

Chapitre 3

Des extensions et des améliorations à la réallocation

Ce chapitre contient une collection de contributions dédiées à la généralisation et à l'amélioration de la méthode de réallocation. L'unité de ce chapitre tient dans la manière avec laquelle sont abordées toutes ces contributions à savoir : l'abandon de la distribution réallouée comme point de départ à tout traitement supplémentaire et son remplacement par la distribution classique, i.e., non réallouée (dans ce chapitre, le spectrogramme) accompagnée de son champ des vecteurs de réallocation. Par conséquent, toutes les propositions faites ici partent d'un point de vue en amont, i.e., avant réallocation, dont l'objectif est l'extraction du maximum d'information des quantités dont on dispose. Les distributions bilinéaires d'énergie temps-fréquence comme le spectrogramme, sont bien connues. Il reste à approfondir notre connaissance du champ de réallocation et en tirer parti. Dans un premier temps, nous décrirons la structure géométrique du champ de vecteurs de réallocation. Nous verrons comment il est possible de l'associer à des réseaux de courbes de niveau (Sect. 3.1.1), ou de lignes de plus grande pente (Sect. 3.1.2), ce qui nous conduira naturellement à introduire une nouvelle forme de réallocation (Sect. 3.2). Fort de cette description géométrique et doté de cette nouvelle méthode, nous verrons que l'on peut déduire des informations supplémentaires sur la structure même du signal (Sect. 3.3). Enfin, nous exploiterons à nouveau le champ des vecteurs de réallocation du spectrogramme, mais cette fois lorsqu'on fait varier la fenêtre, pour remédier au manque de robustesse des vecteurs de réallocation établi au chapitre précédent et obtenir une distribution réallouée plus stable en présence de bruit.

3.1 Géométrie du champ des vecteurs de réallocation

Les champs de vecteurs de réallocation ne sont pas n'importe quels champs de vecteurs. Ils obéissent à des contraintes structurelles en relation avec celles qui asservissent la forme des densités d'énergie temps-fréquence, plus précisément, celles qui déterminent l'admissibilité d'une fonction quelconque de \mathbb{R}^2 en tant que distribution d'énergie temps-fréquence. Ces contraintes se traduisent par une géométrie particulière pour le champ des vecteurs de réallocation que nous allons décrire ici. Nous montrons à ce sujet l'existence deux fonctions particulièrement importantes puisque les vecteurs de réallocation sont tangents aux courbes de niveau de la première, et aux lignes de plus grande pente de la seconde. Ces travaux ont fait l'objet d'une publication [24].

3.1.1 Phase géométrique et courbes de niveau

L'opérateur de Weyl, pris dans sa définition symplectique,

$$\mathbf{W}(t, \omega) = \exp(i(\omega t - t\omega)), \quad (3.1)$$

(où $[th](x) = xh(x)$ et $[\omega h](x) = -idh/dx$ sont les opérateurs associés respectivement au temps et à la fréquence) transforme le signal h en le translatant d'un retard t sur l'axe des temps et d'un doppler ω sur l'axe fréquentiel. Son action sur h , $[\mathbf{W}(t, \omega)h](x) = h(x - t) \exp i(\omega x - t\omega/2)$, fait apparaître un terme de phase pure important, puisqu'il symétrise le traitement des variables temporelle et fréquentielle. Si on définit la FCT comme une corrélation entre le signal x et les translatées de la fenêtre h dans le plan temps-fréquence par $\mathbf{W}(t, \omega)$, elle prend alors indifféremment une forme de corrélation temporelle

$$F_x(t, \omega) = \int x(s)h^*(s - t)e^{-i\omega s} ds e^{it\omega/2}, \quad (3.2)$$

ou fréquentielle, qui s'obtient en remplaçant simplement dans l'équation précédente chaque quantité par sa dual par Fourier et en changeant les signes dans les exponentielles :

$$F_x(t, \omega) = \int X(\xi)H^*(\xi - \omega)e^{it\xi} \frac{d\xi}{2\pi} e^{-it\omega/2}. \quad (3.3)$$

La symétrie formelle s'étend également aux opérateurs de réallocation (cf. éq. (1.32)) :

$$\hat{\mathbf{r}}(t, \omega) = (-t/2 - \partial_\omega \varphi, -\omega/2 + \partial_t \varphi)^t, \quad (3.4)$$

qui se décomposent en un vecteur diagonal dans le plan temps-fréquence ajouté à un vecteur orthogonal au gradient de la phase. L'éq. (3.4) suggère que les vecteurs de réallocation suivent des courbes de niveau d'une fonction de \mathbb{R}^2 dans \mathbb{R} , liée à $\varphi(t, \omega)$. Ces courbes de niveau doivent nécessairement être, comme le champ de vecteurs de réallocation, covariantes aux translations du signal dans le plan temps-fréquence. Or, la phase $\varphi(t, \omega)$ ne respecte aucune de ces covariances.

En effet, si on note par $[\mathbf{T}_s x](t) = [\mathbf{W}(s, 0)x](t) = x(t - s)$, la version translatée en temps de x , sa FCT s'écrit

$$F_{\mathbf{T}_s x}(t, \omega) = F_x(t - s, \omega) e^{-i\omega s/2}. \quad (3.5)$$

Si s prend une valeur quelconque dans \mathbb{R} , les phases de $F_{\mathbf{T}_s x}(t, \omega)$ et de $F_x(t - s, \omega)$ ne sont égales qu'à la condition $\omega = 0$. En conséquence, $\varphi(t, \omega)$ n'est pas covariante aux translations en temps du signal. Des équations similaires peuvent être écrites pour les translations en fréquence.

Par conséquent, nous cherchons une fonction 2D qui liée à la phase mais qui doit s'affranchir de la dépendance de la phase sur le choix de l'origine du plan temps-fréquence. Pour résoudre notre problème, nous introduisons une nouvelle fonction $\Phi_{(t_0, \omega_0)}(t, \omega)$ que l'on appelle *phase géométrique*. La fonction $\Phi_{(t_0, \omega_0)}(t, \omega)$ est la phase que nous mesurons au point temps-fréquence de coordonnée (t, ω) si (t_0, ω_0) est l'origine du référentiel. Déplacer l'origine en temps de 0 à t_0 signifie que la valeur du signal x à la distance t de la nouvelle origine est donnée par $x(t + t_0) = [\mathbf{W}(-t_0, 0)x](t)$. De manière similaire, déplacer l'origine temps-fréquence en (t_0, ω_0) correspond à remplacer x par $\mathbf{W}(-t_0, -\omega_0)x$. Donc¹,

$$\Phi_{(t_0, \omega_0)}(t, \omega) = \arg \langle \mathbf{W}(-t_0, -\omega_0)x, \mathbf{W}(t, \omega)h \rangle \quad (3.6)$$

$$= \arg \langle x, \mathbf{W}(t_0, \omega_0)\mathbf{W}(t, \omega)h \rangle. \quad (3.7)$$

1. $\langle x, y \rangle = \int x(s)y^*(s) ds$ désigne le produit scalaire canonique dans $L^2(\mathbb{R})$ et $\|x\| = \sqrt{\langle x, x \rangle}$, la norme associée.

Puisque nous avons

$$\begin{aligned} [\mathbf{W}(t_0, \omega_0) \mathbf{W}(t, \omega) h](x) \\ = \mathbf{W}(t_0, \omega_0) \left(h(x - t) e^{i\omega x} e^{-it\omega/2} \right) \end{aligned} \quad (3.8)$$

$$= h(x - (t + t_0)) e^{i[2(\omega + \omega_0)x - (\omega + \omega_0)(t + t_0) + (t\omega_0 - t_0\omega)]/2} \quad (3.9)$$

$$= [\mathbf{W}(t_0 + t, \omega_0 + \omega) h](x) e^{i[(t_0, \omega_0)(t, \omega)]/2}, \quad (3.10)$$

on peut exprimer $\Phi_{(t_0, \omega_0)}(t, \omega)$ à l'aide de la forme symplectique $[(t_0, \omega_0)(t, \omega)] = t\omega_0 - t_0\omega$,

$$\Phi_{(t_0, \omega_0)}(t, \omega) = \varphi(t + t_0, \omega + \omega_0) - [(t_0, \omega_0)(t, \omega)]/2, \quad (3.11)$$

ce qui prouve directement que le vecteur de réallocation est tangent en (t_0, ω_0) aux courbes de niveau de $\Phi_{(t_0, \omega_0)}(t, \omega)$:

$$\hat{r}(t_0, \omega_0) = (-\partial_\omega \Phi_{(t_0, \omega_0)}(t, \omega), \partial_t \Phi_{(t_0, \omega_0)}(t, \omega))^t|_{(t, \omega)=(0,0)}. \quad (3.12)$$

Ceci justifie l'appellation de phase “géométrique” pour la fonction $\Phi_{(t_0, \omega_0)}(t, \omega)$, en ce qu'elle fournit une interprétation géométrique locale du champ de vecteurs de réallocation au point (t_0, ω_0) .

3.1.2 Potentiel scalaire et lignes de plus grande pente

Réécrivons la FCT $F(t, \omega)$ du signal x à l'aide du formalisme proposé par Bargmann dans [10], c'est-à-dire en utilisant \mathcal{F} , fonction de la variable complexe $z = \omega + it$ et de sa variable conjuguée z^* :

$$F(t, \omega) = \mathcal{F}(z, z^*) \exp(-|z|^2/4). \quad (3.13)$$

Les phases de F et \mathcal{F} sont évidemment égales. En notant que $\log F(t, \omega) = \log |\mathcal{F}(t, \omega)| + i\varphi(t, \omega)$ et en utilisant l'éq. (3.4), il est donc possible d'exprimer les vecteurs de réallocation avec les dérivées partielles de \mathcal{F} :

$$\hat{t}(t, \omega) - t = -t/2 - \text{Im}\{\partial_\omega \mathcal{F}/\mathcal{F}\} \quad (3.14)$$

$$= -t/2 - \text{Im}\{(\partial_z \mathcal{F} + \partial_{z^*} \mathcal{F})/\mathcal{F}\}, \quad (3.15)$$

$$\hat{\omega}(t, \omega) - \omega = -\omega/2 + \text{Im}\{\partial_t \mathcal{F}/\mathcal{F}\} \quad (3.16)$$

$$= -\omega/2 + \text{Re}\{(\partial_z \mathcal{F} - \partial_{z^*} \mathcal{F})/\mathcal{F}\}. \quad (3.17)$$

La dérivation de $\log F(t, \omega)$,

$$\partial_t F/F = -t/2 + (i\partial_z \mathcal{F} - i\partial_{z^*} \mathcal{F})/\mathcal{F} \quad (3.18)$$

$$\partial_\omega F/F = -\omega/2 + (\partial_z \mathcal{F} - \partial_{z^*} \mathcal{F})/\mathcal{F}, \quad (3.19)$$

fait apparaître un couple d'équations

$$\text{Re}\{\partial_t F/F\} = -t/2 - \text{Im}\{(\partial_z \mathcal{F} - \partial_{z^*} \mathcal{F})/\mathcal{F}\} \quad (3.20)$$

$$= \partial_t |F|/|F|, \quad (3.21)$$

$$\text{Re}\{\partial_\omega F/F\} = -\omega/2 + \text{Re}\{(\partial_z \mathcal{F} + \partial_{z^*} \mathcal{F})/\mathcal{F}\} \quad (3.22)$$

$$= \partial_\omega |F|/|F|, \quad (3.23)$$

duquel, lorsqu'il est combiné avec (3.15) et (3.17), on peut déduire :

$$\hat{r} = \nabla \log|F| - 2(\operatorname{Im}\{\partial_{z^*} \log \mathcal{F}\}, \operatorname{Re}\{\partial_{z^*} \log \mathcal{F}\})^t, \quad (3.24)$$

où l'égalité est valide sur $\mathbf{R}^2 \setminus \{(t, \omega) | F(t, \omega) = 0\}$. L'opérateur ∇ désigne le gradient dans le plan $\nabla = (\partial_t, \partial_\omega)^t$.

Ce résultat peut être interprété comme suit : étant donné la factorisation de Bargmann (3.13) de la FCT, le champ de vecteurs de réallocation peut être décomposé en deux termes, un premier qui n'est autre que le gradient d'un potentiel scalaire (nommément $\log|F|$) et un autre qui est une mesure de la non-analyticité de \mathcal{F} . Cette décomposition peut toujours être faite mais elle prend un intérêt pour certaines fenêtres d'observation.

