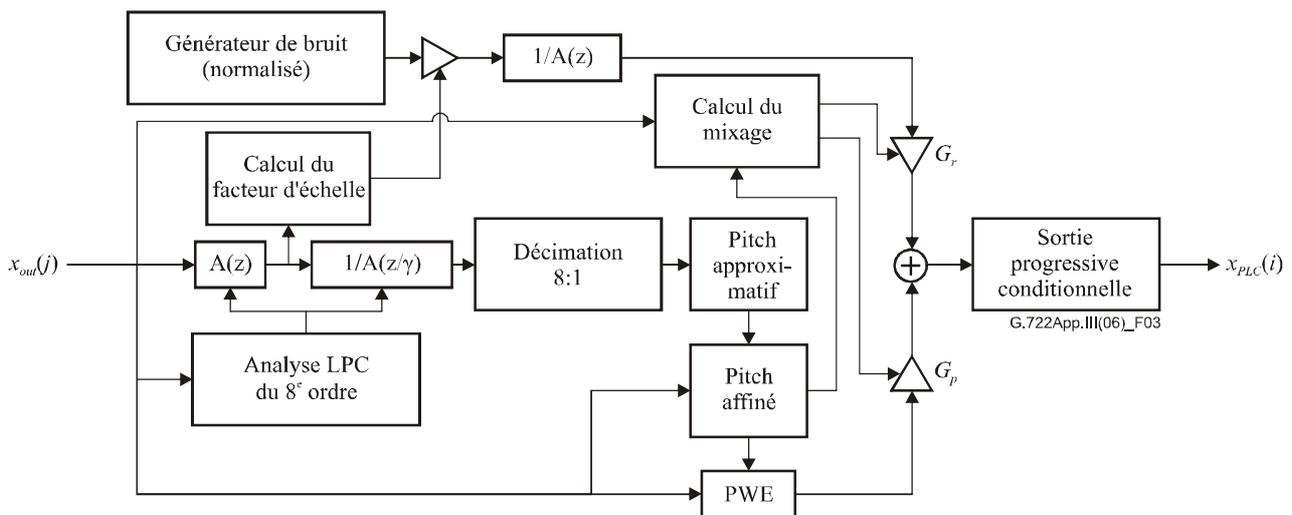


Figure III.3 – Schéma du masquage PLC-MIC-WB (pour les trames des types 2, 3 et 4)

Il ressort de la Figure III.4 que pour les trames des types 2, 3 et 4, on soumet les signaux de sortie du PLC-MIC-WB à la banque des filtres d'analyse QMF G.722 pour obtenir les signaux des sous-bandes correspondants, qui sont ensuite transférés respectivement à des codeurs MICDA en bande basse ou en bande haute, modifiés, afin d'actualiser les états et la mémoire du décodeur. Pour cette opération, seuls sont utilisés des codeurs MICDA en sous-bande simplifiés, partiels.

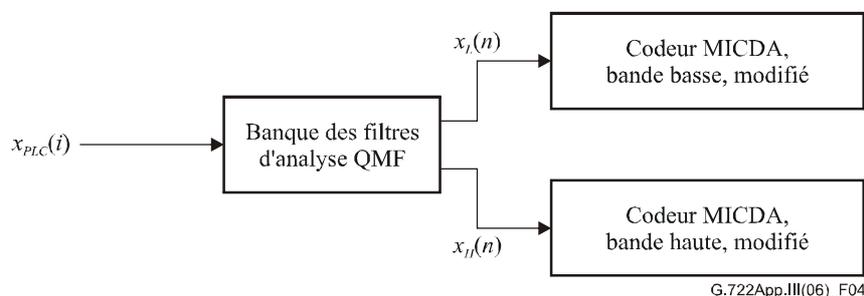


Figure III.4 – Recodage des signaux de sortie PLC pour actualiser les états MICDA des sous-bandes (pour les trames des types 2, 3 et 4)

Dans les Figures III.3 et III.4, le traitement intervient pendant la perte des trames. Les codeurs MICDA, en bandes basse et haute de la Figure III.4, sont dits modifiés car ils sont simplifiés pour

une moindre complexité (voir la description au § III.7). Par rapport aux codeurs normaux, ils ont une fonction différente, par exemple d'une remise à l'état adaptable qui est fonction de la propriété des signaux et de la durée de la perte des paquets.

Type 5

Le traitement le plus complexe concerne les trames de type 5, c'est-à-dire les premières trames à être reçues, pour lesquelles la forme d'onde passe de l'état extrapolé à l'état décodé, les principales techniques employées étant le rephasage et l'alignement temporel (cf. la Figure III.5).

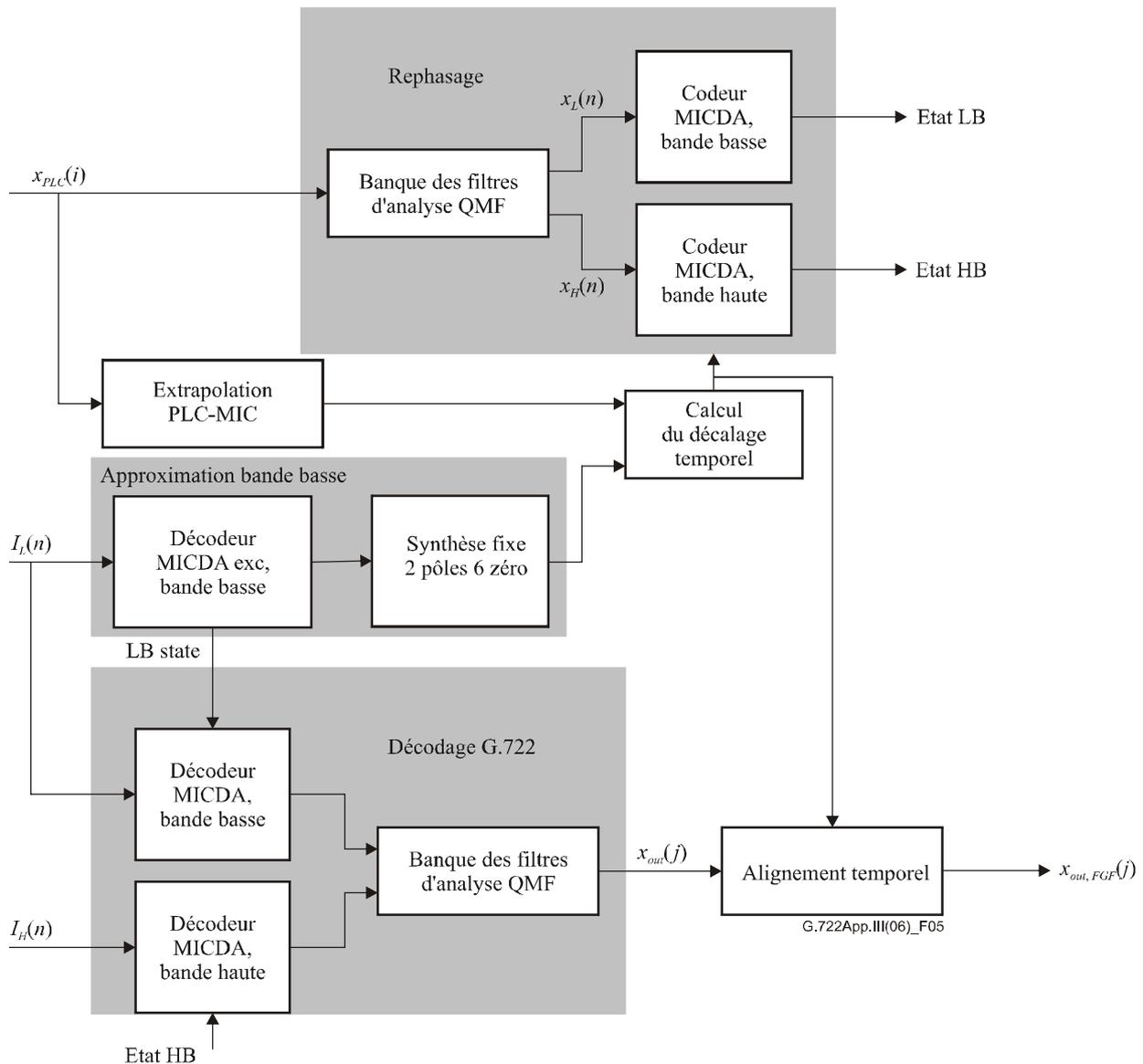


Figure III.5 – Rephasage et alignement temporel (type 5)

Par ailleurs, à la première trame reçue, la mémoire de la banque des filtres de synthèse QMF, au niveau du décodeur, doit être actualisée (voir le § III.9 pour une description détaillée). Les trames du type 5 font l'objet d'un autre traitement important dans lequel sont employées les techniques telles que le réglage adaptatif des facteurs d'échelle logarithmique en bandes basse et haute au début de la première trame reçue (cf. § III.8.1 et III.8.2); en outre, l'actualisation de la mémoire d'états mentionnée pour les trames de type 1 s'applique également aux trames de type 5, les techniques décrites aux § III.8.2.3, III.8.3 et III.8.4 s'appliquant aux trames de type 5 ainsi qu'aux trames de type 6.

Types 5 et 6

Les trames de type 6 sont décodées à l'aide de décodeurs MICDA en sous-bande modifiés et soumis à des restrictions (l'opération est illustrée dans le diagramme de la Figure III.6).

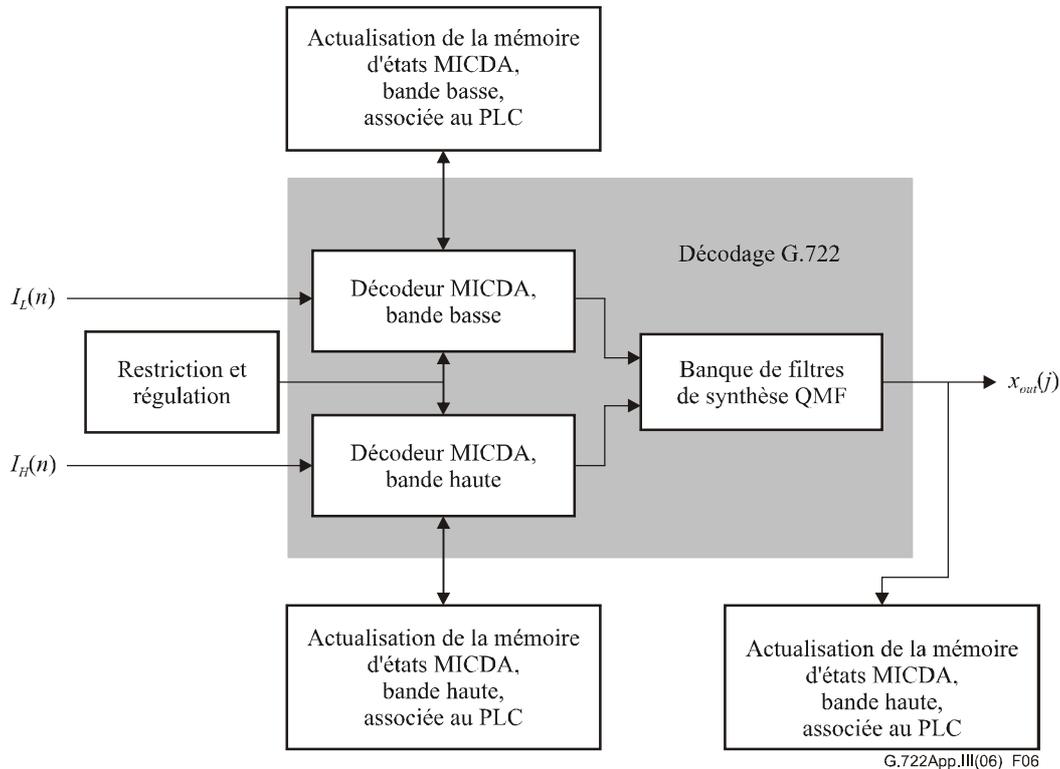


Figure III.6 – Décodage contraint et régulé des premières trames reçues (type 5 et type 6)

La restriction et la régulation des décodeurs MICDA des sous-bandes sont des opérations imposées pour les trames de types 5 et 6, c'est-à-dire pendant les 80 premières ms après la perte des paquets. Certains ne vont pas au-delà de 40 ms, tandis que d'autres sont adaptables en durée et/ou degré (voir § III.8 pour une description détaillée).

Lorsque les canaux sont exempts d'erreurs, l'algorithme est conforme au bit près à [UIT-T G.722]; par contre, en présence d'erreurs, il est identique à [UIT-T G.722] au-delà de la 8ème trame après la perte de paquets, et en l'absence d'erreurs sur les bits on devrait s'attendre à une convergence vers les signaux de sortie exempts d'erreurs de [UIT-T G.722].

L'algorithme PLC supporte toutes longueurs de paquets multiples de 10 ms; pour des longueurs de paquets supérieures à 10 ms, il est simplement appelé autant de fois que le nombre multiple par paquet à intervalles de 10 ms. L'algorithme PLC sera donc décrit ainsi par la suite en longueur de trame intrinsèque de 10 ms.

III.6 Extrapolation de la forme d'onde du PLC-MIC-WB des signaux de sortie G.722

Pour des trames perdues correspondant à une perte de paquets (trames de types 2, 3 et 4), l'opération PLC-MIC-WB illustrée à la Figure III.3 extrapole la forme d'onde des signaux de sortie G.722, $x_{out}(j)$, des trames précédentes pour remplir la trame en cours. Ce type de signal extrapolé en bande large, $x_{PLC}(i)$, est alors utilisé comme forme d'onde de sortie du PLC G.722 pendant les trames de types 2, 3 et 4. Pour faciliter la description des différents blocs représentés dans la Figure III.3, après que le signal $x_{PLC}(i)$ a été calculé par le PLC-MIC-WB pour les trames perdues, on considère qu'il est écrit dans la mémoire tampon pour $x_{out}(j)$, qui est le signal de sortie final de l'ensemble du système décodeur G.722/PLC.

Chaque bloc de la Figure III.3 sera maintenant décrit plus en détail dans les paragraphes ci-dessous.

