



(86) Date de dépôt PCT/PCT Filing Date: 2013/06/28
(87) Date publication PCT/PCT Publication Date: 2014/01/03
(85) Entrée phase nationale/National Entry: 2014/11/27
(86) N° demande PCT/PCT Application No.: FR 2013/051517
(87) N° publication PCT/PCT Publication No.: 2014/001730
(30) Priorité/Priority: 2012/06/29 (FR1256285)

(51) Cl.Int./Int.Cl. *G10L 21/0364* (2013.01),
G10L 19/26 (2013.01)
(71) Demandeur/Applicant:
ORANGE, FR
(72) Inventeurs/Inventors:
KOVESI, BALAZS, FR;
RAGOT, STEPHANE, FR
(74) Agent: GOUDREAU GAGE DUBUC

(54) Titre : ATTENUATION EFFICACE DE PRE-ECHOS DANS UN SIGNAL AUDIONUMERIQUE
(54) Title: EFFECTIVE PRE-ECHO ATTENUATION IN A DIGITAL AUDIO SIGNAL

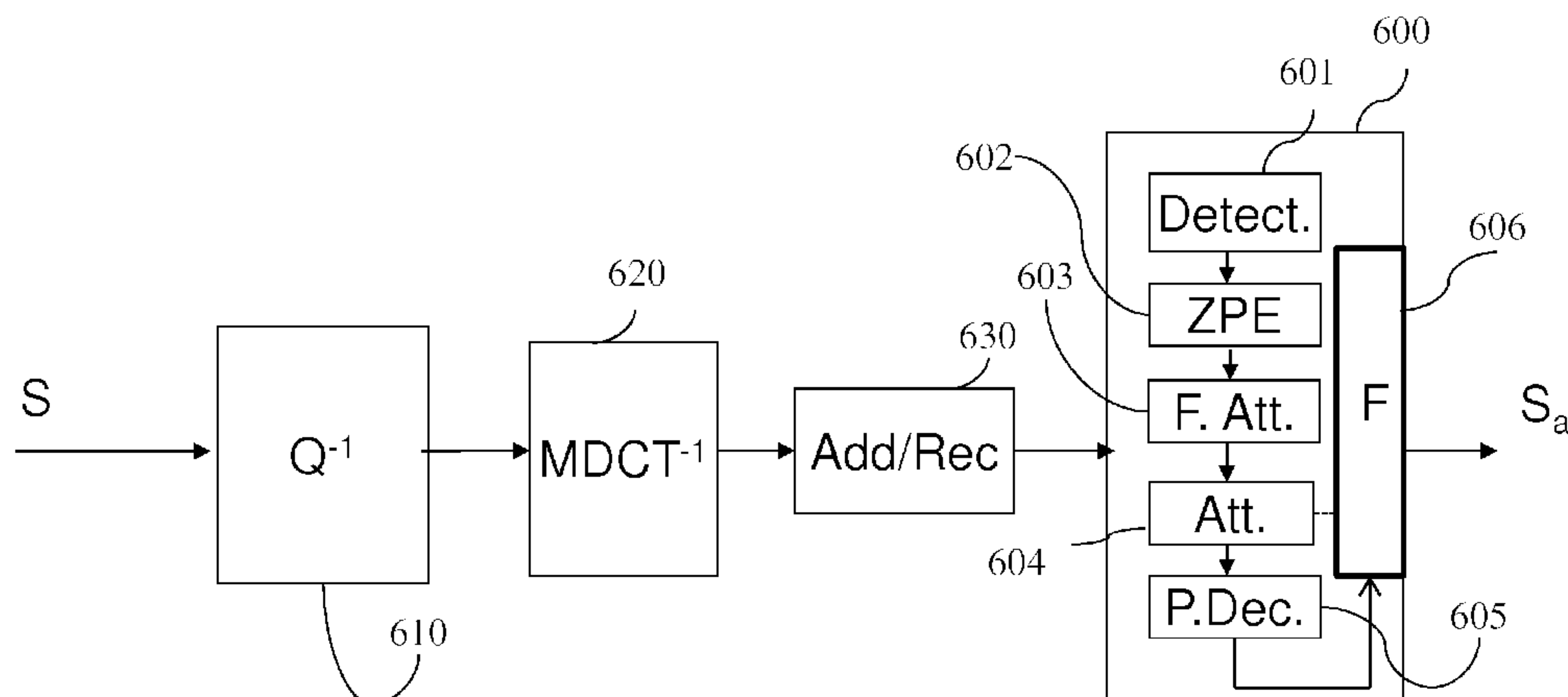
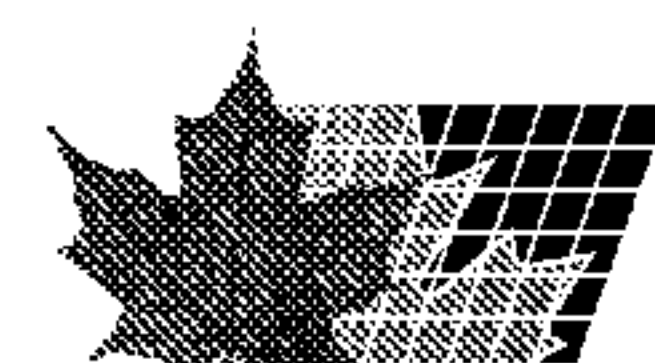


FIG.6

(57) **Abrégé/Abstract:**

L'invention porte sur un procédé de traitement d'atténuation de pré -écho dans un signal audionumérique engendré à partir d'un codage par transformée, dans lequel, au décodage, le procédé comporte les étapes de détection (Detect.) d'une position d'attaque dans le signal décodé, de détermination (ZPE) d'une zone de pré-écho précédant la position d'attaque détectée dans le signal décodé, de calcul (F. Att.) de facteurs d'atténuations par sous-bloc de la zone de pré-écho, en fonction au moins de la trame dans laquelle l'attaque a été détectée et de la trame précédente, d' atténuation (Att.) de pré-écho dans les sous-blocs de la zone de pré -écho par les facteurs d'atténuation correspondants. Le procédé comporte en outre, l'application d'un filtrage (F) de mise en forme spectrale de la zone de pré-écho sur la trame courante jusqu'à la position détectée de l'attaque. L'invention vise également un dispositif mettant en ouvre le procédé ainsi qu'un décodeur comportant un tel dispositif.



(12) DEMANDE INTERNATIONALE PUBLIÉE EN VERTU DU TRAITÉ DE COOPÉRATION EN MATIÈRE DE BREVETS (PCT)

(19) Organisation Mondiale de la
Propriété Intellectuelle
Bureau international(43) Date de la publication internationale
3 janvier 2014 (03.01.2014)

WIPO | PCT

(10) Numéro de publication internationale
WO 2014/001730 A1(51) Classification internationale des brevets :
G10L 21/0364 (2013.01) *G10L 19/26* (2013.01)(21) Numéro de la demande internationale :
PCT/FR2013/051517(22) Date de dépôt international :
28 juin 2013 (28.06.2013)

(25) Langue de dépôt : français

(26) Langue de publication : français

(30) Données relatives à la priorité :
1256285 29 juin 2012 (29.06.2012) FR(71) Déposant : ORANGE [FR/FR]; 78 rue Olivier de Serres,
F-75015 Paris (FR).(72) Inventeurs : KOVESI, Balazs; 16 chemin du Moulin à
Vent, F-LANNION 22300 (FR). RAGOT, Stéphane; Al-
legoat (Serval), F-LANNION 22300 (FR).(74) Mandataire : FRANCE
TELECOM/OLNC/OLPS/IPL/PAT; FROGER Marie-
Hélène, 38-40 rue du Général Leclerc, F-92794 Issy Mou-
lineaux Cedex 9 (FR).(81) États désignés (sauf indication contraire, pour tout titre
de protection nationale disponible) : AE, AG, AL, AM,
AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY,
BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM,
DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT,
HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KN, KP, KR,
KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME,
MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ,
OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC,
SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN,
TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.(84) États désignés (sauf indication contraire, pour tout titre
de protection régionale disponible) : ARIPO (BW, GH,
GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ,
UG, ZM, ZW), eurasien (AM, AZ, BY, KG, KZ, RU, TJ,
TM), européen (AL, AT, BE, BG, CH, CY, CZ, DE, DK,
EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV,
MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM,
TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW,
KM, ML, MR, NE, SN, TD, TG).

Publiée :

— avec rapport de recherche internationale (Art. 21(3))

(54) Title : EFFECTIVE PRE-ECHO ATTENUATION IN A DIGITAL AUDIO SIGNAL

(54) Titre : ATTÉNUATION EFFICACE DE PRÉ-ÉCHOS DANS UN SIGNAL AUDIONUMÉRIQUE

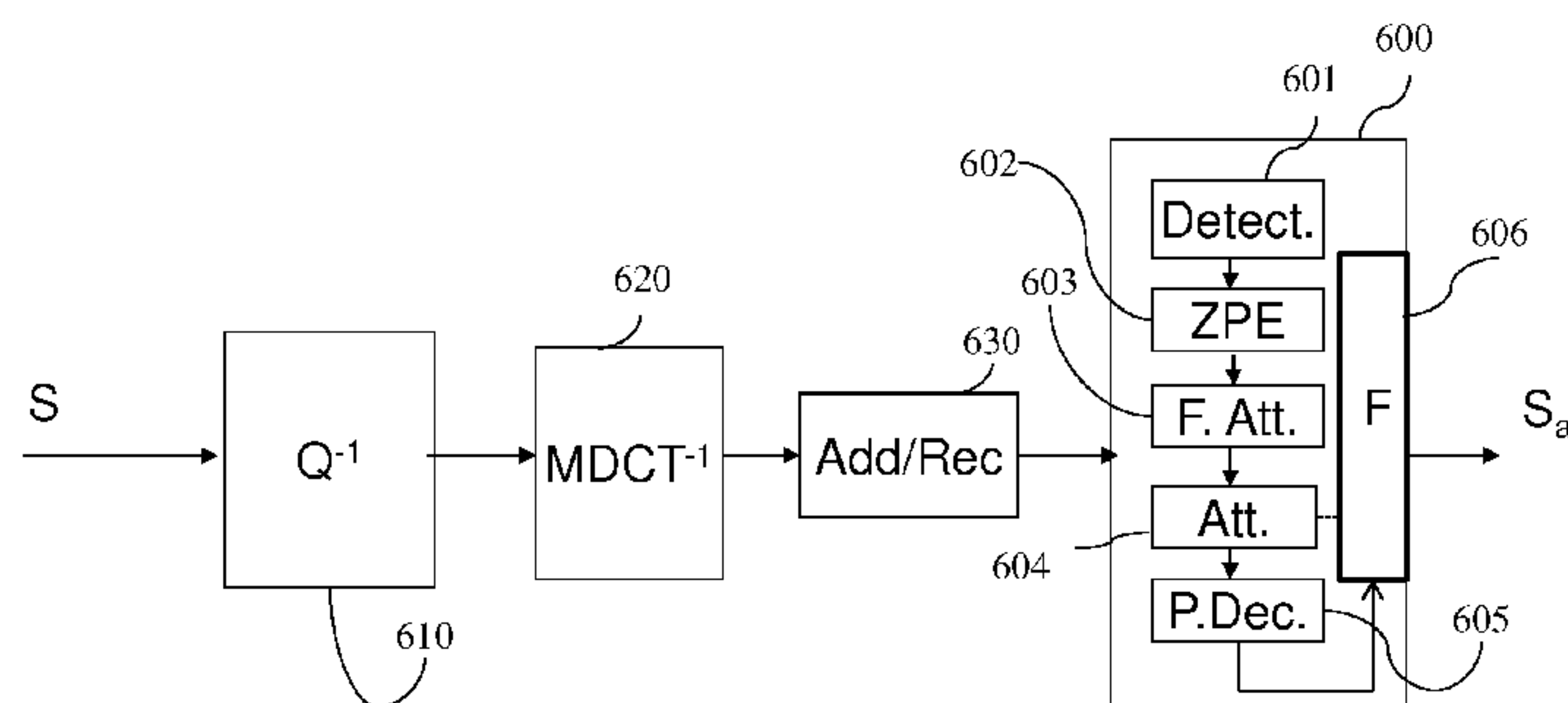


FIG.6

(57) Abstract : The invention relates to a method for processing pre-echo attenuation in a digital audio signal generated from a transform coding, wherein, at the decoding point, the method comprises steps of: detection (Detect.) of a position of attack in the de-coded signal; determination (ZPE) of a pre-echo region preceding the position of attack detected in the decoded signal; calculation (F. Att.) of attenuation factors per sub-block of the pre-echo region, according to at least the frame wherein the attack has been detected and the preceding frame; and pre-echo attenuation (Att.) in the sub-blocks of the pre-echo region by the corresponding damping factors. The method also comprises the application of a filter (F) for the spectral shaping of the pre-echo region on the current frame up to the detected position of the attack. The invention also relates to a device implementing said method and to a decoder comprising such a device.

(57) Abrégé :

[Suite sur la page suivante]

WO 2014/001730 A1

L'invention porte sur un procédé de traitement d'atténuation de pré -écho dans un signal audionumérique engendré à partir d'un codage par transformée, dans lequel, au décodage, le procédé comporte les étapes de détection (Detect.) d'une position d'attaque dans le signal décodé, de détermination (ZPE) d'une zone de pré-écho précédant la position d'attaque détectée dans le signal décodé, de calcul (F. Att.) de facteurs d'atténuations par sous-bloc de la zone de pré-écho, en fonction au moins de la trame dans laquelle l'attaque a été détectée et de la trame précédente, d' atténuation (Att.) de pré-écho dans les sous-blocs de la zone de pré-écho par les facteurs d'atténuation correspondants. Le procédé comporte en outre, l'application d'un filtrage (F) de mise en forme spectrale de la zone de pré-écho sur la trame courante jusqu'à la position détectée de l'attaque. L'invention vise également un dispositif mettant en oeuvre le procédé ainsi qu'un décodeur comportant un tel dispositif.

Atténuation efficace de pré-échos dans un signal audionumérique

L'invention concerne un procédé et un dispositif de traitement d'atténuation des pré-échos lors du décodage d'un signal audionumérique.

5 Pour le transport des signaux audionumériques sur les réseaux de transmission, qu'il s'agisse par exemple de réseaux fixes ou mobiles, ou pour le stockage des signaux, on fait appel à des processus de compression (ou codage source) mettant en œuvre des systèmes de codage du type codage temporel ou codage fréquentiel par transformée.

10 Le procédé et le dispositif, objets de l'invention, ont ainsi comme domaine d'application la compression des signaux sonores, en particulier les signaux audionumériques codés par transformée fréquentielle.

La figure 1 représente à titre illustratif, un schéma de principe du codage et du décodage d'un signal audio numérique par transformée incluant une analyse-synthèse par addition/recouvrement selon l'art antérieur.

15 Certaines séquences musicales, telles que les percussions et certains segments de parole comme les plosives (/k/, /t/, ...), sont caractérisées par des attaques extrêmement brusques qui se traduisent par des transitions très rapides et une variation très forte de la dynamique du signal en l'espace de quelques échantillons. Un exemple de transition est donné à la figure 1 à partir de l'échantillon 410.

20 Pour le traitement de codage/décodage, le signal d'entrée est découpé en blocs d'échantillons de longueur L, représentés sur la **figure 1** par des traits verticaux en pointillés. Le signal d'entrée est noté $x(n)$, où n est l'indice de l'échantillon. La découpe en blocs successifs conduit à définir les blocs $X_N(n) = [x(N.L) \dots x(N.L+L-1)] = [x_N(0) \dots x_N(L-1)]$, où N est l'indice de la trame, L est la longueur de la trame. A la figure 1 on a $L=160$ échantillons. Dans le cas de la transformée modulée en cosinus modifiée MDCT (pour "Modified Discrete Cosine Transform" en anglais), deux blocs $X_N(n)$ et $X_{N+1}(n)$ sont analysés conjointement pour donner un bloc de coefficients transformés associés à la trame d'indice N.

