

HUITIEME COLLOQUE SUR LE TRAITEMENT DU SIGNAL ET SES APPLICATIONS

NICE du 1^{er} au 5 JUIN 1981

MESURE DU RAPPORT SIGNAL A BRUIT DE CODECS
PAR UNE METHODE D'INTERMODULATION GENERALISEE

André GILLOIRE

Centre National d'Etudes des Télécommunications LANNION A

RESUME

La qualité des codecs (codeurs-décodeurs : convertisseurs analogique-numérique-analogique classiques ou autres) est exprimée habituellement en termes de rapport signal à bruit. Les procédés de mesure utilisés actuellement dans le cadre de la téléphonie, s'ils donnent des résultats significatifs sur des codages simples comme le MIC, fournissent des résultats parfois sans rapport avec la qualité subjective sur des systèmes plus complexes comme par exemple les codecs MIC différentiels. Nous présentons un procédé de mesure du type intermodulation, permettant l'accès à la puissance et au spectre du bruit de codage quel que soit le spectre du signal codé, lequel doit cependant être stationnaire. Les résultats obtenus sur des codecs à 32 kbit/s pour le réseau téléphonique sont comparés à un indice significatif de qualité subjective : le rapport signal sur bruit d'isopréférence avec bruit multiplicatif.

SUMMARY

The quality of codecs (classical A/D and D/A converters or others) is usually measured by the signal to noise ratio. The methods presently used for quality measurements of codecs designed for digital telephone links give significant results when working on simple coding-decoding processes like PCM ; the same methods may give results unrelated to the subjective quality when used on more complex processes like ADPCM. We describe here a measurement method called "generalized intermodulation measurement" which furnishes the power spectrum of coding noise whatever the spectrum of the coded signal is (this signal must be stationary). Results obtained on 32 kbit/s ADPCM codecs designed for telephone transmission are compared with a significant indicator of subjective quality : the SNR obtained from an isopreference test with multiplicative noise.



MESURE DU RAPPORT SIGNAL A BRUIT DE CODECS
PAR UNE METHODE D'INTERMODULATION GENERALISEE

1. INTRODUCTION : QUALITE DES SYSTEMES DE CODAGE DE LA PAROLE ET MESURES OBJECTIVES

Nous entendons ici par qualité d'un système de codage de parole ou "Codec" (abréviation de codeur-décodeur) un ensemble de caractéristiques appréciables subjectivement à l'écoute de la parole codée, et pouvant être décrites en termes d'agrément : un codec produira des bruits ou des distorsions perturbant le signal original, plus ou moins audibles et gênants.

Nous nous intéresserons ici à des codecs MIC différentiels au débit de 32 kbit/s destinés à la transmission téléphonique ; la qualité de référence sera donc celle du MIC 64 kbit/s (MIC : Modulation par Impulsions et Codage : conversion analogique numérique avec loi de compression non linéaire pseudo-logarithmique ; échantillonnage à 8 kHz, 8 bits par échantillon). Les dégradations produites sur la parole par les codecs MIC différentiels sont proches d'un bruit multiplicatif, c'est-à-dire d'un bruit dont l'enveloppe instantanée varie comme celle du signal ; l'intelligibilité du message n'est pas affectée, et le timbre de la voix est préservé.

La qualité subjective d'un système de codage de parole est déterminée habituellement au moyen de tests d'écoute ou de conversation. Ces tests sont longs et fastidieux, et de mise en oeuvre souvent délicate. Aussi des travaux ont été entrepris pour tenter de définir un indice mesurable convenablement corrélé à la qualité subjective.

Parmi les indices habituellement mentionnés, outre le Rapport signal à bruit classique, signalons le Rapport signal à bruit par segments, somme de rapports signal à bruit calculés sur des fenêtres temporelles successives de 10 à 20 ms, et d'autres indices dérivés (CROCHIERE et coll. 1978). La méthode décrite ici est basée sur un principe et un modèle bien connus pour les codecs (AARON et coll. 1969) ; elle fournit le spectre du bruit de codage et un rapport signal à bruit calculé, que l'on comparera au rapport signal sur bruit d'isopréférence avec bruit multiplicatif, fourni par un test d'écoute particulier. Précisons tout de suite que cette méthode, dite d'intermodulation généralisée, n'a de sens que pour des signaux d'entrée stationnaires au second ordre, contrairement à des méthodes plus récentes basées sur l'identification en temps réel du modèle du codec, et qui peuvent utiliser un signal quelconque (CROCHIERE et coll., 1980).