- *Fenêtre gaussienne de variance unité*

En effet, si la fenêtre d'observation h est une gaussienne de variance unité (dont les isocontours sont des cercles dans une représentation de Wigner, fenêtre "ronde" dans le plan temps-fréquence), i.e. si nous considérons la représentation de Bargmann d'un espace d'"états cohérents" [10], alors \mathcal{F} est une fonction entière de z , d'où $\partial_{z^*} \mathcal{F} = 0$ (équations de Cauchy). Nous concluons que²

$$\hat{r}(t, \omega) = \nabla \log|F|, \quad (3.25)$$

ce qui prouve que, dans ce cas, *le champ de vecteurs de réallocation dérive du potentiel scalaire $\log|F|$* .

Ce résultat est **important** pour deux raisons :

1. Une conséquence immédiate en est la description géométrique précise du champ \hat{r} . En effet, l'éq. (3.25) prouve que *tous les vecteurs de réallocation pointent tous dans la direction des maxima du module de la FCT*. Cet aspect est fondamental puisqu'il ouvre la porte à des méthodes particulières qui seront décrites dans la suite de ce chapitre,
2. Si l'on compare les éqs. (3.4) et (3.25), on peut de plus remarquer que, à une constante près, *la phase $\varphi(t, \omega)$ est entièrement déterminée par le module, et vice-versa*. Ce qui veut dire que le spectrogramme correspondant (module carré de la FCT) porte autant d'information que la FCT (à valeur complexe).

Ces deux arguments apportent une justification à l'utilisation de méthodes de traitement d'images (recherche de composantes connexes, de contours, de lignes de partage des eaux) sur le spectrogramme (considéré comme une image) pour l'analyse du signal.

- *Fenêtre arbitraire*

Dans les autres cas (fenêtre arbitraire), un terme supplémentaire

$$g(t, \omega) = -2(\operatorname{Im}\{\partial_z \log \mathcal{F}\}, \operatorname{Re}\{\partial_z \log \mathcal{F}\})^t, \quad (3.26)$$

vient modifier la direction indiquée par la ligne de plus grande pente de $-\log|F|$. On peut donner une expression plus simple de ce terme en écrivant la FCT en fonction des variables complexes z et z^*

$$\mathcal{F}(t, \omega) = \int x(s) h^* \left(s - \frac{z - z^*}{2} \right) e^{-i(z+z^*)s/2} ds e^{(z^2 - z^{*2})/8} e^{zz^*/4}, \quad (3.27)$$

2. Remarquons que l'équivalent de l'éq. (3.25) pour l'image r dans le plan complexe de \hat{r} (cf. éq. (1.26)), définie comme fonction des variables complexes z et z^* prend une forme très compacte, parfois utile pour les calculs : $r = 2i\sqrt{2}\partial_z \log|F|$.

puis en dérivant son logarithme par rapport à z^* . On reconnaît alors le quotient de FCT suivant

$$\partial_{z^*} \log(\mathcal{F}) = -\frac{i}{2} \frac{F^{dh/dt+th}}{F}, \quad (3.28)$$

d'où on peut déduire le vecteur

$$\mathbf{g}(t, \omega) = \left(\operatorname{Re} \left\{ \frac{F^{dh/dt+th}}{F} \right\}, -\operatorname{Im} \left\{ \frac{F^{dh/dt+th}}{F} \right\} \right)^t. \quad (3.29)$$

- *Influence de la fenêtre sur $\mathbf{g}(t, \omega)$.*

Si la fenêtre d'observation est gaussienne de variance λ^2 quelconque, $h(t) = \pi^{-1/4}/\sqrt{\lambda} e^{-t^2/(2\lambda^2)}$ alors la colinéarité entre dh/dt et $th(t)$ autorise la réécriture de l'éq. (3.28) selon les deux formes équivalentes

$$\partial_{z^*} \log(\mathcal{F}) = -\frac{i}{2} \left(1 - \frac{1}{\lambda^2} \right) \frac{F^{th}}{F} = -\frac{i}{2} (\lambda^2 - 1) \frac{F^{dh/dt}}{F}, \quad (3.30)$$

avec lesquelles, compte tenu des éqs. (1.37) et (1.38), on prouve la proportionnalité du terme supplémentaire $\mathbf{g}(t, \omega)$ à une version anamorphosée du vecteur de réallocation

$$\mathbf{g}(t, \omega) = (\lambda - 1/\lambda) \left(\frac{\hat{t} - t}{\lambda}, \lambda(\hat{\omega} - \omega) \right)^t. \quad (3.31)$$

Il est clair dans l'éq. précédente que la prépondérance de $\mathbf{g}(t, \omega)$ dans la décomposition (3.24) s'accroît quand la fenêtre devient large (λ tend vers $+\infty$) ou étroite (λ tend vers 0), et ce par la présence du coefficient $(\lambda - 1/\lambda)$ en facteur. Ce résultat suggère que l'importance du terme supplémentaire \mathbf{g} augmente lorsque la fenêtre d'observation et sa transformée de Fourier deviennent différentes l'une de l'autre ; autrement dit, lorsque l'on analyse le signal de manière "inhomogène" en temps et en fréquence. Pour argumenter cette intuition, remarquons d'abord que la norme du vecteur \mathbf{g} ,

$$\|\mathbf{g}\|^2 = \frac{1}{4} \left| \frac{F^{dh/dt+th}}{F} \right|^2, \quad (3.32)$$

est d'autant plus faible que $F^{dh/dt+th}$ prend de petites valeurs. Soit $h(t)$, une fenêtre de norme unité, $\|h\|^2 = 1$ dont la durée $\Delta t = \|th(t)\|$ et la largeur de bande $\Delta\omega = \|dh/dt\|$ existent et sont égales et $x(t)$, un signal quelconque d'énergie unité. On note $h_\lambda(t) = \sqrt{\lambda}h(\lambda t)$, la fenêtre obtenue par une dilatation de h d'un facteur λ en maintenant l'énergie à 1. On se propose de chercher la valeur de λ qui minimise celle de $|F^{dh_\lambda/dt+th_\lambda}|^2$. L'inégalité de Cauchy-Schwarz nous fournit un majorant

$$|F^{dh_\lambda/dt+th_\lambda}(t, \omega)|^2 \leq \|x\|^2 \|dh_\lambda/dt + th_\lambda\|^2, \quad (3.33)$$

qui après développement

$$\|dh_\lambda/dt + th_\lambda\|^2 = \|dh_\lambda/dt\|^2 + \|th_\lambda\|^2 + 2\langle th_\lambda(t), dh_\lambda/dt \rangle, \quad (3.34)$$

s'évalue par intégration par parties,

$$\left\| \frac{dh_\lambda}{dt} + th_\lambda \right\|^2 = \Delta t_\lambda^2 + \Delta\omega_\lambda^2 - 1. \quad (3.35)$$

Il suffit ensuite d'utiliser les expressions des durée $\Delta t_\lambda = \Delta t/\lambda$ et largeur de bande $\Delta\omega_\lambda = \lambda\Delta\omega$ de la fenêtre dilatée $h_\lambda(t)$ en fonction de la fenêtre de référence pour obtenir la majoration suivante :

$$|F^{dh_\lambda/dt+th_\lambda}(t, \omega)|^2 \leq \Delta t^2/\lambda^2 + \lambda^2\Delta\omega^2 - 1. \quad (3.36)$$

On note $f(\lambda^2)$, la fonction majorante dans le membre de gauche. Sa dérivée $f'(\lambda^2) = -\Delta t^2/\lambda^4 + \Delta\omega^2$ s'annule en $\lambda^2 = \Delta\omega/\Delta t = 1$, minimum de f . On en conclut qu'à forme de fenêtre d'observation donnée, le terme supplémentaire g est de norme minimale lorsqu'on observe le signal sans privilégié ni le temps, ni la fréquence. Inversement lorsque les résolutions en temps et en fréquence deviennent très différentes l'une de l'autre, la norme de ce terme supplémentaire devient de plus en plus grande.

Si $\lambda = 1$, l'éq. (3.36) s'écrit

$$|F^{dh/dt+th}(t, \omega)|^2 \leq \Delta t^2 + \Delta\omega^2 - 1, \quad (3.37)$$

et dépend donc du choix de la fenêtre $h(t)$. Parmi tous ceux qui sont possibles, le majorant atteint une valeur minimale, pour la gaussienne à variance unité, cas unique où il s'annule, ce que nous avons déjà trouvé plus directement.

Les Figs. 3.1 et 3.3 complètent cette argumentation en montrant la décomposition (3.24) sur un chirp linéaire pour différentes fenêtres gaussiennes (pour la première figure) et triangulaire (pour la deuxième). En conclusion, la norme de g dépend de deux critères :

1. le produit durée-bande $\Delta t_h^2 \Delta\omega_h^2$, qui mesure l'encombrement de fenêtre conjointement en temps et en fréquence. Ce produit dépend naturellement du type de fenêtre employée et est minimum pour les gaussiennes,
2. le rapport durée/bande $\Delta t_h^2 / \Delta\omega_h^2$ qui mesure les différences de tailles entre $h(t)$ et $H(\omega)$. Pour un type de fenêtre donné, la norme est minimale si la durée et la bande d'analyse sont égales.

La fenêtre qui réalise simultanément ces deux critères est la fenêtre gaussienne "ronde" pour laquelle la norme de g est uniformément nulle. Remarquons parallèlement que la fonction \mathcal{F} , qui fait apparaître la factorisation de Bargmann, s'éloigne de l'analyticité (au sens où les conditions de Cauchy deviennent de moins en moins vraies) au fur et à mesure que l'un de ses critères s'écarte de la valeur obtenue avec la fenêtre gaussienne de variance unité.

- *En résumé*

Nous venons montrer que le champ de vecteurs de réallocation peut être décrit géométriquement par deux ensembles de courbes : d'une part, les courbes de niveaux d'une fonction, la phase géométrique, liée à la phase de la FCT, et d'autre part, les lignes de plus grande pente d'une fonction liée au module de la FCT, qui alors peut s'interpréter comme un potentiel dont le champ de réallocation dérive (sous certaines conditions, exactement). C'est ce dernier point que nous allons exploiter maintenant.

3.2 Réallocation différentielle

3.2.1 Principe

Le fait qu'un lien existe entre le champ de vecteurs de réallocation et un potentiel scalaire suggère de considérer un système dynamique dont le fonctionnement serait gouverné par ce potentiel.