30 La division en blocs, aussi appelés trames, opérée par le codage par transformée est totalement indépendante du signal sonore et les transitions peuvent donc apparaître en un point quelconque de la fenêtre d'analyse. Or après décodage par transformée, le signal reconstruit est entaché de "bruit" (ou distorsion) engendré par l'opération quantification (Q)-quantification inverse (Q^{-1}). Ce bruit de codage est réparti temporellement de façon relativement uniforme sur tout le support temporel du bloc transformé, c'est-à-dire sur toute

- 2 -

la longueur de la fenêtre de longueur $2L$ d'échantillons (avec recouvrement de L échantillons). L'énergie du bruit de codage est en général proportionnelle à l'énergie du bloc et est fonction du débit de codage/décodage.

Pour un bloc comportant une attaque (comme le bloc 320-480 de la figure 1) l'énergie du signal est élevée, le bruit est donc également de niveau élevé.

En codage par transformée, le niveau du bruit de codage est typiquement inférieur à celui du signal pour les segments de forte énergie qui suivent immédiatement la transition, mais le niveau est supérieur à celui du signal pour les segments d'énergie plus faible, notamment sur la partie précédant la transition (échantillons 160 - 410 de la figure 1). Pour la partie précitée, le rapport signal à bruit est négatif et la dégradation résultante peut apparaître très gênante à l'écoute. On appelle pré-écho le bruit de codage antérieur à la transition et post-écho le bruit postérieur à la transition.

On peut observer sur la figure 1 que le pré-écho affecte la trame précédant la transition ainsi que la trame où se produit la transition.

Les expériences psycho-acoustiques ont montré que l'oreille humaine effectue un pré-masquage temporel des sons assez limité, de l'ordre de quelques millisecondes. Le bruit précédant l'attaque, ou pré-écho, est audible lorsque la durée du pré-écho est supérieure à la durée du pré-masquage.

L'oreille humaine effectue également un post-masquage d'une durée plus longue, de 5 à 60 millisecondes, lors du passage de séquences de forte énergie à des séquences de faible énergie. Le taux ou niveau de gêne acceptable pour les post-échos est donc plus important que pour les pré-échos.

Le phénomène des pré-échos, plus critique, est d'autant plus gênant que la longueur des blocs en nombre d'échantillons est importante. Or, en codage par transformée, il est bien connu que pour les signaux stationnaires plus la longueur de la transformée augmente, plus le gain de codage est important. A fréquence d'échantillonnage fixée et à débit fixé, si on augmente le nombre de points de la fenêtre (donc la longueur de la transformée) on disposera de plus de bits par trame pour coder les raies fréquentielles jugées utiles par le modèle psychoacoustique, d'où l'avantage d'utiliser des blocs de grande longueur. Le codage MPEG AAC (Advanced Audio Coding), par exemple, utilise une fenêtre de grande longueur qui contient un nombre fixe d'échantillons, 2048, soit sur une durée de 64 ms à une fréquence d'échantillonnage de 32 kHz; le problème des pré-échos y est géré en permettant de commuter de ces fenêtres longues à 8 fenêtres courtes par le biais de fenêtres intermédiaires (de transition), ce qui nécessite un certain retard au codage pour détecter la présence d'une

- 3 -

transition et adapter les fenêtres. La longueur de ces fenêtres courtes est donc de 8 ms. A bas débit on peut toujours avoir un pré-écho audible de quelques ms. La commutation des fenêtres permet d'atténuer le pré-écho mais pas de le supprimer. Les codeurs par transformée utilisés pour les applications conversationnelles comme UIT-T G.722.1, G.722.1C ou G.719
5 utilisent souvent une fenêtre de durée 40 ms à 16, 32 ou 48 kHz (respectivement) et une longueur de trame de 20 ms. On peut noter que le codeur UIT-T G.719 intègre un mécanisme de commutation de fenêtres avec détection de transitoire, cependant le pré-écho n'est pas complètement réduit à bas débit (typiquement à 32 kbit/s).

Dans le but de réduire l'effet gênant précité du phénomène des pré-échos,
10 différentes solutions ont été proposées au niveau du codeur et/ou du décodeur.

La commutation de fenêtres a été citée précédemment. Une autre solution consiste à appliquer un filtrage adaptatif. Dans la zone précédant l'attaque, le signal reconstruit est vu comme la somme du signal original et du bruit de quantification.

Une technique de filtrage correspondante a été décrite dans l'article intitulé High
15 Quality Audio Transform Coding at 64 kbits, IEEE Trans. on Communications Vol 42, No. 11, November 1994, publié par Y. Mahieux et J. P. Petit.

La mise en œuvre d'un tel filtrage nécessite la connaissance de paramètres dont certains, comme les coefficients de prédiction et la variance du signal corrompu par le pré-écho, sont estimés au décodeur à partir des échantillons bruités. Par contre, des informations
20 telles que l'énergie du signal d'origine ne peuvent être connues qu'au codeur et doivent par conséquent être transmises. Ceci nécessite de transmettre des informations supplémentaires, ce qui à débit contraint diminue le budget relatif alloué au codage par transformée. Lorsque le bloc reçu contient une variation brusque de dynamique, le traitement de filtrage lui est appliqué.

25 Le processus de filtrage précité ne permet pas de retrouver le signal d'origine, mais procure une forte réduction des pré-échos. Il nécessite toutefois de transmettre les paramètres supplémentaires au décodeur.

Différentes techniques de réduction de pré-écho sans transmission spécifique de l'information ont été proposées. Par exemple, une revue de la réduction de pré-échos dans le
30 contexte du codage hiérarchique est présentée dans l'article B. Kövesi, S. Ragot, M. Gartner, H. Taddei, "Pre-echo reduction in the ITU-T G.729.1 embedded coder," EUSIPCO, Lausanne, Suisse, Août 2008.

Un exemple typique de procédé d'atténuation de pré-échos est décrit dans la demande de brevet français FR 08 56248. Dans cet exemple, on détermine des facteurs

- 4 -

d'atténuation par sous-bloc, dans les sous-blocs de faible énergie précédant un sous-bloc dans lequel une transition ou attaque a été détectée.

Le facteur d'atténuation par sous-bloc $g(k)$ est calculé par exemple en fonction du rapport $R(k)$ entre l'énergie du sous-bloc de plus forte énergie et l'énergie du k-ième sous-bloc en question :

$$g(k) = f(R(k))$$

où f est une fonction décroissante à valeurs entre 0 et 1 et k est le numéro du sous-bloc.

D'autres définitions du facteur $g(k)$ sont possibles, par exemple en fonction de l'énergie $En(k)$ dans le sous-bloc courant et de l'énergie $En(k-1)$ dans le sous-bloc précédent.

Si la variation de l'énergie par rapport à l'énergie maximale est faible, aucune atténuation n'est alors nécessaire. Le facteur $g(k)$ est alors fixé à une valeur d'atténuation inhibant l'atténuation, c'est-à-dire 1. Sinon, le facteur d'atténuation est compris entre 0 et 1.

Dans la plupart des cas, surtout quand le pré-écho est gênant, la trame qui précède la trame de pré-écho a une énergie homogène qui correspond à l'énergie d'un segment de faible énergie (typiquement un bruit de fond). Selon l'expérience il n'est pas utile ni même souhaitable qu'après le traitement d'atténuation de pré-écho l'énergie du signal devienne inférieure à l'énergie moyenne par sous-bloc du signal précédant la zone de traitement (typiquement celle de la trame précédente $\overline{En}$ ou celle de la deuxième moitié de la trame précédente $\overline{En'}$).

Pour le sous-bloc k à traiter on peut calculer la valeur limite du facteur $\lim_g(k)$ afin d'obtenir exactement la même énergie que l'énergie moyenne par sous-bloc du segment précédant le sous-bloc à traiter. Cette valeur est bien sûr limitée à un maximum de 1 puisqu'on s'intéresse ici aux valeurs d'atténuation. Plus précisément :

$$\lim_g(k) = \min \left(\sqrt{\frac{\max(\overline{En}, \overline{En'})}{En(k)}}, 1 \right)$$

où l'énergie moyenne du segment précédent est approximée par $\max(\overline{En}, \overline{En'})$.

La valeur $\lim_g(k)$ ainsi obtenue sert de limite inférieure dans le calcul final du facteur d'atténuation du sous-bloc :

- 5 -

$$g(k) = \max(g(k), \lim_g(k))$$

Les facteurs d'atténuation (ou gains) $g(k)$ déterminés par sous-blocs sont ensuite lissés par une fonction de lissage appliquée échantillon par échantillon pour éviter des variations brusques du facteur d'atténuation aux frontières des blocs.

5 Par exemple, on peut d'abord définir le gain par échantillon comme une fonction constante par morceaux :

$$g_{pre}(n) = g(k), n = kL', \dots, (k+1)L'-1$$

où L' représente la longueur d'un sous-bloc.

La fonction est ensuite lissée suivant l'équation suivante:

$$10 \quad g_{pre}(n) := \alpha g_{pre}(n-1) + (1-\alpha) g_{pre}(n), n = 0, \dots, L-1$$

avec la convention que $g_{pre}(-1)$ est le dernier facteur d'atténuation obtenu pour le dernier échantillon du sous-bloc précédent, α est le coefficient de lissage, typiquement $\alpha=0.85$.

D'autres fonctions de lissage sont également possibles. Une fois les facteurs $g_{pre}(n)$ ainsi calculés, l'atténuation de pré-écho est faite sur le signal reconstruit de la

15 trame courante, $x_{rec}(n)$, en multipliant chaque échantillon par le facteur correspondant :

$$x_{rec,g}(n) = g_{pre}(n)x_{rec}(n), n = 0, \dots, L-1$$

où $x_{rec,g}(n)$ est le signal décodé et post-traité par la réduction de pré-écho.

Les **figures 2 et 3** illustrent la mise en œuvre du procédé d'atténuation tel que décrit dans la demande de brevet de l'état de l'art, précitée, et résumé précédemment.

20 Dans ces exemples le signal est échantillonné à 32 kHz, la longueur de la trame est $L=640$ échantillons et chaque trame est divisée en 8 sous-blocs de $K=80$ échantillons.

Dans la partie a) de la figure 2, une trame d'un signal original échantillonné à 32 kHz, est représentée. Une attaque (ou transition) dans le signal est située dans le sous-bloc commençant à l'indice 320. Ce signal a été codé par un codeur par transformée de type
25 MDCT à bas débit (24 kbit/s).

Dans la partie b) de la figure 2, le résultat du décodage sans traitement de pré-écho est illustré. On peut observer le pré-écho à partir de l'échantillon 160, dans les sous-blocs précédents celui contenant l'attaque.

La partie c) montre l'évolution du facteur d'atténuation de pré-écho (ligne
30 continue) obtenu par le procédé décrit dans la demande de brevet de l'état de l'art précitée. La

- 6 -

ligne pointillée représente le facteur avant lissage. On remarque ici que la position de l'attaque est estimée autour de l'échantillon 380 (dans le bloc délimité par les échantillons 320 et 400).

La partie d) illustre le résultat du décodage après application du traitement de pré-écho (multiplication du signal b) avec le signal c)). On voit que le pré-écho a bien été atténué. La figure 2 montre également que le facteur lissé ne remonte pas à 1 au moment de l'attaque, ce qui implique une diminution de l'amplitude de l'attaque. L'impact perceptible de cette diminution est très faible mais peut néanmoins être évité. La figure 3 illustre le même exemple que la figure 2, dans lequel, avant lissage, la valeur de facteur d'atténuation est forcée à 1 pour les quelques échantillons du sous-bloc précédant le sous-bloc où se situe l'attaque. La partie c) de la figure 3 donne un exemple d'une telle correction.

Dans cet exemple on a affecté la valeur de facteur 1 aux 16 derniers échantillons du sous-bloc précédant l'attaque, à partir de l'indice 364. Ainsi la fonction de lissage accroît progressivement le facteur pour avoir une valeur proche de 1 au moment de l'attaque. L'amplitude de l'attaque est alors préservée, comme illustré dans la partie d) de la figure 3, par contre quelques échantillons de pré-écho ne sont pas atténués.

Dans l'exemple de la figure 3 la réduction de pré-écho par atténuation ne permet pas de réduire le pré-écho jusqu'au niveau de l'attaque, à cause du lissage du gain.

Un autre exemple avec le même réglage que celui de la figure 3 est illustré sur la **figure 4**. Cette figure représente 2 trames pour mieux montrer la nature du signal avant l'attaque. Ici, l'énergie du signal original avant l'attaque est plus forte (partie a)) que dans le cas illustré par la figure 3, et le signal avant l'attaque est audible (échantillons 0 - 850). Sur la partie b) on peut observer le pré-écho sur le signal décodé sans traitement de pré-écho dans la zone 700-850. Selon la procédure de limitation de l'atténuation expliquée précédemment on atténue l'énergie du signal de la zone de pré-écho jusqu'à l'énergie moyenne du signal précédant la zone de traitement. On observe sur la partie c) que le facteur d'atténuation calculé en tenant compte de la limitation d'énergie est proche de 1 et que le pré-écho est toujours présent sur la partie d) après application du traitement de pré-écho (multiplication du signal b) avec le signal c)), malgré la bonne mise à niveau du signal dans la zone pré-écho. On peut en effet bien distinguer ce pré-écho sur la forme d'onde où on remarque qu'une composante haute fréquence est superposée au signal dans cette zone.

Cette composante haute fréquence est bien audible et gênante, et l'attaque est moins nette (partie d) figure 4).

- 7 -

L'explication de ce phénomène est la suivante : dans le cas d'une attaque très brusque, impulsive (comme illustrée sur la figure 4) le spectre du signal (dans la trame contenant l'attaque) est plutôt blanc et donc contient également beaucoup de hautes fréquences. Ainsi le bruit de quantification est également blanc et composé de hautes
5 fréquences, ce qui n'est pas le cas du signal précédant la zone de pré-écho. Il y a donc un changement brusque dans le spectre d'une trame à l'autre, qui résulte en un pré-écho audible malgré le fait que l'énergie a été mise au bon niveau.

Ce phénomène est à nouveau représenté sur les figures 5a et 5b qui montrent respectivement les spectrogrammes du signal original en 5a, correspondant au signal
10 représenté en partie a) de la figure 4 et le spectrogramme du signal avec atténuation de pré-échos selon l'état de l'art, en 5b, correspondant au signal représenté en partie d) de la figure 4.

On remarque bien un pré-écho encore audible dans la partie encadrée à la figure 5b.

15 Il existe donc un besoin d'une technique d'atténuation améliorée de pré-échos au décodage, qui permet d'atténuer également les hautes fréquences indésirables ou pré-échos parasites et sans qu'aucune information auxiliaire ne soit transmise par le codeur.

La présente invention améliore la situation de l'état de l'art.

A cet effet, la présente invention traite d'un procédé de traitement d'atténuation de
20 pré-écho dans un signal audionumérique engendré à partir d'un codage par transformée, dans lequel, au décodage, le procédé comporte les étapes suivantes:

- détection d'une position d'attaque dans le signal décodé;
- détermination d'une zone de pré-écho précédant la position d'attaque détectée dans le signal décodé;
- 25 - calcul de facteurs d'atténuations par sous-bloc de la zone de pré-écho, en fonction au moins de la trame dans laquelle l'attaque a été détectée et de la trame précédente;
- atténuation de pré-écho dans les sous-blocs de la zone de pré-écho par les facteurs d'atténuation correspondants. Le procédé est tel qu'il comporte en outre:
- l'application d'un filtrage adaptatif de mise en forme spectrale de la zone de pré-
30 écho sur la trame courante jusqu'à la position détectée de l'attaque.