2. PRINCIPE DE LA METHODE DE MESURE D'INTERMODULATION GENERALISEE

La méthode de mesure d'intermodulation généralisée est basée sur la modélisation du codec par un filtre linéaire à la sortie duquel est superposé un bruit non corrélé avec le signal d'entrée : voir figure 1.

Pour un signal d'entrée $e(t)$ stationnaire, la densité spectrale de puissance de la sortie $s(t)$ est donnée par :

$$S_{SS}(j\omega) = |G(j\omega)|^2 S_{ee}(j\omega) + S_{bb}(j\omega) \quad (1)$$

où $S_{ee}(j\omega)$ et $S_{bb}(j\omega)$ sont les densités spectrales de puissance respectives des signaux $e(t)$ et $b(t)$.

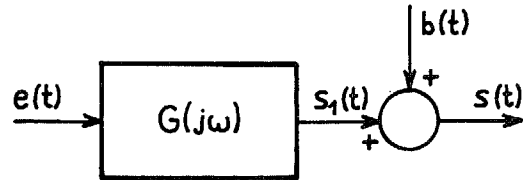


Figure 1 : modèle du codec.

La fonction de transfert $G(j\omega)$ du filtre équivalent est définie comme celle donnant la meilleure approximation linéaire $s_1(t)$ au sens des moindres carrés du signal $s(t)$, c'est-à-dire que :

$$\langle b^2(t) \rangle = \langle [s(t) - s_1(t)]^2 \rangle$$

est minimum, ou encore, si $g(t)$ est la réponse impulsionnelle du filtre :

$$\langle [s(t) - g(t) * e(t)]^2 \rangle \text{ minimum}$$

(la notation $\langle \rangle$ symbolise une opération de moyenne temporelle).

La minimisation de cette expression conduit à la condition de non-corrélation entre $b(t)$ et $e(t)$, et à l'expression de $G(j\omega)$:

$$G(j\omega) = \frac{S_{se}(j\omega)}{S_{ee}(j\omega)} \quad (2)$$

où $S_{se}(j\omega)$ est la transformée de FOURIER de la fonction d'intercorrélacion entre l'entrée $e(t)$ et la sortie $s(t)$ du codec. Cette fonction d'intercorrélacion est en général peu accessible, et on ne peut pas alors calculer la fonction de transfert $G(j\omega)$. Par contre, on accède au bruit $b(t)$ par la méthode dite d'élimination de bande. Soit en effet une bande étroite $[\omega_1, \omega_2]$ telle que :

$$S_{ee}(j\omega) = 0, \quad \omega \in [\omega_1, \omega_2] \quad (3)$$

Alors :

$$S_{SS}(j\omega) = S_{bb}(j\omega), \quad \omega \in [\omega_1, \omega_2] \quad (4)$$

et à condition que la bande éliminée soit suffisamment étroite pour que sa contribution au bruit total puisse être négligée, on mesure directement de cette façon la densité spectrale de puissance du bruit $b(t)$. Il s'agit bien ici d'une mesure d'intermodulation, (LACHAISE, 1970) puisqu'on mesure l'énergie produite dans une bande de fréquences par toutes les autres composantes spectrales du signal ; elle est généralisée parce que la bande éliminée est déplacée dans tout le spectre du signal d'entrée.