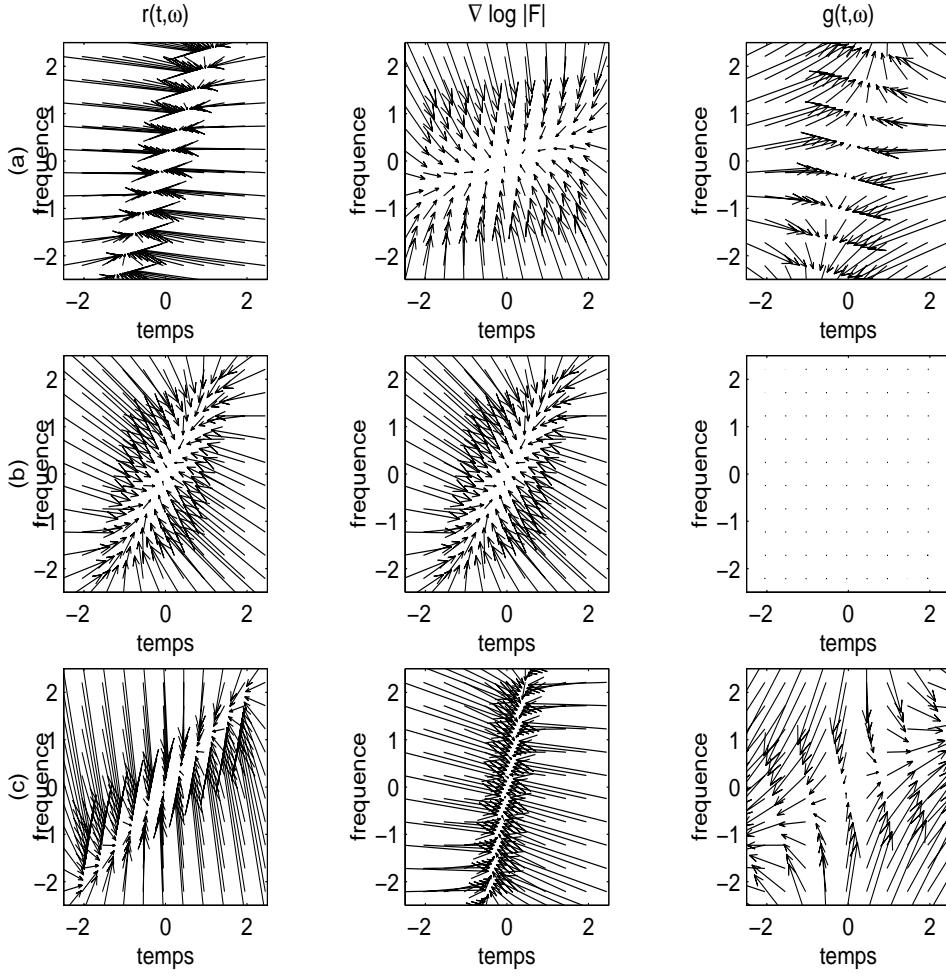


FIG. 3.1 – *Décomposition de Bargmann du champ de vecteurs de réallocation pour une fenêtre gaussienne.* Le champ de vecteurs de réallocation peut se décomposer (cf. éq (3.24)) en la somme du gradient du potentiel $\log |F(t, \omega)|$ et d'un terme supplémentaire $g(t, \omega)$. Nous avons montré que $g(t, \omega)$ est prépondérant lorsque la fenêtre et sa transformée de Fourier sont de tailles (resp. durée et largeur de bande) différentes l'une de l'autre. C'est ce que montre ce tableau de figures où, à chaque ligne, on effectue la décomposition (3.24) pour la fenêtre correspondante en Fig. 3.2. La ligne centrale correspond au cas où on observe le signal (ici, un chirp linéaire d'enveloppe gaussienne, $s(t) = e^{-(1-i)t^2/2}$) identiquement en temps et en fréquence (fenêtre gaussienne à variance unité). Le terme $g(t, \omega)$ s'annule alors uniformément. On dissymétrise cette situation particulière en rétrécissant (première ligne) ou en élargissant (troisième ligne) la fenêtre. Le terme $g(t, \omega)$ n'est alors plus négligeable devant $\nabla \log |F(t, \omega)|$.

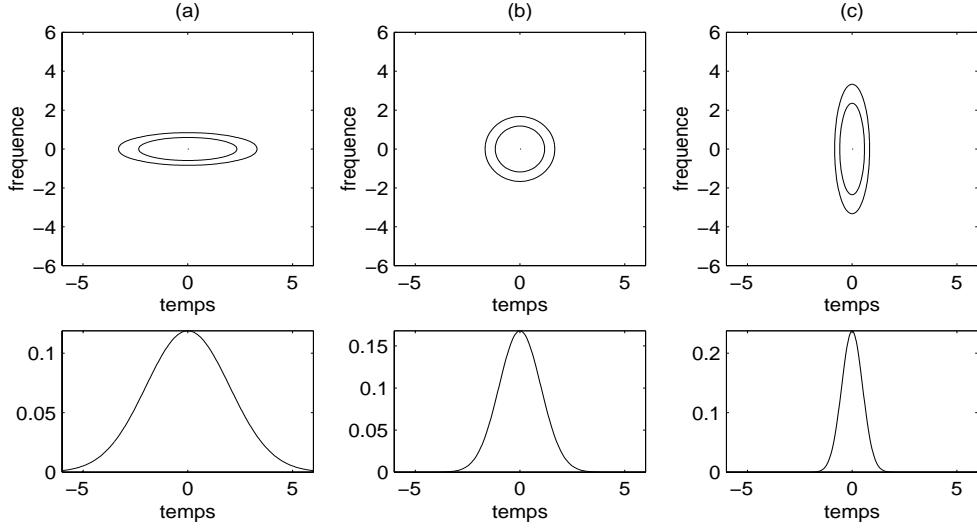


FIG. 3.2 – **Fenêtres utilisées en Fig. 3.1.** Cette figure présente les trois fenêtres gaussiennes utilisées pour la Fig. 3.1. Leur distribution de Wigner-Ville sont indiquées sur la ligne du haut par deux courbes de niveau. La fenêtre en (b) est une gaussienne de variance unité, seule fonction superposable à sa transformée de Fourier. Sa distribution de Wigner-Ville est à symétrie circulaire dans le plan temps-fréquence, d'où l'appellation de fenêtre “ronde”.

En partant de cette perspective, on propose d'introduire une méthode de *réallocation différentielle* [24] qui généralise la version originale. Dans la méthode originale de la réallocation, chaque valeur du spectrogramme est déplacée d'une quantité finie, éventuellement grande, sans tenir compte des régions du plan temps-fréquence qu'elle traverse. L'idée est maintenant d'assimiler chaque point du plan temps-fréquence à une particule dont le mouvement est complètement déterminé par sa position initiale et par sa vitesse que l'on fixe en tout point égale au vecteur de réallocation correspondant. Chaque valeur du spectrogramme est alors déplacée par sauts infinitésimaux et de façon continue jusqu'à convergence vers un équilibre.

Plus formellement, considérons le champ de vecteurs de réallocation comme un champ de vitesse qui contrôle le mouvement de chaque contribution temps-fréquence $F(t, \omega)$, assimilée à une particule de position de départ, le point (t, ω) . Nous obtenons les équations du mouvement suivantes,

$$\begin{cases} t(0) &= t \\ \omega(0) &= \omega \\ dt(s)/ds &= \hat{t}(t(s), \omega(s)) - t(s) \\ d\omega(s)/ds &= \hat{\omega}(t(s), \omega(s)) - \omega(s), \end{cases} \quad (3.38)$$

qui définissent le processus que l'on appellera *réallocation différentielle*.

Dans le cas fenêtre gaussienne à variance unité, ces équations prennent une forme particulière liée au fait que le champ de vecteurs de réallocation dérive d'un potentiel scalaire (cf. éq. (3.25))

$$(t(0), \omega(0)) = (t, \omega) \quad \frac{d(t, \omega)^t}{ds} = \nabla \log |F|. \quad (3.39)$$

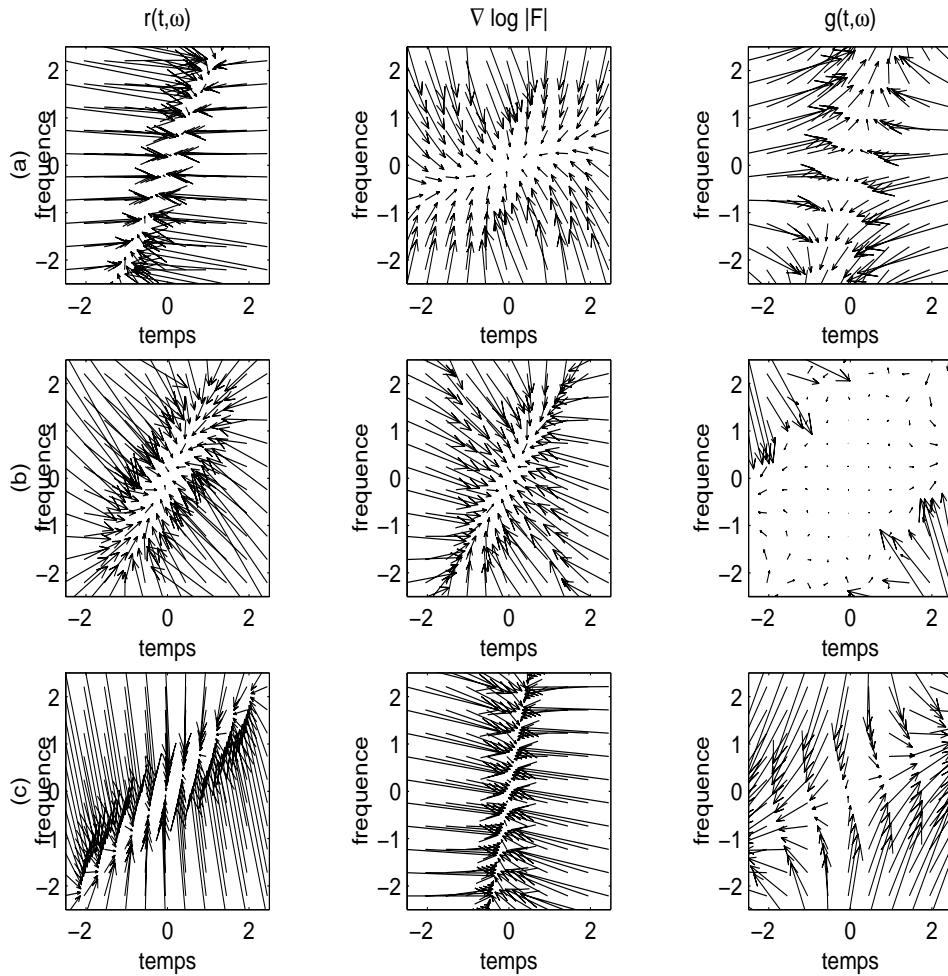


FIG. 3.3 – **Décomposition de Bargmann du champ de vecteurs de réallocation pour une fenêtre triangulaire.** Cette figure est identique à la Fig. 3.1, exception faite que la fenêtre est cette fois triangulaire (cf. 3.4). La ligne centrale est associée à la fenêtre triangulaire dont la durée est égale à la largeur de bande. Cette fenêtre n'est pas de produit durée-bande minimum (comme c'était le cas pour la gaussienne). Le vecteur $g(t, \omega)$ n'a plus de raison de s'annuler. Néanmoins au sein de cette famille de fenêtre, elle conduit à un terme $g(t, \omega)$ de norme L^2 minimale.

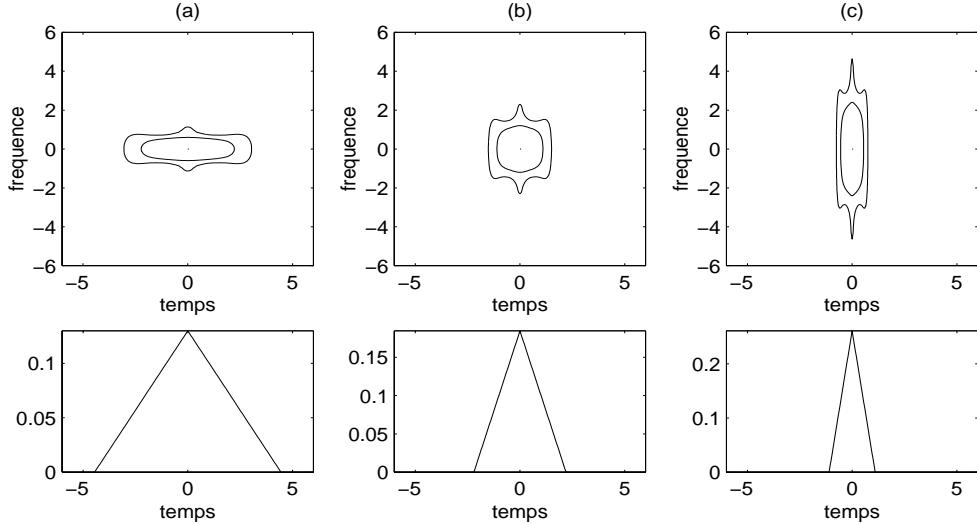


FIG. 3.4 – **Fenêtres utilisées en Fig. 3.3.** Cette figure présente les trois fenêtres triangulaires utilisées pour la Fig. 3.3 avec, sur la ligne du haut, leur distribution de Wigner-Ville indiquée à l'aide de deux courbes de niveau.

Le système dynamique ainsi décrit est un système purement dissipatif dans lequel chaque particule converge vers les maxima de $\log|F|$. Nous reviendrons plus en détail sur ce point.

Si la fenêtre est quelconque, la présence d'un terme supplémentaire dans la décomposition (3.24) du champ de vecteurs de réallocation ne nous permet pas de conclure directement sur la nature de la dynamique.

La réallocation différentielle peut être vue comme un post-traitement des “images” temps-fréquence à base d’EDP. Dans cette famille de traitements, il existe déjà un algorithme similaire. En effet, les méthodes “arête et squelette”³ qui extraient les lignes de maxima dans les transformations de Gabor (ou les décompositions en ondelettes) en explorant le plan temps-fréquence (ou temps-échelle) à l’aide d’une recherche itérative le long d’une direction nécessairement parallèle à l’axe des temps ou des fréquences. Dans le cas de la fenêtre gaussienne à variance unité, la réallocation différentielle peut s’interpréter comme une généralisation d’un tel algorithme dans la mesure où la recherche s’effectue sans contrainte de direction. En effet, nous avons montré ici que la réallocation différentielle utilisait alors le plus court chemin pour aller chercher l’“arête”.