III.6.1 Analyse LPC du huitième ordre

L'analyse LPC du 8^e ordre (voir la Figure III.3) est exécutée vers la fin de la boucle du traitement des trames, après que la trame en cours du signal $x_{out}(j)$ a été calculée et stockée dans la mémoire tampon. Cette analyse est le type commun d'analyse LPC d'autocorrélation, une fenêtre d'analyse asymétrique de 10 ms étant appliquée au signal $x_{out}(j)$ de la trame en cours reçue; cette fenêtre est égale à:

$$w(j) = \begin{cases} \frac{1}{2} \left[1 - \cos\left(\frac{(j+1)\pi}{121}\right) \right], & \text{pour } j = 0, 1, 2, \dots, 119 \\ \cos\left(\frac{(j-120)\pi}{80}\right), & \text{pour } j = 120, 121, \dots, 159 \end{cases} \quad (\text{III-1})$$

Si $x_{out}(0), x_{out}(1), \dots, x_{out}(159)$ représente dans la trame en cours reçue les échantillons des signaux large bande en sortie décodeur G.722 (PLC), l'opération de fenêtrage s'effectue comme suit:

$$x_w(j) = x_{out}(j)w(j), \quad j = 0, 1, 2, \dots, 159 \quad (\text{III-2})$$

Ensuite, on calcule comme suit les coefficients de corrélation:

$$r(i) = \sum_{j=i}^{159} x_w(j)x_w(j-i), \quad i = 0, 1, 2, \dots, 8 \quad (\text{III-3})$$

Les opérations de lissage spectral et de correction du bruit blanc sont alors appliquées aux coefficients d'autocorrélation comme suit:

$$\hat{r}(i) = \begin{cases} 1,0001 \times r(0), & i = 0 \\ r(i) e^{\frac{-(2\pi i \sigma / f_s)^2}{2}}, & i = 1, 2, \dots, 8 \end{cases} \quad (\text{III-4})$$

où $f_s = 16000$ est le taux d'échantillonnage du signal d'entrée et $\sigma = 40$.

On utilise ensuite la récursion Levinson-Durbin pour convertir les coefficients d'autocorrélation $\hat{r}(i)$ aux coefficients du prédicteur LPC $\hat{a}_i, i = 0, 1, \dots, 8$. Si la récursion est abandonnée prématurément avant d'être terminée (par exemple, parce que l'énergie résiduelle de prédiction $E(i)$ est inférieure à zéro), les coefficients du prédicteur sur le court terme de la dernière trame sont alors utilisés eux aussi dans la trame en cours. Pour exécuter ainsi le traitement d'exception, il faut une valeur initiale de l'ensemble \hat{a}_i qui est établie à $\hat{a}_0 = 1$ et $\hat{a}_i = 0$ pour $i = 1, 2, \dots, 8$. L'algorithme de récursion Levinson-Durbin est indiqué ci-dessous.

- 1) Si $\hat{r}(0) \leq 0$, utiliser l'ensemble \hat{a}_i de la dernière trame, et quitter la récursion Levinson-Durbin.
- 2) $E(0) = \hat{r}(0)$
- 3) $k_1 = -\hat{r}(1) / \hat{r}(0)$
- 4) $\hat{a}_1^{(1)} = k_1$
- 5) $E(1) = (1 - k_1^2)E(0)$
- 6) Si $E(1) \leq 0$, utiliser l'ensemble \hat{a}_i de la dernière trame, et quitter la récursion Levinson-Durbin.

7) Pour $i = 2, 3, 4, \dots, 8$, procéder comme suit:

$$k_i = \frac{-\hat{r}(i) - \sum_{j=1}^{i-1} \hat{a}_j^{i-1} \hat{r}(i-j)}{E(i-1)}$$

$$\hat{a}_i^{(i)} = k_i$$

$$\hat{a}_j^{(i)} = \hat{a}_j^{i-1} + k_i \hat{a}_{i-j}^{i-1}, \quad \text{pour } j = 1, 2, \dots, i-1$$

$$E(i) = (1 - k_i^2)E(i-1)$$

Si $E(i) \leq 0$, utiliser l'ensemble \hat{a}_i de la dernière trame, et quitter la récursion Levinson-Durbin.

Si la récursion est quittée prématurément, l'ensemble \hat{a}_i de la dernière trame est utilisé. Si par contre la récursion est menée à terme (ce qui est normalement le cas), les coefficients du prédicteur LPC ont les valeurs suivantes:

$$\hat{a}_0 = 1 \quad \text{(III-5)}$$

$$\hat{a}_i = \hat{a}_i^{(8)}, \quad \text{pour } i = 1, 2, \dots, 8 \quad \text{(III-6)}$$

En appliquant aux coefficients dérivés ci-dessus l'opération d'expansion de la largeur de bande, on obtient l'ensemble final des coefficients du prédicteur LPC, qui se présente comme suit:

$$a_i = (0,96852)^i \hat{a}_i^{(8)}, \quad \text{pour } i = 0, 1, \dots, 8 \quad \text{(III-7)}$$

III.6.2 Calcul du signal résiduel de la prédiction sur le court terme

Dans la Figure III.3, le bloc avec la légende "A(z)" représente un filtre d'erreur de prédiction linéaire sur le court terme, pourvu des coefficients de a_i pour $i = 0, 1, \dots, 8$ selon les calculs indiqués ci-dessus; exécuté après l'analyse LPC du 8ème ordre, il calcule le signal résiduel de la prédiction sur le court terme, $d(j)$, comme suit:

$$d(j) = x_{out}(j) + \sum_{i=1}^8 a_i \cdot x_{out}(j-i), \quad \text{pour } j = 0, 1, 2, \dots, 159 \quad \text{(III-8)}$$

Comme de coutume, l'indice temporel n de la trame en cours se poursuit à partir de l'indice temporel de la dernière trame; autrement dit, si la fourchette de l'indice temporel de 0, 1, 2, ..., 159, représente la trame en cours, la fourchette de l'indice temporel de -160, -159, ..., -1 représente alors la dernière trame. En conséquence, dans l'équation ci-dessus, si l'indice $(j-i)$ est négatif, il désigne juste l'échantillon du signal vers la fin de la dernière trame.

III.6.3 Calcul du facteur d'échelle

Dans la Figure III.3, le bloc avec la légende "Calcul du facteur d'échelle" calcule l'ampleur moyenne du signal résiduel de la prédiction sur le court terme dans la trame en cours; il intervient après qu'a été calculé le signal résiduel de la prédiction sur le court terme, $d(j)$, comme il est indiqué au § III.6.2. Cette ampleur moyenne, avm , se calcule comme suit:

$$avm = \frac{1}{160} \sum_{j=0}^{159} |d(j)| \quad (\text{III-9})$$

Si la trame suivante est une trame perdue (c'est-à-dire correspondant à une perte de paquets), cette ampleur moyenne, avm , peut être utilisée comme facteur d'échelle pour réduire une séquence de bruit gaussien blanc si la trame en cours est suffisamment non voisée.

III.6.4 Calcul des signaux vocaux pondérés

Dans la Figure III.3, le bloc avec la légende " $1/A(z/\gamma)$ " représente un filtre de synthèse sur le court terme pondéré; il intervient après le calcul du signal résiduel de la prédiction sur le court terme, $d(j)$, pour la trame en cours. Les coefficients de ce filtre, a'_i pour $i = 0, 1, \dots, 8$, sont calculés comme suit, avec $\gamma_1 = 0,75$.

$$a'_i = \gamma_1^i a_i, \quad \text{pour } i = 1, 2, \dots, 8 \quad (\text{III-10})$$

Le signal résiduel de la prédiction sur le court terme, $d(j)$, est soumis à ce filtre de synthèse; le signal vocal pondéré en sortie correspondant, $xw(j)$, se calcule comme suit:

$$xw(j) = d(j) - \sum_{i=1}^8 a'_i \cdot xw(j-i), \quad \text{pour } j = 0, 1, 2, \dots, 159 \quad (\text{III-11})$$

III.6.5 Décimation huit pour un

Tableau III.1 – Coefficients pour le filtre FIR du 60^e ordre

Décalage, i	b_i dans Q15	Décalage, i	b_i dans Q15	Décalage, i	b_i dans Q15
0	1209	20	-618	40	313
1	728	21	-941	41	143
2	1120	22	-1168	42	-6
3	1460	23	-1289	43	-126
4	1845	24	-1298	44	-211
5	2202	25	-1199	45	-259
6	2533	26	-995	46	-273
7	2809	27	-701	47	-254
8	3030	28	-348	48	-210
9	3169	29	20	49	-152
10	3207	30	165	50	-89
11	3124	31	365	51	-30
12	2927	32	607	52	21
13	2631	33	782	53	58
14	2257	34	885	54	81
15	1814	35	916	55	89

Tableau III.1 – Coefficients pour le filtre FIR du 60^e ordre

Décalage, i	b_i dans Q15	Décalage, i	b_i dans Q15	Décalage, i	b_i dans Q15
16	1317	36	881	56	84
17	789	37	790	57	66
18	267	38	654	58	41
19	-211	39	490	59	17

Le signal vocal pondéré est amené à traverser un filtre passe-bas FIR à phase minimale du 60^e ordre, puis la décimation 8:1 est exécutée pour réduire par échantillonnage le signal vocal pondéré, après le filtre passe-bas, à 16 kHz en un signal vocal pondéré de 2 kHz, $xwd(n)$; cette opération de décimation est réalisée après qu'a été calculé le signal vocal pondéré $xw(j)$. Pour réduire la complexité, l'opération de filtrage par filtre passe-bas FIR est effectuée uniquement lorsqu'on a besoin d'un nouvel échantillon de signal $xwd(n)$. Par conséquent, le signal vocal pondéré, réduit par échantillonnage, $xwd(n)$, se calcule comme suit:

$$xwd(n) = \sum_{i=0}^{59} b_i \cdot xw(8n + 7 - i), \quad \text{pour } n = 0, 1, 2, \dots, 19 \quad (\text{III-12})$$

où b_i , $i = 0, 1, 2, \dots, 59$, sont les coefficients de filtre pour le filtre passe-bas FIR du 60^e ordre, comme il est indiqué au Tableau III.1.

III.6.6 Extraction de la période fondamentale approximative

Pour réduire la complexité des calculs, on extrait dans PLC-MIC-WB la période fondamentale en deux étapes:

- 1) détermination de la période fondamentale approximative avec une résolution temporelle du signal décimé à 2 kHz;
- 2) affinement de la période fondamentale avec une résolution temporelle du signal non décimé à 16 kHz. Cette extraction de la période fondamentale n'est effectuée que pendant les trames reçues après qu'a été calculé le signal vocal pondéré, réduit par échantillonnage, $xwd(n)$.

Le présent paragraphe décrit l'algorithme d'extraction de la période fondamentale approximative (première étape) représenté par le bloc portant la légende "Pitch approximatif" dans la Figure III.3; cet algorithme est basé sur le principe d'une maximalisation de la corrélation croisée normalisée avec un peu de logique de décision additionnelle.

Une fenêtre d'analyse du pitch de 15 ms est utilisée dans l'opération d'extraction de la période fondamentale approximative; la fin de cette fenêtre est alignée avec la fin de la trame en cours. A un taux d'échantillonnage de 2 kHz, 15 ms correspondent à 30 échantillons. Sans perte de généralité, la fourchette de l'indice $n = 0$ à $n = 29$ correspond à la fenêtre d'analyse du pitch pour $xwd(n)$. L'algorithme d'extraction de la période fondamentale approximative débute par le calcul des valeurs ci-après:

$$c(k) = \sum_{n=0}^{29} xwd(n)xwd(n-k) \quad (\text{III-13})$$

$$E(k) = \sum_{n=0}^{29} [xwd(n-k)]^2 \quad (\text{III-14})$$

$$c2(k) = \begin{cases} c^2(k), & \text{si } c(k) \geq 0 \\ -c^2(k), & \text{si } c(k) < 0 \end{cases} \quad (\text{III-15})$$

pour tous les nombres entiers de $k = MINPPD - 1$ à $k = MAXPPD + 1$, où $MINPPD = 5$ et $MAXPPD = 33$ sont respectivement la période fondamentale minimale et maximale dans le domaine décimé. L'algorithme d'extraction de la période fondamentale approximative balaie ensuite toute la gamme de $k = MINPPD, MINPPD + 1, MINPPD + 2, \dots, MAXPPD$ pour trouver tous les pics locaux de l'ensemble $\{c2(k)/E(k)\}$ pour lequel $c(k) > 0$. (Une valeur est assimilée à un pic local si ses deux valeurs adjacentes sont plus petites.) Supposons que N_p indique le nombre des pics locaux positifs de ce type; supposons que $k_p(j), j = 1, 2, \dots, N_p$ sont les indices où $c2(k_p(j))/E(k_p(j))$ est un pic local et $c(k_p(j)) > 0$, et supposons que $k_p(1) < k_p(2) < \dots < k_p(N_p)$. Dans un souci de facilité, le terme $c2(k)/E(k)$ sera désigné comme le "carré de la corrélation normalisée".

Si $N_p = 0$, c'est-à-dire s'il n'y a pas de pic local positif pour la fonction $c2(k)/E(k)$, l'algorithme recherche alors le pic local négatif le plus large pourvu de l'ampleur la plus large de $|c2(k)/E(k)|$; si un pic de ce type est trouvé, l'indice correspondant k est utilisé comme la période fondamentale approximative de sortie, c_{pp} , et le traitement de ce bloc s'arrête. Par contre, si la fonction carrée de la corrélation normalisée $c2(k)/E(k)$ n'a ni pic local positif, ni pic local négatif, la période fondamentale approximative de sortie est alors mise à la valeur $c_{pp} = MINPPD$, et le traitement de ce bloc s'arrête. Si $N_p = 1$, la période fondamentale approximative de sortie est mise à la valeur $c_{pp} = k_p(1)$, et le traitement de ce bloc s'arrête.

S'il existe deux pics locaux ($N_p \geq 2$) ou plus, ce bloc utilise alors les *Algorithmes 1, 2, 3 et 4* (voir ci-dessous), dans cet ordre, pour déterminer la période fondamentale approximative de sortie c_{pp} . Les variables calculées dans les premiers des quatre algorithmes seront transférées et utilisées dans les algorithmes suivants.

L'*Algorithme 1* ci-dessous sert à déterminer le plus large pic interpolé quadratiquement autour des pics locaux du carré de la corrélation normalisée $c2(k_p)/E(k_p)$. Une interpolation quadratique est effectuée pour $c(k_p)$, tandis qu'une interpolation linéaire est effectuée pour $E(k_p)$. Cette interpolation est effectuée avec la résolution temporelle du signal vocal non décimé à 16 kHz. Dans l'algorithme ci-dessous, D indique le facteur de décimation utilisé de $xw(n)$ à $xwd(n)$; en conséquence, ici $D = 8$.