Ainsi, la mise en forme spectrale appliquée, permet d'améliorer l'atténuation de pré-écho. Le traitement permet d'atténuer les composantes de pré-écho qui pourraient subsister à la mise en œuvre de l'atténuation de pré-écho telle que décrite dans l'état de l'art.

Le filtrage étant appliqué jusqu'à la position détectée de l'attaque, il permet de traiter l'atténuation du pré-écho jusqu'au plus près de l'attaque. Cela compense donc le désavantage de la réduction d'écho par atténuation temporelle qui est limitée à une zone n'allant pas jusqu'à la position de l'attaque (marge de 16 échantillons par exemple).

5 Ce filtrage ne nécessite pas d'informations en provenance du codeur.

Cette technique de traitement d'atténuation de pré-écho peut être mise en œuvre avec ou sans connaissance d'un signal issu d'un décodage temporel et pour le codage d'un signal monophonique ou d'un signal stéréophonique.

10 L'adaptation du filtrage permet de s'adapter au signal et de n'enlever que les composantes parasites gênantes.

Les différents modes particuliers de réalisation mentionnés ci-après peuvent être ajoutés indépendamment ou en combinaison les uns avec les autres, aux étapes du procédé défini ci-dessus.

15 Dans un mode de réalisation particulier, le procédé comporte en outre le calcul d'au moins un paramètre de décision sur le filtrage à appliquer à la zone de pré-écho et l'adaptation des coefficients du filtrage en fonction dudit au moins un paramètre de décision.

Ainsi, le traitement n'est alors appliqué que quand cela est nécessaire à un niveau de filtrage adapté.

20 Dans un mode de réalisation, ledit au moins un paramètre de décision est une mesure de la force de l'attaque détectée.

La force de l'attaque détermine en effet la présence de composantes hautes fréquences audibles dans la zone de pré-écho. Lorsque l'attaque est brusque, le risque d'avoir une composante parasite gênante dans la zone de pré-écho est grand et le filtrage à mettre en œuvre selon l'invention est alors à prévoir.

25 Dans un mode de calcul possible de ce paramètre, la mesure de la force de l'attaque détectée est de la forme:

$$P = \max (EN(k), EN(k+1) / \min(EN(k-1), EN(k-2)))$$
 avec k, le numéro du sous-bloc dans lequel l'attaque a été détectée et EN(k) l'énergie du k^{ième} sous-bloc.

30 Ce calcul est de moindre complexité et permet de bien définir la force de l'attaque détectée.

Le dit au moins un paramètre de décision peut aussi être la valeur du facteur d'atténuation dans le sous-bloc précédant celui contenant la position de l'attaque.

En effet, une attaque peut être considérée comme brusque si cette atténuation est significative.

Dans un autre mode de réalisation, ledit au moins un paramètre de décision est basé sur une analyse de répartition spectrale du signal de la zone de pré-écho et/ou du signal précédant la zone de pré-écho.

Ceci permet par exemple de déterminer l'importance des composantes hautes fréquences dans le signal de pré-écho et de savoir également si ces composantes hautes fréquences étaient déjà présentes dans le signal avant la zone de pré-écho.

Ainsi, dans le cas où des composantes hautes fréquences étaient déjà présentes avant la zone de pré-écho, il n'est alors pas nécessaire d'effectuer un filtrage pour atténuer ces composantes hautes fréquences, l'adaptation des coefficients de filtrage s'effectue alors par la mise à 0 ou à une valeur proche de 0 des coefficients de filtrage.

Ainsi, l'adaptation des coefficients du filtrage peut s'effectuer de façon discrète en fonction de la comparaison d'au moins un paramètre de décision à un seuil prédéterminé.

Les coefficients de filtrage peuvent prendre des valeurs prédéterminées selon un jeu de valeurs. Le jeu de valeurs le plus petit étant celui où seulement deux valeurs sont possibles, c'est un dire par exemple le choix entre un filtrage et pas de filtrage.

Dans une variante de réalisation, l'adaptation des coefficients du filtrage s'effectue de façon continue en fonction dudit au moins un paramètre de décision.

L'adaptation est alors plus précise et plus progressive.

Dans un mode particulier de réalisation, le filtrage est à réponse impulsionnelle finie à phase nulle de fonction de transfert:

$$c(n)z^{-1} + (1 - 2c(n)) + c(n)z$$

avec $c(n)$ un coefficient compris entre 0 et 0.25.

Ce type de filtrage est de faible complexité et permet de plus un traitement sans retard (le traitement s'arrêtant avant la fin de la trame courante). Grâce à son retard nul, le filtrage peut atténuer les hautes fréquences avant l'attaque sans modifier l'attaque elle-même.

Ce type de filtrage permet d'éviter les discontinuités et permet de passer d'un signal non-filtré à un signal filtré de façon progressive.

Selon un mode de réalisation, l'étape d'atténuation est effectuée en même temps que le filtrage de mise en forme spectrale en intégrant les facteurs d'atténuation aux coefficients définissant le filtrage.

La présente invention vise également un dispositif de traitement d'atténuation de pré-échos dans un signal audionumérique engendré à partir d'un codeur par transformée, dans lequel, le dispositif associé à un décodeur comprend:

- 10 -

- un module de détection pour détecter une position d'attaque dans le signal décodé;
- un module de détermination pour déterminer une zone de pré-écho précédant la position d'attaque détectée dans le signal décodé;

5 - un module de calcul de facteurs d'atténuations par sous-bloc de la zone de pré-écho, en fonction au moins de la trame dans laquelle l'attaque a été détectée et de la trame précédente;

- un module d'atténuation pour atténuer les pré-échos dans les sous-blocs de la zone de pré-écho par les facteurs d'atténuation correspondants. Le dispositif est tel qu'il comprend en outre:

10 - un module de filtrage adaptatif pour effectuer une mise en forme spectrale de la zone de pré-écho sur la trame courante jusqu'à la position détectée de l'attaque.

L'invention vise un décodeur d'un signal audionumérique comportant un dispositif tel que décrit précédemment.

15 Enfin, l'invention vise un programme informatique comportant des instructions de code pour la mise en œuvre des étapes du procédé de traitement d'atténuation tel que décrit, lorsque ces instructions sont exécutées par un processeur.

20 Enfin l'invention se rapporte à un support de stockage, lisible par un processeur, intégré ou non au dispositif de traitement, éventuellement amovible, mémorisant un programme informatique mettant en œuvre un procédé de traitement tel que décrit précédemment.

D'autres caractéristiques et avantages de l'invention apparaîtront plus clairement à la lecture de la description suivante, donnée uniquement à titre d'exemple non limitatif, et faite en référence aux dessins annexés, sur lesquels:

25 - la figure 1 décrite précédemment illustre un système de codage-décodage par transformée selon l'état de l'art;

- la figure 2 décrite précédemment illustre un exemple de signal audionumérique pour lequel une méthode d'atténuation selon l'état de l'art est effectuée;

- la figure 3 décrite précédemment illustre un autre exemple de signal audionumérique pour lequel une méthode d'atténuation selon l'état de l'art est effectuée;

30 - la figure 4 décrite précédemment illustre encore un autre un exemple de signal audionumérique pour lequel une méthode d'atténuation selon l'état de l'art est effectuée;

- les figures 5a et 5b illustrent respectivement le spectrogramme du signal original et le spectrogramme du signal avec atténuation de pré-échos selon l'état de l'art (correspondant respectivement aux parties a) et d) de la figure 4);

- 11 -

-la figure 6 illustre un dispositif de traitement d'atténuation de pré-échos dans un décodeur de signal audionumérique, ainsi que les étapes mises en œuvre par le procédé de traitement selon un mode de réalisation de l'invention;

5 - la figure 7 illustre la réponse fréquentielle d'un filtre de mise en forme spectrale mis en œuvre selon un mode de réalisation de l'invention, en fonction du paramètre du filtre;

- la figure 8 illustre un exemple de signal audionumérique pour lequel le traitement selon l'invention a été mis en œuvre;

- la figure 9 illustre le spectrogramme du signal correspondant au signal d) de la figure 4, pour lequel le traitement selon l'invention est mis en œuvre;

10 - la figure 10 illustre un exemple de signal présentant des composantes hautes fréquences à l'origine pour lequel une méthode d'atténuation des pré-échos selon l'état de l'art est mise en œuvre;

- La figure 11 illustre le même signal que la figure 11, présentant des composantes hautes fréquences à l'origine pour lequel le traitement selon l'invention a été mis en œuvre sans la prise en compte d'un critère de décision du niveau de filtrage à appliquer;

15 - la figure 12 illustre un exemple matériel de dispositif de traitement d'atténuation selon l'invention.

En référence à la **figure 6**, un dispositif 600 de traitement d'atténuation de pré-écho est décrit. Dans un mode de réalisation, ce dispositif met en œuvre une méthode d'atténuation des pré-échos dans le signal décodé comme par exemple celle décrite dans la demande de brevet FR 08 56248. Il met en outre en œuvre un filtrage de mise en forme spectrale de la zone de pré-écho.

Ainsi, le dispositif 600 comporte un module de détection 601 apte à mettre en œuvre une étape de détection (Detect.) de la position d'une attaque dans un signal audio décodé.

Une attaque (ou *onset* en anglais) est une transition rapide et une variation brusque de la dynamique (ou amplitude) du signal. On peut désigner ce type de signaux par le terme plus général de "transitoire". Dans la suite et sans perte de généralité, on utilisera uniquement les termes d'attaque ou de transition pour désigner également des transitoires.

30 Dans un mode de réalisation, chaque trame de L échantillons du signal décodé $x_{rec}(n)$ est divisée en K sous-blocs de longueur L', avec par exemple L = 640 échantillons (20 ms) à 32 kHz, L' = 80 échantillons (2.5 ms) et K = 8.

Des fenêtres d'analyse-synthèse spéciales à bas retard similaires à celles décrites dans la norme UIT-T G.718 sont utilisées pour la partie analyse et pour la partie synthèse de la transformation MDCT. Ainsi la fenêtre de synthèse MDCT ne contient que 415 échantillons non nuls contrairement aux 640 échantillons dans le cas d'utilisation des
 5 fenêtres sinusoïdales classiques. Dans une variante de ce mode de réalisation, d'autres fenêtres d'analyse/synthèse peuvent être utilisées, ou des commutations entre fenêtres longues et courtes peuvent être utilisées.

Par ailleurs, on utilise la mémoire MDCT $x_{MDCT}(n)$ qui donne une version avec repliement temporel ("folding" en anglais) du signal futur. Cette mémoire est aussi divisée en
 10 sous-blocs de longueur L' et on ne retient – en fonction de la fenêtre MDCT utilisée – que les K' premiers sous-blocs, où K' dépend de la fenêtre utilisée – par exemple $K' = 4$ pour une fenêtre sinusoïdale. En effet, la figure 1 montre que le pré-écho influence la trame précédant celle où se situe l'attaque, et il est souhaitable de détecter une attaque dans la trame future qui est en partie contenue dans la mémoire MDCT.

15 La réduction de pré-échos dépend ici de plusieurs paramètres:

- Le signal décodé dans la trame courante (qui contient potentiellement des pré-échos) de longueur L ,
- La mémoire de la transformation inverse MDCT qui correspond au signal partiellement décodé dans la trame suivante avant addition-
 20 recouvrement.
- Le niveau moyen d'énergie dans la trame (ou demi-trame) précédente.

On peut noter que le signal contenu dans la mémoire MDCT inclut un repliement temporel (qui est compensé lorsque la trame suivante est reçue). Comme expliqué ci-dessous, la mémoire MDCT sert ici essentiellement à estimer l'énergie par sous-blocs du
 25 signal dans la trame suivante (future) et on considère que cette estimation est suffisamment précise pour les besoins de la détection et réduction de pré-écho lorsqu'elle est réalisée avec la mémoire MDCT disponible à la trame courante au lieu du signal complètement décodé à la trame future.

La trame courante et la mémoire MDCT peuvent être vues comme des signaux
 30 concaténés formant un signal de longueur $(K + K')L'$ découpé en $(K+K')$ sous-blocs consécutifs. Dans ces conditions, on définit l'énergie dans le k -ième sous-bloc comme :

$$En(k) = \sum_{n=kL'}^{(k+1)L'-1} x_{rec}(n)^2, \quad k = 0, \dots, K-1$$

- 13 -

quand le k-ième sous-bloc se situe dans la trame courante et, comme:

$$En(k) = \sum_{n=(k-K)L'}^{(k-K+1)L'-1} x_{MDCT}(n)^2, \quad k = K, \dots, K + K'$$

quand le sous-bloc est dans la mémoire MDCT (qui représente le signal disponible pour la trame future).

5 L'énergie moyenne des sous-blocs dans la trame courante s'obtient donc comme :

$$\overline{En} = \frac{1}{K} \sum_{k=0}^{K-1} En(k)$$

On définit également l'énergie moyenne des sous-blocs dans la deuxième partie de la trame courante comme :

$$\overline{En}' = \frac{2}{K} \sum_{k=K/2}^{K-1} En(k)$$

10 Une transition associée à un pré-écho est détectée si le rapport

$$R(k) = \frac{\max_{k=0, K+K'} (En(k))}{En(k)} \text{ dépasse un seuil prédéfini, dans un des sous-blocs considérés.}$$

D'autres critères de détection de pré-écho sont possibles sans changer la nature de l'invention.

Par ailleurs, on considère que la position de l'attaque est définie comme

$$15 \quad pos = \min \left(L', \left(\arg \max_{k=0, K+K'} (En(k)) \right), L \right)$$

où la limitation à L assure que la mémoire MDCT n'est jamais modifiée. D'autres méthodes d'estimation plus précise de la position de l'attaque sont également possibles.

20 Dans des variantes de réalisation avec commutation des fenêtres, d'autres méthodes donnant la position de l'attaque peuvent être utilisées avec une précision allant de l'échelle d'un sous-bloc jusqu'à une position à l'échantillon près.

Le dispositif 600 comporte également un module 602 de détermination mettant en œuvre une l'étape de détermination (ZPE) d'une zone de pré-écho précédant la position d'attaque détectée.

25 Les énergies $En(k)$ sont concaténées en ordre chronologique, avec d'abord l'enveloppe temporelle du signal décodé, puis l'enveloppe du signal de la trame suivante estimée à partir de la mémoire de la transformée MDCT. En fonction de cette enveloppe

- 14 -

temporelle concaténée et des énergies moyennes $\overline{En}$ et $\overline{En'}$ de la trame précédente, la présence de pré-écho est détectée si le rapport $R(k)$ est suffisamment fort.

Les sous-blocs dans lequel un pré-écho a été détecté constituent ainsi une zone de pré-écho, qui en général couvre les échantillons $n = 0, \dots, pos - 1$, soit du début de la trame
5 courante à la position de l'attaque (pos).

Dans des variantes de réalisation, la zone de pré-écho ne commence pas nécessairement au début de la trame, et peut faire intervenir une estimation de la longueur du pré-écho. Si une commutation de fenêtres est utilisée, la zone de pré-écho devra être définie pour prendre en compte les fenêtres utilisées.

10 Un module 603 du dispositif 600 met en œuvre une étape de calcul de facteurs d'atténuation par sous-blocs de la zone de pré-écho déterminée, en fonction de la trame dans laquelle l'attaque a été détectée et de la trame précédente.

Conformément à la description de la demande de brevet FR 08 56248, les atténuations $g(k)$ sont estimées par sous-bloc.

15 Le facteur d'atténuation par sous-bloc $g(k)$ est calculé par exemple, en fonction du rapport $R(k)$ entre l'énergie du sous-bloc de plus forte énergie et l'énergie du k-ième sous-bloc en question :

$$g(k) = f(R(k))$$

où f est une fonction décroissante à valeurs entre 0 et 1. D'autres définitions du facteur

20 $g(k)$ sont possibles, par exemple en fonction de $En(k)$ et de $En(k-1)$.