MESURE DU RAPPORT SIGNAL A BRUIT DE CODECS
 PAR UNE METHODE D'INTERMODULATION GENERALISEE

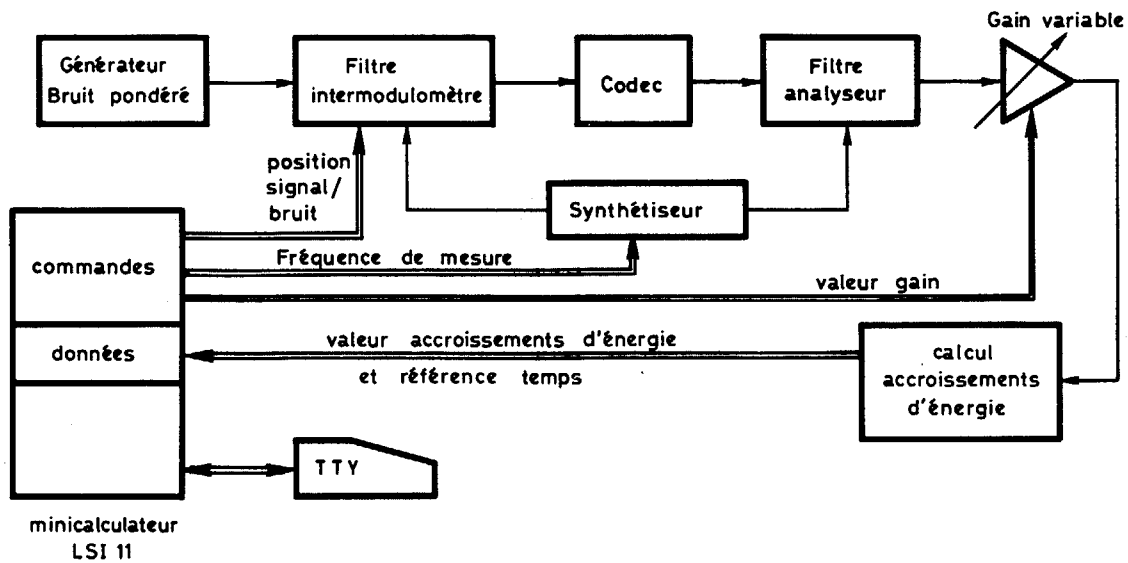


Figure 2 : Schéma synoptique de l'appareillage de mesure d'intermodulation.

3. MISE EN OEUVRE DE LA METHODE

3.1. La réalisation d'un appareillage de mesure spécifique autonome a été entreprise afin de mesurer les caractéristiques de maquettes de codecs. Le spectre du signal d'entrée a été limité à la bande 300-3400 Hz, et on s'est efforcé de préserver une bonne dynamique (au moins 30 dB). L'ensemble de l'appareillage est représenté schématiquement figure 2.

Le signal de mesure est produit par un générateur de bruit pondéré par une enveloppe spectrale particulière (plusieurs spectres ont été essayés).

Afin d'annuler la densité spectrale de puissance du signal d'entrée du codec dans une bande étroite mobile, il est nécessaire de disposer d'un filtre réjecteur de bande, la position de cette bande dans le spectre étant variable. Ce filtre réjecteur variable est réalisé par hétérodynage : le spectre du signal d'entrée est transposé par modulation autour de 10 kHz, fréquence centrale de la bande éliminée par un filtre fixe. La bande latérale inutile provenant de la modulation est éliminée par un filtrage passe-bande de largeur 3100 Hz, puis le signal est récupéré dans la bande de base par une dernière modulation. Les caractéristiques de réjection du filtre coupe-bande sont données figure 3.

Le filtre réjecteur est mis hors circuit sur la position "Signal", ce qui permet de mesurer la puissance de ce dernier en vue du calcul du rapport signal à bruit.

Le bruit de codage dans la bande rejetée est isolé au moyen d'un filtre passe-bande très étroit (analyseur de signal HEWLETT-PACKARD 3581 C). Pour une bande rejetée de largeur utile 100 Hz comme indiqué figure 3, la largeur de bande à -3 dB du filtre passe-bande est 10 Hz.

Après multiplication par un gain adéquat pour mise à un niveau suffisant, le signal à la sortie du filtre passe-bande est échantillonné, élevé au carré et intégré dans un appareil qui fournit chaque milliseconde un accroissement d'énergie. Les accroissements d'énergie sont introduits dans le minicalculateur LSI 11, où ils sont exploités dans des calculs

statistiques en vue de déterminer pour chaque fréquence la durée de la mesure (qui est d'autant plus longue que les fluctuations de l'énergie sont importantes), puis la puissance du signal et la puissance du bruit à la fréquence considérée ; après balayage de la totalité du spectre du signal d'entrée, le rapport signal sur bruit est calculé.

Une phase préalable de tarage élimine l'effet du bruit propre du filtre intermodulomètre, que l'on suppose non corrélé avec le bruit de codage.