Algorithme 1 – Trouver le plus large pic interpolé quadratiquement autour de $c2(k_p)/E(k_p)$:

- i) Poser $c2_{max} = -1, E_{max} = 1$ et $j_{max} = 0$.
- ii) Pour $j = 1, 2, \dots, N_p$, procéder selon les 12 étapes suivantes:
 - 1) poser $a = 0,5 [c(k_p(j) + 1) + c(k_p(j) - 1)] - c(k_p(j))$
 - 2) poser $b = 0,5 [c(k_p(j) + 1) - c(k_p(j) - 1)]$
 - 3) poser $j_i = 0$
 - 4) poser $e_i = E(k_p(j))$
 - 5) poser $c2_m = c2(k_p(j))$
 - 6) poser $E_m = E(k_p(j))$
 - 7) si $c2(k_p(j) + 1)E(k_p(j) - 1) > c2(k_p(j) - 1)E(k_p(j) + 1)$, effectuer la partie restante de l'étape 7:
$$\Delta = [E(k_p(j) + 1) - e_i]/D$$

Pour $k = 1, 2, \dots, D/2$, effectuer la partie en retrait suivante de l'étape 7:

$$c_i = a (k/D)^2 + b(k/D) + c(k_p(j))$$

$$e_i \leftarrow e_i + \Delta$$

Si $(ci)^2 Em > (c2m) ei$, effectuer les trois lignes en retrait ci-dessous:

$$ji = k$$

$$c2m = (ci)^2$$

$$Em = ei$$

- 8) si $c2(k_p(j)+1)E(k_p(j)-1) \leq c2(k_p(j)-1)E(k_p(j)+1)$, effectuer la partie restante de l'étape 8:

$$\Delta = [E(k_p(j) - 1) - ei]/D$$

Pour $k = -1, -2, \dots, -D/2$, effectuer la partie en retrait suivante de l'étape 8:

$$ci = a(k/D)^2 + b(k/D) + c(k_p(j))$$

$$ei \leftarrow ei + \Delta$$

Si $(ci)^2 Em > (c2m) ei$, effectuer les trois lignes en retrait suivantes:

$$ji = k$$

$$c2m = (ci)^2$$

$$Em = ei$$

- 9) poser $lag(j) = k_p(j) + ji/D$

- 10) poser $c2i(j) = c2m$

- 11) poser $Ei(j) = Em$

- 12) si $c2m \times Emax > c2max \times Em$, effectuer les trois lignes en retrait suivantes:

$$jmax = j$$

$$c2max = c2m$$

$$Emax = Em$$

- iii) Donner au premier candidat pour la période fondamentale approximative la valeur suivante:
 $cpp = lag(jmax)$.

Le symbole \leftarrow indique que le paramètre à gauche est actualisé avec la valeur indiquée à droite.

Pour éviter de prendre une période fondamentale approximative qui se trouve autour d'un multiple de nombre entier de la période fondamentale approximative vraie, on procède à un balayage des intervalles de temps correspondant aux pics locaux de $c2(k_p)/E(k_p)$ pour voir si l'un d'entre eux est suffisamment près de la période fondamentale approximative de sortie de la dernière trame, dénotée $cpplast$. (Pour la première trame, $cpplast$ est initialisé à 12.) Si un décalage temporel se trouve à moins de 25% de $cpplast$, on considère qu'il est suffisamment proche. Pour tous les décalages temporels de ce type se trouvant à moins de 25% $cpplast$, on compare les valeurs de crête (interpolation quadratique) correspondantes du carré de la corrélation normalisée, $c2(k_p)/E(k_p)$ et on sélectionne aux fins d'approfondissement le décalage temporel interpolé correspondant au carré de la corrélation normalisée maximale. L'*Algorithme 2* ci-dessous exécute la tâche décrite ci-dessus. Les ensembles interpolés $c2i(j)$ et $Ei(j)$ calculés dans l'*Algorithme 1* sont utilisés dans le présent algorithme.

Algorithme 2 – Trouver le décalage temporel maximalisant le décalage $c2(k_p)/E(k_p)$ interpolé parmi tous les décalages temporels proches de la période fondamentale approximative de sortie de la dernière trame:

- i) Poser l'indice $im = -1$
- ii) Poser l'indice $c2m = -1$
- iii) Poser l'indice $Em = 1$

- iv) Pour $j = 1, 2, \dots, N_p$, procéder comme suit:
 si $|k_p(j) - cpplast| \leq 0,25 \times cpplast$, procéder comme suit:
 si $c2i(j) \times Em > c2m \times Ei(j)$, effectuer les trois opérations suivantes:
- $$im = j$$
- $$c2m = c2i(j)$$
- $$Em = Ei(j)$$

A noter que, si aucun décalage temporel, $k_p(j)$, ne se trouve à 25% de $cpplast$, la valeur de l'indice im restera alors à -1 après exécution de l'Algorithme 2. Par contre, dans le cas où un ou plusieurs décalages temporels se trouvent à moins de 25% de $cpplast$, l'indice im correspond au carré de la corrélation normalisée le plus grand parmi les décalages en question.

Ensuite, l'Algorithme 3 permet de déterminer s'il convient de choisir comme période fondamentale approximative de sortie un autre décalage temporel dans la première moitié de la fourchette des valeurs du pitch; en gros, il balaie tous les décalages temporels interpolés, $lag(j)$, qui sont moins de 16, et vérifie si l'un d'entre eux a un pic local du carré de la corrélation normalisée suffisamment grand à proximité de tout multiple entier de sa valeur (comprise) jusqu'à 32. Si un ou plusieurs de ces décalages temporels satisfont à cette condition, c'est le plus petit d'entre eux qui est retenu comme période fondamentale approximative de sortie.

Là encore, les variables calculées aux Algorithmes 1 et 2 ci-dessus conservent leurs valeurs finales jusqu'à l'Algorithme 3 ci-dessous. Dans ce qui suit le paramètre $MPDTH$ est de 0,06, et l'ensemble seuil $MPTH(k)$ s'exprime comme suit: $MPTH(2) = 0,7$, $MPTH(3) = 0,55$, $MPTH(4) = 0,48$, $MPTH(5) = 0,37$ et $MPTH(k) = 0,30$, pour $k > 5$.

Algorithme 3 – Vérifier si un autre décalage temporel dans la première moitié de la fourchette de la période fondamentale approximative devrait être sélectionné comme période fondamentale approximative de sortie:

Pour $j = 1, 2, 3, \dots, N_p$, dans cet ordre, procéder comme suit pendant que $lag(j) < 16$:

- i) Si $j \neq im$, poser $threshold = 0,73$; autrement, poser $threshold = 0,4$.
- ii) Si $c2i(j) \times Emax \leq threshold \times c2max \times Ei(j)$, disqualifier ce j , sauter l'étape iii) pour ce j , incrémenter j de 1 et revenir à l'étape i).
- iii) Si $c2i(j) \times Emax > threshold \times c2max \times Ei(j)$, procédez comme suit:
 - a) Pour $k = 2, 3, 4, \dots$, procéder comme suit pendant que: $k \times lag(j) < 32$:
 - 1) $s = k \times lag(j)$
 - 2) $a = (1 - MPDTH) s$
 - 3) $b = (1 + MPDTH) s$
 - 4) Parcourir $m = j + 1, j + 2, j + 3, \dots, N_p$, dans cet ordre, et voir si un des décalages temporels $lag(m)$ se trouvent entre a et b . Si tel n'est pas le cas, disqualifier ce j , arrêtez l'étape iii), incrémenter j de 1 et revenir à l'étape i). Si au moins un de ces m satisfait la condition $a < lag(m) \leq b$ et $c2i(m) \times Emax > MPTH(k) \times c2max \times Ei(m)$, considérer alors qu'un pic du carré de la corrélation normalisée suffisamment large se trouve à proximité du k ième multiple entier de $lag(j)$; dans ce cas, arrêter l'étape iii) a) 4), incrémenter k de 1 et revenir à l'étape iii) a) 1).

- b) Si l'étape iii) a) se termine sans s'arrêter prématurément, c'est-à-dire que s'il existe un pic interpolé suffisamment large du carré de la corrélation normalisée à $\pm 100 \times MPDTH\%$ de tout multiple entier de $lag(j)$ qui est inférieur à 32, alors arrêter cet algorithme, sauter l'Algorithme 4 et poser $c_{pp} = lag(j)$ comme période fondamentale approximative de sortie finale.

Si l'Algorithme 3 ci-dessus se termine sans qu'il soit trouvé une période fondamentale approximative de sortie valable, c_{pp} , alors l'Algorithme 4 examine le pic local le plus large du carré de la corrélation normalisée autour de la période fondamentale approximative de la dernière trame, trouvé dans l'Algorithme 2 ci-dessus, et prend une décision finale en ce qui concerne la période fondamentale approximative de sortie, c_{pp} . Là encore, les variables calculées dans les Algorithmes 1 et 2 ci-dessus conservent leurs valeurs finales jusqu'à l'Algorithme 4 ci-dessous. Dans la suite, les paramètres sont $SMDTH = 0,095$ et $LPTH1 = 0,78$.

Algorithme 4 – Choix final de la période fondamentale approximative de sortie:

- i) Si $im = -1$, c'est-à-dire s'il n'existe pas de pic local suffisamment large du carré de la corrélation normalisée autour de la période fondamentale approximative de la dernière trame, utiliser alors la c_{pp} calculée à la fin de l'Algorithme 1 comme période fondamentale approximative de sortie finale, et quitter cet algorithme.
- ii) Si $im = jmax$, c'est-à-dire si le pic local le plus large du carré de la corrélation normalisée autour de la période fondamentale approximative de la dernière trame se trouve être également la somme maximale de tous les pics interpolés du carré de la corrélation normalisée à l'intérieur de cette trame, utiliser alors la c_{pp} calculée à la fin de l'Algorithme 1 comme période fondamentale approximative de sortie finale, et quitter cet algorithme.
- iii) Si $im < jmax$, effectuer la partie en retrait suivante:
 Si $c_{2m} \times E_{max} > 0,43 \times c_{2max} \times E_m$, effectuer la partie en retrait suivante de l'étape iii):
 a) Si $lag(im) > MAXPPD/2$, poser comme sortie $c_{pp} = lag(im)$ et quitter cet algorithme.
 b) Autrement, pour $k = 2, 3, 4, 5$, effectuer la partie en retrait suivante:
 1) $s = lag(jmax) / k$
 2) $a = (1 - SMDTH) s$
 3) $b = (1 + SMDTH) s$
 4) Si $lag(im) > a$ et $lag(im) < b$, poser comme sortie $c_{pp} = lag(im)$ et quitter cet algorithme.
- iv) Si $im > jmax$, effectuer la partie en retrait suivante:
 Si $c_{2m} \times E_{max} > LPTH1 \times c_{2max} \times E_m$, poser comme sortie $c_{pp} = lag(im)$ et quitter cet algorithme.
- v) Si l'exécution des algorithmes se poursuit jusqu'à ce stade, c'est qu'aucune des étapes ci-dessus n'a permis de sélectionner une période fondamentale approximative de sortie finale; dans ce cas, accepter seulement la c_{pp} calculée à la fin de l'Algorithme 1 comme période fondamentale approximative de sortie finale.

III.6.7 Affinement de la période fondamentale

Dans la Figure III.3, le bloc avec la légende "Affiner le pitch" exécute le traitement en seconde étape de l'algorithme d'extraction de la période fondamentale en cherchant dans la proximité de la période fondamentale approximative, en pleine résolution temporelle à 16 kHz, à l'aide du signal vocal de sortie décodé G.722. Ce bloc transforme d'abord la période fondamentale approximative,

cpp , en domaine des signaux non décimés en la multipliant par le facteur de décimation D , où $D = 8$. La taille des fenêtres d'analyse en vue de l'affinement du pitch, WSZ , est choisie comme étant le plus petit des échantillons $cpp \times D$ et de 160 échantillons (correspondant à 10 ms): $WSZ = \min(cpp \times D, 160)$.

Ensuite, la limite inférieure de la fourchette de recherche est calculée comme étant $lb = \max(MINPP, cpp \times D - 4)$, où $MINPP = 40$ échantillons est la période fondamentale minimale. La limite supérieure de la fourchette de recherche est calculée comme étant $ub = \min(MAXPP, cpp \times D + 4)$, où $MAXPP = 265$ échantillons est la période fondamentale maximale.

Ce bloc gère une mémoire tampon de signaux vocaux décodés G.722 à 16 kHz, $x_{out}(j)$, avec un total de $XQOFF = MAXPP + 1 + FRSZ$ échantillons, où $FRSZ = 160$ est la longueur de trame. Les derniers échantillons $FRSZ$ de ce tampon contiennent les signaux vocaux décodés G.722 de la trame en cours. Les premiers $MAXPP + 1$ échantillons sont pourvus de signaux de sortie décodeur G.722/PLC dans les précédentes trames immédiatement avant la trame en cours. Le dernier échantillon de la fenêtre d'analyse est aligné sur le dernier échantillon de la trame en cours. On suppose que la gamme de l'indice de $j = 0$ à $j = WSZ - 1$ correspond à la fenêtre de l'analyse, soit aux derniers échantillons WSZ dans le tampon $x_{out}(j)$, et on suppose que les indices négatifs désignent les échantillons avant la fenêtre d'analyse. Les termes de corrélation et d'énergie ci-après dans le domaine des signaux non décimés sont calculés pour des décalages temporels k à l'intérieur de la fourchette de recherche $[lb, ub]$.

$$\tilde{c}(k) = \sum_{j=0}^{WSZ-1} x_{out}(j)x_{out}(j-k) \quad (III-16)$$

$$\tilde{E}(k) = \sum_{j=0}^{WSZ-1} x_{out}(j-k)^2 \quad (III-17)$$

Le décalage temporel $k \in [lb, ub]$ qui maximalise le rapport $\tilde{c}^2(k)/\tilde{E}(k)$ est choisi comme période fondamentale affinée finale pour l'effacement de trame, ou $ppfe$, c'est-à-dire:

$$ppfe = \arg \max_{k \in [lb, ub]} \left[\frac{\tilde{c}^2(k)}{\tilde{E}(k)} \right] \quad (III-18)$$

Ensuite, le bloc avec la légende "Affiner le pitch" calcule de son côté deux nouveaux facteurs d'échelle se rapportant au pitch: le premier est appelé $ptfe$, pour *pitch tap for frame erasure*, soit dérivation de pitch pour effacement de trame; c'est le facteur d'échelle utilisé pour l'extrapolation périodique d'ondes, facteur qui est rapport de l'ampleur moyenne du signal $x_{out}(j)$ dans la fenêtre d'analyse sur l'ampleur moyenne de la partie du signal $x_{out}(j)$ correspondant aux échantillons $ppfe$ précédemment évoqués et pourvus du même signe que la corrélation entre ces deux parties du signal.

$$ptfe = \text{sign}(\tilde{c}(ppfe)) \left[\frac{\sum_{j=0}^{WSZ-1} |x_{out}(j)|}{\sum_{j=0}^{WSZ-1} |x_{out}(j - ppfe)|} \right] \quad (III-19)$$

Dans le cas dégradé où $\sum_{j=0}^{WSZ-1} |x_{out}(j - ppfe)| = 0$, $ptfe$ est mis à 0; après ce calcul de $ptfe$, la valeur de $ptfe$ se limite à la fourchette $[-1, 1]$.