Si la variation de l'énergie par rapport à l'énergie maximale est faible, aucune atténuation n'est alors nécessaire. Le facteur est alors fixé à une valeur d'atténuation inhibant l'atténuation, c'est-à-dire 1. Sinon, le facteur d'atténuation est compris entre 0 et 1.

Ces atténuations sont limitées en fonction de l'énergie moyenne de la trame
25 précédente.

Pour le sous-bloc à traiter on peut calculer la valeur limite du facteur $\lim_g(k)$ afin d'obtenir exactement la même énergie que l'énergie moyenne du segment précédant le sous-bloc à traiter. Cette valeur est bien sûr limitée à un maximum de 1 puisqu'on s'intéresse ici aux valeurs d'atténuation. Plus précisément :

- 15 -

$$\lim_g(k) = \min \left(\sqrt{\frac{\max(\overline{En}, \overline{En'})}{En(k)}}, 1 \right)$$

La valeur $\lim_g(k)$ ainsi obtenue sert de limite inférieure dans le calcul final du facteur d'atténuation du sous-bloc :

$$g(k) = \max(g(k), \lim_g(k))$$

5 Les facteurs d'atténuation $g(k)$ déterminés par sous-blocs sont ensuite lissés par une fonction de lissage appliquée échantillon par échantillon pour éviter des variations brusques du facteur d'atténuation aux frontières des blocs.

Le gain par échantillon est d'abord défini comme une fonction constante par morceaux :

$$10 \quad g_{pre}(n) = g(k), \quad n = kL', \dots, (k+1)L'-1$$

La fonction de lissage est par exemple définie par les équations suivantes:

$$g_{pre}(n) := \alpha g_{pre}(n-1) + (1-\alpha) g_{pre}(n), \quad n = 0, \dots, L-1$$

avec la convention que $g_{pre}(-1)$ est le dernier facteur d'atténuation obtenu pour le dernier échantillon du sous-bloc précédent, α est le coefficient de lissage, typiquement $\alpha=0.85$.

15 D'autres fonctions de lissage sont possibles.

Le module 604 du dispositif 600 de la figure 6 met en œuvre l'atténuation (Att.) dans les sous-blocs de la zone de pré-écho, par les facteurs d'atténuation obtenus.

Ainsi, une fois les facteurs $g_{pre}(n)$ calculés, l'atténuation de pré-écho est faite sur le signal reconstruit de la trame courante, $x_{rec}(n)$, en multipliant chaque échantillon par le

20 facteur correspondant :

$$x_{rec,g}(n) = g_{pre}(n) x_{rec}(n), \quad n = 0, \dots, L-1$$

où $x_{rec,g}(n)$ est le signal décodé et post-traité pour la réduction de pré-écho.

Le dispositif 600 comporte un module de filtrage 606 apte à effectuer l'étape (F) d'application d'un filtrage de mise en forme spectrale de la zone de pré-écho sur la trame

25 courante du signal décodé, jusqu'à la position détectée de l'attaque.

Typiquement, le filtre de mise en forme spectrale utilisé est un filtre linéaire. Comme l'opération de multiplication par un gain est également une opération linéaire leur ordre peut être inversé : on peut également faire d'abord le filtrage de mise en forme

- 16 -

spectrale de la zone de pré-écho puis l'atténuation de pré-écho en multipliant chaque échantillon de la zone de pré-écho par le facteur correspondant.

Dans un exemple de réalisation le filtre utilisé pour atténuer les hautes fréquences dans la zone de pré-écho est un filtre FIR (filtre à réponse impulsionnelle finie) à 3
5 coefficients et à phase nulle de fonction de transfert $c(n)z^{-1} + (1 - 2c(n)) + c(n)z$ avec $c(n)$ une valeur comprise entre 0 et 0.25, où $[c(n), 1 - 2c(n), c(n)]$ sont les coefficients du filtre de mise en forme spectrale ; ce filtre est mis en œuvre avec l'équation aux différences :

$$x_{rec,f}(n) = c(n)x_{rec,g}(n-1) + (1 - 2c(n))x_{rec,g}(n) + c(n)x_{rec,g}(n+1)$$

avec par exemple $c(n) = 0.25$ sur la zone $n = 5, \dots, pos - 5$.

10 La réponse fréquentielle de ce filtre est illustrée sur la **figure 7**, en fonction du coefficient $c(n)$, pour $c(n) = 0.05, 0.1, 0.15, 0.2$ et 0.25 . La motivation pour utiliser ce filtre est sa faible complexité, sa phase nulle et donc son retard nul (possible car le traitement s'arrête avant la fin de trame courante) mais également sa réponse fréquentielle qui correspond bien aux caractéristiques de passe-bas désirées pour ce filtre.

15 L'application de ce filtre peut compenser le fait que l'atténuation temporelle du pré-écho est typiquement limitée à une zone n'allant pas jusqu'à la position de l'attaque (avec une marge de par exemple 16 échantillons), alors que le filtrage de mise en forme spectrale tel que défini par la fonction de transfert $c(n)z^{-1} + (1 - 2c(n)) + c(n)z$ peut être appliqué jusqu'à la position de l'attaque, avec éventuellement quelques échantillons d'interpolation
20 des coefficients du filtre.

Pour passer d'un signal non-filtré à un signal filtré et éviter des discontinuités il est préférable d'introduire le filtrage de façon progressive. Le filtre FIR proposé permet facilement de passer en douceur du domaine non-filtré vers le domaine filtré et vice-versa, par interpolation ou variation lente de ses coefficients. Par exemple, si la position de l'attaque
25 est $pos = 16$, le filtrage des 16 échantillons dans la zone de pré-écho $n = 0, \dots, pos - 1$ peut être effectué de la façon suivante :

$$x_{rec,f}(0) = x_{rec}(0)$$

$$x_{rec,f}(1) = 0.1x_{rec}(0) + 0.8x_{rec}(1) + 0.1x_{rec}(2)$$

$$x_{rec,f}(2) = 0.1x_{rec}(1) + 0.8x_{rec}(2) + 0.1x_{rec}(3)$$

30 $x_{rec,f}(3) = 0.15x_{rec}(2) + 0.7x_{rec}(3) + 0.15x_{rec}(4)$

- 17 -

$$x_{rec,f}(4) = 0.2x_{rec}(3) + 0.6x_{rec}(4) + 0.2x_{rec}(5) =$$

$$x_{rec,f}(n) = 0.25x_{rec}(n-1) + 0.5x_{rec}(n) + 0.25x_{rec}(n+1), n = 5, \dots, 11$$

$$x_{rec,f}(12) = 0.2x_{rec}(11) + 0.6x_{rec}(12) + 0.2x_{rec}(13)$$

$$x_{rec,f}(13) = 0.15x_{rec}(12) + 0.7x_{rec}(13) + 0.15x_{rec}(14)$$

$$5 \quad x_{rec,f}(14) = 0.1x_{rec}(13) + 0.8x_{rec}(14) + 0.1x_{rec}(15)$$

$$x_{rec,f}(15) = 0.05x_{rec}(14) + 0.9x_{rec}(15) + 0.05x_{rec}(16)$$

On observe, que grâce à son retard nul, le filtre $c(n)z^{-1} + (1-2c(n)) + c(n)z$ peut atténuer les hautes fréquences avant l'attaque sans modifier l'attaque elle-même.

Un exemple de signal audionumérique, pour lequel le traitement tel que décrit ici
10 est effectué, est illustré en partie d) de la figure 8. Les parties a), b) et c) de cette figure reprennent les mêmes signaux que ceux décrits en référence à la figure 4 précédemment. La partie d) diffère par la mise en œuvre du filtrage selon l'invention. On peut ainsi remarquer que la composante haute fréquence gênante est fortement diminuée, si bien que le signal décodé après filtrage a une meilleure qualité que celui décrit en partie d) de la figure 4.

15 Le spectrogramme représentant ce signal filtré est représenté en figure 9. On observe bien par rapport à la figure 5b représentant le même signal sans filtrage de mise en forme, l'atténuation des hautes fréquences gênantes avant l'attaque. L'attaque devient alors plus nette au décodage.

Bien entendu, d'autres types de filtre de mise en forme spectrale peuvent être
20 envisagés pour remplacer le filtre $c(n)z^{-1} + (1-2c(n)) + c(n)z$. Par exemple, il est possible d'utiliser un filtre FIR d'ordre différent ou avec des coefficients différents. Alternativement le filtre de mise en forme spectrale peut être à réponse impulsionnelle infinie (IIR). De plus, la mise en forme spectrale peut être différente d'un filtrage passe-bas, par exemple un filtre passe-bande pourrait être mis en œuvre.

25 Un filtre d'ordre 1, de la forme $c(n)z^{-1} + (1-c(n))$ peut également être utilisé dans un mode de réalisation de l'invention.

Dans un mode de réalisation particulier, le filtrage mis en œuvre selon le procédé décrit, est un filtrage adaptatif. Il peut ainsi être adapté aux caractéristiques du signal audio décodé.

- 18 -

Dans ce mode de réalisation, une étape de calcul d'un paramètre (P) de décision sur le filtrage à appliquer à la zone de pré-écho est mise en œuvre dans le module de calcul 605 de la figure 6.

En effet, il existe des cas comme celui illustré par exemple à la figure 10 où il est
5 préférable de ne pas appliquer un tel filtrage dans la zone de pré-écho.

En effet, dans le cas, plus rare, illustré à la figure 10, partie a) les hautes fréquences sont déjà présentes dans le signal à coder. Dans ce cas l'atténuation des hautes fréquences pourrait causer une dégradation audible qu'il faut donc éviter. Dans cet exemple de signal, on observe que l'attaque est moins brusque que dans les exemples précédents.

10 Il est alors intéressant de déterminer au moins un paramètre qui permet de décider s'il faut mettre en forme spectralement la zone du signal contenant un pré-écho, en atténuant (ou non) les hautes fréquences.

Dans un exemple de réalisation, ce paramètre de décision est représentatif de la présence de composantes hautes fréquences dans la zone de pré-écho.

15 Ce paramètre peut être par exemple une mesure de la force de l'attaque (brusque ou non). Si l'attaque est localisée dans le sous-bloc numéro k, le paramètre peut-être calculé comme :

$$P = \frac{\max(En(k), En(k+1))}{\min(En(k-1), En(k-2))}$$

où k le numéro du sous-bloc et $En(k)$ l'énergie dans le k-ième sous-bloc.

20 Selon un réglage expérimental, dans cet exemple de réalisation, $P \geq 32$ indique une attaque brusque (très impulsive).

La mesure de force de l'attaque peut être complétée en tenant en compte également de l'atténuation déterminée pour le sous-bloc précédant l'attaque $g(k-1)$. Une attaque peut être considérée comme brusque si cette atténuation est significative, par exemple si $g(k-1) \leq 0.5$. Ceci montre que l'énergie dans la zone de pré-écho est considérablement augmentée (plus que doublée) à cause du pré-écho, ce qui signale également une attaque brusque.
25

Si $P < 32$ et $g(k-1) > 0.5$, où k est l'indice du sous-bloc contenant le début de l'attaque, le filtrage n'est pas nécessaire. En effet, si $g(k-1) > 0.5$, $\lim_g(k) > 0.5$, ce qui signifie que la zone de pré-écho a une énergie comparable avec celle de la trame précédente et comme l'attaque qui génère le pré-écho n'est pas brusque, le risque d'avoir une
30 composante parasite gênante est faible.

- 19 -

Ainsi, dans ce mode de réalisation avec les conditions ($P < 32$ et $g(k-1) > 0.5$), aucun filtrage ne sera fait sur la zone de pré-écho.

Dans les autres cas ($g(k-1) \leq 0.5$ ou $P > 32$) le filtre de mise en forme spectrale est appliqué, selon l'invention, du début de la trame courante jusqu'à la position pos de position de l'attaque.

Dans l'exemple de réalisation décrit ci-dessus la mise en forme spectrale de la zone de pré-écho par filtrage selon l'invention est adaptative en fonction du paramètre P et des valeurs d'atténuation. Ainsi, le filtrage est soit appliqué avec des coefficients $[0.25, 0.5, 0.25]$, soit désactivé avec des coefficients $[0, 1, 0]$.

L'adaptation des coefficients de filtrage s'effectue alors de façon discrète limitée à un jeu de valeurs prédéfini.

L'adaptation des coefficients de filtrage (permettant d'adapter le niveau d'atténuation des hautes fréquences) est déterminé donc par des paramètres de décision qui mesurent la force de l'attaque comme les paramètres P et $g(k-1)$.

Il s'agit dans ce cas d'une d'adaptation des coefficients du filtre de façon discrète suivant deux jeux de valeurs possibles ($[0.25, 0.5, 0.25]$ ou $[0, 1, 0]$). On peut noter que le jeu de coefficients $[0, 1, 0]$ correspond à une désactivation du filtrage.

Une transition progressive entre ces deux filtres peut être effectuée en utilisant également par exemple les filtres intermédiaires de coefficient $[0.05, 0.9, 0.05]$, $[0.1, 0.8, 0.1]$, $[0.15, 0.7, 0.15]$ et $[0.2, 0.6, 0.2]$.

Il s'agit dans ce cas d'une adaptation des coefficients du filtre de façon discrète suivant plusieurs jeux de valeurs possibles, si on tient compte de la variation lente (ou interpolation).

Dans des variantes de réalisations, d'autres méthodes d'interpolation peuvent être utilisées.

Par exemple, le filtrage peut être encore plus finement adaptatif avec $c(n) = f(P)$ par exemple en utilisant un filtre intermédiaire avec $c(n) = [0.15, 0.7, 0.15]$ si $16 < P < 32$. $c(n)$ peut être également calculé de façon continue en fonction de P , par exemple avec le

$$\text{formule } c(n) = \frac{\arctan(P/10)}{2\pi}$$

Il s'agit dans ce cas d'une d'adaptation des coefficients du filtre de façon continue suivant des valeurs possibles où $c(n)$ est dans l'intervalle $[0, 0.25]$.

- 20 -

D'autres paramètres de décision peuvent être également utilisés dans la décision du choix et de l'adaptation du filtre, comme par exemple le taux de passage à zéro ("zero crossing rate" en anglais) du signal décodé de la zone pré-écho de la trame courante et/ou de la trame précédente. Le taux de passage par zéro peut être calculé de la façon suivante si on

5 considère la zone $n = 0, \dots, L-1$ à titre d'exemple :

$$zC = \frac{1}{2} \sum_{n=0}^{L-1} \left| \operatorname{sgn} \left[x_{rec,g}(n-1) \right] - \operatorname{sgn} \left[x_{rec,g}(n) \right] \right|$$

où

$$\operatorname{sgn}(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ -1 & \text{if } x < 0 \end{cases}$$

En effet, un taux élevé de passage à zéro zC dans la trame précédente (donc sans

10 pré-écho) signale la présence de hautes fréquences dans le signal. Dans ce cas, par exemple quand $zC > L/2$ sur la trame précédente, il est préférable de ne pas appliquer le filtrage $c(n)z^{-1} + (1-2c(n)) + c(n)z$.

Afin d'éliminer le biais de la composante continue, un préfiltrage du signal décodé est également possible avant calcul du taux de passage par zéro, ou bien le nombre de

15 passage par zéro de la dérivée estimée $x_{rec,g}(n) - x_{rec,g}(n-1)$ peut être utilisé.

Dans une variante, une analyse spectrale du signal peut être également faite pour aider à la décision. Par exemple, l'enveloppe spectrale dans le domaine MDCT issue du codage/décodage MDCT peut être exploitée dans le choix du filtre à utiliser, cependant cette

20 variante suppose que les fenêtres d'analyse/synthèse MDCT sont suffisamment courtes pour que les statistiques locales du signal avant l'attaque restent stables sur la longueur d'une fenêtre.