Le balayage en fréquence est assuré par un synthétiseur commandé par le programme du calculateur.

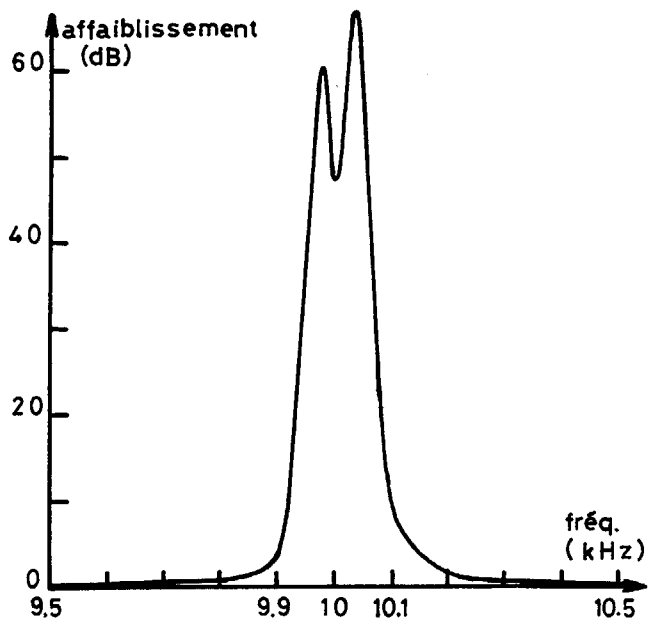


Figure 3 : réponse en fréquence du filtre coupe-bande



MESURE DU RAPPORT SIGNAL A BRUIT DE CODECS
PAR UNE METHODE D'INTERMODULATION GENERALISEE

3.2. Principales difficultés rencontrées dans la mise au point de l'appareillage

La réalisation du filtre coupe-bande à 10 kHz, qui a été confiée à un service spécialisé du CNET, a pratiquement imposé une largeur de bande utile de 100 Hz, qui s'est avérée trop importante, et a conduit à l'utilisation de signaux à spectre peu formé (sans accidents marqués). C'est une limitation importante à l'utilisation de l'appareillage actuel. Il faut remarquer d'autre part que si on diminue la largeur de la bande rejetée, on allonge d'autant la durée de la mesure, car il faut intégrer des fluctuations d'énergie plus lentes.

4. RESULTATS D'UNE EXPERIMENTATION ET COMPARAISON AVEC LES RESULTATS D'UN TEST SUBJECTIF D'ISOPREFERENCE

4.1. Description de l'expérimentation

Plusieurs maquettes de codecs ont été soumises à la mesure d'intermodulation ; nous en avons retenu deux comme codecs-types :

- le codec n° 1 est un MIC différentiel à prédicteur à 4 coefficients fixes de compromis parole-données : les coefficients ont été calculés pour assurer une prédiction convenable à la fois sur la parole et sur les données provenant de modems analogiques.

- le codec n° 2 est un MIC différentiel à prédicteur adaptatif comportant 5 cellules du second ordre en cascade, la fonction de transfert de chaque cellule étant limitée à une portion de la bande 0-4000 Hz ; le coefficient variable de chaque cellule est adapté par un algorithme du type gradient simplifié. (JACKSON, 1978 ; BONNEROT, 1981).

Les deux codecs qui fournissent des codes de 4 bits à un rythme de 8 kHz, possèdent un quantificateur adaptatif à CAG ajusté à chaque échantillon, et ont été conçus de telle sorte que la dégradation produite par plusieurs transcodages code MIC 8 bits → MIC différentiel → code MIC 8 bits soit théoriquement réduite à celle produite par un seul transcodage (BONNEROT et coll, 1979).

Trois configurations simulant autant de cas d'exploitation possibles ont été examinées :

- un codage-décodage ou bond analogique (1 BA) ;
- quatre codages-décodages, avec retour au signal analogique entre chaque opération de codage-décodage MIC différentiel (4 BA) ;
- quatre transcodages ou bonds numériques MIC 8 bits → MIC différentiel → MIC 8 bits, le signal analogique étant présent seulement au début et à la fin de la chaîne (4 BN).