Le deuxième facteur d'échelle se rapportant au pitch est appelé *ppt*, pour *pitch predictor tap*, soit dérivation du prédicteur de pitch; il est utilisé pour calculer le signal annonçant les filtres sur le long terme (description à suivre), ce qui donne $ppt = 0,75 \times ptf$.

III.6.8 Calculer le rapport de mélange

Dans la Figure III.3, le bloc avec la légende "Calculer le mélange" permet de calculer un facteur de mérite pour déterminer le rapport de mélange entre les signaux périodiquement extrapolés et les contributions de bruit filtrées pendant les trames perdues; ce calcul se fait uniquement pendant la toute première trame perdue dans chaque occurrence de perte de paquets, et le rapport de mélange qui en résulte est utilisé pendant toute la durée de cette perte de paquets particulière. Le facteur de mérite est une somme pondérée des trois caractéristiques du signal suivantes: le gain logarithmique, la première autocorrélation normalisée et le gain de prédiction du pitch, chacune se calculant comme suit.

A l'aide de la même convention d'indexation que celle utilisée pour $x_{out}(j)$ au § III.6.7, on obtient l'énergie du signal $x_{out}(j)$ dans la fenêtre d'analyse d'affinement du pitch qui est:

$$sige = \sum_{j=0}^{WSZ-1} x_{out}^2(j) \quad (III-20)$$

et le gain logarithmique en base -2, *lg*, qui se calcule comme suit:

$$lg = \begin{cases} \log_2(sige) & \text{si } sige \neq 0 \\ 0 & \text{si } sige = 0 \end{cases} \quad (III-21)$$

Si $\tilde{E}(ppfe) \neq 0$, l'énergie résiduelle de prédiction du pitch se calcule comme suit:

$$rese = sige - \tilde{c}^2(ppfe) / \tilde{E}(ppfe) \quad (III-22)$$

et le gain de prédiction du pitch, *pg*, se calcule comme suit:

$$pg = \begin{cases} 10 \log_{10} \left(\frac{sige}{rese} \right) & \text{si } rese \neq 0 \\ 20 & \text{si } rese = 0 \end{cases} \quad (III-23)$$

Si $\tilde{E}(ppfe) = 0$, poser $pg = 0$. Si $sige = 0$, poser également $pg = 0$.

La première autocorrélation normalisée ρ_1 se calcule comme suit:

$$\rho_1 = \begin{cases} \left[\frac{\sum_{j=0}^{WSZ-2} x_{out}(j)x_{out}(j+1)}{sige} \right] & \text{si } sige \neq 0 \\ 0 & \text{si } sige = 0 \end{cases} \quad (III-24)$$

Après obtention de ces trois caractéristiques du signal, le facteur de mérite se calcule comme suit:

$$merit = lg + pg + 12\rho_1 \quad (III-25)$$

Le *merit* calculé ci-dessus permet de déterminer les deux facteurs d'échelle G_p et G_r , qui de fait déterminent le rapport de mélange entre le signal périodiquement extrapolé et la contribution de bruit filtré. Il existe deux seuils pour le *merit*: un seuil élevé MHI , et un seuil bas MLO , qui sont mis aux valeurs suivantes: $MHI = 28$ et $MLO = 20$. Le facteur d'échelle G_r pour la composante aléatoire (bruit filtré) se calcule comme suit:

$$G_r = \frac{MHI - merit}{MHI - MLO} \quad (III-26)$$

et le facteur d'échelle G_p pour la composante périodique se calcule comme suit:

$$G_p = 1 - G_r \quad (III-27)$$

III.6.9 Extrapolation périodique d'ondes

Le bloc avec la légende "Extrapolation périodique d'ondes" dans la Figure III.3 extrapole périodiquement le précédent signal vocal de sortie pendant les trames perdues si $merit > MLO$.

A la toute première trame perdue de chaque perte de paquets, on calcule l'incrément de la période fondamentale moyenne par trame. Un tampon chronologique de la période fondamentale, $pph(m)$, $m = 1, 2, \dots, 5$, mémorise la période fondamentale $ppfe$ pendant les 5 trames précédentes. L'incrément de la période fondamentale moyenne s'obtient comme suit: débiter avec la dernière trame immédiate, puis calculer l'incrément de la période fondamentale à partir de sa trame précédente jusqu'à cette trame (une valeur négative indique une diminution de la période fondamentale); si l'incrément est nul, l'algorithme vérifie l'incrément de la période à la trame précédente. Ce processus se poursuit jusqu'à la première trame dotée d'un incrément non nul, ou jusqu'à ce que la quatrième trame précédente ait été examinée. Si les cinq trames précédentes ont toutes une période fondamentale identique, l'incrément moyen est mis à zéro; autrement, si le premier incrément non nul est constaté à la $mième$ trame précédente, et si l'ampleur de l'incrément est inférieure à 5% de la période fondamentale à cette trame-là, alors l'incrément de la période fondamentale moyenne, $ppinc$, est égal à l'incrément à cette trame divisé par m , la valeur résultante se limitant alors à la fourchette $[-1, 2]$.

Dans la seconde trame perdue consécutive lors d'une perte de paquets, l'incrément de la période fondamentale moyenne, $ppinc$, vient s'ajouter à la période fondamentale, $ppfe$, le résultat étant arrondi au nombre entier le plus proche, puis limité à la fourchette $[MINPP, MAXPP]$.

Si la trame en cours est la toute première trame perdue lors d'une perte de paquets, un signal dit de "sonnerie" est calculé pour être utilisé aux fins de superposition/admission (*overlap-add*) pour garantir une bonne transition de la forme d'onde au début de la trame. La longueur de cette opération pour le signal de sonnerie et pour la forme d'onde extrapolée périodiquement est de 20 échantillons pour la première trame perdue. On suppose que la fourchette des indices de $j = 0, 1, 2, \dots, 19$, correspond aux 20 premiers échantillons de la première trame perdue en cours, ce qui est la période de superposition/admission, et que les indices négatifs correspondent aux trames précédentes. Le signal de sonnerie sur le long terme qui s'obtient est une version réduite du signal résiduel de prédiction sur le court terme qui intervient une période fondamentale plus tôt que la période de superposition/admission.

$$ltring(j) = x_{out}(j - ppfe) + \sum_{i=1}^8 a_i \cdot x_{out}(j - ppfe - i), \quad \text{pour } j = 0, 1, 2, \dots, 19 \quad (III-28)$$

Après que ces 20 échantillons de $ltring(j)$ sont calculés, ils sont encore réduits par le facteur d'échelle ppt calculé au § III.6.7.

$$ltring(j) \leftarrow ppt \cdot ltring(j), \quad \text{pour } j = 0, 1, 2, \dots, 19 \quad (III-29)$$

Une fois la mémoire à filtre $ring(j), j = -8, -7, \dots, -1$ initialisée aux 8 derniers échantillons du signal $x_{out}(j)$ de la dernière trame, le signal de sonnerie final s'obtient comme suit:

$$ring(j) = ltring(j) - \sum_{i=1}^8 a_i \cdot ring(j-i), \quad \text{pour } j = 0, 1, 2, \dots, 19 \quad (\text{III-30})$$

On suppose que la fourchette des indices de $j = 0, 1, 2, \dots, 159$ correspond à la première trame perdue en cours, et que la fourchette des indices de $j = 160, 161, 162, \dots, 209$ correspond aux 50 premiers échantillons de la trame suivante; supposons en outre que $wi(j)$ et $wo(j), j = 0, 1, \dots, 19$, soient les fenêtres triangulaires respectivement d'ouverture en fondu et de fondu de fermeture, de sorte que $wi(j) + wo(j) = 1$; l'extrapolation périodique d'ondes s'effectue alors en deux étapes comme suit:

Etape 1:

$$x_{out}(j) = wi(j) \cdot ptfe \cdot x_{out}(n - ppfe) + wo(j) \cdot ring(j), \quad \text{pour } j = 0, 1, 2, \dots, 19 \quad (\text{III-31})$$

Etape 2:

$$x_{out}(j) = ptfe \cdot x_{out}(j - ppfe), \quad \text{pour } j = 20, 21, 22, \dots, 209 \quad (\text{III-32})$$

III.6.10 Générateur de bruit normalisé

Si $merit < MHI$, le bloc avec la légende "Générateur de bruit (normalisé)" dans la Figure III.3 génère une séquence de bruit blanc gaussien aléatoire avec une ampleur d'unité moyenne. Pour réduire la complexité des calculs, le bruit blanc gaussien aléatoire est précalculé et stocké dans un tableau. Pour éviter que le tableau soit trop long, et inversement pour éviter d'avoir à répéter la même structure de bruit s'il était trop court, on utilise un système d'indexation spécial, dans lequel le tableau du bruit blanc gaussien, $wn(j)$, compte 127 entrées, et la version réduite de la sortie de ce bloc de générateur de bruit est:

$$wgn(j) = avm \times wn(\text{mod}(cfecount \times j, 127)), \quad \text{pour } j = 0, 1, 2, \dots, 209 \quad (\text{III-33})$$

où $cfecount$ est le compteur de trames avec $cfecount = k$, pour la k ème trame perdue consécutive dans la perte de paquets en cours, et $\text{mod}(m, 127) = m - 127 \times \lfloor m / 127 \rfloor$ est l'opération par modulo.

III.6.11 Filtrage de la séquence de bruit

Dans la Figure III.3, le bloc avec la légende " $1/A(z)$ " représente un filtre de synthèse sur le court terme. Si $merit < MHI$, il filtre le bruit blanc gaussien réduit pour lui procurer la même enveloppe spectrale que celle du signal $x_{out}(j)$ dans la dernière trame. Le bruit filtré $fn(j)$ s'obtient comme suit:

$$fn(j) = wgn(j) - \sum_{i=1}^8 a_i \cdot fn(j-i), \quad \text{pour } j = 0, 1, 2, \dots, 209 \quad (\text{III-34})$$

III.6.12 Mélange des composantes périodiques et aléatoires

Si $merit > MHI$, seul le signal extrapolé périodiquement $x_{out}(j)$, calculé au § III.6.9, est utilisé comme sortie du PLC-MIC-WB. Si $merit < MLO$, seul le signal de bruit filtré, $fn(j)$, est utilisé comme sortie du PLC-MIC-WB. Si $MLO \leq merit \leq MHI$, les deux composantes se mélangent alors comme suit:

$$x_{out}(j) \leftarrow G_p \cdot x_{out}(j) + G_r \cdot fn(j), \quad \text{pour } j = 0, 1, 2, \dots, 209 \quad (\text{III-35})$$

Les 40 premiers échantillons du signal $x_{out}(j)$ extrapolés pour $j = 160, 161, 162, \dots, 199$, deviendront le signal de sonnerie $ring(j), j = 0, 1, 2, \dots, 39$ de la trame suivante. Si cette dernière

est, là encore, une trame perdue, seuls les 20 premiers échantillons de ce signal de sonnerie seront utilisés pour la superposition/admission.

III.6.13 Réduction progressive conditionnelle

Si la perte de paquets dure 20 ms, ou moins, le signal $x_{out}(j)$ calculé au § III.6.12 est utilisé comme signal de sortie PLC-MIC-WB; si par contre elle dure plus de 60 ms, le signal de sortie PLC-MIC-WB est complètement atténué. Si la perte de paquets dure plus de 20 ms mais moins de 60 ms, le signal $x_{out}(j)$, calculé au § III.6.12, est linéairement réduit (atténué vers 0 de façon linéaire); cette réduction conditionnelle est effectuée comme il est indiqué dans l'algorithme ci-dessous pendant les trames perdues lorsque $cfecount > 2$. Le dispositif $gawd()$ est donné par $\{-52, -69, -104, -207\}$ dans le format Q15. Là encore la fourchette de l'indice de $j = 0, 1, 2, \dots, 159$, correspond à la trame en cours du signal $x_{out}(j)$.

Si $cfecount \leq 6$, effectuer les neuf lignes en retrait suivantes:

$$\delta = gawd(cfecount - 3)$$

$$gaw = 1$$

Pour $j = 0, 1, 2, \dots, 159$, faire les deux lignes suivantes:

$$x_{out}(j) = gaw \cdot x_{out}(j)$$

$$gaw = gaw + \delta$$

Si $cfecount < 6$, faire les trois lignes suivantes:

Pour $j = 160, 161, 162, \dots, 209$, faire les deux lignes suivantes:

$$x_{out}(j) = gaw \cdot x_{out}(j)$$

$$gaw = gaw + \delta$$

Autrement (si $cfecount > 6$), poser $x_{out}(j) = 0$, pour $j = 0, 1, 2, \dots, 209$.

III.6.14 Superposition/admission dans la première trame reçue

Dans les trames du type 5, la sortie du décodeur G.722 $x_{out}(j)$ est après transition ajoutée au signal de sonnerie de la dernière trame perdue, $ring(j)$ (voir le § III.6.12):

$$x_{out}(j) = w_i(j) \cdot x_{out}(j) + w_o(j) \cdot ring(j) \quad j = 0 \dots L_{OLA} - 1 \quad (\text{III-36})$$

où:

$$L_{OLA} = \begin{cases} 8 & \text{si } G_p = 0 \\ 40 & \text{autrement} \end{cases} \quad (\text{III-37})$$

III.7 Recodage de la sortie PLC

Pour actualiser la mémoire et les paramètres des décodeurs MICDA G.722 pendant les trames perdues (trames des types 2, 3 et 4), d'un point de vue conceptuel on passe la sortie PLC au travers du codeur G.722, ce qui suppose de:

- 1) passer la sortie PLC au travers de la banque des filtres d'analyse QMF;
- 2) coder le signal en bande basse de la sous-bande avec le codeur MICDA pour bande basse;
- 3) coder le signal en bande haute de sous-bande avec le codeur MICDA de bande haute.