3.2.2 Points fixes par la réallocation

Le rôle des *points fixes* c’est-à-dire les points qui se réallouent en eux-mêmes, est crucial. On peut les classer en trois catégories :

1. *point fixe stable* : toutes les particules comprises dans un certain voisinage de ce point convergent vers ce point,
2. *point fixe col* : on ne peut converger vers ce point qu’en empruntant certaines trajectoires,
3. *point fixe instable*: aucune particule ne vient finir sa trajectoire en ce point, exception faite de celle dont la position initiale est ce point exactement.

3. Les méthodes “arête et squelette” [33, 53] sont présentées en Sect. 1.4.

Dans le cas de la fenêtre gaussienne à variance unité, les particules temps-fréquence se meuvent sur un potentiel. Le fait que ce potentiel soit le logarithme du spectrogramme (et donc pas n'importe quelle fonction 2D) va contraindre la nature des points fixes et leur position. C'est ce qui va nous intéresser maintenant. Nous allons préciser également ce qu'il advient au lien qui unit les points fixes aux extrema du potentiel $\log|F|$ lorsque la fenêtre est gaussienne, mais n'est plus “ronde” (à variance quelconque).

Points fixes et maxima de $\log|F|$

Un point fixe par la réallocation est un point où le vecteur de réallocation est nul, soit

$$\hat{t}(t, \omega) - t = \operatorname{Re} \left\{ \frac{F^{th}}{F} \right\} = 0 \quad \hat{\omega}(t, \omega) - \omega = -\operatorname{Im} \left\{ \frac{F^{dh/dt}}{F} \right\} = 0. \quad (3.40)$$

Si la fenêtre est gaussienne, $h_\lambda(t) = \pi^{-1/4}/\sqrt{\lambda} e^{-t^2/(2\lambda^2)}$, il devient simple de caractériser ces points. Les fenêtres $th_\lambda(t)$ et dh_λ/dt étant proportionnelles, l'éq. (3.40) implique

$$\operatorname{Re} \left\{ \frac{F^{dh_\lambda/dt}}{F} \right\} = \operatorname{Im} \left\{ \frac{F^{th_\lambda}}{F} \right\} = 0. \quad (3.41)$$

Les grandeurs complexes $F^{th_\lambda(t)}(t, \omega)$ et $F^{dh_\lambda/dt}(t, \omega)$ sont donc nulles au point fixe. En utilisant le fait que $\partial \log|F| = \operatorname{Re}\{\partial F/F\}$ (la dérivée partielle opérant sur le temps ou la fréquence indifféremment) et les relations suivantes

$$\partial_t F = -F^{dh_\lambda/dt} + i\omega F/2 \quad \partial_\omega F = -iF^{th_\lambda} - itF/2, \quad (3.42)$$

qui s'obtiennent simplement par dérivation de l'éq. (1.3), le gradient de $\log|F|$ s'exprime avec des quotients de FCT

$$\partial_t \log|F| = -\operatorname{Re} \left\{ \frac{F^{dh_\lambda/dt}}{F} \right\} \quad (3.43)$$

$$\partial_\omega \log|F| = \operatorname{Im} \left\{ \frac{F^{th_\lambda}}{F} \right\}, \quad (3.44)$$

nuls aux points fixes selon l'éq. (3.40). Les points fixes sont donc également les points où s'annule le gradient de $\log|F|$. En ces points, le terme $\mathbf{g}(t, \omega)$ dans la décomposition (3.24) s'annule à son tour, nécessairement.

Il y a donc identité entre le lieu des points fixes par la réallocation pour les FCT de Gabor et celui des extrema du module de la FCT (ou du spectrogramme).

Géométrie des maxima de $\log|F|$

Cette identité est très importante puisqu'il existe des résultats forts concernant la structure des courbes de maxima du module de la FCT, qui s'appliquent donc désormais aux ensembles de points fixes par la réallocation. En effet, si $h(t) = h_1(t)$ est gaussienne “ronde” (états cohérents), on peut montrer [91] que les maxima de la FCT peuvent s'aligner uniquement selon deux types de courbes

1. les droites qui correspondent aux modulations linéaires de fréquence d'enveloppe constante,

2. les cercles qui correspondent aux fonctions d’Hermite,

sinon ils forment un ensemble de points isolés. Cela limite donc considérablement les agencements possibles de points fixes par la réallocation.

Il est possible d’étendre ces résultats aux FCT calculées avec une fenêtre gaussienne de variance quelconque en invoquant la correspondance entre les FCT

$$F_x^{h\lambda}(t, \omega) = F_{x_1/\lambda}^h(\lambda t, \omega/\lambda), \quad (3.45)$$

où $x_{1/\lambda}(t) = x(t/\lambda)/\sqrt{\lambda}$, qui se démontre par un simple changement de variable (cf. Sect. 1.2.4). Les résultats de [91] appliqués sur le membre de droite de l’éq. (3.45) où la fenêtre d’analyse est adéquate, sont alors transférés sur le membre de gauche. La famille de courbes d’extrema autorisées pour $F_x^{h\lambda}$ inclut donc des droites, des ellipses d’excentricité λ^2 , ou des points isolés.

Stabilité des points fixes, attracteurs

Les points fixes stables sont centraux pour la dynamique engendrée par la réallocation différentielle puisqu’ils constituent les points d’équilibre vers où convergent toutes les particules temps-fréquence : nous les qualifierons donc d’*attracteurs*. On peut se poser la question de savoir si le terme supplémentaire ne vient pas entraver cette convergence. La réponse peut être donnée par le calcul du produit scalaire entre le vecteur de réallocation et celui de la direction de la plus grande pente

$$\hat{r} \cdot \nabla \log|F| = \left(\operatorname{Re} \left\{ \frac{F^{th}}{F} \right\}, -\operatorname{Im} \left\{ \frac{F^{dh}}{F} \right\} \right)^t \cdot \left(-\operatorname{Re} \left\{ \frac{F^{dh}}{F} \right\}, \operatorname{Im} \left\{ \frac{F^{th}}{F} \right\} \right)^t, \quad (3.46)$$

qui, dans le cas à fenêtre gaussienne qui nous intéresse, se réduit alors à

$$\hat{r} \cdot \nabla \log|F| = \frac{1}{\lambda^2} \left| F^{th} \right|^2. \quad (3.47)$$

Ce produit scalaire est donc toujours positif. L’angle entre \hat{r} et $\nabla \log|F|$ reste compris entre $-\pi/2$ et $\pi/2$ ce qui garantit que g ne perturbe pas la course des particules vers les points fixes.

La réallocation différentielle va sélectionner parmi tous les extrema, seulement ceux où l’équilibre est stable. Pour prouver la stabilité d’un point fixe, il suffit de trouver le signe de la partie réelle des valeurs propres de la matrice dérivée du champ de réallocation, $D(\hat{r})$. Dans le cas gaussien à variance unité, cela revient à étudier le spectre de la matrice

$$D(\hat{r})(t, \omega) = \begin{pmatrix} \partial_t^2 \log|F| & \partial_{t\omega}^2 \log|F| \\ \partial_{\omega t}^2 \log|F| & \partial_\omega^2 \log|F| \end{pmatrix}, \quad (3.48)$$

qui, au point fixe où $\partial_t|F| = \partial_\omega|F| = 0$, est aussi égale

$$D(\hat{r})(t, \omega) = \frac{1}{|F|^2} \begin{pmatrix} \partial_t^2|F| & \partial_{t\omega}^2|F| \\ \partial_{\omega t}^2|F| & \partial_\omega^2|F| \end{pmatrix}. \quad (3.49)$$

La stabilité est donc directement liée à la courbure locale du module de la FCT. Nous ne pouvons conclure sur la stabilité des points fixes pour tous les signaux. En revanche, il est montré dans [91] que si le point en question appartient à une droite ou à un cercle de points fixes, il est nécessairement un maximum local de $|F|$ (la concavité de $\log|F|$ y est tournée vers le bas), ce qui prouve la stabilité de cette position d’équilibre. En ce qui concerne les extrema isolés, il est possible que certains d’entre eux correspondent à des points *col* ou à des minima locaux en lesquels l’équilibre est instable.

- **En résumé**

En s'appuyant sur les résultats de la section précédente et sur une analogie aux systèmes physiques, nous avons introduit une nouvelle méthode de réallocation, la réallocation différentielle, généralisation de la version classique. Nous avons étudié quelques unes de ses propriétés. En particulier, nous montrons l'importance des points fixes par la réallocation qui s'interprètent comme des attracteurs vers lesquels convergent asymptotiquement toutes les contributions temps-fréquence ainsi réallouées. Nous établissons, dans le cas à fenêtre gaussienne, la correspondance entre ces points fixes et les extrema locaux du spectrogramme.

3.3 Partition

3.3.1 Différentes approches

De nombreux problèmes de traitement du signal comme le débruitage, la classification de signaux, etc., peuvent se reformuler dans le contexte de l'analyse temps-fréquence par la recherche d'un moyen de partager le plan temps-fréquence en zones marquant chacune des composantes d'un signal et permettant leur extraction de manière indépendante. Dessiner une carte temps-fréquence, c'est essentiellement décomposer le signal en composantes modulées en fréquence et/ou en amplitude.

Cette question n'est pas nouvelle et vouloir y donner une réponse définitive est un défi perdu d'avance. En effet, pour apporter une solution à ce problème, il faut donner une définition à ce qu'est une composante d'un signal. Le concept de composante, même si on peut en avoir une intuition assez précise, est complètement ambigu. Si on le maintient au rang d'idée philosophique, aucune preuve mathématique n'est possible. Il faut alors le formaliser ; il perd alors, inévitablement, de son universalité. On dresse une liste de quelques contributions relatives à ce problème. Le propos n'est pas l'exhaustivité mais plutôt d'insister sur les diverses conceptions que l'on peut avoir de la nature d'une composante.

Cohérence de phase sous l'enveloppe

Dans cette approche, on applique au signal le modèle explicite suivant (déjà évoqué au chapitre 1)

$$x(t) = \sum_{n=1}^N A_n(t) e^{i\theta_n(t)}, \quad (3.50)$$

où l'amplitude $A_n(t)$ présente de faibles variations par rapport à celles de la phase $\theta_n(t)$. Les différences entre les méthodes attachées à cette approche sont dans la manière de traduire cette dernière phrase. Une composante du signal est donc clairement décrite comme une modulation de fréquence oscillant de façon régulière sous une enveloppe aux variations lentes.

Cette idée a été exploitée par exemple en traitement de la parole dans [76] sous le nom de “modèle sinusoïdal” en imposant à l'amplitude d'être linéaire par morceaux et à la dérivée de la phase d'être cubique par morceaux.

Pour les méthodes “arête et squelette” [33, 53] qui l'utilise également, chaque composante est dite “asymptotique”, ce qui se traduit par la contrainte suivante sur la phase et l'amplitude

$$\left| \frac{1}{A_n} \frac{dA_n}{dt} \right| \ll \left| \frac{d\theta_n}{dt} \right|. \quad (3.51)$$

Cette hypothèse permet par la suite d'invoquer le principe de phase stationnaire (pour le calcul des transformées temps-fréquence ou temps-échelle, cf. Sect. 1.4).

Cette approche atteint ses limites lorsque l'on essaie d'analyser une impulsion, ou plus généralement lorsque l'enveloppe varie très rapidement. Il faut alors faire appel à l'“asymptotisme” en fréquence [33, 53], version duale du précédent, qui contraint les variations relatives de l'amplitude et de la phase de la transformée de Fourier du signal de la même manière qu'en éq. (3.51). Pour les signaux qui mélagent les deux types de comportement, le problème est alors de savoir quand passer d'un modèle à l'autre, dans l'hypothèse que ce passage soit possible.

Encombrement dans le plan temps-fréquence

Aucun modèle n'est, cette fois, appliqué au signal. On s'appuie sur le fait que, dans le plan temps-fréquence, les signaux ne peuvent pas être arbitrairement concentrés simultanément en temps et en fréquence. C'est ce que traduit par exemple l'inégalité bien connue d'Heisenberg-Gabor. L'idée est de faire une mesure de l'encombrement du signal dans le plan temps-fréquence puis de comparer cette mesure à celle obtenue avec un signal de référence (celui d'encombrement minimum). Le nombre de composantes présentes dans le signal se déduit dans le principe, de la manière suivante : si l'encombrement mesuré est, par exemple, le double de l'encombrement de référence, alors on conclut qu'il y a deux composantes dans le signal. La difficulté principale réside dans la recherche d'une bonne mesure de l'encombrement. Les approches les plus prometteuses sont basées sur l'évaluation de la quantité d'information portée par une distribution d'énergie temps-fréquence de la classe de Cohen considérée comme une densité de probabilité (à deux dimensions).