Alternativement, on pourra filtrer le signal dans la zone de pré-écho et dans la trame passée par un filtre complémentaire passe-haut comme $-c(n)z^{-1} + (1-2c(n)) - c(n)z$, avec par exemple $c(n) = 0.25$, et ensuite on choisira la

25 valeur de $c(n)$ de sorte que l'énergie moyenne des signaux filtrés dans la zone de pré-écho et sur la trame passée soient les plus proches possibles; le choix de $c(n)$ pourra se faire sur un jeu limité de valeurs possibles montrées à la figure 7 ou à partir du ratio d'énergie (ou d'une quantité équivalente comme la racine carrée de l'énergie) du signal après filtrage passe-haut dans la zone de pré-écho et dans la trame passée.

- 21 -

A noter que le filtrage passe-haut peut également être mis en œuvre de façon alternative en calculant la différence entre le signal $x_{rec,g}(n)$ et le signal filtré par le filtre passe-bas $c(n)z^{-1} + (1-2c(n)) + c(n)z$ quand $c(n) = 0.25$.

Dans une autre variante, quand le filtrage de mise en forme est de type $c(n)z^{-1} + (1-c(n))$, on pourra fixer la valeur de $c(n)$ en fonction du coefficient de prédiction $-r(1)/r(0)$ issu d'une analyse par prédiction linéaire (LPC pour "Linear Predictive Coding" en anglais) à l'ordre 1 du signal dans la zone de pré-écho et du signal dans la trame passée.

Dans toutes ces dernières variantes (taux de passage à zéro, enveloppe spectrale MDCT, filtrage passe haut, analyse LPC), le paramètre de décision sur le filtrage à appliquer à la zone de pré-écho est basé sur une analyse de répartition spectrale du signal de la zone pré-écho et/ou du signal précédant de la zone pré-écho ; si le signal précédant la zone pré-écho contient déjà beaucoup de hautes fréquences ou si la quantité des hautes fréquences du signal dans la zone pré-écho et du signal précédant la zone pré-écho est sensiblement identique, le filtrage selon l'invention n'est pas nécessaire et peut même causer une légère dégradation. Dans ces cas il faut désactiver ou atténuer le filtrage selon l'invention en fixant $c(n)$ à 0 ou à une faible valeur proche de 0.

Dans une variante de l'invention l'ordre entre l'étape d'atténuation et de filtrage pourra être inversé.

Il se peut en effet que le filtrage (F) de mise en forme spectrale se fasse avant l'atténuation (Att.). Ainsi, après avoir effectué le filtrage adaptatif des échantillons de la zone de pré-écho du signal reconstruit de la trame courante, ces échantillons sont alors pondérés en multipliant chaque échantillon par le facteur d'atténuation correspondant calculé précédemment :

$$x_{rec,f,g}(n) = g_{pre}(n)x_{rec,f}(n), \quad n = 0, \dots, L-1$$

L'atténuation des amplitudes peut également être combinée (ou intégrée) en définissant un jeu de coefficients de filtre "conjoint", par exemple si pour l'échantillon n le filtre a des coefficients $[c(n), 1-2c(n), c(n)]$ et le facteur d'atténuation est $g(n)$, on peut directement utiliser le filtre $[g_{pre}(n)c(n), g_{pre}(n)2g_{pre}(n)c(n), g_{pre}(n)c(n)]$.

La figure 11 illustre l'avantage de rendre le filtrage adaptatif. Elle reprend les mêmes signaux parties a), b) et c) que la figure 10 et illustre le fait que la mise en œuvre du

- 22 -

filtrage non-adaptatif représenté en partie d), modifie inutilement le signal dans le cas où les composantes hautes-fréquences sont déjà présentes dans le signal à coder. On observe qu'à partir de l'échantillon 640 les hautes fréquences sont inutilement atténués ce qui pourrait poser une légère dégradation de qualité. L'utilisation d'un filtrage adaptatif comme décrit ci-dessus permet d'inhiber ou d'atténuer le filtrage dans ces conditions, de pas enlever des hautes fréquences déjà présentes dans le signal à coder et d'éviter ainsi une éventuelle dégradation dû au filtrage.

Pour revenir à la figure 6, le dispositif de traitement d'atténuation 600 tel que décrit est ici compris dans un décodeur comportant un module 610 de quantification inverse (Q^{-1}) recevant un signal S, un module 620 de transformée inverse ($MDCT^{-1}$), un module 630 de reconstruction du signal par addition/recouvrement (add/rec) comme décrit en référence à la figure 1 et délivrant un signal reconstruit au dispositif de traitement d'atténuation selon l'invention.

En sortie du dispositif 600, un signal traité Sa est fourni dans lequel une atténuation de pré-écho a été effectué. Le traitement effectué a permis d'améliorer l'atténuation de pré-écho par l'atténuation, le cas échéant, des composantes hautes fréquences, dans la zone de pré-écho.

Un exemple de réalisation d'un dispositif de traitement d'atténuation selon l'invention est maintenant décrit en référence à la **figure 12**.

Matériellement, ce dispositif 100 au sens de l'invention comporte typiquement, un processeur μP coopérant avec un bloc mémoire BM incluant une mémoire de stockage et/ou de travail, ainsi qu'une mémoire tampon MEM précitée en tant que moyen pour mémoriser toutes données nécessaire à la mise en œuvre du procédé de traitement d'atténuation tel que décrit en référence à la figure 6. Ce dispositif reçoit en entrée des trames successives du signal numérique Se et délivre le signal Sa reconstruit avec atténuation de pré-écho et filtrage de mise en forme spectrale, le cas échéant.

Le bloc mémoire BM peut comporter un programme informatique comportant les instructions de code pour la mise en œuvre des étapes du procédé selon l'invention lorsque ces instructions sont exécutées par un processeur μP du dispositif et notamment une étape de détection d'une position d'attaque dans le signal décodé, de détermination d'une zone de pré-écho précédant la position d'attaque détectée dans le signal décodé, de calcul de facteurs d'atténuations par sous-bloc de la zone de pré-écho, en fonction de la trame dans laquelle l'attaque a été détectée et de la trame précédente, d'atténuation de pré-écho dans les sous-blocs de la zone de pré-écho par les facteurs d'atténuation correspondants et en outre, une

- 23 -

étape d'application d'un filtrage de mise en forme spectrale de la zone de pré-écho sur la trame courante jusqu'à la position détectée de l'attaque. La figure 6 peut illustrer l'algorithme d'un tel programme informatique.

5 Ce dispositif d'atténuation selon l'invention peut être indépendant ou intégré dans un décodeur de signal numérique.

REVENDICATIONS

1. Procédé de traitement d'atténuation de pré-écho dans un signal audionumérique engendré à partir d'un codage par transformée, dans lequel, au décodage, le procédé comporte les étapes suivantes:

- détection (Detect.) d'une position d'attaque dans le signal décodé;
- détermination (ZPE) d'une zone de pré-écho précédant la position d'attaque détectée dans le signal décodé;
- calcul (F. Att.) de facteurs d'atténuations par sous-bloc de la zone de pré-écho, en fonction au moins de la trame dans laquelle l'attaque a été détectée et de la trame précédente;
- atténuation (Att.) de pré-écho dans les sous-blocs de la zone de pré-écho par les facteurs d'atténuation correspondants;

le procédé étant caractérisé en ce qu'il comporte en outre:

- l'application d'un filtrage (F) adaptatif de mise en forme spectrale de la zone de pré-écho sur la trame courante jusqu'à la position détectée de l'attaque.

2. Procédé selon la revendication 1, caractérisé en ce qu'il comporte en outre le calcul d'au moins un paramètre de décision sur le filtrage à appliquer à la zone de pré-écho et l'adaptation des coefficients du filtrage en fonction dudit au moins un paramètre de décision.

3. Procédé selon la revendication 2, caractérisé en ce que ledit au moins un paramètre de décision est une mesure de la force de l'attaque détectée.

4. Procédé selon la revendication 2, caractérisé en ce que ledit au moins un paramètre de décision est la valeur du facteur d'atténuation dans le sous-bloc précédant celui contenant la position de l'attaque.

5. Procédé selon la revendication 2, caractérisé en ce que ledit au moins un paramètre de décision est basé sur une analyse de répartition spectrale du signal de la zone de pré-écho et/ou du signal précédant la zone de pré-écho.

6. Procédé selon la revendication 3, caractérisé en ce que la mesure de la force de l'attaque détectée est de la forme:

- 25 -

$P = \max (EN(k), EN(k+1)) / \min(EN(k-1), EN(k-2))$ avec k , le numéro du sous-bloc dans lequel l'attaque a été détectée et $EN(k)$ l'énergie du $k^{\text{ième}}$ sous-bloc.

7. Procédé selon la revendication 2, caractérisé en ce que l'adaptation des coefficients du filtrage s'effectue de façon discrète en fonction de la comparaison d'au moins un paramètre de décision à un seuil prédéterminé.

8. Procédé selon la revendication 2, caractérisé en ce que l'adaptation des coefficients du filtrage s'effectue de façon continue en fonction dudit au moins un paramètre de décision.

9. Procédé selon la revendication 1, caractérisé en ce que le filtrage est à réponse impulsionnelle finie à phase nulle de fonction de transfert:

$$c(n)z^{-1} + (1 - 2c(n)) + c(n)z$$

avec $c(n)$ un coefficient compris entre 0 et 0.25.

10. Procédé selon la revendication 1, caractérisé en ce que l'étape d'atténuation est effectuée en même temps que le filtrage de mise en forme spectrale en intégrant les facteurs d'atténuation aux coefficients définissant le filtrage.

11. Dispositif de traitement d'atténuation de pré-écho dans un signal audionumérique engendré à partir d'un codeur par transformée, dans lequel, le dispositif associé à un décodeur comprend:

- un module de détection (601) pour détecter une position d'attaque dans le signal décodé;
 - un module de détermination (602) pour déterminer une zone de pré-écho précédant la position d'attaque détectée dans le signal décodé;
 - un module de calcul (603) de facteurs d'atténuations par sous-bloc de la zone de pré-écho, en fonction au moins de la trame dans laquelle l'attaque a été détectée et de la trame précédente;
 - un module d'atténuation (604) pour atténuer les pré-échos dans les sous-blocs de la zone de pré-écho par les facteurs d'atténuation correspondants;
- le dispositif comprenant en outre:

- 26 -

- un module de filtrage (606) adaptatif pour effectuer une mise en forme spectrale de la zone de pré-écho sur la trame courante jusqu'à la position détectée de l'attaque.

12. Décodeur d'un signal audionumérique comportant un dispositif selon la revendication 11.

13. Programme informatique comportant des instructions de code pour la mise en œuvre des étapes du procédé selon l'une des revendications 1 à 10, lorsque ces instructions sont exécutées par un processeur.