Pour éviter toute complexité supplémentaire, on conçoit des codeurs de sous-bande MICDA simplifiés. La Figure III.4 illustre les étapes 1 à 3.

III.7.1 Passage de la sortie PLC par la banque des filtres d'analyse QMF

La mémoire de la banque des filtres d'analyse QMF est initialisée pour fournir des signaux de sous-bande qui soient continus avec les signaux de sous-bande décodés. Les 22 premiers échantillons de la sortie PLC-MIC-WB constituent la mémoire filtre, et les signaux de sous-bande sont calculés en fonction de:

$$x_L(n) = \sum_{i=0}^{11} h_{2i} \cdot x_{PLC}(23 + j - 2i) + \sum_{i=0}^{11} h_{2i+1} \cdot x_{PLC}(22 + j - 2i) \quad (\text{III-38})$$

et:

$$x_H(n) = \sum_{i=0}^{11} h_{2i} \cdot x_{PLC}(23 + j - 2i) - \sum_{i=0}^{11} h_{2i+1} \cdot x_{PLC}(22 + j - 2i) \quad (\text{III-39})$$

où $x_{PLC}(0)$ correspond au premier échantillon de la sortie PLC-MIC-WB à 16 kHz de la trame en cours, $x_L(n=0)$ et $x_H(n=0)$ correspondent aux premiers échantillons des signaux de sous-bande à 8 kHz, respectivement en bande basse et en bande haute, de la trame en cours. Le filtrage est identique au QMF d'émission du codeur G.722 à l'exception des 22 échantillons supplémentaires de compensation, et sauf que la sortie PLC-MIC-WB (par opposition à l'entrée) est passée à la banque des filtres. Par ailleurs, pour générer une trame complète (80 échantillons ~ 10 ms) de signaux en sous-bande, le PLC-MIC-WB doit dépasser la trame en cours de 22 échantillons et générer (182 échantillons ~ 11,375 ms). Les signaux de sous-bande $x_L(n)$, $n = 0, 1, \dots, 79$, et $x_H(n)$, $n = 0, 1, \dots, 79$, sont respectivement générés selon les équations III-38 et III-39.

III.7.2 Recodage des signaux en bande basse

Le signal en bande basse $x_L(n)$ est codé avec un codeur MICDA bande basse simplifié (voir la Figure III.7). Le quantificateur inverse a été éliminé, l'erreur de prédiction quantifiée étant remplacée par l'erreur de prédiction non quantifiée. En outre, étant donné que la mise à jour du quantificateur adaptable est basée uniquement sur un sous-ensemble à 8 membres de l'ensemble à 64 membres représenté par l'indice du codeur bande basse à 6 bits, $I_L(n)$, l'erreur de prédiction n'est quantifiée que pour l'ensemble à 8 membres, ce qui donne une mise à jour identique du quantificateur adaptable, tout en simplifiant la quantification. Le Tableau III.2 présente les niveaux des décisions, le code de sortie et les multiplicateurs pour le quantificateur simplifié à 8 niveaux, fondé sur la valeur absolue de $e_L(n)$.

Tableau III.2 – Niveaux de décision, code de sortie et multiplicateurs du quantificateur simplifié à 8 niveaux

m_L	Seuil inférieur	Seuil supérieur	I_L	Multiplicateur, W_L
1	0,00000	0,14103	3c	-0,02930
2	0,14103	0,45482	38	-0,01465
3	0,45482	0,82335	34	0,02832
4	0,82335	1,26989	30	0,08398
5	1,26989	1,83683	2c	0,16309
6	1,83683	2,61482	28	0,26270
7	2,61482	3,86796	24	0,58496
8	3,86796	∞	20	1,48535

La propriété de $p_{Lt}(n)$ en tant que signal constant est contrôlée sur la base des trames pour les trames perdues, et est donc recalée au début de chaque trame perdue; elle est actualisée comme suit:

$$\text{cnst}[p_{Lt}(n)] = \begin{cases} \text{cnst}[p_{Lt}(n-1)]+1 & p_{Lt}(n) = p_{Lt}(n-1) \\ \text{cnst}[p_{Lt}(n-1)] & p_{Lt}(n) \neq p_{Lt}(n-1) \end{cases} \quad (\text{III-46})$$

A la fin des trames perdues 3 à 5 le décodeur bande basse est recalé si la condition ci-après est satisfaite:

$$\left| \frac{\text{sgn}[p_{Lt}(n)]}{N_{lost}} \right| > 36 \quad \text{OU} \quad \text{cnst}[p_{Lt}(n)] > 40 \quad (\text{III-47})$$

où N_{lost} est le nombre de trames perdues, c'est-à-dire 3, 4 ou 5.

III.7.3 Recodage des signaux en bande haute

Le signal en bande haute, $x_H(n)$, est codé avec un codeur MICDA bande haute simplifié (voir la Figure III.8); le quantificateur adaptable a été éliminé car l'algorithme écrase le facteur de l'échelle logarithmique à la première trame reçue avec une moyenne mobile avant la perte, et n'a donc pas besoin du facteur d'échelle logarithmique recodé en bande haute. L'erreur de prédiction quantifiée du codeur MICDA bande haute est remplacée par l'erreur de prédiction non quantifiée.

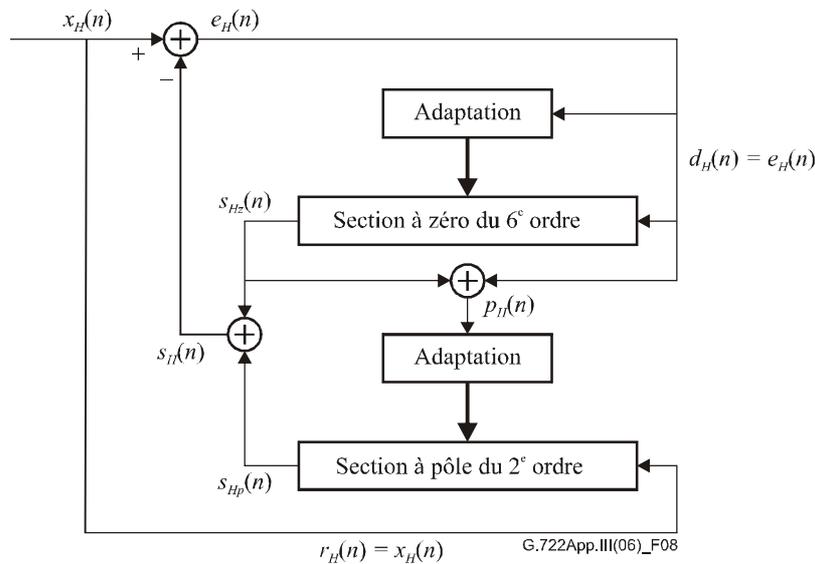


Figure III.8 – Recodage en sous-bande MICDA bande haute

Les entités de la Figure III.8 se calculent suivant leurs équivalents du codeur en sous-bande MICDA bande haute G.722:

$$s_{Hz}(n) = \sum_{i=1}^6 b_{H,i}(n-1) \cdot e_H(n-i) \quad (\text{III-48})$$

$$s_{Hp}(n) = \sum_{i=1}^2 a_{H,i}(n-1) \cdot x_H(n-i) \quad (\text{III-49})$$

$$s_H(n) = s_{Hp}(n) + s_{Hz}(n) \quad (\text{III-50})$$

$$e_H(n) = x_H(n) - s_H(n) \quad (\text{III-51})$$

$$p_H(n) = s_{Hz}(n) + e_H(n) \quad (\text{III-52})$$

L'adaptation des sections à zéro et à pôle se déroule comme dans le codeur G.722, décrit aux § 3.6.3 et 3.6.4 de [UIT-T G.722].

Comme pour le recodage bande basse, le décodeur bande haute est automatiquement recalé après 60 ms de perte de trames, mais il peut se recalculer de façon adaptative au bout de 30 ms pendant la perte de trames. Pendant le recodage du signal bande haute, les propriétés du signal partiel reconstruit, $p_H(n)$, sont contrôlées et gèrent le recalage adaptatif du décodeur MICDA bande haute. Le signe de $p_H(n)$ est contrôlé pendant toute la durée de la perte et est ainsi remis à zéro à la première trame perdue:

$$\text{sgn}[p_H(n)] = \begin{cases} \text{sgn}[p_H(n-1)] + 1 & p_H(n) > 0 \\ \text{sgn}[p_H(n-1)] & p_H(n) = 0 \\ \text{sgn}[p_H(n-1)] - 1 & p_H(n) < 0 \end{cases} \quad (\text{III-53})$$

La propriété de $p_H(n)$ en tant que signal constant est contrôlée sur la base des trames pour les trames perdues et est donc recalée au début de chaque trame perdue; elle est actualisée comme suit:

$$\text{cnst}[p_H(n)] = \begin{cases} \text{cnst}[p_H(n-1)] + 1 & p_H(n) = p_H(n-1) \\ \text{cnst}[p_H(n-1)] & p_H(n) \neq p_H(n-1) \end{cases} \quad (\text{III-54})$$

A la fin des trames perdues 3 à 5 le décodeur bande haute est recalé si la condition suivante est satisfaite:

$$\left| \frac{\text{sgn}[p_H(n)]}{N_{lost}} \right| > 36 \quad \text{OR} \quad \text{cnst}[p_H(n)] > 40 \quad (\text{III-55})$$

III.8 Contrôle des caractéristiques des signaux et de leur utilisation pour le PLC

III.8.1 Facteur d'échelle logarithmique bande basse

Les caractéristiques du facteur d'échelle logarithmique bande basse, $\nabla_L(n)$, sont actualisées pendant les trames reçues et utilisées à la première trame reçue après une perte de trames pour fixer de manière adaptative l'état des quantificateurs adaptables pour le facteur d'échelle. Une mesure de la stationnarité du facteur d'échelle logarithmique bande basse est dérivée et utilisée pour déterminer la bonne remise à la valeur initiale de l'état.

III.8.1.1 Stationnarité du facteur d'échelle logarithmique bande basse

La stationnarité du facteur d'échelle logarithmique bande basse, $\nabla_L(n)$, est calculée et actualisée pendant les trames reçues; elle est basée sur une moyenne mobile du premier ordre, $\nabla_{L,m1}(n)$, de $\nabla_L(n)$ avec fuite à la terre constante:

$$\nabla_{L,m1}(n) = 7/8 \cdot \nabla_{L,m1}(n-1) + 1/8 \cdot \nabla_L(n) \quad (\text{III-56})$$

La mesure du suivage, $\nabla_{L,trck}(n)$, de la moyenne mobile du premier ordre se calcule comme suit:

$$\nabla_{L,trck}(n) = 127/128 \cdot \nabla_{L,trck}(n-1) + 1/128 \cdot \left| \nabla_{L,m1}(n) - \nabla_{L,m1}(n-1) \right| \quad (\text{III-57})$$

La moyenne mobile du deuxième ordre, $\nabla_{L,m2}(n)$, avec une fuite à la terre adaptative se calcule suivant l'équation III-58:

$$\nabla_{L,m2}(n) = \begin{cases} 7/8 \cdot \nabla_{L,m2}(n-1) + 1/8 \cdot \nabla_{L,m1}(n) & \nabla_{L,trck}(n) < 3277 \\ 3/4 \cdot \nabla_{L,m2}(n-1) + 1/4 \cdot \nabla_{L,m1}(n) & 3277 \leq \nabla_{L,trck}(n) < 6554 \\ 1/2 \cdot \nabla_{L,m2}(n-1) + 1/2 \cdot \nabla_{L,m1}(n) & 6554 \leq \nabla_{L,trck}(n) < 9830 \\ \nabla_{L,m2}(n) = \nabla_{L,m1}(n) & 9830 \leq \nabla_{L,trck}(n) \end{cases} \quad (\text{III-58})$$

La stationnarité du facteur d'échelle logarithmique bande basse s'exprime sous forme de degré de changement selon:

$$\nabla_{L,chg}(n) = 127/128 \cdot \nabla_{L,chg}(n-1) + 1/128 \cdot 256 \cdot |\nabla_{L,m2}(n) - \nabla_{L,m2}(n-1)| \quad (\text{III-59})$$

Pendant les trames perdues il n'y a pas d'actualisation, c'est-à-dire:

$$\begin{aligned} \nabla_{L,m1}(n) &= \nabla_{L,m1}(n-1) \\ \nabla_{L,trck}(n) &= \nabla_{L,trck}(n-1) \\ \nabla_{L,m2}(n) &= \nabla_{L,m2}(n-1) \\ \nabla_{L,chg}(n) &= \nabla_{L,chg}(n-1) \end{aligned} \quad (\text{III-60})$$

III.8.1.2 Recalage du facteur d'échelle logarithmique du quantificateur adaptable bande basse

A la première trame reçue après une perte de trames le facteur d'échelle logarithmique bande basse est recalé (écrasé) de façon adaptative en fonction de la stationnarité avant la perte de trames:

$$\nabla_L(n-1) \leftarrow \begin{cases} \nabla_{L,m2}(n-1) & \nabla_{L,chg}(n-1) < 6554 \\ \frac{\nabla_L(n-1)}{3276} [\nabla_{L,chg}(n-1) - 6554] + \frac{\nabla_{L,m2}(n-1)}{3276} [9830 - \nabla_{L,chg}(n-1)] & 6554 \leq \nabla_{L,chg}(n-1) \leq 9830 \\ \nabla_L(n-1) & 9830 < \nabla_{L,chg}(n-1) \end{cases} \quad (\text{III-61})$$

III.8.2 Facteur d'échelle logarithmique bande haute

Les caractéristiques du facteur d'échelle logarithmique bande haute, $\nabla_H(n)$, sont actualisées pendant les trames reçues et utilisées à la trame reçue après une perte de trames pour établir l'état du facteur d'échelle de quantification adaptable. En outre, les caractéristiques gèrent de manière adaptative la convergence du facteur d'échelle logarithmique bande haute après une perte de trames.