On a utilisé deux signaux de mesure dont le spectre couvre la bande 300-3400 Hz :

- un "signal moyen de parole" dont le spectre résulte de la superposition des spectres expérimentaux de 20 locuteurs parlant sur une longue période ;

- un signal "haute fréquence" dont le spectre a un maximum large autour de 2500 Hz (simulation d'un son comme /Ch/).

Ces signaux sont stationnaires et ont une distribution des amplitudes sensiblement gaussienne, compte tenu des importants filtrages linéaires intervenant dans l'appareillage de mesure.

Le niveau des signaux est d'environ 6 dB en dessous du niveau de saturation à l'entrée des codecs.

4.2. Résultats

Les résultats sont présentés sous forme de spectres sur les figures 4a, 4b, 5a et 5b. Les spectres des signaux de mesure, notés S, proviennent de l'expérimentation dans la configuration 1BA ; en raison de la procédure de mesure, en position signal on obtient la somme des puissances du signal et du bruit, et lorsque ce dernier est important par rapport au signal la courbe S est sensiblement relevée (voir particulièrement figure 4b). Le temps de mesure est calculé de sorte que la précision sur chaque valeur efficace soit d'environ 5 %, d'où des irrégularités pas toujours significatives dans la forme des spectres. On a donné également les valeurs des rapports signal à bruit dans les trois configurations.

Pour le codec 1, le bruit sur le signal moyen de parole (figure 4a) est faible ; par contre le codage du signal haute fréquence (figure 4b) est mauvais, ce qui est normal compte tenu du prédicteur de compromis qui favorise les basses fréquences. D'autre part, la quasi-identité des spectres des bruits dans les configurations 1BA et 4BN correspond à la non-accumulation des dégradations dans les bonds numériques.

Pour le codec 2, les bruits sur le signal moyen de parole (figure 5a) et sur le signal haute fréquence (figure 5b) sont voisins : l'insensibilité au spectre du signal provient de l'adaptation du prédicteur. On observe un spectre de bruit intermédiaire dans la configuration 4BN ; la propriété de non-accumulation des dégradations n'est donc pas vérifiée ici.

Les spectres des bruits sont plats ou légèrement décroissants dans les hautes fréquences ; avec les réserves qui s'imposent en égard à la bande de réjection trop large, on peut expliquer ce fait par une faible proportion de bruit de saturation, dont le spectre se rapproche de celui du signal, et donc une prédominance du bruit de quantification, à spectre plat : en effet, le quantificateur adaptatif à CAG très rapide n'est pratiquement jamais saturé.

4.3. Validité des résultats

Comme on l'a dit plus haut, la largeur trop grande de la bande rejetée (100 Hz) introduit une erreur systématique en diminuant l'énergie du signal, particulièrement lorsque la bande rejetée coïncide avec un maximum du spectre ; ainsi dans le cas du signal moyen de parole, la tendance des spectres de bruit à décroître vers 300 Hz (figures 4a et 5a) semble due à ce phénomène. Dans ce cas, si on mesure la valeur efficace du signal à l'entrée du codec on observe en position bruit (c'est-à-dire lorsque le filtre réjecteur centré sur 300 Hz agit) une réduction de cette valeur à 0.62 V eff alors qu'elle est de 1V eff en position signal (le filtre réjecteur est hors-circuit) ; dans le cas du signal haute fréquence, lorsque le filtre réjecteur est centré sur 2500 Hz on observe une réduction de la valeur efficace de 1V en position signal à 0.76 V en position bruit. Les rapports signal à bruit indiqués sont donc probablement à revoir en baisse ; la correction (pouvant être de 1 à 3 dB suivant le spectre du signal) serait calculable à condition de faire une hypothèse sur la forme du spectre du bruit.

4.4. Comparaison avec les résultats d'un test d'isopréférence

Dans le test d'isopréférence avec bruit multiplicatif (LE GUYADER et PISSARD, 1979) pratiqué ici, des sujets entraînés équilibrent manuellement le bruit multiplicatif calibré avec le bruit de codage



MESURE DU RAPPORT SIGNAL A BRUIT DE CODECS
PAR UNE METHODE D'INTERMODULATION GENERALISEE

sur des séries de phrases prononcées par 10 à 12 locuteurs ; les signaux d'origine proviennent d'un microphone linéaire, l'écoute est faite sur casque binaural type HI-FI. Les résultats de ce test (Rapports signal à bruit d'isopréférence RSB ISO) sont regroupés dans le tableau 1 avec les rapports signal à bruit calculés provenant de la mesure d'intermodulation.