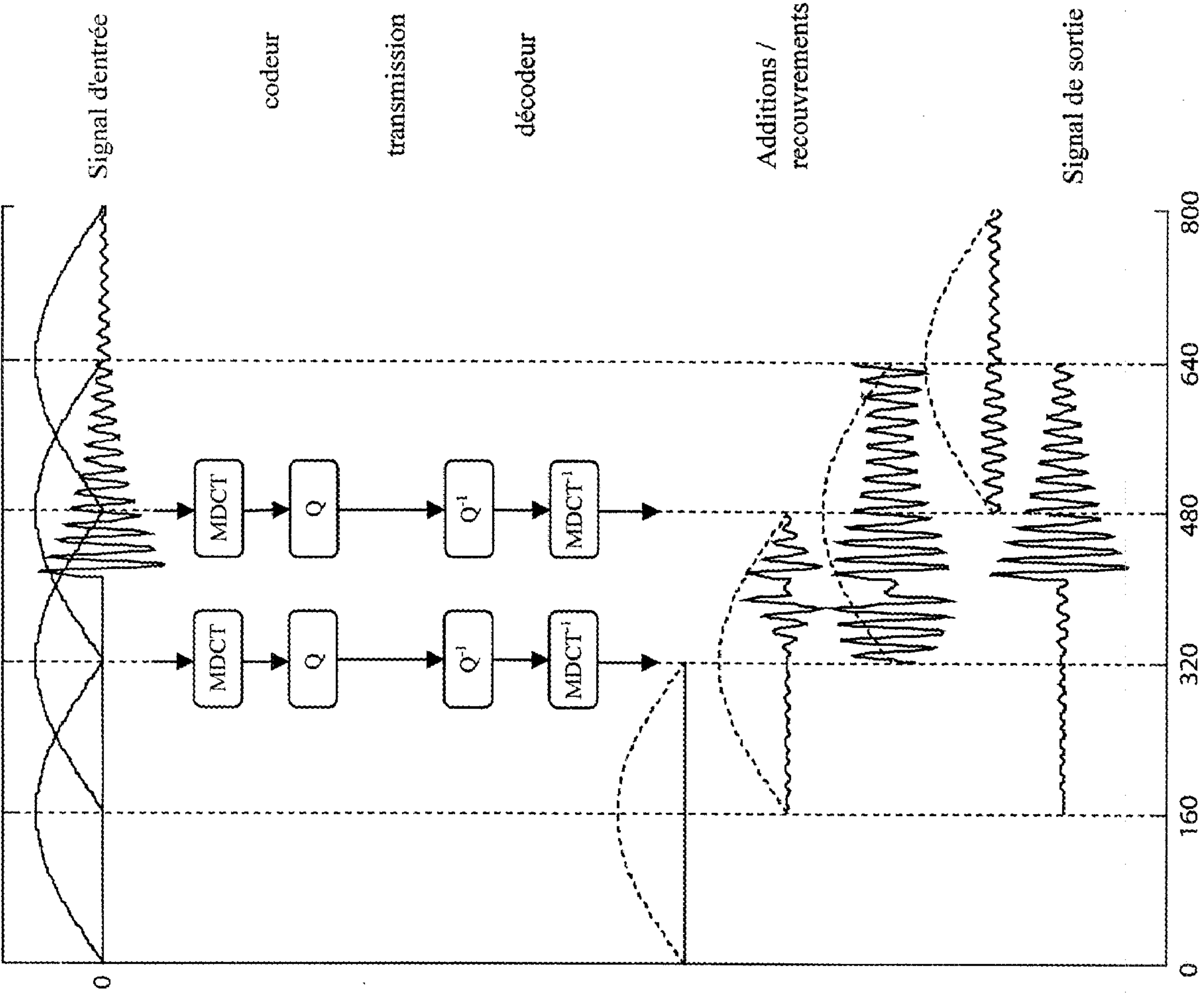


FIG.1

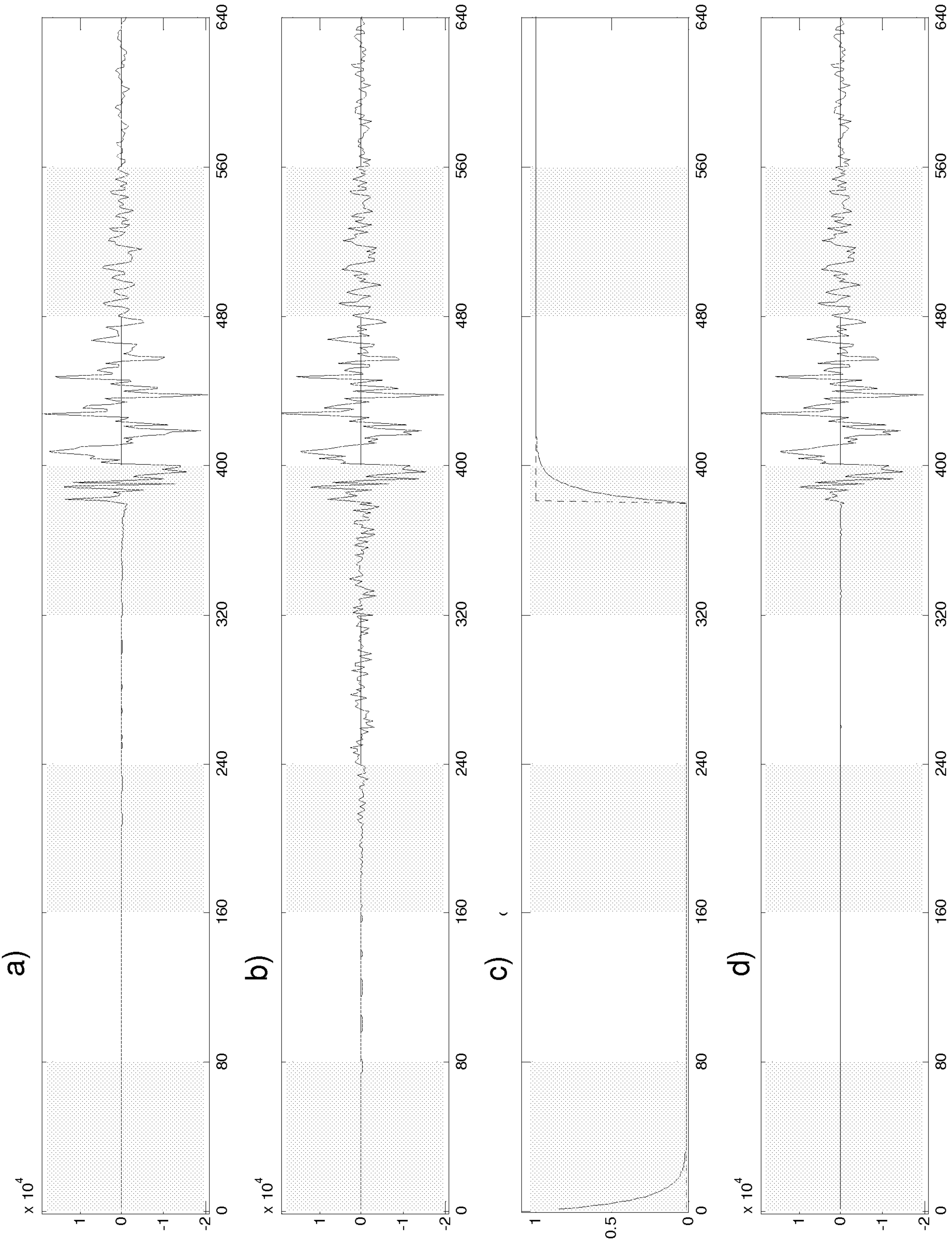


FIG.2

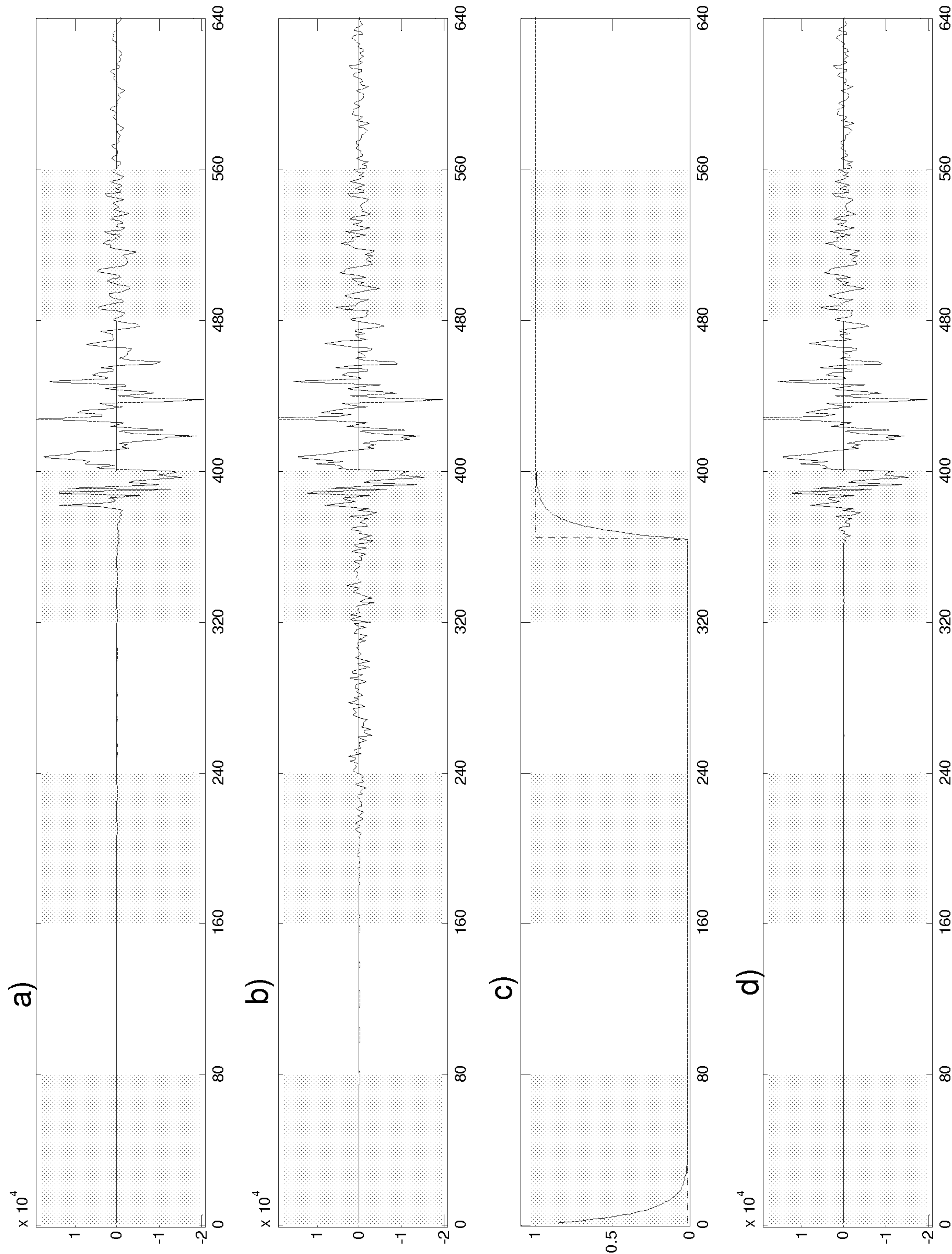


FIG.3

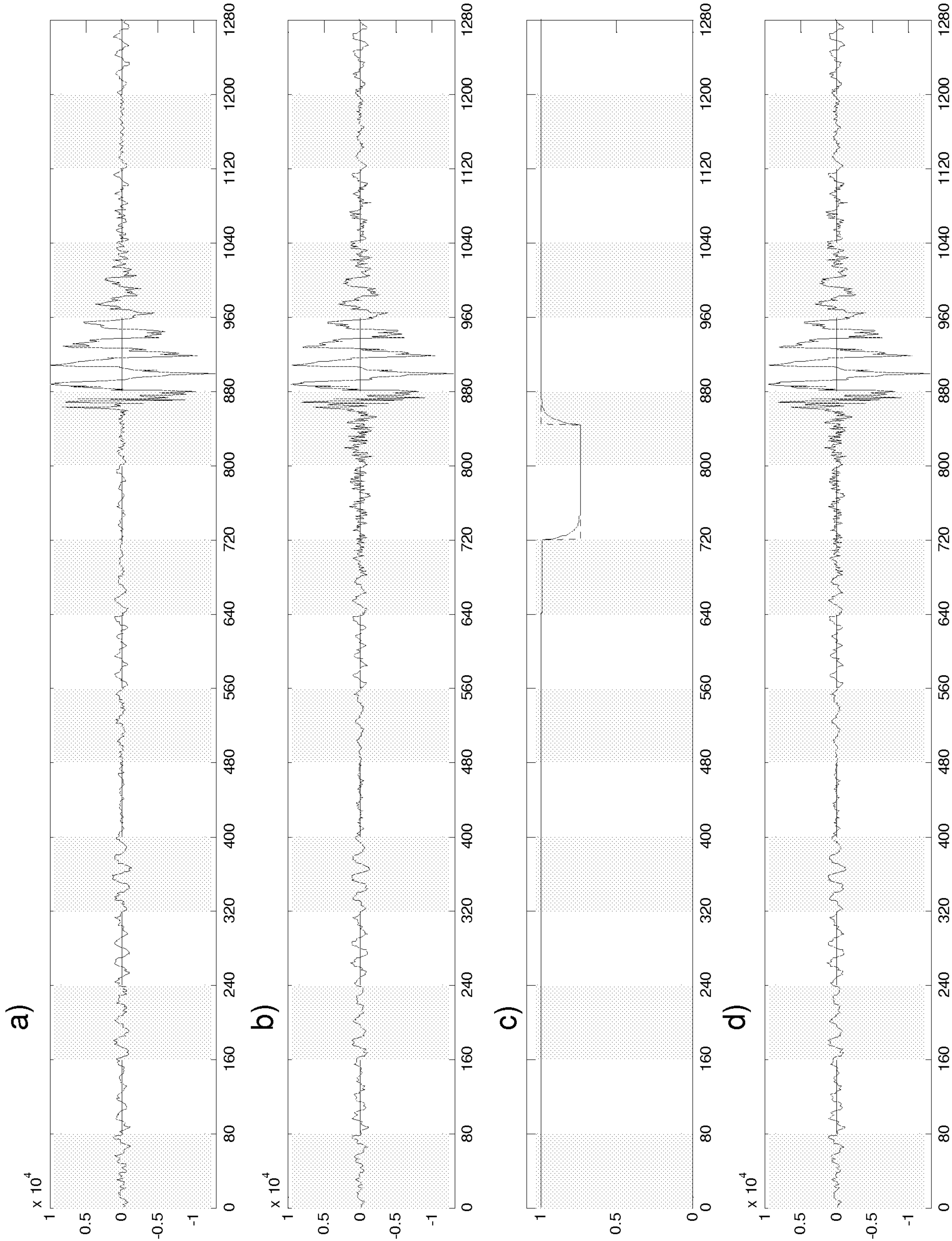


FIG.4

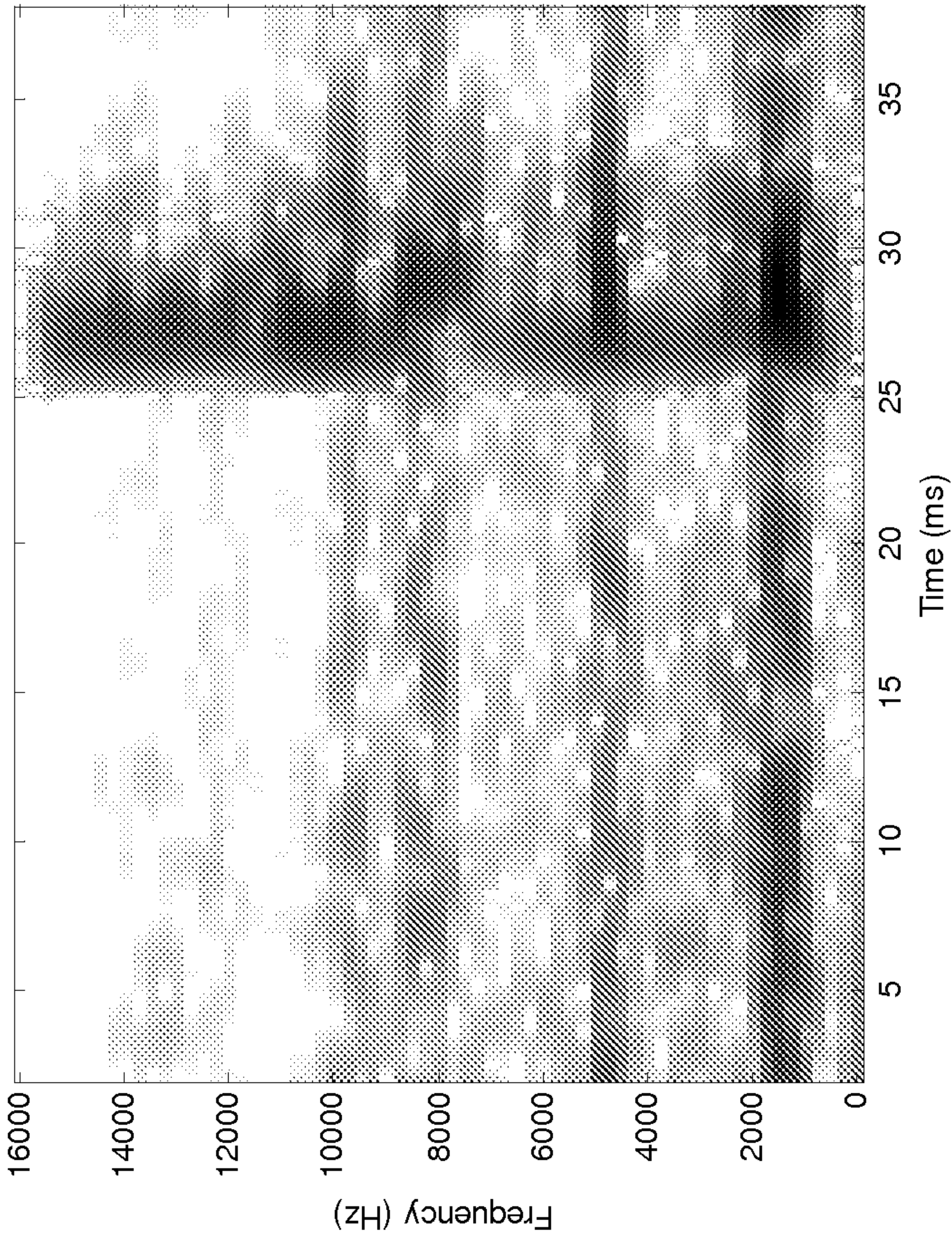