III.8.2.1 Moyenne mobile et stationnarité du facteur d'échelle logarithmique bande haute

Le suivage de $\nabla_H(n)$ se calcule en fonction de:

$$\nabla_{H,trck}(n) = 0.97 \cdot \nabla_{H,trck}(n-1) + 0.03 \cdot [\nabla_{H,m}(n-1) - \nabla_H(n)] \quad (\text{III-62})$$

En fonction du suivage, la moyenne mobile se calcule avec la fuite de terre adaptable comme suit:

$$\nabla_{H,m}(n) = \begin{cases} 255/256 \cdot \nabla_{H,m}(n-1) + 1/256 \cdot \nabla_H(n) & |\nabla_{H,trck}(n)| < 1638 \\ 127/128 \cdot \nabla_{H,m}(n-1) + 1/128 \cdot \nabla_H(n) & 1638 \leq |\nabla_{H,trck}(n)| < 3277 \\ 63/64 \cdot \nabla_{H,m}(n-1) + 1/64 \cdot \nabla_H(n) & 3277 \leq |\nabla_{H,trck}(n)| < 4915 \\ 31/32 \cdot \nabla_{H,m}(n-1) + 1/32 \cdot \nabla_H(n) & 4915 \leq |\nabla_{H,trck}(n)| \end{cases} \quad (\text{III-63})$$

La moyenne mobile est utilisée pour recalculer le facteur d'échelle logarithmique bande haute à la première trame reçue, suivant la description du § III.8.2.2.

La mesure de la stationnarité du facteur d'échelle logarithmique bande haute se calcule à partir de la valeur médiane en fonction de:

$$\nabla_{H,chg}(n) = 127/128 \cdot \nabla_{H,chg}(n-1) + 1/128 \cdot 256 \cdot |\nabla_{H,m}(n) - \nabla_{H,m}(n-1)| \quad (\text{III-64})$$

La mesure de la stationnarité est utilisée pour gérer la reconvergence de $\nabla_H(n)$ après une perte de trames (voir le § III.8.2.3).

Pendant les trames perdues il n'y a pas d'actualisation, c'est-à-dire:

$$\begin{aligned} \nabla_{H,trck}(n) &= \nabla_{H,trck}(n-1) \\ \nabla_{H,m}(n) &= \nabla_{H,m}(n-1) \\ \nabla_{H,chg}(n) &= \nabla_{H,chg}(n-1) \end{aligned} \quad (\text{III-65})$$

III.8.2.2 Recalage du facteur d'échelle logarithmique du quantificateur adaptable bande haute

À la première trame reçue le facteur d'échelle logarithmique bande haute est recalé à la moyenne mobile des trames reçues avant la perte:

$$\nabla_H(n-1) \leftarrow \nabla_{H,m}(n-1) \quad (\text{III-66})$$

III.8.2.3 Convergence du facteur d'échelle logarithmique du quantificateur adaptable bande haute

La convergence du facteur d'échelle logarithmique bande haute après une perte de trames est gérée par la mesure de stationnarité, $\nabla_{H,chg}(n)$, avant la perte de trames. En cas de stationnarité, un filtre passe-bas adaptatif est appliqué à $\nabla_H(n)$ après une perte de paquets, et ce pendant 0 ms, 40 ms ou 80 ms, durées pendant lesquelles le degré de filtrage passe-bas est graduellement réduit. La durée des échantillons, N_{LP,∇_H} est déterminée en fonction de:

$$N_{LP,\nabla_H} = \begin{cases} 640 & \nabla_{H,chg} < 819 \\ 320 & \nabla_{H,chg} < 1311 \\ 0 & \nabla_{H,chg} \geq 1311 \end{cases} \quad (\text{III-67})$$

Le filtrage passe-bas s'exprime ainsi:

$$\nabla_{H,LP}(n) = \alpha_{LP}(n) \nabla_{H,LP}(n-1) + (1 - \alpha_{LP}(n)) \nabla_H(n) \quad (\text{III-68})$$

où le coefficient est égal à:

$$\alpha_{LP}(n) = 1 - \left(\frac{n+1}{N_{LP,\nabla_H} + 1} \right)^2 \quad n = 0, 1, \dots, N_{LP,\nabla_H} - 1 \quad (\text{III-69})$$

Le filtrage passe-bas réduit donc échantillon après échantillon en fonction du temps n . Le facteur d'échelle logarithmique ainsi filtré remplace simplement le facteur d'échelle logarithmique normal pendant les échantillons N_{LP,∇_H} .

III.8.3 Section à pôle bande basse

Une entité appelée marge de stabilité (de la section à pôle) est actualisée pendant les trames reçues pour le décodeur MICDA bande basse et utilisée pour restreindre la section à pôle suivant la perte des trames.

III.8.3.1 Marge de stabilité de la section à pôle bande basse

La marge de la stabilité de la section à pôle bande basse se définit ainsi:

$$\beta_L(n) = 1 - |a_{L,1}(n)| - a_{L,2}(n) \quad (\text{III-70})$$

où $a_{L,1}(n)$ et $a_{L,2}(n)$ sont les deux coefficients de pôle. Une moyenne mobile de la marge de stabilité est actualisée selon:

$$\beta_{L,MA}(n) = 15/16 \cdot \beta_{L,MA}(n-1) + 1/16 \cdot \beta_L(n) \quad (\text{III-71})$$

pendant les trames reçues.

Pendant les trames perdues, cette moyenne mobile n'est pas actualisée:

$$\beta_{L,MA}(n) = \beta_{L,MA}(n-1) \quad (\text{III-72})$$

III.8.3.2 Restriction de la section à pôle bande basse

Pendant les opérations normales de codage et de décodage MICDA en bande basse (et en bande haute) une marge de stabilité minimale de $\beta_{L,min} = 1/16$ est maintenue selon la Rec. UIT-T G.722. Pendant les 40 premières ms après une perte de trames, une marge de stabilité minimale accrue est maintenue pour le décodeur MICDA bande basse; elle est fonction de la durée écoulée depuis la perte des trames et de la moyenne mobile de la marge de stabilité.

Pour les trois premières trames de 10 ms, une marge de stabilité minimale de:

$$\beta_{L,min} = \min\{3/16, \beta_{L,MA}(n-1)\} \quad (\text{III-73})$$

est fixée à la frontière des trames et mise en œuvre tout au long de la trame. A la frontière des trames dans la quatrième trame de 10 ms, une marge de stabilité minimale de:

$$\beta_{L,min} = \min\left\{2/16, \frac{1/16 + \beta_{L,MA}(n-1)}{2}\right\} \quad (\text{III-74})$$

est mise en œuvre, tandis que la marge de stabilité minimale normale de $\beta_{L,min} = 1/16$ est appliquée pour toutes les autres trames.

III.8.4 Signaux partiels reconstruits et signaux reconstruits bande haute

Pendant toute la durée des trames, qu'elles soient perdues ou reçues, les versions filtrées passe-haut du signal partiel reconstruit bande haute, $p_H(n)$, et du signal reconstruit, $r_H(n)$, sont maintenues:

$$p_{H,HP}(n) = 0,97[p_H(n) - p_H(n-1) + p_{H,HP}(n-1)] \quad (\text{III-75})$$

$$r_{H,HP}(n) = 0,97[r_H(n) - r_H(n-1) + r_{H,HP}(n-1)] \quad (\text{III-76})$$

Ce qui correspond à une diminution de 3 dB d'environ 40 Hz, soit en gros à une suppression de la composante directe.

Pendant les 40 premières ms après une perte de trames le signal partiel normal reconstruit et le signal normal construit sont remplacés par leurs versions filtrées passe-haut respectives aux fins respectivement d'adaptation de la section à pôle bande haute et de la sortie du signal reconstruit bande haute.

III.9 Calcul des décalages temporels

Les techniques de rephasage (§ III.10) et d'alignement temporel (§ III.11) exigent le nombre d'échantillons nécessaire pour indiquer que la forme d'onde de masquage des trames perdues, $x_{PLC}(j)$, et le signal de la première trame reçue ne sont pas bien alignés.

III.9.1 Estimation peu complexe du signal reconstruit en sous-bande inférieure

Le signal utilisé dans la première trame reçue pour calculer le décalage temporel s'obtient en filtrant le signal de différence tronqué dans la sous-bande inférieure, $d_{Lt}(n)$, voir Equation 3-11 de [UIT-T G.722] avec les coefficients du filtre à pôle zéro ($a_{Lpwe,i}(159)$, $b_{Lpwe,i}(159)$) et d'autres renseignements d'état nécessaires émanant de $STATE_{159}$ (voir § III.10.1):

$$r_{Le}(n) = \sum_{i=1}^2 a_{Lpwe,i}(159) \cdot r_{Le}(n-i) + \sum_{i=1}^6 b_{Lpwe,i}(159) \cdot d_{Lt}(n-i) + d_{Lt}(n), \quad \text{pour } n = 0, 1, \dots, 79 \quad (\text{III-77})$$

III.9.2 Détermination des besoins de rephasage et d'alignement temporel

Si la dernière trame reçue est non voisée, selon l'indication du facteur de mérite, le décalage temporel T_L est mis à zéro:

$$IF \text{ merit} \leq MLO, T_L = 0 \quad (\text{III-78})$$

En outre, si la première trame reçue est non voisée, selon l'indication du premier coefficient d'autocorrélation normalisé,

$$r(1) = \frac{\sum_{n=0}^{78} r_{Le}(n) \cdot r_{Le}(n)}{\sum_{n=0}^{78} r_{Le}(n) \cdot r_{Le}(n+1)} \quad (\text{III-79})$$

le décalage temporel est mis à zéro:

$$IF r(1) < 0,125, T_L = 0 \quad (\text{III-80})$$

Autrement, le décalage temporel est calculé comme il est expliqué ci-dessous.

III.9.3 Calcul du décalage temporel

Le calcul du décalage temporel se décompose comme suit:

- 1) génération du signal extrapolé;
- 2) recherche du décalage temporel approximatif;
- 3) recherche du décalage temporel affiné.

Ces étapes sont décrites dans les paragraphes ci-après.

III.9.3.1 Génération du signal extrapolé

Un décalage temporel représente un mauvais alignement entre $x_{PLC}(j)$ et $r_{Le}(n)$; pour calculer ce mauvais alignement, on prolonge $x_{PLC}(j)$ sur la première trame reçue et on maximalise une fonction de corrélation croisée normalisée. Le présent paragraphe décrit comment est extrapolé $x_{PLC}(j)$, qui indique la longueur du signal nécessaire. Comme au § III.6, on part du principe que $x_{PLC}(j)$ est copié dans la mémoire tampon $x_{out}(j)$. Etant donné qu'il s'agit d'une trame du type 5 (première trame reçue), la correspondance supposée est:

$$x_{out}(j-160) = x_{PLC}(j), \quad j = 0, 1, \dots, 159 \quad (\text{III-81})$$

La fourchette à l'intérieur de laquelle est recherchée la corrélation s'établit ainsi:

$$\Delta_{TL} = \min(\lfloor ppfe \cdot 0,5 + 0,5 \rfloor + 3, \Delta_{TLMAX}) \quad (\text{III-82})$$

où $\Delta_{TLMAX} = 28$, et $ppfe$ est la période fondamentale pour l'extrapolation périodique d'ondes utilisée dans les générations de $x_{PLC}(j)$.

La taille de la fenêtre (pour un échantillonnage à 16 kHz) pour la recherche du décalage temporel se calcule comme suit:

$$LSW_{16k} = \begin{cases} 80 & \lfloor ppfe \cdot 1,5 + 0,5 \rfloor < 80 \\ 160 & \lfloor ppfe \cdot 1,5 + 0,5 \rfloor > 160 \\ \lfloor ppfe \cdot 1,5 + 0,5 \rfloor & \text{autrement} \end{cases} \quad (\text{III-83})$$

Il n'est pas inutile de spécifier la fenêtre de recherche du décalage, LSW , pour un échantillonnage à 8 kHz, c'est-à-dire:

$$LSW = \lfloor LSW_{16k} \cdot 0,5 \rfloor \quad (\text{III-84})$$

Etant donné ce qui précède, la longueur totale du signal extrapolé qui doit être dérivé de $x_{PLC}(j)$ est égale à:

$$L = 2 \cdot (LSW + \Delta_{TL}) \quad (\text{III-85})$$

La position de départ du signal extrapolé par rapport au premier échantillon dans la trame reçue est la suivante:

$$D = 12 - \Delta_{TL} \quad (\text{III-86})$$

Le signal extrapolé $es(j)$ est construit de la manière suivante:

Si $D < 0$

$$es(j) = x_{out}(D + j) \quad j = 0, 1, \dots, -D - 1$$

Si $(L + D \leq ppfe)$

$$es(j) = x_{out}(-ppfe + D + j) \quad j = -D, -D + 1, \dots, L - 1$$

Autrement

$$es(j) = x_{out}(-ppfe + D + j) \quad j = -D, -D + 1, \dots, ppfe - D - 1$$

$$es(j) = es(j - ppfe) \quad j = ppfe - D, ppfe - D + 1, \dots, L - 1$$

Autrement

$$ovs = ppfe \cdot \lceil D / ppfe \rceil - D$$

Si $(ovs \geq L)$

$$es(j) = x_{out}(-ovs + j) \quad j = 0, 1, \dots, L - 1$$

Autrement

Si $(ovs > 0)$

$$es(j) = x_{out}(-ovs + j) \quad j = 0, 1, \dots, ovs - 1$$

Si $(L - ovs \leq ppfe)$

$$es(j) = x_{out}(-ovs - ppfe + j) \quad j = ovs, ovs + 1, \dots, L - 1$$

Autrement

$$es(j) = x_{out}(-ovs - ppfe + j) \quad j = ovs, ovs + 1, \dots, ovs + ppfe - 1$$

$$es(j) = es(j - ppfe) \quad j = ovs + ppfe, ovs + ppfe + 1, \dots, L - 1$$

III.9.3.2 Recherche de décalage temporel approximatif

Un décalage temporel approximativement estimé, T_{LSUB} , est d'abord calculé en recherchant le pic de la fonction de corrélation croisée normalisée, sous-échantillonnée, $R_{SUB}(k)$:

$$R_{SUB}(k) = \frac{\sum_{i=0}^{LSW/2-1} es(4i - k + \Delta_{TL}) \cdot r_{Le}(2i)}{\sqrt{\sum_{i=0}^{LSW/2-1} es^2(4i - k + \Delta_{TL}) \sum_{i=0}^{LSW-1} r_{Le}^2(2i)}}, \quad k = -\Delta_{TL}, -\Delta_{TL} + 4, -\Delta_{TL} + 8, \dots, \Delta_{TL} \quad (III-87)$$