		RSB ISO	RSB signal moyen Parole	RSB Signal HF
Codec 1	1BA	28.1	28.7	14.8
	4BN	27.3	27.9	15.3
	4BA	22.4	22.6	9.7
Codec 2	1BA	32.8	27.8	27.3
	4BN	31.3	25.7	23.9
	4BA	25.9	22.6	21.2

Tableau 1 : résultats - rapports signal à bruit en dB.

Les valeurs du rapport signal à bruit d'isopréférence RSB_{ISO} pour le codec 1 sont très proches des valeurs du rapport signal à bruit calculé sur le signal moyen de parole, qui serait donc dans ce cas un bon indice pour mesurer la qualité subjective du codec (avec les réserves à faire sur l'exactitude des valeurs du rapport signal à bruit calculé).

Pour le codec 2, RSB_{ISO} dépasse de 5 dB environ les rapports signal à bruit sur les deux signaux de mesure : on peut expliquer cet écart par deux phénomènes : d'une part la prédiction adaptative est plus efficace sur les sons à spectre possédant des maxima ou formants marqués (voyelles) que sur les signaux de mesure dont la dynamique spectrale est plus réduite ; d'autre part, le bruit de codage dans les transitions entre sons peut être atténué par un effet de masquage temporel (FASTL, 1976, 1977).

Enfin, les écarts respectifs entre les trois configurations (1BA, 4BN et 4BA) obtenus avec le test d'isopréférence et les mesures sont très proches, ceci pour les deux codecs.

Quantitativement, une correspondance entre RSB_{ISO} et les rapports signal à bruit calculés pourrait être établie, par exemple au moyen d'une combinaison linéaire de ces derniers :

$$RSB_{ISO} = k_1 \times RSB (\text{Signal moyen parole}) + k_2 \times RSB (\text{signal HF}).$$

Prenant comme références les configurations 1BA et 4BA, on trouve :

$$\text{Codec 1 : } k_1 = 1.05, \quad k_2 = -0.14$$

$$\text{Codec 2 : } k_1 = 0.424, \quad k_2 = 0.77$$

donc des coefficients très différents.

Nous pensons que pour calculer un indice significatif de la qualité subjective à partir des résultats de la mesure d'intermodulation il faudrait un nombre important de signaux de mesure avec des spectres variés, représentatifs des divers sons de la parole ; l'indice serait une combinaison linéaire des rapports signal à bruit mesurés, affectés

d'un poids correspondant à la fréquence d'occurrence de ces sons dans la parole continue (à ce point de vue, les coefficients k_1 et k_2 relatifs au codec 1 sont plus proches de ces fréquences que ceux relatifs au codec 2).

De plus, la forme du spectre du bruit de codage (si celui-ci diffère sensiblement d'un bruit blanc) pourrait être prise en compte dans le calcul d'un facteur de masquage fréquentiel du bruit par le signal (ZWICKER et Coll, 1967).

Tout cela suppose que les limitations de l'appareillage actuel soient levées (surtout la largeur de bande trop importante du filtre réjecteur).

5. CONCLUSION

Nous avons réalisé un appareillage de mesure d'intermodulation généralisée, permettant d'accéder au spectre du bruit de codage des codecs. Les imperfections de cet appareillage nous ont limité dans le choix des signaux de mesure ; la validité des relations que nous pouvons établir entre les résultats des mesures et ceux du test subjectif d'isopréférence sont donc discutables.

Par l'utilisation de signaux de mesure à spectres représentatifs des principaux sons de la parole et par la prise en compte de phénomènes subjectifs comme le masquage fréquentiel, nous pensons que cette méthode est perfectible et peut donner des résultats significatifs. Basée sur une procédure indépendante du codec, elle nous semble au moins aussi sûre que les méthodes d'identification dynamique dont les performances dépendent étroitement de l'algorithme d'identification choisi.