Nous allons nous intéresser, en particulier, à deux mesures d'information : à l'information de Rényi d'ordre 3 sur la distribution de Wigner-Ville⁴ (cf. éq. (1.2)) [7, 6]

$$I_{R3}^W(x) = -\frac{1}{2} \log_2 \iint W_x^3(t, \omega) \frac{dt d\omega}{2\pi}, \quad (3.52)$$

et également à l'information de Shannon portée par le spectrogramme [98]

$$I_{SH}^S(x) = - \iint S_x(t, \omega) \log_2 S_x(t, \omega) \frac{dt d\omega}{2\pi}, \quad (3.53)$$

où dans les deux cas $x(t)$ est un signal d'énergie unité.

Ces fonctionnelles peuvent s'interpréter comme des mesures de concentration puisque, par analogie à la théorie des probabilités, une variable aléatoire présentant une densité de probabilité très concentrée (“piquée”) est porteuse de peu d'information. On peut montrer [6] par ailleurs qu'elles possèdent la propriété de compter les composantes. En effet, si un signal est la somme de $x(t)$ et de $y(t) = x(t - \Delta t) \exp(i\Delta\omega t)$, son translaté en temps et en fréquence, alors pour Δt et $\Delta\omega$ suffisamment grands [6]

$$I_{R3}^W(x + y) = I_{R3}^W(x) + 1 \quad ; \quad I_{SH}^S(x + y) = I_{SH}^S(x) + 1. \quad (3.54)$$

Il existe néanmoins des différences entre ces solutions qui peuvent faire préférer l'une plutôt que l'autre. En effet, la première $I_{R3}^W(x)$ hérite de la distribution de Wigner-Ville de quelques propriétés d'invariance (en particulier, par les changements d'échelle et par les modulations et convolutions par les chirps linéaires) que ne possède pas la deuxième $I_{SH}^S(x)$.

4. Il est important de noter que cette première mesure n'est pas toujours définie [6].

Les quantités d'information (3.52) et (3.53) sont exprimées en nombre de bits que l'on peut convertir en nombre de composantes (ou dimension de Rényi [6]) respectivement par :

$$N_{R3}^W(x) = 2^{I_{R3}^W(x) - I_{R3}^W(h)}, \quad (3.55)$$

où $h(t)$ est le signal d'information minimale (à savoir n'importe quelle modulation linéaire de fréquence, d'enveloppe gaussienne) et par :

$$N_{SH}^S(x) = 2^{I_{SH}^S(x) - I_{SH}^S(h)}, \quad (3.56)$$

où $h(t)$ est la fenêtre d'observation (et, dans ce cas, le seul signal d'information minimum).

Remarquons cependant que ces méthodes nous donne au final le nombre de composantes présentes dans le signal mais aucune information n'est donnée sur leur position, ni même de leur structure.

3.3.2 Une proposition basée sur la réallocation différentielle

Notre contribution [21] à ce problème se distingue des autres dans la mesure où elle définit la notion de composante indirectement par un outil de représentation du signal, le champ de vecteurs de réallocation du spectrogramme, plutôt que par des arguments physiques comme précédemment.

Fort des résultats obtenus en Sect. 3.2, il devient naturel de décrire le signal dans le plan temps-fréquence en termes d'attracteurs, de bassins d'attraction et de lignes de partage des eaux. Une variété de caractérisation des signaux peut être déduite d'un tel paramétrage. Pour préciser la terminologie, on appelle le *bassin d'attraction* le plus grand voisinage de points à partir desquels il y a convergence vers un point fixe stable. Un bord séparant deux bassins ne peut être affecté à l'un ou l'autre de ces bassins. Les bords regroupent tous les endroits depuis lesquels on converge vers un point fixe col, ou vers l'infini. Notons que l'ensemble de ces lignes, habituellement appelées *lignes de partage des eaux*, est de mesure nulle. La réunion de tous les bassins d'attraction recompose le plan temps-fréquence dans son entier. Le partage du plan temps-fréquence en zones distinctes, chacune associée à une composante, peut donc être obtenu par ce biais. Conformément à l'intuition, une composante se trouve ainsi décrite par un *centre* temps-fréquence (l'attracteur) et un *domaine* essentiel d'existence dans le plan (le bassin d'attraction).

Si on utilise une fenêtre gaussienne, les composantes que l'on détecte sont des ensembles de points connexes (on peut tracer une ligne continue entre chacun de ces points et l'attracteur) contenant un maximum local du spectrogramme (ou plusieurs à condition qui s'alignent de manière dense, le long d'une droite, ou d'une ellipse). À l'intérieur d'une composante, on peut imaginer le spectrogramme comme un "dôme" d'énergie. En cela, cette méthode procède de la même philosophie que celle exposée dans [84].

3.3.3 Algorithme pour la partition

La mise en œuvre du calcul de la partition temps-fréquence passe par plusieurs étapes.

La réallocation différentielle en pratique

On calcule les *points de réallocation asymptotiques*, i.e., les coordonnées de la position de chaque particule en fin de trajectoire en intégrant le système d'équations différentielles (éq. (3.38)) liant la position de la particule à son champ de vitesse. Chaque trajectoire est évaluée numériquement par

une méthode Runge-Kutta d'ordre 2 à pas k fixe dont la précision est en $O(k^3)$. Si t_n et ω_n représentent respectivement les coordonnées temporelles et fréquentielles de la particule au cours de son mouvement, on a

$$(\bar{t}_n, \bar{\omega}_n)^t = (t_n, \omega_n)^t + k \hat{\mathbf{r}}(t_n, \omega_n) \quad (3.57)$$

$$(t_{n+1}, \omega_{n+1})^t = (t_n, \omega_n)^t + k/2 (\hat{\mathbf{r}}(t_n, \omega_n) + \hat{\mathbf{r}}(\bar{t}_n, \bar{\omega}_n)), \quad (3.58)$$

l'arrêt de l'itération étant déterminé par un critère spécifique. Il ne s'agit pas en effet d'arrêter la progression de l'algorithme alors que la particule est encore loin de sa position asymptotique. Réciproquement, il est inutile d'itérer la méthode de Runge-Kutta si la convergence est déjà effective. Le critère d'arrêt que nous proposons nous garantit que l'approximation que nous obtenons est dans un carré d'arête d (donnée⁵) centré sur la valeur asymptotique vraie. Ce critère est obtenu en faisant l'hypothèse d'une approximation quadratique locale du potentiel (ce qui revient à approcher localement le signal par une modulation linéaire de fréquence d'enveloppe gaussienne). Si le potentiel est quadratique, cela implique que son gradient est linéaire. Donc les équations différentielles (3.38) de la réallocation différentielle admettent au voisinage de la position asymptotique, $(t_\infty, \omega_\infty)$ une solution exponentielle à savoir

$$t_n = t_\infty + (t_0 - t_\infty) \exp(-n/T) \quad (3.59)$$

$$\omega_n = \omega_\infty + (\omega_0 - \omega_\infty) \exp(-n/\Omega), \quad (3.60)$$

où T et Ω sont les temps (algorithmiques) de relaxation qui règlent la vitesse de convergence vers la valeur asymptotique selon chaque axe.

On veut un critère simple qui nous indique lorsque

$$|t_n - t_\infty| < d \quad |\omega_n - \omega_\infty| < d. \quad (3.61)$$

Concentrons-nous sur la variable temps, et observons sa variation locale

$$|t_n - t_{n-1}| = |(t_0 - t_\infty) \exp(-n/T)| |1 - \exp(1/T)|. \quad (3.62)$$

En comparant à (3.59), on voit apparaître la distance à la valeur asymptotique

$$|t_n - t_\infty| = \frac{|t_n - t_{n-1}|}{|1 - \exp(1/T)|}. \quad (3.63)$$

On obtient une estimation d'un paramètre T généralement inconnu, en notant que le rapport de deux variations successives est constant et égal à

$$\frac{|t_{n-1} - t_{n-2}|}{|t_n - t_{n-1}|} = \exp(1/T), \quad (3.64)$$

résultat que l'on reporte dans éq. (3.63), pour obtenir une forme exploitable de la première inéquation dans (3.61)

$$\frac{|t_n - t_{n-1}|}{\left|1 - |t_{n-1} - t_{n-2}|/|t_n - t_{n-1}|\right|} < d. \quad (3.65)$$

5. Pour garantir une convergence acceptable, on fixe d à une valeur négligeable par rapport au pas de la grille du plan temps-fréquence. Pour un plan échantillonné au pas unité en temps et en fréquence, un bon choix est $d = 10^{-2}$. Choisir une valeur plus petite augmente, bien sûr, le temps de calcul.

Ces calculs sont directement transposables à la variable fréquence. Le critère complet prend la forme de deux conditions à respecter simultanément :

$$\begin{cases} |t_n - t_{n-1}|^2 > d \left| |t_n - t_{n-1}| - |t_{n-1} - t_{n-2}| \right| \\ |\omega_n - \omega_{n-1}|^2 > d \left| |\omega_n - \omega_{n-1}| - |\omega_{n-1} - \omega_{n-2}| \right|. \end{cases} \quad (3.66)$$

Notons qu'il nécessite seulement la connaissance de trois points de calculs successifs.

Pré-traitement

Il s'agit d'effectuer le tri parmi les opérateurs de réallocation asymptotiques de manière à repérer les positions exactes des attracteurs. Le nombre de points étant à ce stade généralement très grand, on effectue un premier passage que l'on appelle *pré-traitement*. Il consiste à regrouper tous les points situés à une distance inférieure à $\sqrt{2}d$, i.e., la diagonale d'un carré de côté d . Tous les points qui convergent vers la même valeur asymptotique sont résumés en un seul (que l'on fixe égal à leur barycentre). Notons que cette étape concerne en premier lieu aux signaux modulés en amplitude pour lesquels les points de réallocation asymptotiques forment un ensemble de points isolés.

Classification ascendante hiérarchique

Il nous faut maintenant associer les points qui appartiennent à la même ligne de modulation de fréquence. Un algorithme plus complexe (d'où l'importance de la réduction du nombre de points réalisée lors de l'étape précédente), la Classification Ascendante Hiérarchique [11] (CAH), va nous le permettre. La CAH des points de réallocation prétraités nous fournit une structure arborescente (le dendrogramme) où sont ordonnés les points suivant la distance qui les sépare. Il suffit alors d'appliquer un seuil (égal à $d + D$, la distance maximale entre deux attracteurs appartenant à la même ligne de modulation de fréquence dans le pire des cas, où D est le pas de la grille de discréétisation du plan temps-fréquence) dans le dendrogramme pour mettre ensemble tous les points qui s'alignent sur une même ligne⁶ du plan. Chaque classe ainsi obtenue est alors affiliée à une même composante.

Carte temps-fréquence.

La dernière étape consiste à affecter à tous les antécédents des points de réallocation asymptotiques le numéro de la composante correspondante.

3.3.4 Quelques exemples

L'algorithme que nous proposons peut servir de point de départ à de nombreuses applications en traitement du signal. La classification de signaux, leur reconnaissance, l'estimation de paramètres peuvent s'appuyer sur la partition temps-fréquence. Voici quelques exemples d'illustration aux objectifs multiples : les premiers sont dédiés à la comparaison de cet algorithme avec d'autres, le dernier montre comment il peut s'appliquer au débruitage de signaux.

6. Nous savons, qu'en théorie, les ensembles denses d'attracteurs ne peuvent être que des droites ou des ellipses. La CAH devrait donc être destinée à traiter ces deux cas. Mais, en pratique, les contributions ne convergent pas complètement vers leur position asymptotique (erreur d'arrondi, convergence trop lente). Par conséquent, il arrive parfois que l'on se retrouve avec un ensemble de contributions amassées le long d'une ligne (pas nécessairement une droite ou un cercle) du plan alors qu'elles auraient dû toutes être regroupées en un seul point. Lors de cette étape, on rattrape les erreurs commises en affectant à toutes ces contributions la même étiquette.