Pour éviter de chercher à l'extérieur des limites pendant l'opération d'affinement, T_{LSUB} peut être ajusté comme suit:

$$\text{Si } (T_{LSUB} > \Delta_{TLMAX} - 4) \quad T_{LSUB} = \Delta_{TLMAX} - 4 \quad (III-88)$$

$$\text{Si } (T_{LSUB} < -\Delta_{TLMAX} + 4) \quad T_{LSUB} = -\Delta_{TLMAX} + 4 \quad (III-89)$$

III.9.3.3 Recherche du décalage temporel affiné

La recherche est ensuite affinée pour trouver le décalage temporel, T_L , en recherchant le pic de $R(k)$ donné par:

$$R(k) = \frac{\sum_{i=0}^{LSW-1} es(2i - k + \Delta_{TL}) \cdot r_{Le}(i)}{\sqrt{\sum_{i=0}^{LSW-1} es^2(2i - k + \Delta_{TL}) \sum_{i=0}^{LSW-1} r_{Le}^2(i)}}, \quad k = -4 + T_{LSUB}, -2 + T_{LSUB}, \dots, 4 + T_{LSUB} \quad (III-90)$$

Enfin, on vérifie les conditions ci-après:

Si:

$$\sum_{i=0}^{LSW-1} r_{Le}^2(i) = 0 \quad (III-91)$$

Ou:

$$\sum_{i=0}^{LSW-1} es(2i - T_L + \Delta_{TL}) \cdot r_{Le}(i) \leq 0,25 \cdot \sqrt{\sum_{i=0}^{LSW-1} r_{Le}^2(i)} \quad (III-92)$$

Ou:

$$(T_L > \Delta_{TLMAX} - 2) \parallel (T_L < -\Delta_{TLMAX} + 2) \quad (III-93)$$

Alors:

$$T_L = 0$$

III.10 Rephasage

Le rephasage consiste à établir les états internes en un point du temps où la forme d'ondes de masquage des trames perdues, $x_{PLC}(j)$, est en phase avec le dernier échantillon de signaux d'entrée précédent immédiatement la première trame reçue; le processus peut se décomposer comme suit:

- 1) stocker les états G.722 intermédiaires pendant le recodage des trames perdues;
- 2) ajuster le recodage en fonction du décalage temporel;
- 3) mettre à jour la mémoire des filtres de synthèse QMF.

Ces trois étapes vont être maintenant décrites plus en détail dans les paragraphes ci-dessous.

III.10.1 Stockage des états G.722 intermédiaires pendant le recodage

Comme il est indiqué au § III.7, le signal reconstruit $x_{PLC}(j)$ est recodé pendant les trames perdues pour que soit mise à jour la mémoire des états du décodeur G.722. Si $STATE_j$ est l'état G.722 et l'état PLC après recodage du j ème échantillon de $x_{PLC}(j)$, alors en plus de l'état G.722 à la frontière des trames qui normalement seraient maintenues (c'est-à-dire $STATE_{159}$), l'état $STATE_{159-\Delta_{TLMAX}}$ est également mémorisé. Pour faciliter le processus de rephasage, les signaux de sous-bande $x_L(n)$, $x_H(n)$, $n = 69 - \Delta_{TLMAX}/2 \dots 79 + \Delta_{TLMAX}/2$ sont eux aussi sauvegardés.

III.10.2 Ajustement du recodage en fonction du décalage temporel

Suivant le signe du décalage temporel, l'ajustement du recodage se fait selon la procédure suivante:

Si $\Delta_{TL} > 0$:

- 1) rétablir l'état G.722 et l'état PLC à $STATE_{159-\Delta_{TLMAX}}$.
- 2) recoder $x_L(n), x_H(n)$, $n = 80 - \Delta_{TLMAX}/2 \dots 79 - \Delta_{TL}/2$, suivant les dispositions des § III.7.2 et III.7.3.

Si $\Delta_{TL} < 0$:

- 1) rétablir l'état G.722 et l'état PLC à $STATE_{159}$.
- 2) recoder $x_L(n), x_H(n)$, $n = 80 \dots 79 + |\Delta_{TL}/2|$ en fonction des dispositions des § III.7.2 et III.7.3.

A noter que, pour faciliter le recodage de $x_L(n)$ et de $x_H(n)$ jusqu'à $n = 79 + |\Delta_{TL}/2|$, il faut disposer d'échantillons jusqu'à $\Delta_{TLMAX} + 182$ de $x_{PLC}(j)$.

III.10.3 Mise à jour de la mémoire des filtres de synthèse QMF

A la première trame reçue, la mémoire des filtres de synthèse QMF doit être calculée étant donné que la banque des filtres de synthèse QMF est inactive pendant les trames perdues, l'opération PLC intervenant dans le domaine des signaux vocaux à 16 kHz. Dans la séquence temporelle, la mémoire correspondrait généralement aux derniers échantillons de la dernière trame perdue, mais il convient de tenir compte du processus de rephasage. Selon les dispositions de [UIT-T G.722], la mémoire des filtres de synthèse QMF est donnée par:

$$x_d(i) = r_L(n-i) - r_H(n-i), \quad i = 1, 2, \dots, 11 \quad (\text{III-94})$$

$$x_s(i) = r_L(n-i) + r_H(n-i), \quad i = 1, 2, \dots, 11 \quad (\text{III-95})$$

alors que les deux premiers échantillons de sortie de la première trame reçue sont obtenus par le calcul suivant:

$$x_{out}(j) = 2 \sum_{i=0}^{11} h_{2i} \cdot x_d(i) \quad (\text{III-96})$$

$$x_{out}(j+1) = 2 \sum_{i=0}^{11} h_{2i+1} \cdot x_s(i) \quad (\text{III-97})$$

La mémoire des filtres, c'est-à-dire $x_d(i)$ et $x_s(i)$, $i = 1, 2, \dots, 11$, est calculée à partir des 11 derniers échantillons des signaux d'entrée rephasés jusqu'au codeur MICDA des sous-bandes simplifié pendant l'opération de recodage, $x_L(n)$ et $x_H(n)$, $n = 69 - \Delta_{TL}/2, 69 - \Delta_{TL}/2 + 1, \dots, 79 - \Delta_{TL}/2$, c'est-à-dire les derniers échantillons jusqu'au point de rephasage:

$$x_d(i) = x_L(80 - \Delta_{TL}/2 - i) - x_H(80 - \Delta_{TL}/2 - i), \quad i = 1, 2, \dots, 11 \quad (\text{III-98})$$

$$x_s(i) = x_L(80 - \Delta_{TL}/2 - i) + x_H(80 - \Delta_{TL}/2 - i), \quad i = 1, 2, \dots, 11 \quad (\text{III-99})$$

où $x_L(n)$ et $x_H(n)$ ont été sauvegardés dans la mémoire d'états pendant la trame perdue.

III.11 Alignement temporel

L'alignement temporel consiste à étirer ou à contracter un signal le long de l'axe temporel. Le présent paragraphe décrit la procédure d'alignement temporel du signal $x_{out}(j)$ pour en améliorer l'alignement avec le signal extrapolé de la forme d'onde périodique $x_{PLC}(j)$. L'algorithme décrit dans les paragraphes ci-après est exécuté uniquement à condition que $T_L \neq 0$.

III.11.1 Affinement du décalage temporel

Le décalage temporel, T_L , est affiné aux fins d'alignement temporel, par maximalisation de la corrélation croisée dans la fenêtre superposition/admission. Le point de départ estimé de cette fenêtre, à l'intérieur de la première trame reçue, en fonction de T_L est donné par:

$$SP_{OLA} = \max(0, MIN_UNSTBL - T_L) \quad (\text{III-100})$$

où $MIN_UNSTBL = 16$.

Le point de départ du signal extrapolé par rapport à SP_{OLA} est donné par:

$$D_{ref} = SP_{OLA} - T_L - RSR \quad (\text{III-101})$$

où $RSR = 4$ est la fourchette de recherche pour l'affinement.

La longueur requise du signal extrapolé est donnée par:

$$L_{ref} = OLALG + RSR \quad (\text{III-102})$$

Pour un signal extrapolé, $es_{tw}(j)$, on utilise les mêmes procédures que celles décrites au § III.9.3.1, sauf pour $LSW = OLALG$, $L = L_{ref}$, et $D = D_{ref}$.

Un décalage d'affinement, T_{ref} , est calculé par recherche du pic de l'élément suivant:

$$R(k) = \frac{\sum_{i=0}^{OLALG-1} es_{tw}(i - k + RSR) \cdot x_{out}(i + SP_{OLA})}{\sqrt{\left(\sum_{i=0}^{OLALG-1} es_{tw}^2(i - k + RSR) \right) \cdot \left(\sum_{i=0}^{OLALG-1} x_{out}^2(i + SP_{OLA}) \right)}}, \quad k = -RSR, -RSR + 1, \dots, RSR \quad (\text{III-103})$$

Le décalage temporel final utilisé pour l'alignement temporel est alors calculé comme suit:

$$T_{Lwarp} = T_L + T_{ref} \quad (\text{III-104})$$

III.11.2 Calcul du signal $x_{out}(j)$ aligné dans le temps

Le signal $x_{out}(j)$ est aligné dans le temps par des échantillons T_{Lwarp} pour former le signal $x_{warp}(j)$ qui ensuite est ajouté après transition au signal extrapolé en forme d'onde, $es_{ola}(j)$. Selon la valeur de T_{Lwarp} , il existe trois cas, qui sont illustrés à la Figure III.9. Dans le cas a), $T_{Lwarp} < 0$ et $x_{out}(j)$ subit un rétrécissement ou une compression. Les premiers échantillons MIN_UNSTBL de $x_{out}(j)$ ne sont pas utilisés dans l'alignement $x_{out}(j)$ pour la génération de $x_{warp}(j)$, et on a

$xstart = MIN_UNSTBL$. Dans le cas b, on a $0 \leq T_{Lwarp} < MIN_UNSTBL$, et $x_{out}(j)$ est rallongé de T_{Lwarp} échantillons. Là encore, les premiers échantillons MIN_UNSTBL de $x_{out}(j)$ ne sont pas utilisés, et $xstart = MIN_UNSTBL$. Dans le cas c), $T_{Lwarp} \geq MIN_UNSTBL$, et $x_{out}(j)$ est une fois de plus rallongé de T_{Lwarp} échantillons mais les premiers T_{Lwarp} échantillons de $x_{out}(j)$ ne sont pas nécessaires dans ce cas puisque des T_{Lwarp} échantillons supplémentaires seront créés pendant l'alignement, de sorte que $xstart = T_{Lwarp}$.

Dans chaque cas, le nombre d'échantillons par superposition/admission est donné par:

$$spad = \frac{(160 - xstart)}{|T_{Lwarp}|} \quad (III-105)$$

L'alignement est implémenté par un décalage d'un seul échantillon en pas à pas et une superposition/admission triangulaire, à partir de $x_{out}[xstart]$. Pour procéder au rétrécissement, on supprime un échantillon à intervalles périodiques; de ce point de vue, le signal original et le signal décalé à gauche (par suite de la suppression de l'échantillon) sont ajoutés après transition. Pour procéder à l'étirement, un échantillon est répété à intervalles périodiques; de ce point de vue, le signal original et le signal décalé vers la droite (par suite de la répétition de l'échantillon) sont ajoutés après transition. La longueur de la fenêtre de superposition/adjonction, $L_{olawarp}$, (noter que différente de la région OLA illustrée à la Figure III.9) dépend de la périodicité de la superposition/adjonction de l'échantillon et est calculée comme suit:

$$\begin{aligned} \text{Si } T_{Lwarp} < 0, L_{olawarp} &= \frac{(160 - xstart - |T_{Lwarp}|)}{|T_{Lwarp}|} \\ \text{Autrement } L_{olawarp} &= \lceil spad \rceil \\ L_{olawarp} &= \min(8, L_{olawarp}) \end{aligned} \quad (III-106)$$

La durée du signal d'entrée aligné, x_{warp} est égale à:

$$L_{xwarp} = \min(160, 160 - MIN_UNSTBL + T_{Lwarp}) \quad (III-107)$$

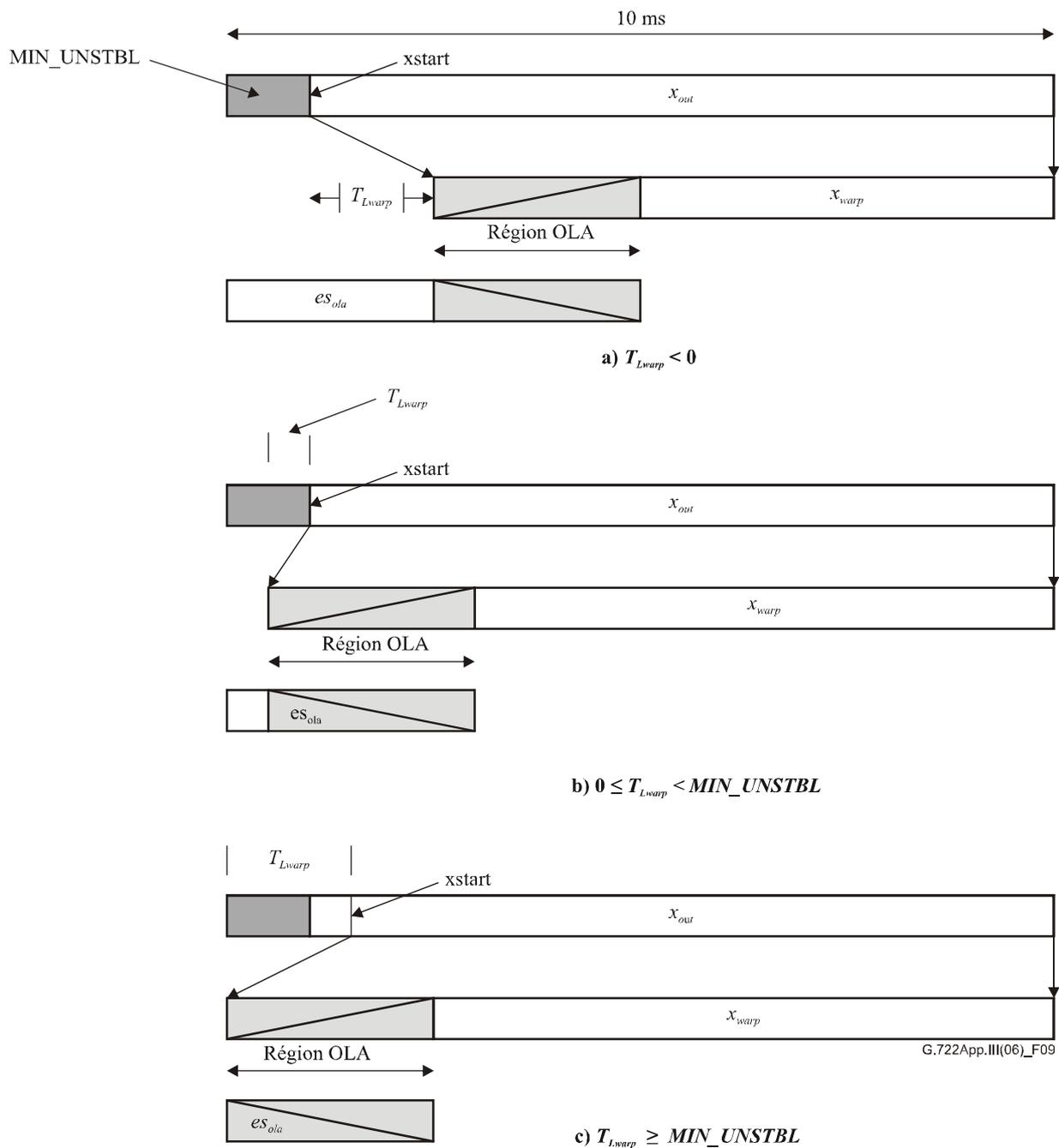


Figure III.9 – Trois cas d'alignement de x_{out}

III.11.3 Calcul du signal extrapolé en forme d'onde

Le signal aligné $x_{warp}(j)$ et le signal extrapolé $es_{ola}(j)$ sont ajoutés après transition dans la première trame reçue (voir la Figure III.9). Le signal extrapolé $es_{ola}(j)$ est généré directement à l'intérieur de la mémoire tampon du signal $x_{out}(j)$ en deux étapes:

Etape 1:

$$es_{ola}(j) = x_{out}(j) = ptfe \cdot x_{out}(j - ppfe), \quad j = 0, 1, \dots, 160 - L_{xwarp} + 39 \quad (\text{III-108})$$

Etape 2:

$$x_{out}(j) = x_{out}(j) \cdot w_i(j) + ring(j) \cdot w_o(j), \quad j = 0, 1, \dots, 39 \quad (\text{III-109})$$

où $w_i(j)$ et $w_o(j)$ sont des fenêtres de réduction ascendante et descendante, triangulaires, par superposition/adjonction de longueur 40 et $ring(j)$ est le signal de sonnerie calculé conformément au § III.6.12.

III.11.4 Superposition/admission du signal aligné dans le temps avec le signal extrapolé en forme d'onde

Le signal extrapolé calculé au § III.11.3 est ajouté par transition au signal aligné $x_{warp}(j)$ en fonction de:

$$x_{out}(160 - L_{xwarp} + j) = x_{out}(160 - L_{xwarp} + j) \cdot w_o(j) + x_{warp}(j) \cdot w_i(j), \quad j = 0, 1, \dots, 39 \quad (\text{III-110})$$

La partie restante de $x_{warp}(j)$ est ensuite simplement copiée dans la mémoire tampon du signal:

$$x_{out}(160 - L_{xwarp} + j) = x_{warp}(j), \quad j = 40, 41, \dots, L_{xwarp} - 1 \quad (\text{III-111})$$

III.12 Description au bit près de l'algorithme PLC G.722

Le masquage PLC G.722 est spécifié au bit près par le code ANSI C à virgule fixe. En cas de divergence, c'est la spécification du code C qui l'emporte sur la spécification textuelle de la présente Recommandation. Le code ANSI C à virgule fixe est implémenté à l'aide du code C G.722 de la Liste des outils logiciels (STL2005) dressée dans la Recommandation UIT-T G.191 ainsi que la version 2.3 de la Bibliothèque des opérateurs à virgule fixe à 16 bits, qui fait également partie de la liste STL2005.

III.12.1 Utilisation du logiciel de simulation

Le decg722 exécutable (y compris le PLC) est appelé par

```
decg722 [-fsize N] g192_input_file speech_output_file
```

où N est la durée de la trame, tout multiple de 160, c'est-à-dire tout multiple de 10 ms.

A noter que le mode et la longueur de trame sont indirectement intégrés dans le train de bits G.192, de sorte que tout conflit avec une longueur de trame indiquée par une ligne de commande amènera le décodeur à utiliser une longueur de trame conforme au train de bits G.192. Pour des longueurs de trame de 10 ms et de 20 ms, le décodeur peut déterminer de façon unique aussi bien le mode que la longueur de trame à partir du train de bits G.192; toutefois, pour des longueurs de 30 ms et plus, une longueur de trame de ligne de commande doit être spécifiée pour permettre au décodeur de déterminer correctement et le mode et la longueur de la trame.

III.12.2 Organisation du logiciel de simulation

Le code source est contenu dans le répertoire dénommé "src". Un fichier d'espace de travail Microsoft Visual C 6.0 est situé dans "workspace/VC6.0/". Ouvrir g722_plc_g192.dsw ouvrira le code source C et le projet PLC G.722.

Après compilation, il est possible de procéder à un simple test de fonctionnement au bit près dans le répertoire dénommé "testplc"; exécuter le perl script intitulé "testplc.pl" dans le répertoire lancera l'exécution du test.

A noter que l'exécution de ce simple test pour vérifier l'exactitude au bit près permet seulement de procéder à une simple vérification, et non pas de vérifier une implémentation pour voir si elle fonctionne bien.

Le Tableau III.3 présente la nouvelle mémoire des états du décodeur G.722 nécessaire à la mise en œuvre de la procédure PLC, tandis que le Tableau III.4 est une récapitulation des nouveaux tableaux.

**Tableau III.3 – Nouvelle mémoire des états du décodeur G.722
(structure WB_PLC_state)**

Membre	Mots (16 bits)	Description
energymax32	2	Energie
cormax	2	Corrélation
wsz	1	Taille de la fenêtre
scaled_flag	1	Drapeau d'échelle
xq	638	Tampon en sortie des signaux vocaux à 16 kHz
stsym1	8	Mémoire des filtres de synthèse sur le court terme
al	9	Coefficients des filtres LPC
alast	9	Coefficients des filtres LPC passés
pPpt	1	Source du prédicteur de pitch
stwpml	8	Mémoire des filtres tout-pôle, pondérée, sur le court terme
xwd	45	Tampon des signaux vocaux, pondéré, échantillonnage descendant
xwd_exp	1	Exposant de tampon de signaux vocaux, pondéré, échantillon descendant
dfm	60	Echantillon descendant, mémoire à filtre
scaler	1	Facteur d'échelle pour composante aléatoire
merit	1	Facteur de mérite pour ratio de mélange
ptfe	1	Source de pitch pour effacement de trame
ppf	1	Période fondamentale – valeur à virgule "flottante"
ppinc	1	Incrément de période fondamentale
pweflag	1	Drapeau d'extrapolation périodique d'onde
cpplast	1	Dernier pitch approximatif
pph	5	Historique de la période fondamentale
pp	1	Période fondamentale
cfecount	1	Compteur d'effacement de trames consécutives
ngfae	1	Nombre de bonnes trames après effacement
nfle	1	Nombre de trames dans le dernier effacement
avm	1	Ampleur moyenne
lag	1	Décalage temporel
psml_mean	1	Marge de la section des pôles, moyenne bande basse
nbpl_mean1	1	nbpl première moyenne (bande basse)
nbpl_mean2	1	nbpl seconde moyenne (bande basse)
nbpl_trck	1	nbpl suivage (bande basse)
nbpl_chng	1	nbpl changement (bande basse)
pl_postn	1	Mesure positive-négative du signal pl (bande basse)
lb_reset	1	Drapeau de reset du décodeur bande basse

**Tableau III.3 – Nouvelle mémoire des états du décodeur G.722
(structure WB_PLC_state)**

Membre	Mots (16 bits)	Description
nbph_mean	1	nbph moyenne (bande haute)
nbph_trck	1	nbph suivage (bande haute)
nbph_chng	1	nbph changement (bande haute)
nbh_mode	1	mode nbh pour convergence (bande haute)
hp_flag	1	Drapeau pour filtre hp sur signaux rh et ph (bande haute)
nbph_lp	1	nbph après filtre passe-bas (bande haute)
ph_postn	1	mesure positive-négative du signal ph (bande haute)
hb_reset	1	Drapeau de reset du décodeur bande haute
rhhp_m1	1	Echantillon passé de signal rh après filtre passe-haut (bande haute)
rh_m1	1	Echantillon passé de signal rh (bande haute)
phhp_m1	1	Echantillon passé de signal ph après filtre passe-haut (bande haute)
ph_m1	1	Echantillon passé de signal ph (bande haute)
sb_sample	1	Nombre d'échantillons par sous-bande
cpl_postn	1	Copie de pl_postn
cph_postn	1	Copie de ph_postn
crhhp_m1	1	Copie de rhhp_m1
crh_m1	1	Copie de rh_m1
cphp_m1	1	Copie de phhp_m1
cph_m1	1	Copie de ph_m1
ds	104	Copie d'états de décodeur G.722 normal
lb	39	Signal bande basse
hb	39	Signal bande haute

Tableau III.4 – Nouveau tableau ROM de décodeur G.722

Tableau	Mots (16 bits)	Description
inv_frm_size	3	Inverse de longueur de trame
wlil4rilil	9	Tableau pour mise à jour de facteur d'échelle bande basse
q4	8	Tableau pour mise à jour de facteur d'échelle bande basse
NGFAEOFFSET_P1	8	Compensation d'échantillon en trames de 10 ms
div_n	20	Tableau de division
gawd	4	Tableau d'atténuation graduelle
olaup	16	Tableau de superposition/admission
oladown	16	Tableau de superposition/admission
wn	127	Tableau d'échantillons de bruit normalisés
bdf	60	Filtre pour décimation 8:1

Tableau III.4 – Nouveau tableau ROM de décodeur G.722

Tableau	Mots (16 bits)	Description
x	4	Pour pitch approximatif
x2	4	Pour pitch approximatif
invk	4	Pour pitch approximatif
MPTH	4	Pour pitch approximatif
sstwin_h	8	16 bits supérieurs pour lissage spectral
sstwin_l	8	16 bits inférieurs pour lissage spectral
bwel	9	Expansion de largeur de bande
STWAL	9	Pour filtre de pondération court terme
win	160	Fenêtre pour analyse LPC
tablog	33	Tableau pour fonction log2
olaug	40	Fenêtre pour superposition/admission
oladg	40	Fenêtre pour superposition/admission
nbphtab	8	Tableau pour nbph
nbpltab	6	Tableau pour nbpl
ola3	3	Tableau pour superposition/admission
ola4	4	Tableau pour superposition/admission
ola5	5	Tableau pour superposition/admission
ola6	6	Tableau pour superposition/admission
ola7	7	Tableau pour superposition/admission
ola8	8	Tableau pour superposition/admission

Le Tableau III.5 énumère les fichiers source du code C de [UIT-T G.722], émanant [UIT-T G.191], qui reste identiques; le Tableau III.6 indique les fichiers de [UIT-T G.722] qui sont modifiés et le Tableau III.7 désigne les nouveaux fichiers de l'algorithme PLC G.722.

Tableau III.5 – Fichiers source identiques G.722

Nom du fichier	Description
decg722.c	Fonction de décodeur principal G.722
softbit.c/h	Fonctions softbit G.722
g722_com.h	Définitions G.722 communes
ugstdemo.h	Définitions pour programmes de démonstration UGST

Tableau III.6 – Fichiers source modifiés G.722

Nom de fichier	Description
funcg722.c/h	Fonctions G.722
g722.c	Fonctions de codage/décodage de trames G.722

Tableau III.7 – Nouveaux fichiers source PLC G.722

Nom de fichier	Description
apfilter.c	Fonctions de filtre tout-pôle
autocor.c	Fonction d'autocorrélation
azfilter.c	Fonction de filtre tout-zéro
coarptch.c	Analyse de pitch approximative
decim.c	Fonction de décimation 8:1
dspfunc.c	Fonctions DSP
g722plc.c/h	Fonctions PLC G.722
levinson.c	Fonction de récursion Levinson-Durbin
memutil.c/h	Fonctions d'utilité mémoire permettant une utilisation automatique de la mémoire scratch
merit.c	Fonction de calcul du mérite
ppchange.c/h	Fonctions relatives au rephasage et à l'alignement temporel
prfn.c	Fonction d'affinement du pitch
table.c/h	Tableau ROM
utility.c/h	Fonctions utilitaires

SÉRIES DES RECOMMANDATIONS UIT-T

Série A	Organisation du travail de l'UIT-T
Série D	Principes généraux de tarification
Série E	Exploitation générale du réseau, service téléphonique, exploitation des services et facteurs humains
Série F	Services de télécommunication non téléphoniques
Série G	Systèmes et supports de transmission, systèmes et réseaux numériques
Série H	Systèmes audiovisuels et multimédias
Série I	Réseau numérique à intégration de services
Série J	Réseaux câblés et transmission des signaux radiophoniques, télévisuels et autres signaux multimédias
Série K	Protection contre les perturbations
Série L	Construction, installation et protection des câbles et autres éléments des installations extérieures
Série M	Gestion des télécommunications y compris le RGT et maintenance des réseaux
Série N	Maintenance: circuits internationaux de transmission radiophonique et télévisuelle
Série O	Spécifications des appareils de mesure
Série P	Qualité de transmission téléphonique, installations téléphoniques et réseaux locaux
Série Q	Commutation et signalisation
Série R	Transmission télégraphique
Série S	Equipements terminaux de télégraphie
Série T	Terminaux des services télématiques
Série U	Commutation télégraphique
Série V	Communications de données sur le réseau téléphonique
Série X	Réseaux de données, communication entre systèmes ouverts et sécurité
Série Y	Infrastructure mondiale de l'information, protocole Internet et réseaux de prochaine génération
Série Z	Langages et aspects généraux logiciels des systèmes de télécommunication