L'auteur remercie Lucien MACE et Marcel CLEMENT qui ont conçu et construit le système de filtrage à bande de réjection mobile, et Yasser SARHAN qui a réalisé l'interface avec le calculateur LSI 11.

BIBLIOGRAPHIE

AARON M.R., FLEISCHMAN J.S., Mc DONALD R.W., PROTONOTARIOS E.N., 1969 : Response of Delta Modulation to gaussian signals - B.S.T.J. Mai-juin 1969, pp. 1167-1195.

BONNEROT G., MAITRE X., ZURCHER J.F., 1979 : Accroissement de la capacité du réseau numérique par des techniques de réduction de débit - 3ème forum mondial des Télécommunications, Genève, actes Vol. 2, pp. 3.5.9.1-3.5.9.9.

BONNEROT G., ROLLIN J.M., 1981 : Digital ADPCM converter for Digital Tandem connections - communication à ICASSP 81, à paraître.

CROCHIERE R.E., RABINER L.R., JAYANT N.S., TRIBOLET JM, 1978 : A study of objective measures for speech waveform coders - ZURICH seminar on Digital Communications, actes pp. H1.1-H1.7.

CROCHIERE R.E., TRIBOLET J.M., RABINER L.R., 1980 : On the measurement of waveform coder distortion using the log likelihood ratio - communication à ICASSP 80, actes pp. 340-343.



MESURE DU RAPPORT SIGNAL A BRUIT DE CODECS
 PAR UNE METHODE D'INTERMODULATION GENERALISEE

FASTL H., 1976 : Temporal masking effects : I broad band noise masker - ACUSTICA 35, p. 287.

FASTL H., 1977 : Temporal masking effects : II cut off band noise masker - ACUSTICA 36, pp. 317-331.

JACKSON L.B., WOOD S.L., 1978 : Linear prediction in cascade form - Trans. ASSP 26/6, pp. 518-528.

LACHAISE J., 1970 : Méthode de mesure du rapport signal à bruit de quantification - Câbles et Transmissions 24/1, pp. 26-31.

LE GUYADER A., PISSARD A., 1979 : Codage différentiel adaptatif de la parole pour le réseau téléphonique - 7ème colloque du GRETSI, pp. 33/1-33/7.

ZWICKER, FELDTEKKER, 1967 : Das Ohr als Nachrichtenempfänger - HIRZEL Verlag, STUTTGART.

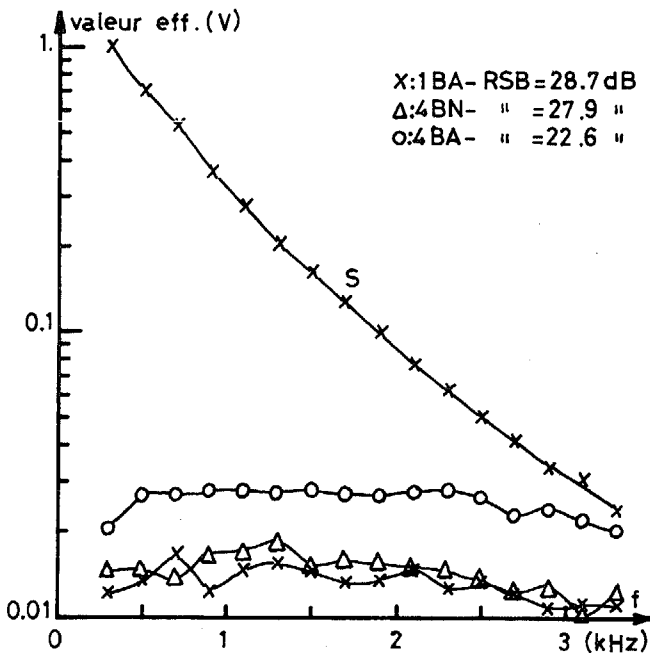


Figure 4a : codec 1, signal moyen de parole.