Partition de Voronoï du plan temps-fréquence

Le but de ce premier exemple, plutôt de l'ordre de l'exercice, est de donner un éclairage sur le résultat de la partition. On considère le signal $x(t)$, somme de N atomes gaussiens de variance unité $x_n(t) = \pi^{-1/4} \exp((t - t_n)^2/2 + i\omega_n t)$ aux centres temps-fréquence (t_n, ω_n) suffisamment éloignés les uns des autres (nous donnerons plus loin une signification à ces derniers mots)

$$x(t) = \sum_{n=1}^N x_n(t). \quad (3.67)$$

(Notons que la référence de phase de chacun de ces atomes se situe en leur centre t_n .)

La FCT d'un tel signal se déduit par linéarité, $F(t, \omega) = \sum_{n=1}^N F_n(t, \omega)$. Si $h(t)$, la fenêtre d'observation, est gaussienne, son champ de vecteurs de réallocation \hat{r} s'obtient par une généralisation pour N signaux de l'éq. (2.22). Plus précisément, l'image complexe de r (champ de vecteurs normalisé défini en éq. (1.39)) s'écrit

$$r = \sum_{n=1}^N \frac{F_n}{F} r_n. \quad (3.68)$$

On se place en (t, ω) au voisinage de (t_j, ω_j) et (t_k, ω_k) de tel sorte que pour $n \neq j$ et $n \neq k$, $|F_n(t, \omega)| \ll \|x\|^2$. En ce point, tout se passe comme si le champ de réallocation résultait de l'interaction des atomes j et k

$$r \approx \frac{F_j}{F_j + F_k} r_j + \frac{F_k}{F_j + F_k} r_k. \quad (3.69)$$

Si la fenêtre $h(t)$ est "ronde" (de variance unité), le calcul explicite de r pour un tel signal peut être ramené, par une translation suivie d'une rotation dans le plan temps-fréquence, à celui plus simple de la somme $y(t) = y_1(t) + y_2(t)$ de deux atomes $y_1(t) = \pi^{-1/4} \exp(-(t + t_0)^2/2)$ et $y_2(t) = \pi^{-1/4} \exp(-(t - t_0)^2/2)$ centrés sur l'axe des temps, de part et d'autre de l'origine. Remarquons que cette équivalence est possible uniquement parce que la distribution de Wigner Ville des atomes et celle de la fenêtre sont invariantes par toute rotation dans le plan temps-fréquence.

Le calcul de r suit la même trame que pour le signal à deux impulsions et commence à l'éq. (2.22)

$$r = \frac{r_1}{1 + F_2/F_1} + \frac{r_2}{1 + F_1/F_2}, \quad (3.70)$$

où le quotient des FCT s'obtient avec (2.10)

$$F_1/F_2(t, \omega) = \exp(-t_0(t - i\omega)). \quad (3.71)$$

On en déduit que le champ de réallocation de y s'écrit

$$r = \frac{1}{1 + \exp(t_0(t - i\omega))} \frac{(\omega + i(t + t_0))}{-\sqrt{2}} + \frac{1}{1 + \exp(-t_0(t - i\omega))} \frac{(\omega + i(t - t_0))}{-\sqrt{2}}, \quad (3.72)$$

ou bien, réduit sous une forme plus compacte,

$$\hat{r} = \left(-\frac{t}{2} + \frac{t_0}{2} \frac{\sinh(t_0 t)}{\cosh(t_0 t) + \cos(t_0 \omega)}, -\frac{\omega}{2} - \frac{t_0}{2} \frac{\sin(t_0 \omega)}{\cosh(t_0 t) + \cos(t_0 \omega)} \right)^t. \quad (3.73)$$

Les points fixes par la réallocation sont les points tels que

$$\omega = -t_0 \frac{\sin(t_0\omega)}{\cosh(t_0t) + \cos(t_0\omega)}, \quad t = t_0 \frac{\sinh(t_0t)}{\cosh(t_0t) + \cos(t_0\omega)}. \quad (3.74)$$

Nous allons séparer la recherche systématique de ces points en deux étapes : d'abord, le long des axes du plan temps-fréquence, puis dans le reste du plan.

- *Axe des temps, $\omega = 0$.*

La première équation de (3.74) est vérifiée sur cet axe. La deuxième équation qui devient

$$t = t_0 \frac{\sinh(t_0t)}{\cosh(t_0t) + 1}, \quad (3.75)$$

possède un nombre différent de solutions selon la valeur de t_0 . Si $t_0 \leq \sqrt{2}$, l'éq. (3.75) admet une seule solution en $t = 0$ (la fonction $\sinh(t_0t)/(\cosh(t_0t) + 1)$ croît plus lentement que t au voisinage de 0). Les deux logons sont trop proches pour pouvoir être distingués avec la fenêtre h ; on détectera alors une seule composante. Dans le cas contraire $t_0 > \sqrt{2}$ (condition de Rayleigh), l'éq. (3.74) admet trois solutions : 0 et un couple de solutions non algébriques de signe opposé t_1 (dont la valeur proche de t_0 peut être évaluée numériquement) et $-t_1$.

Ces solutions sont de natures différentes, ce que l'on peut vérifier en calculant la dérivée du champ de vecteurs de réallocation

$$\partial_t(\hat{t} - t) = -\frac{1}{2} \left(1 - t_0^2 \frac{1 + \cos(t_0\omega) \cosh(t_0t)}{(\cosh(t_0t) + \cos(t_0\omega))^2} \right), \quad \partial_t(\hat{\omega} - \omega) = -\frac{t_0}{2} \frac{\sin(t_0\omega) \sinh(t_0t)}{(\cosh(t_0t) + \cos(t_0\omega))^2} \quad (3.76)$$

$$\partial_\omega(\hat{t} - t) = \partial_t(\hat{\omega} - \omega), \quad \partial_\omega(\hat{\omega} - \omega) = -\frac{1}{2} \left(1 + t_0^2 \frac{1 + \cos(t_0\omega) \cosh(t_0t)}{(\cosh(t_0t) + \cos(t_0\omega))^2} \right). \quad (3.77)$$

Des simplifications interviennent sur l'axe $\omega = 0$:

$$\partial_\omega(\hat{\omega} - \omega) = -\frac{1}{2} \left(1 + \frac{t_0^2}{\cosh(t_0t) + 1} \right) \quad \partial_t(\hat{t} - t) = -\frac{1}{2} \left(1 - \frac{t_0^2}{\cosh(t_0t) + 1} \right), \quad (3.78)$$

de plus, les dérivées croisées s'annulent. On peut conclure directement que $\partial_\omega(\hat{\omega} - \omega)$ est strictement négatif, $\cosh(t_0t) + 1$ étant toujours positif.

Au point $(t_1, 0)$, il en est de même pour l'autre dérivée $\partial_t(\hat{t} - t) = -(1 - t_0 t_1 / \sinh(t_0 t_1))$ dans la mesure où \sinh est toujours au-dessus de la première bissectrice des axes. En ce point, la matrice dérivée de \hat{r} est définie négative ; c'est donc un maximum. On peut faire le même raisonnement pour $(-t_1, 0)$.

Au point $(0, 0)$, la situation est différente dans la mesure où $\partial_t(\hat{t} - t) = -(1 - t_0^2/2)/2$ est strictement positive sous la condition de Rayleigh ($t_0 > \sqrt{2}$) ce qui rend donc ce point instable (point col).

- *Axe des fréquences, $t = 0$.*

La deuxième équation de (3.74) est vérifiée sur cet axe. La première équation qui s'écrit

$$\omega = -t_0 \frac{\sin(t_0\omega)}{1 + \cos(t_0\omega)}, \quad (3.79)$$

admet, quelque soit t_0 , une infinité de solutions. Sur l'axe $t = 0$, l'équilibre est instable : la matrice dérivée du champ de vecteurs de réallocation est diagonale et la dérivée partielle

$$\partial_t(\hat{t} - t) = -\frac{1}{2} \left(1 - \frac{t_0^2}{\cos(t_0\omega) + 1} \right), \quad (3.80)$$

est minorée en tout point par une quantité positive $\partial_t(\hat{t} - t) > -(1 - t_0^2/2)/2 > 0$ si $t_0 > \sqrt{2}$. Tous les points solutions sont donc des points cols.

- *Reste du plan, $t \neq 0$ et $\omega \neq 0$.*

L'existence d'un extremum dans le reste du plan est soumise à la condition nécessaire

$$f(t, \omega) = t \sin(t_0\omega) + \omega \sinh(t_0t) = 0, \quad (3.81)$$

obtenue en faisant le quotient, membre à membre des deux éqs (3.74). Pour un temps $t > 0$ fixé, la dérivée partielle $\partial_\omega f = t_0 t \cos(\omega t_0) + \sinh(t_0 t)$ est strictement positive dans la mesure où $\sinh(t)$ est toujours au-dessus de la première bissectrice si $t > 0$. La fonction $f(t, \omega)$ est donc strictement croissante à t fixé. Puisque $f(t, 0) = 0$, $f(t, \omega)$ ne s'annule pas ailleurs. On peut faire le même raisonnement pour un temps $t < 0$. La condition nécessaire (3.81) est donc vérifiée uniquement sur les axes $t = 0$ ou $\omega = 0$, excluant de cette façon l'existence d'extrema ailleurs.

Il existe donc deux maxima parmi tous les extrema. Il est intéressant de savoir quelles sont les particules qui vont converger vers l'un ou l'autre, autrement dit comment la partition se fait. Les champs de vecteurs de réallocation et les spectrogrammes associés à chacun de ces cas envisagés précédemment sont montrés en Fig. 3.5.

A partir de (3.73), on peut vérifier que, conformément à l'intuition, $t = 0$ est la ligne de partage des eaux selon laquelle s'effectue la découpe du plan temps-fréquence. Une ligne de partage des eaux d'un potentiel V est une ligne de champ composée de points (t, ω) stationnaires et instables selon au moins une direction u . Plus précisément, les points d'une ligne de partage des eaux doivent respecter trois conditions, soit

$$\exists u \in \mathbb{R}^2, \frac{\partial V}{\partial u} = 0 \quad (3.82)$$

$$\frac{\partial^2 V}{\partial u^2} > 0 \quad (3.83)$$

$$\nabla V \wedge (dt, d\omega)^t = 0, \quad (3.84)$$

où $(dt, d\omega)^t$ est la tangente à la ligne de partage des eaux au point (t, ω) (notons bien que, pour une analogie avec des systèmes physiques, le sens de la gravité est ici inversé) et $u \wedge v$ désigne le produit vectoriel des vecteurs u et v .

L'éq. (3.82) impose que le point (t, ω) soit un extrema local et par conséquent une position d'équilibre selon la direction donnée par u . L'instabilité de cet équilibre est fixée par l'éq. (3.83). La ligne de partage des eaux est la seule ligne de champ (ligne tangente en tout point à la ligne de plus grande pente) parmi l'ensemble de points définis par les deux conditions précédentes, ce que sélectionne l'éq. (3.84).

En ce qui concerne notre problème, regardons en $t = 0$ et en choisissant la direction $u = (1, 0)^t$,

ce que deviennent ces équations

$$\partial(\log|F|)/\partial \mathbf{u}(0, \omega) = \hat{\mathbf{r}} \cdot \mathbf{u} = \hat{t}(0, \omega) = 0 \quad (3.85)$$

$$\partial^2(\log|F|)/\partial \mathbf{u}^2(0, \omega) = (1, 1)\mathbf{u}\mathbf{u}^t\partial\hat{\mathbf{r}} > 0 \quad \text{si } t_0 > \sqrt{2} \text{ (voir éq. (3.80))} \quad (3.86)$$

$$\nabla(\log|F|)(0, \omega) \wedge (0, 1)^t = \hat{\mathbf{r}} \wedge (1, 0)^t = 0, \quad (3.87)$$

où $\partial\hat{\mathbf{r}}$ est la matrice dérivée du champs de vecteurs de réallocation.

La ligne $t = 0$ est donc une ligne de partage des eaux. L'algorithme pratique une coupure à la médiatrice entre $(t_0, 0)$ et $(-t_0, 0)$, centres effectifs des deux logons. Il sépare donc le plan temps-fréquence en deux ensembles : l'un qui regroupe les points plus proches de $(t_0, 0)$ que de $(-t_0, 0)$, et son réciproque. Formulé de cette manière, ce que fait l'algorithme de partition est donc *a posteriori* identique à ce qu'on obtiendrait en pratiquant la partition du plan temps-fréquence par des polygônes de Voronoï⁷ [86] associés à chacun des centres des logons.