FIG.5a

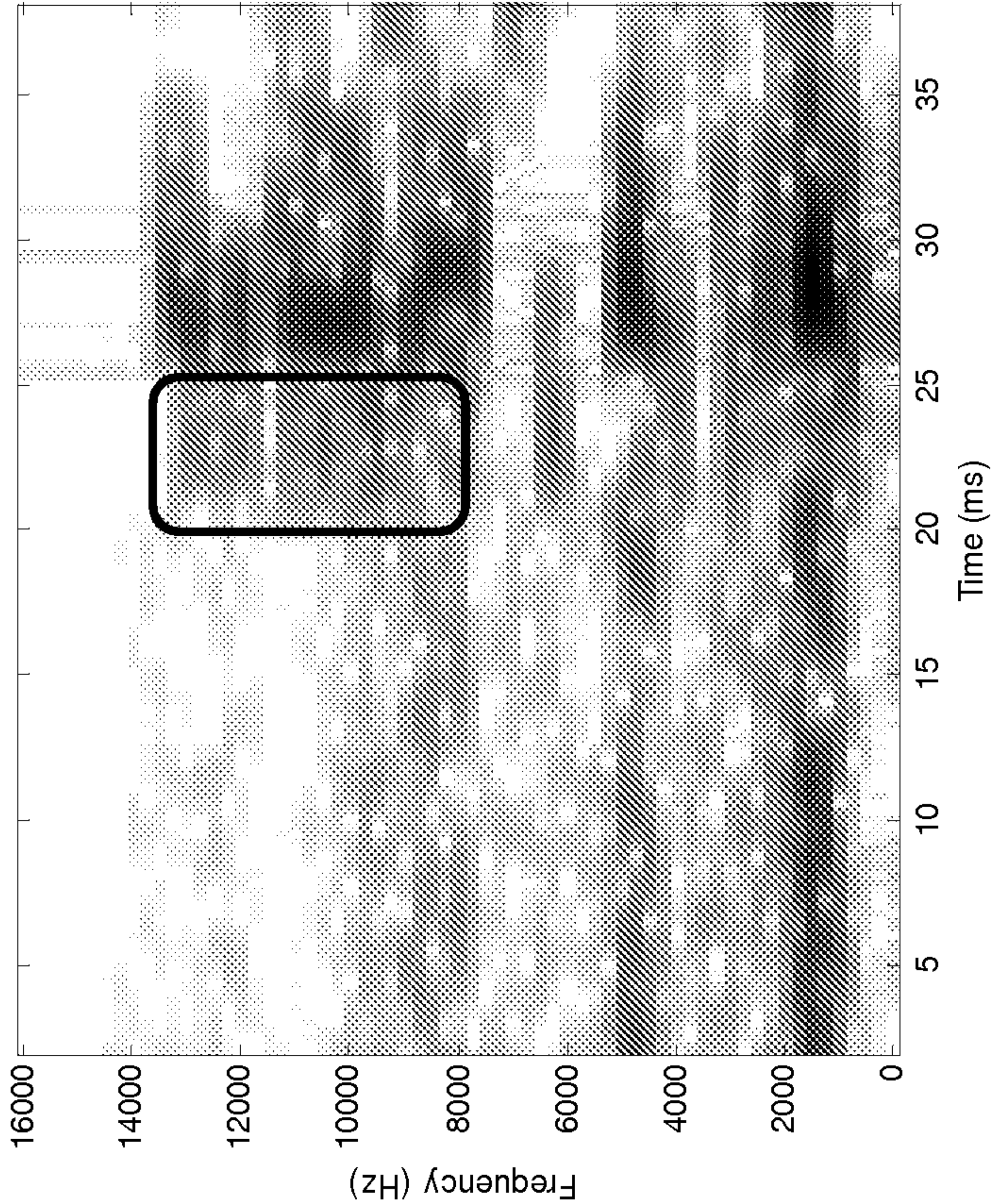


FIG.5b

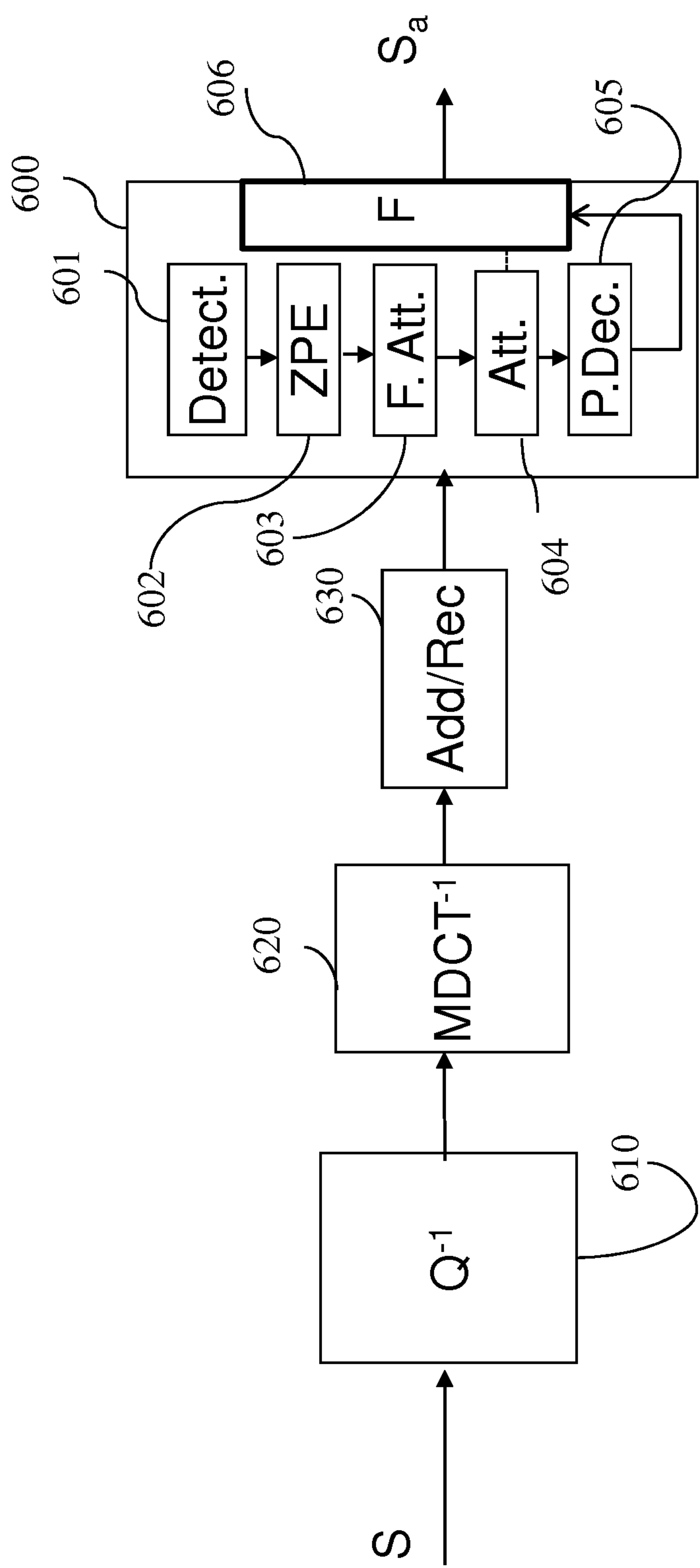


FIG.6

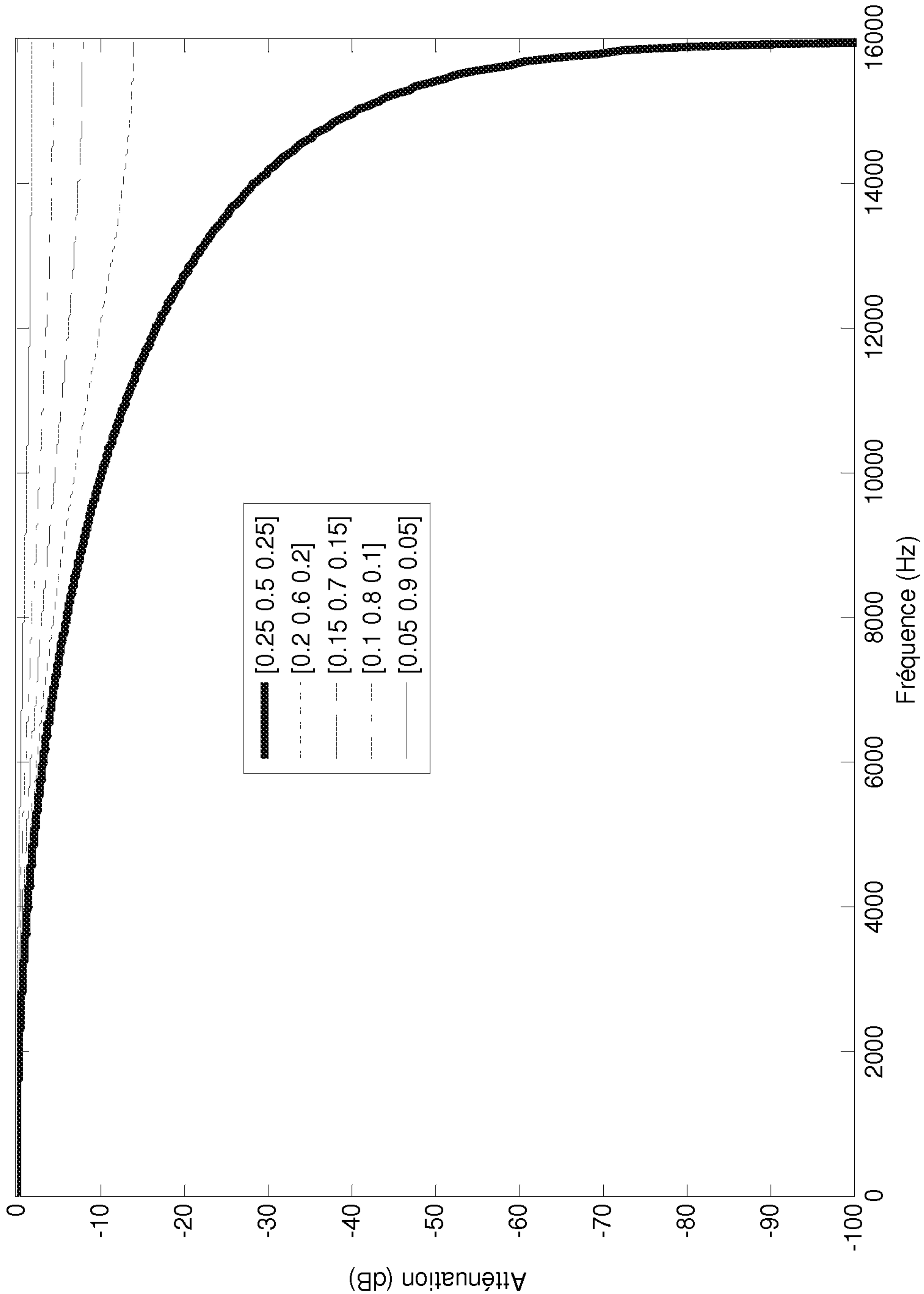


FIG.7

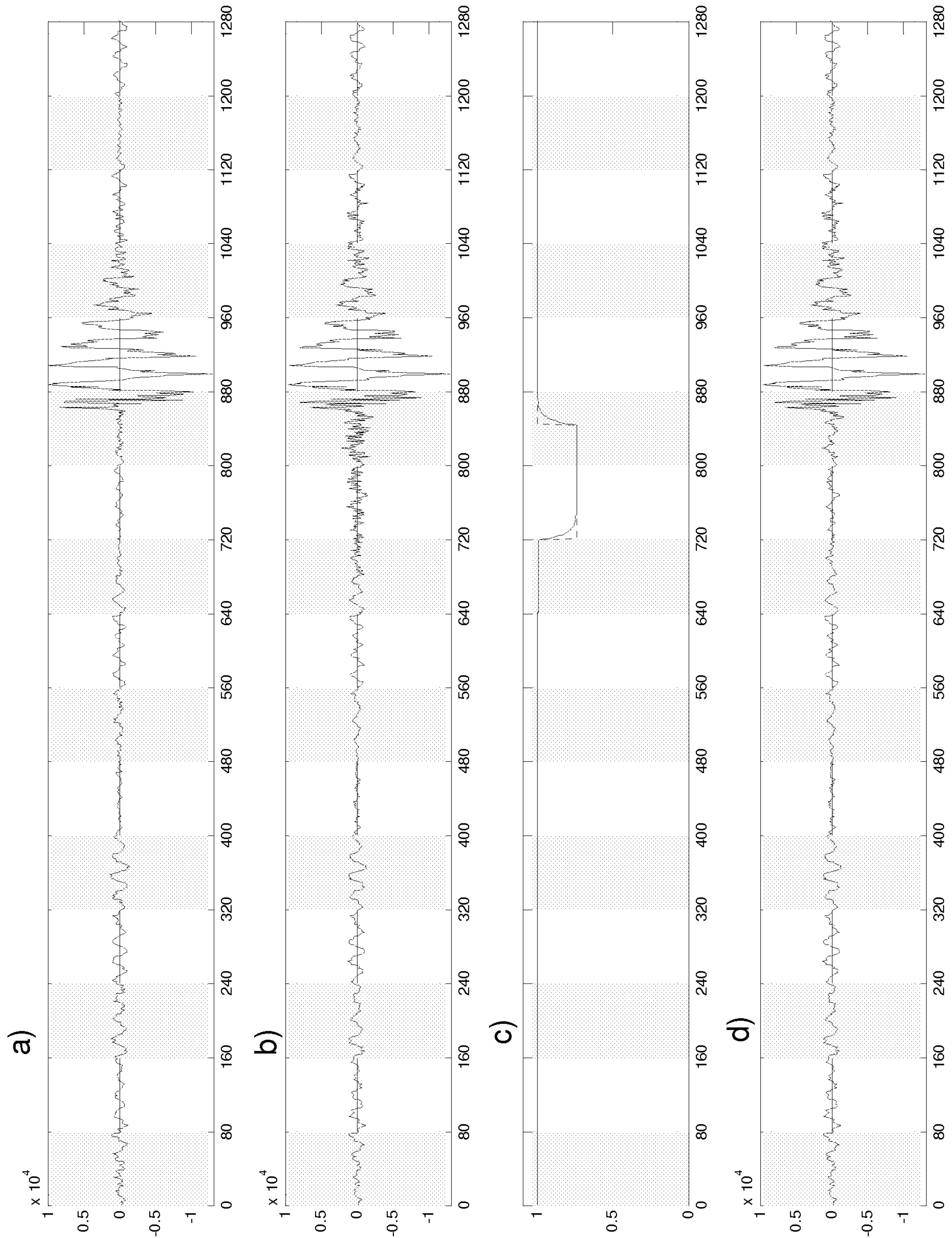


FIG.8