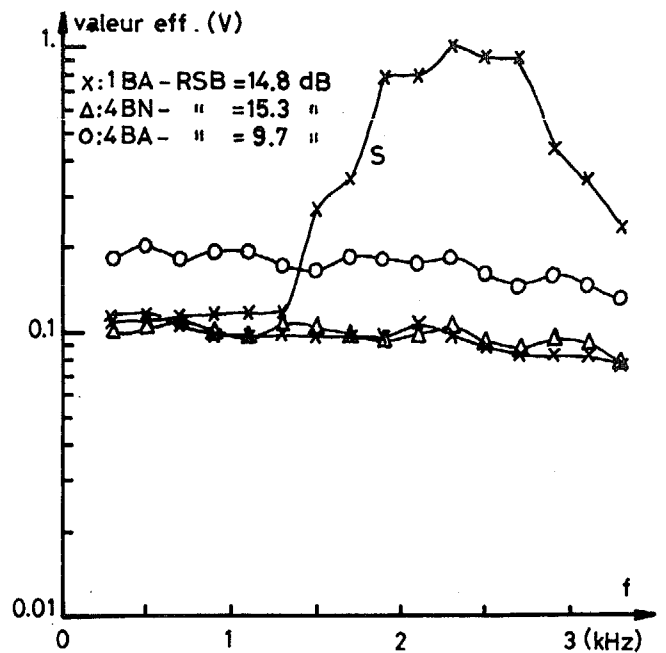


Figure 4b : codec 1, signal haute fréquence.

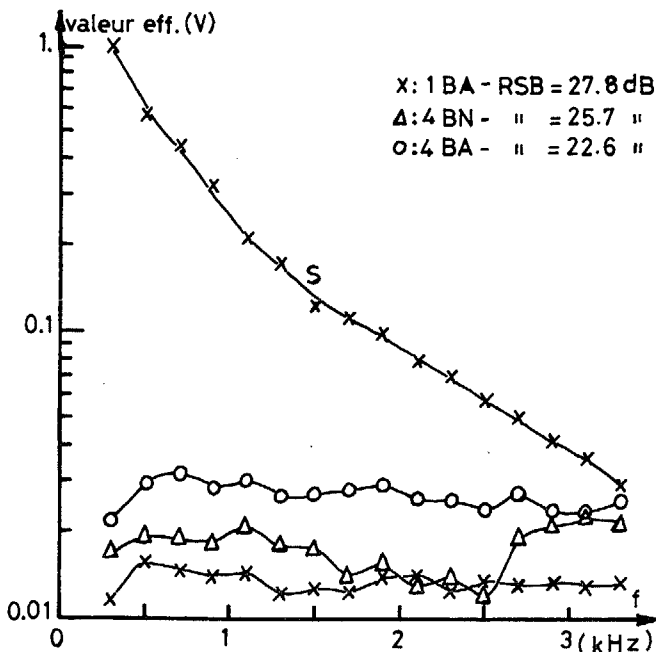


Figure 5a : codec 2, signal moyen de parole.

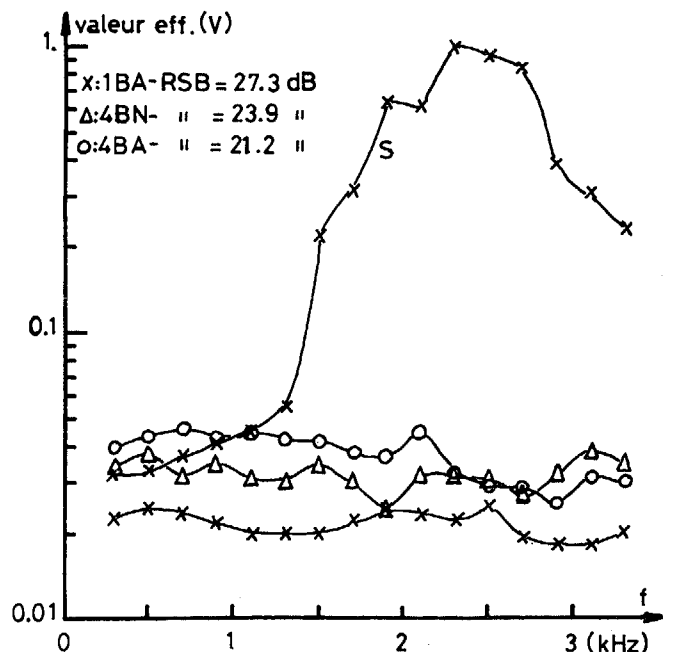


Figure 5b : codec 2, signal haute fréquence.