On vient d'établir l'équivalence *a posteriori* entre deux méthodes de partition du plan temps-fréquence, ce que l'on peut vérifier sur la Fig. 3.6 pour un signal composé de trois logons gaussiens centrés, en $(-t_0, 0)$, $(t_0, 0)$ et $(0, t_0)$, sommets d'un triangle isocèle. La Fig. 3.7 illustre que l'équivalence fonctionne toujours dans le cas plus général où les logons sont en nombre plus important et sont positionnés en (t_n, ω_n) choisis aléatoirement (densité uniforme sur $[-t_0, t_0] \times [-t_0, t_0]$) dans le plan temps-fréquence.

Si la fenêtre n'est plus de variance unité, l'équivalence avec la partition de Voronoï semble toujours possible en modifiant la distance entre points du plan temps-fréquence. Élargir ou rétrécir la fenêtre "ronde" revient à appliquer au plan temps-fréquence une anamorphose. Pour conserver la superposition des frontières fournies par l'algorithme de partition sur celles données par les polygônes de Voronoï, il s'agit alors de changer la définition de la distance pour l'adapter au plan temps-fréquence anamorphosé.

Influence de l'amplitude relative sur la partition

L'éloignement n'est pas le seul paramètre qui affecte la capacité à séparer deux composantes. Entrent également en ligne de compte, leur amplitude relative et leur relation de phase.

On reprend le problème précédent à deux logons en y ajoutant de nouveaux degrés de liberté. On autorise maintenant des amplitudes et des phases à l'origine arbitraires. Le traitement gagne alors considérablement en complexité.

Regardons d'abord l'influence de l'amplitude en conservant la référence de phase de chaque logon en leur centre. Notre signal $y(t)$ est maintenant la somme de $y_1(t) = \exp(-(t + t_0)^2/2)$ et $y_2(t) = a \exp(-(t - t_0)^2/2)$, où $a \in \mathbb{R}^{+*}$. Le calcul des opérateurs de réallocation suit une trame identique au cas $a = 1$, à la différence qu'un terme en $\log(a)$ vient s'insérer dans l'exponentielle de l'éq. (3.71) qui devient

$$F_1/F_2(t, \omega) = \exp(-\log a - t_0(t - i\omega)). \quad (3.88)$$

Les vecteurs de réallocation s'écrivent alors

$$\hat{\mathbf{r}} = \left(-\frac{t}{2} + \frac{t_0}{2} \frac{\sinh(t_0 t + \log a)}{\cosh(t_0 t + \log a) + \cos(t_0 \omega)}, -\frac{\omega}{2} - \frac{t_0}{2} \frac{\sin(t_0 \omega)}{\cosh(t_0 t + \log a) + \cos(t_0 \omega)} \right)^t. \quad (3.89)$$

7. Soit $P = \{p_n\}_{n=1..N}$ un ensemble de N points du plan, le polygone de Voronoï associé à un point p_i de P est le lieu du plan qui contient tous les points plus proches de p_i que des $N - 1$ autres points de P .

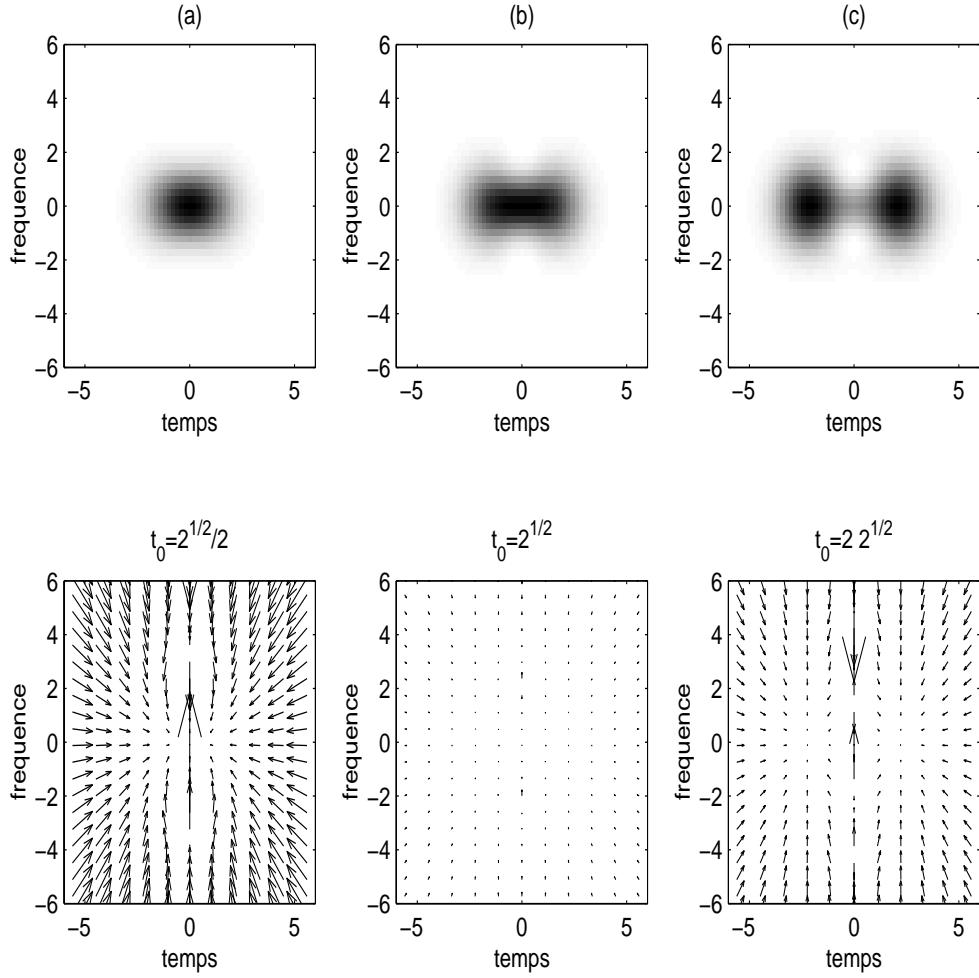


FIG. 3.5 – *Spectrogramme et champ de vecteurs de réallocation de la somme de deux logons gaussiens.* Cette figure répertorie les différentes situations possibles pour la structure du spectrogramme (fenêtre d’observation gaussienne à variance unité) du signal composé de deux logons distants de t_0 défini en Section 3.3.4. (a) Au dessous de la distance critique de $\sqrt{2}$, le spectrogramme ne présente qu’un seul maximum (le champ de vecteur de réallocation ne s’annule qu’en un seul point) $(0, 0)$. (b) Ceci est toujours vrai au point critique $t_0 = \sqrt{2}$. (c) Au dessus de cette distance, le spectrogramme admet trois extrema, dont deux maxima. Il devient possible de séparer les deux logons. (Le champ des vecteurs de réallocation n’est pas représenté en vraie grandeur pour faciliter la lecture. Signaux échantillonnés à 5Hz)

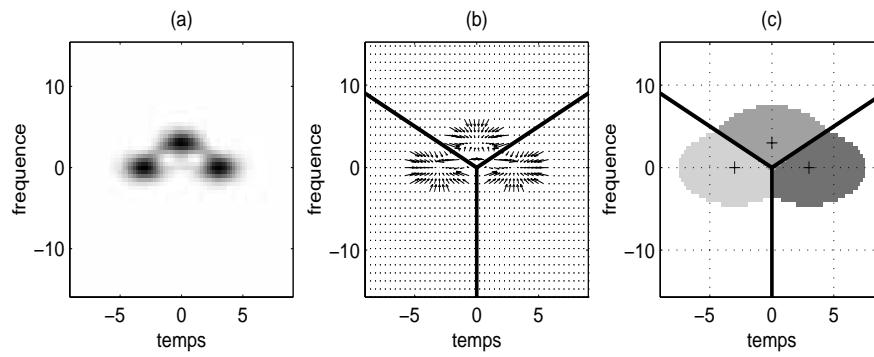


FIG. 3.6 – **Partition temps-fréquence à base de réallocation et partition de Voronoï: trois logons gaussiens.** Les algorithmes de partition du plan temps-fréquence à base de réallocation et de partition du plan temps-fréquence par des polygônes de Voronoï aboutissent pour le signal jouet (éq. (3.67), $N = 3$) formé de trois logons centrés en $(t_1, \omega_1) = (-3, 0)$, $(t_2, \omega_2) = (+3, 0)$ et $(t_3, \omega_3) = (0, +3)$ à des résultats identiques. Il s’agit ici d’une simple vérification car nous savons que ce signal respecte les hypothèses des approximations introduites en Section 3.3.4. (a) spectrogramme, (b) champ des vecteurs de réallocation et lignes de partage des eaux, (c) partition temps-fréquence (niveau de gris) et partition (lignes frontière). (Signal échantillonné à 5Hz)

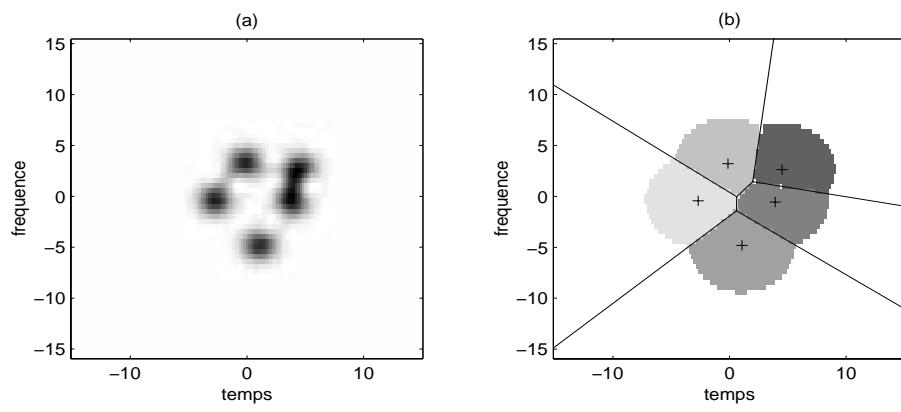


FIG. 3.7 – **Partition temps-fréquence à base de réallocation et partition de Voronoï: cinq logons gaussiens.** On compare sur cette figure, pour un signal (spectrogramme en (a)) composé de 5 logons positionnés aux points (t_n, ω_n) choisis aléatoirement (densité uniforme sur $[-5, 5] \times [-5, 5]$), la partition temps-fréquence obtenue avec la réallocation (zones en niveaux de gris dans (b)) et la partition par les polygônes de Voronoï (lignes frontières dans (b)). Les limites trouvées par les deux méthodes se superposent exactement. (Signal échantillonné à 5Hz)

Les points fixes de la réallocation respectent les deux conditions suivantes

$$t = t_0 \frac{\sinh(t_0 t + \log a)}{\cosh(t_0 t + \log a) + \cos(t_0 \omega)}, \quad (3.90)$$

$$\omega = -t_0 \frac{\sin(t_0 \omega)}{\cosh(t_0 t + \log a) + \cos(t_0 \omega)}. \quad (3.91)$$

Nous n'en effectuerons pas l'étude complète. Il est cependant important de savoir comment l'amplitude influe sur la capacité de séparation des deux composantes. Pour cela, nous nous concentrerons sur l'axe des temps. Sur cet axe, l'éq. (3.91) étant toujours vérifiée, le système se réduit de deux conditions de stationnarité à une seule :

$$f(t) = t - t_0 \frac{\sinh(t_0 t + \log a)}{\cosh(t_0 t + \log a) + 1} = 0. \quad (3.92)$$

L'étude des variations de la fonction $f(t)$ montre qu'elle admet trois zéros distincts sous la condition de Rayleigh, $t_0 > \sqrt{2}$, à laquelle s'ajoute une contrainte sur l'amplitude relative $(1/a_0) < a < a_0$, où l'amplitude critique s'exprime de la manière suivante

$$a_0 = \frac{\exp(t_0 \sqrt{t_0^2 - 2})}{t_0^2 - 1 + t_0 \sqrt{t_0^2 - 2}}. \quad (3.93)$$

Similairement au cas $a = 1$, deux de ces zéros sont des maxima et attracteurs, le dernier situé entre les deux précédents, est un point col. Il indique la position de la ligne de partage des eaux qui est, cette fois, plus difficile à obtenir : elle est une solution non triviale de l'éq. (3.92) qu'il faut résoudre numériquement. C'est ce qu'illustrent les figures 3.8 et 3.9.