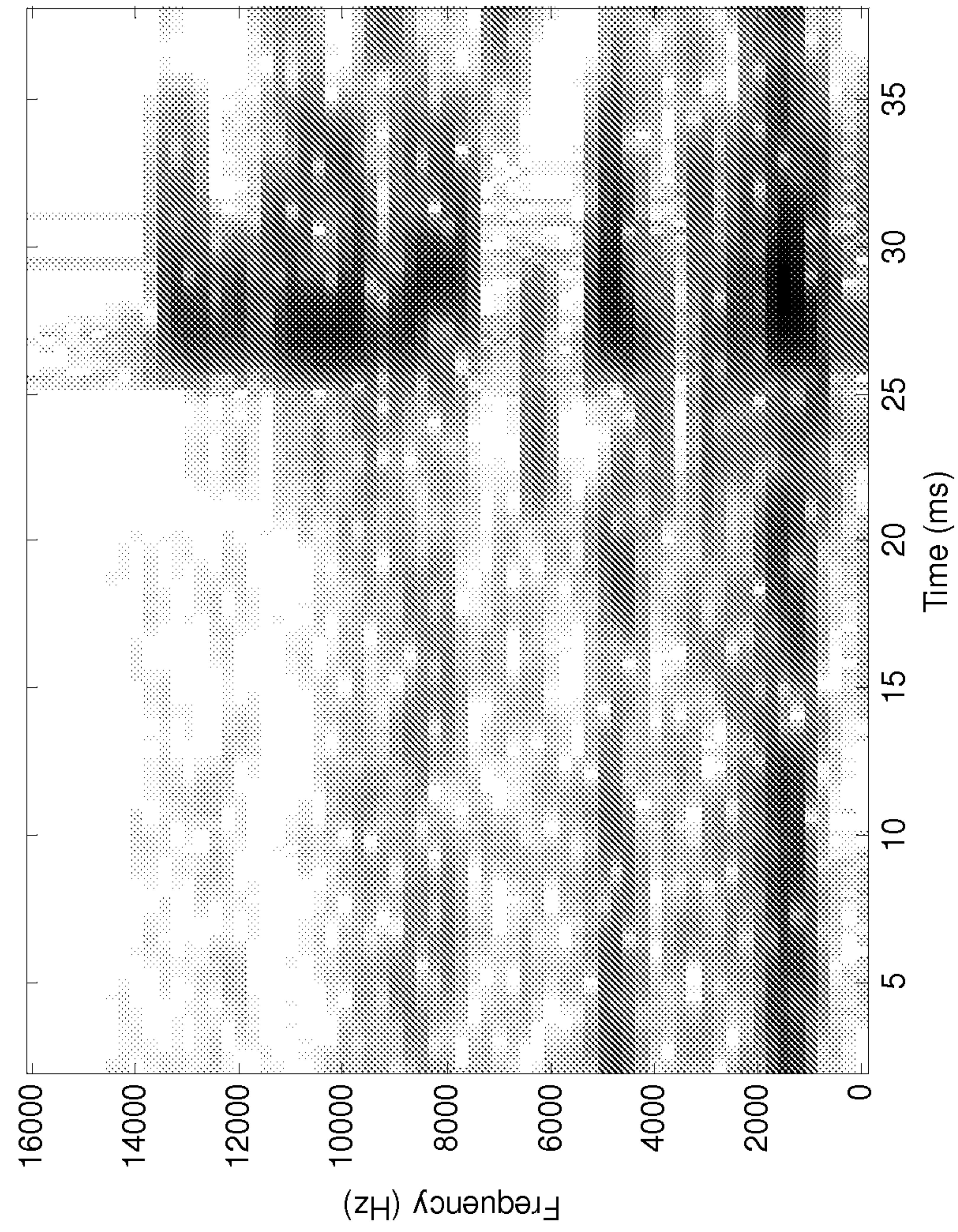


FIG.9

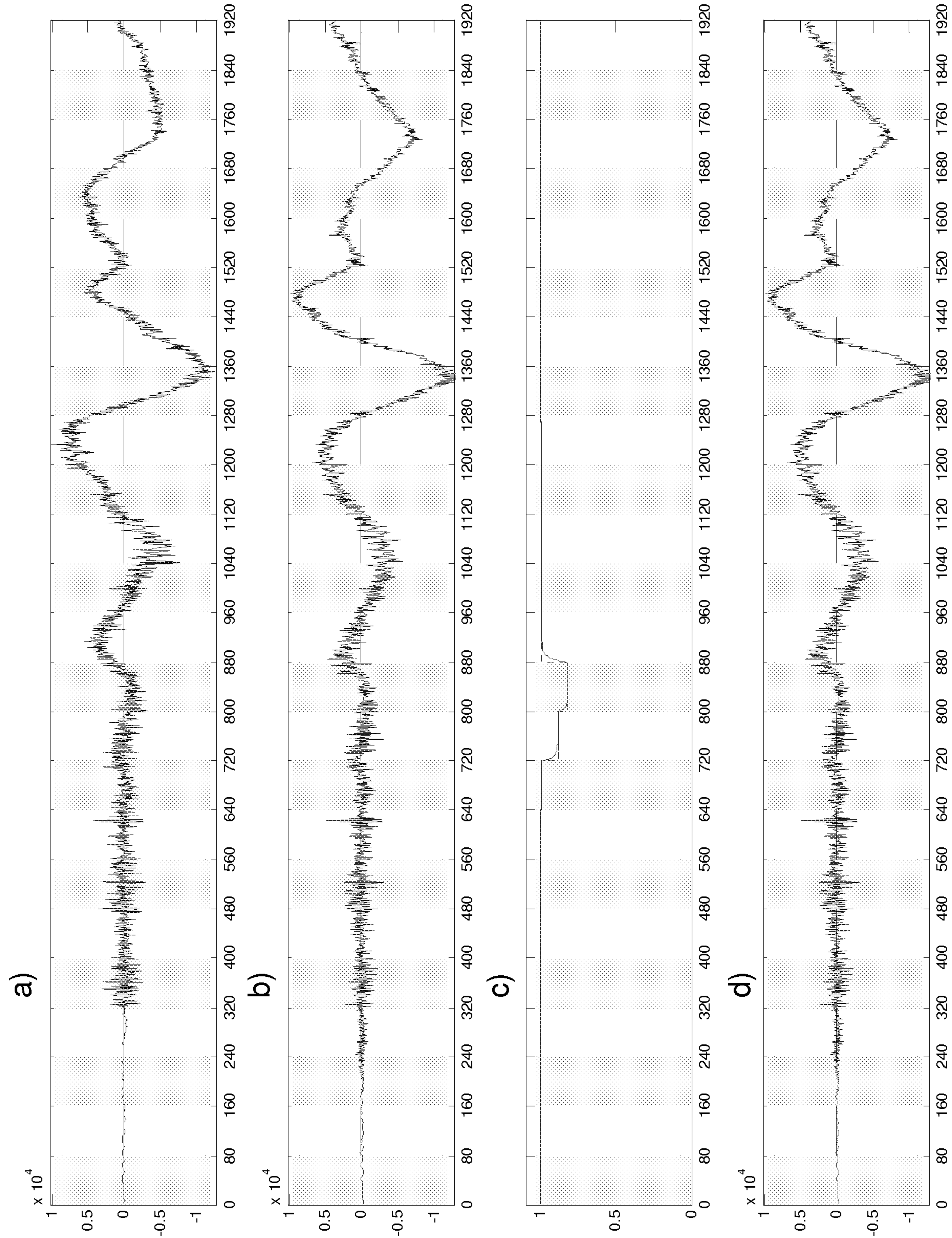


FIG.10

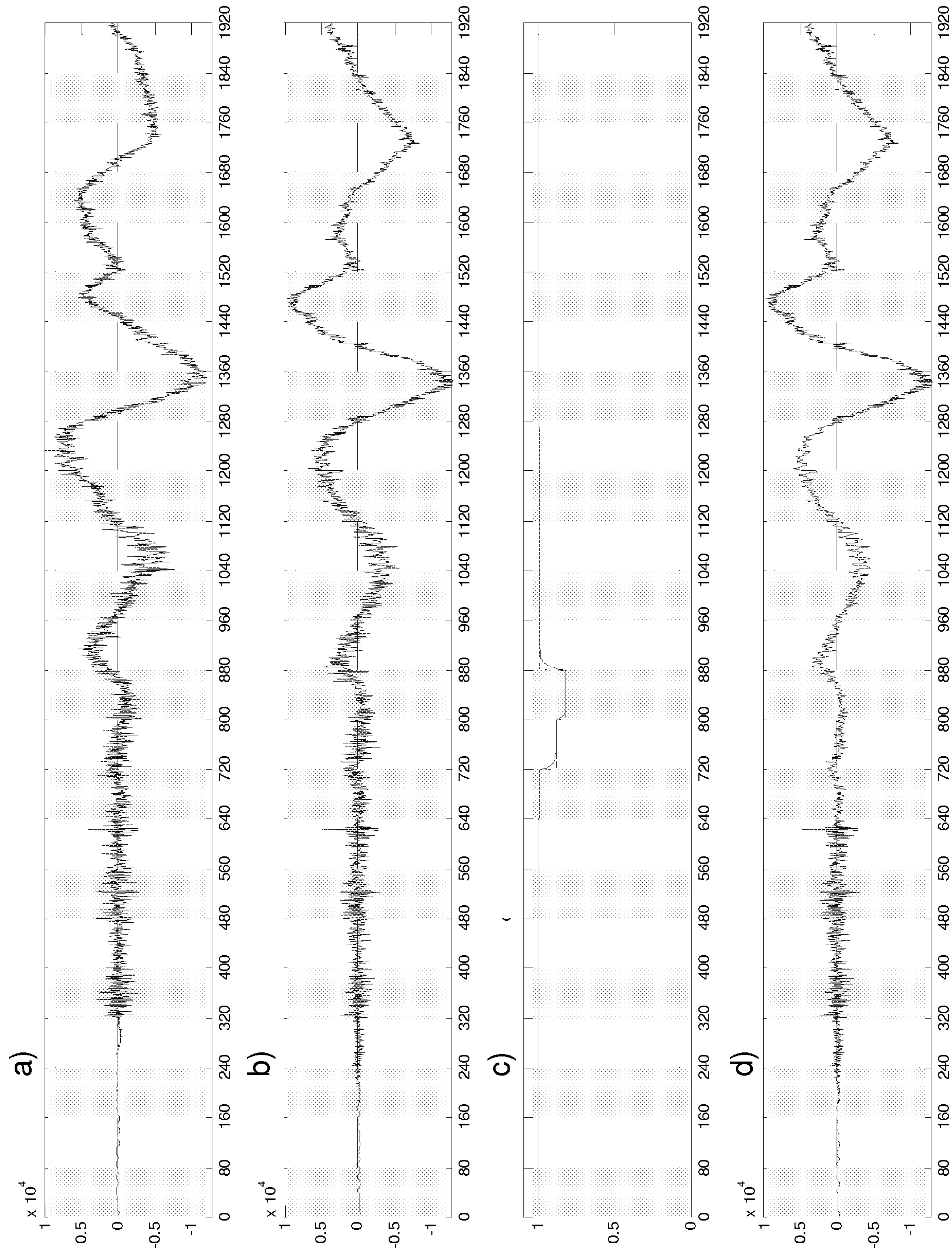


FIG.11

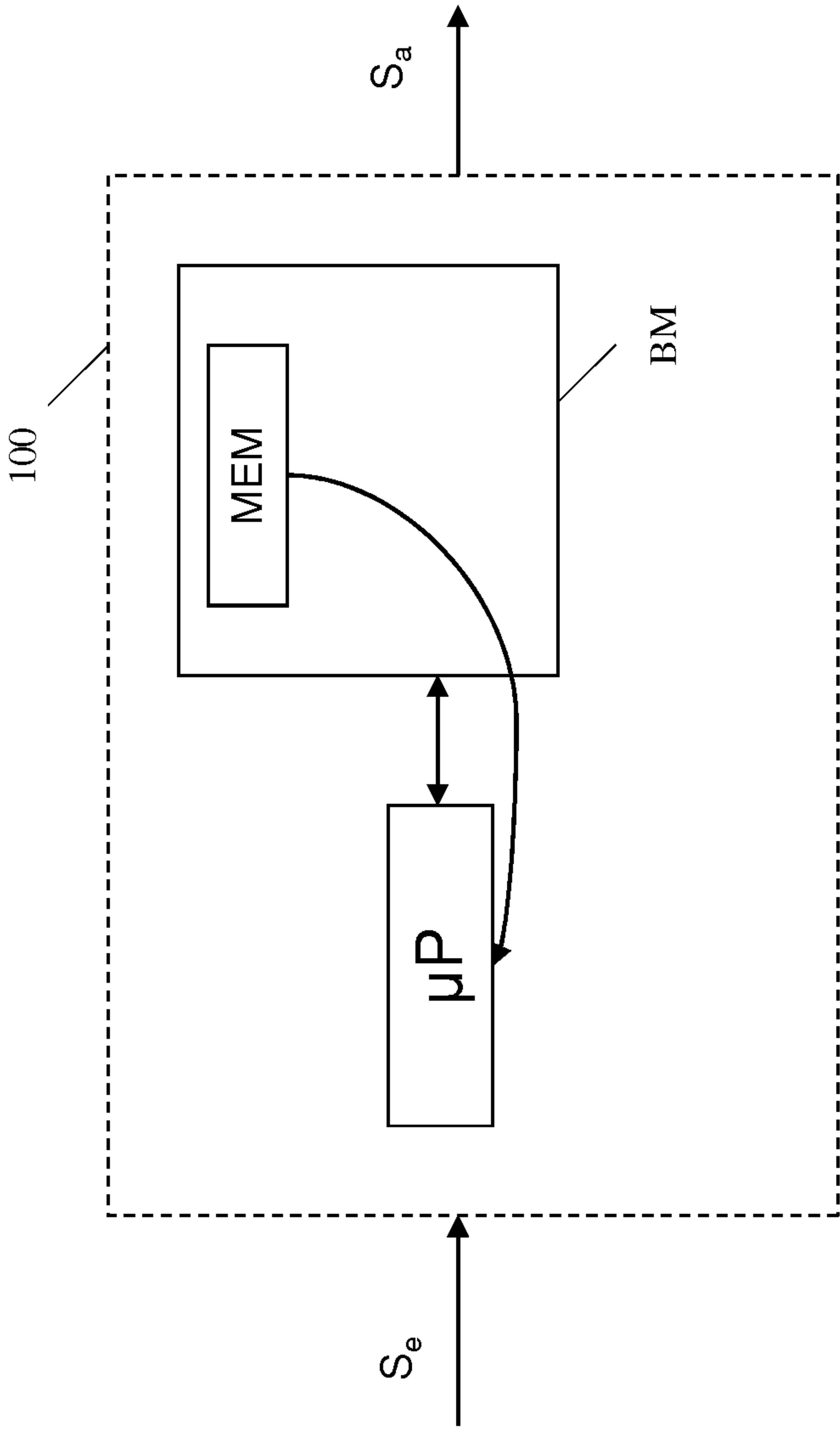


FIG.12

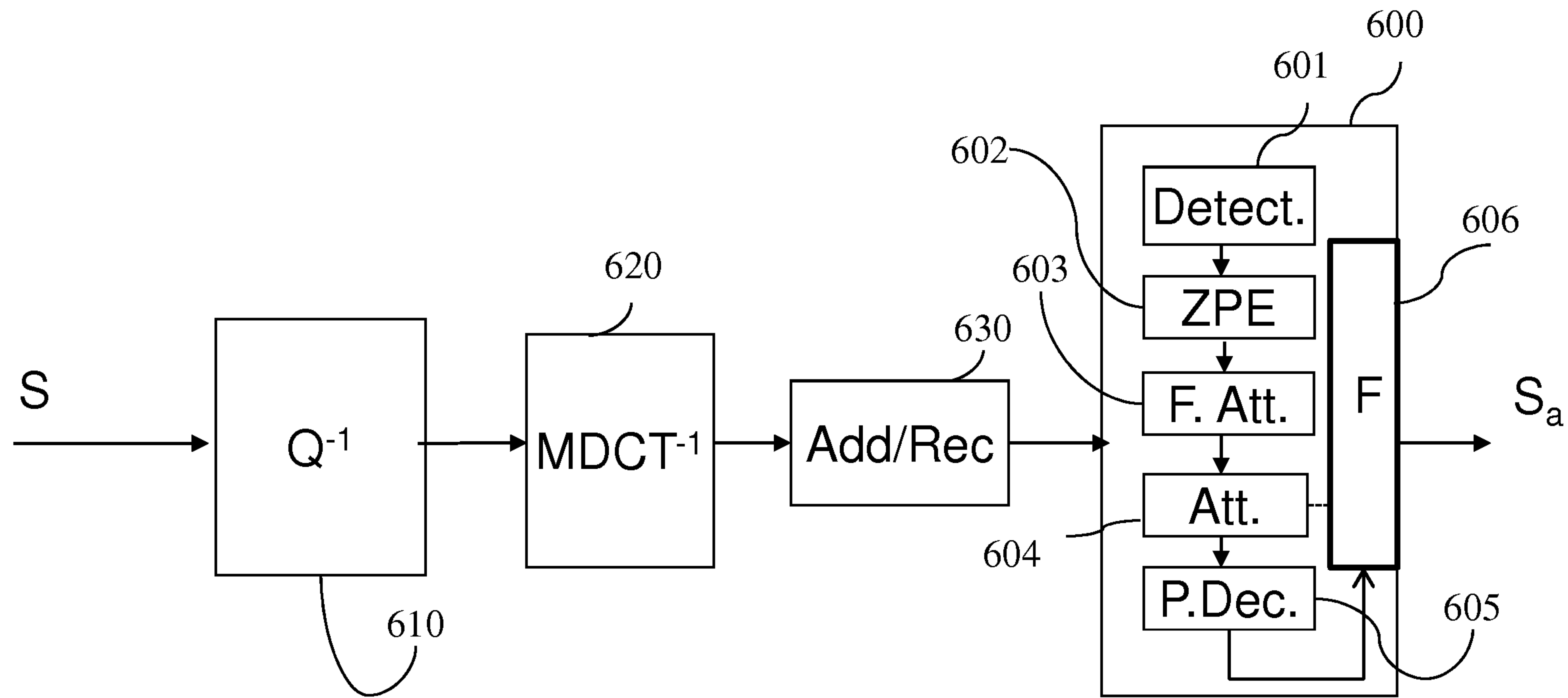


FIG.6