Méthodes s'appuyant sur une mesure d'information

Lorsque les composantes sont proches, leur séparation devient extrêmement délicate. L'ambiguïté est telle qu'il est parfois possible d'envisager plusieurs solutions physiquement acceptables et néanmoins, totalement différentes. Qui plus est, à faible éloignement, la relation de phase devient cruciale : il est naturellement plus difficile de séparer des composantes en phase qu'en quadrature. Aucune réponse définitive ne pouvant nous servir de référence, on se propose de faire une comparaison entre notre algorithme et celui proposé dans [7, 6] qui est basé sur les mesures d'information. Nous nous servirons des deux mesures d'information exposées en Sect. 3.3.1 à savoir l'information de Rényi d'ordre 3 sur la distribution de $I_{R3}^W(x)$ définie en éq. (3.52) et l'information de Shannon portée par le spectrogramme $I_{SH}^S(x)$ définie en éq. (3.53).

Dans les deux cas, on convertit le résultat obtenu exprimé en bits vers une unité homogène à un nombre de composante par les éqs. (3.55) et (3.56).

Les comparaisons entre les mesures d'information et notre méthode sont faites en Fig. 3.10 sur un signal composé de deux logons dont on fait varier la distance et la phase relative (amplitudes égales et constantes). Les résultats sont très comparables. En particulier, la frontière entre les configurations distance-phase où il y séparation des logons par la partition et celle où elle n'est pas réalisée, se superpose assez bien à la courbe de niveau 1, 5 de $N_{SH}^S(x)$.

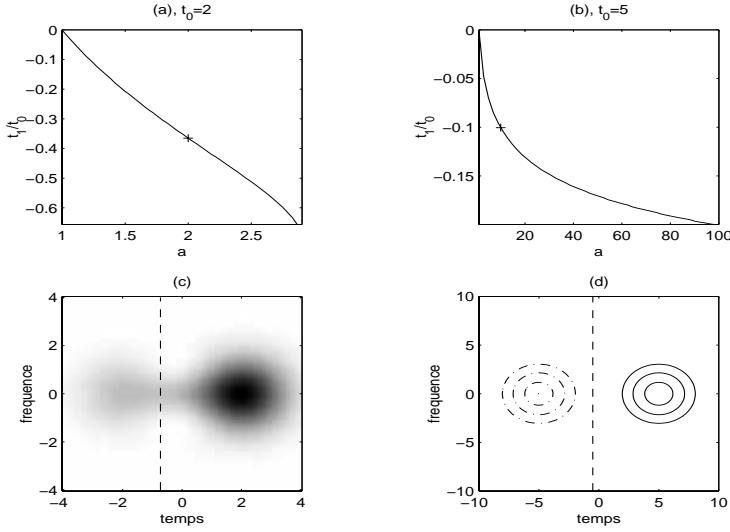


FIG. 3.8 – *Influence de l'amplitude relative et de la distance sur la partition temps-fréquence : distance constante, amplitude variable.* Si le signal est composé de deux logons d'amplitude différentes (cf. 3.3.4), la partition ne sectionne plus le plan temps-fréquence à la médiatrice des centres des deux logons, mais en à un temps t_1 qui varie non linéairement en fonction de la demi-distance t_0 et de l'amplitude relative a . Les graphes (a) et (b) montrent, respectivement pour une distance courte et longue, l'évolution de t_1 en fonction de a choisi dans $[1, a_0]$ (la deuxième moitié de la gamme de valeurs admissibles pour a , $[1/a_0, 1]$, donne des résultats symétriques). Dans le premier cas ($t_0 = 2$), la partition est très sensible à l'amplitude relative (t_1 varie quasi linéairement en fonction de a). À titre d'exemple, on a superposé en (c) le spectrogramme du signal (échantillonné à 13Hz environ) pour $a = 2$ et la ligne de partage des eaux (en pointillé). Dans le deuxième cas ($t_0 = 5$), l'amplitude critique a_0 devient très grande. Dans des gammes d'amplitude raisonnable (ici, a décrit $[1, 100]$), il est remarquable que la différence d'amplitude n'influe plus sur la partition. Tout se passe alors comme si les deux logons étaient de même amplitude. Le graphe (d) en montre un exemple pour $a = 10$ (signal échantillonné à 5Hz). Les deux logons étant d'amplitude très différente, le spectrogramme est indiqué par des courbes de niveaux ($\{50, 10, 1\}$ en trait continu, $\{0, 5, 0.1, 0.001\}$ en trait mixte).

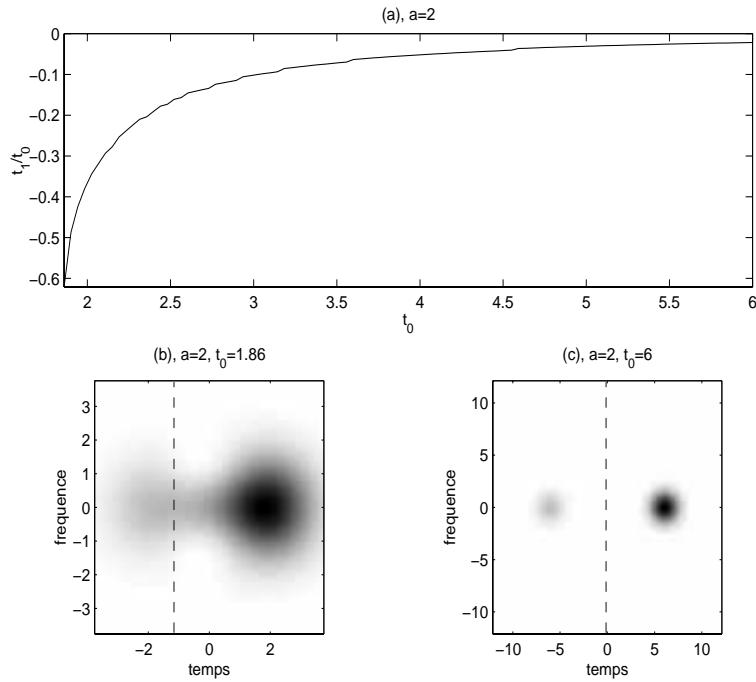


FIG. 3.9 – *Influence de l'amplitude relative et de la distance sur la partition temps-fréquence : amplitude constante, distance variable.* Cette figure complète la Fig. 3.8 en présentant, en (a) et pour le même signal, l'évolution de la position de la ligne de partage des eaux en fonction de t_0 à amplitude relative constante ($a = 2$). On constate qu'à courte distance, l'amplitude influe beaucoup sur la partition. La zone associée par la partition temps-fréquence au logon de plus forte amplitude, s'étend largement au-delà de l'axe $t = 0$ dans le demi plan $t < 0$. Ceci est illustré en (b) pour la demi-distance critique ($t_0 = 1.86$) en deçà de laquelle on ne détecte plus qu'une seule composante (signal échantillonné à 13.6Hz). Ce n'est plus le cas pour les grandes distances (graphe (c), $t_0 = 6$) où cette fois le partage se fait approximativement à la médiatrice (signal échantillonné à 4.2Hz).

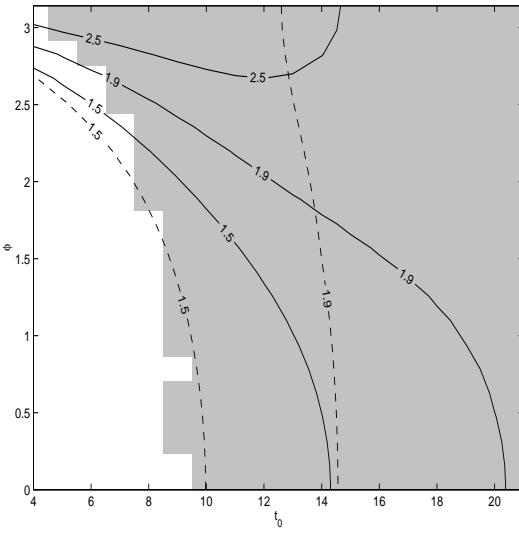


FIG. 3.10 – *Comparaison entre le nombre de composantes détectées par la partition temps-fréquence et celui donné par les méthodes à base de mesure d'information.* Cette figure regroupe les réponses que donnent trois méthodes différentes à la question : “combien y-a t'il de composantes dans le signal $x(t) = \exp(-t^2/2) + \exp(-(t - t_0)^2/2) \exp(i\varphi)$?” pour diverses valeurs de la distance t_0 et de la phase relative φ . En trait plein : information de Rényi d'ordre 3 calculée sur la distribution de Wigner-Ville (éq. (3.55)). En trait pointillé : information de Shannon calculée sur le spectrogramme (éq. (3.56)). En niveau de gris : la méthode de partition par réallocation différentielle (blanc=1 composante, gris=2 composantes). La frontière donnée par la partition temps-fréquence correspond approximativement à la courbe de niveau 1,5 sur N_{SH}^S .

Extraction-reconstruction, application au débruitage

Nous nous intéressons ici à la reconstruction des différentes composantes d'un signal. Il existe de nombreuses méthodes pour reconstruire une partie de signal contenue dans un domaine du plan temps-fréquence. Ce problème est équivalent à celui de la synthèse d'un filtre temps-fréquence de gabarit fixé. Le filtrage de Weyl est une solution qui réalise un bon compromis entre précision de reconstruction et complexité algorithmique [69]. C'est donc à l'aide de cette méthode que nous reconstruisons les différentes parties du signal dans leur représentation temporelle.

La Fig. 3.11 illustre ce que nous proposons sur un signal composé de deux "chirps" linéaires de taux de modulation différents et d'enveloppes gaussiennes d'amplitudes différentes. L'algorithme de partition détecte les deux composantes et les situe sur une carte temps-fréquence. On présente ensuite la reconstruction de chacun des deux "chirps".

Disposer d'une carte temps-fréquence est également utile pour le débruitage de signaux. À titre d'exemple, on veut extraire un signal (une modulation linéaire de fréquence d'enveloppe gaussienne en l'occurrence) noyé dans un bruit additif (blanc et gaussien). La carte peut nous servir à sélectionner la zone temps-fréquence qui correspond au signal et à rejeter celles qui sont associées au bruit. Par ce biais, la carte temps-fréquence nous permet, en quelque sorte, de trouver un filtrage non-stationnaire bien adapté au signal pour son extraction du bruit. En reconstruisant, avec les méthodes présentées dans le paragraphe précédent, ce que contient la zone "signal + bruit", on réhausse le rapport signal sur bruit (RSB). On évalue la qualité du débruitage en mesurant le gain en RSB obtenu après traitement. C'est ce que nous montre la Fig. 3.12 pour un RSB initial variant entre -15dB et 10dB . Les simulations que nous avons réalisées nous indiquent que, pour un RSB de l'ordre de -5 à 0dB , l'atténuation du bruit est d'environ 5dB .

La question cruciale reste tout de même de savoir reconnaître la zone "signal + bruit" sur la carte. Dans l'exemple que nous proposons (que l'on doit considérer simplement comme une illustration), nous avons opté pour un critère simple : nous choisissons parmi toutes les composantes détectées celles de plus grande énergie. Ce critère est bien évidemment voué à l'échec dès que le RSB initial devient faible. Lorsqu'on augmente la puissance du bruit, la quantité d'énergie augmente dans les partitions "bruit seul" jusqu'à, pour certaines d'entre elles, dépasser celle de la partition "signal + bruit". Dans ce cas, le critère se trompe inévitablement de partition. On pourrait alors penser à des sophistications telles que l'utilisation de critères entropiques ou l'intégration de connaissances *a priori* sur le signal si elles sont disponibles.

3.3.5 Fusion de partition

Lorsque deux signaux sont proches dans le plan temps-fréquence, l'algorithme de partition devient particulièrement sensible à la relation de phase. Cela se manifeste par des coupures parasites comme par exemple celle qui est présentée en Fig. 3.13 pour un mélange de deux chirps. Pour donner une intuition de ce qui se passe, plaçons-nous dans le cas d'une fenêtre gaussienne à variance unité pour lequel le champ de vecteurs de réallocation dérive exactement du potentiel $\log|F_x|$. Pour certaines relations de phase, il apparaît des maxima locaux secondaires dans le potentiel qui viennent piéger quelques particules et créer des partitions parasites. En particulier, cela est vrai en présence de bruit. En effet, si l'on considère que $x(t) = s(t) + b(t)$ provient de la perturbation du signal $s(t)$ par un bruit additif $b(t)$ de faible amplitude, le potentiel peut se mettre sous la forme suivante :

$$\log|F_x| = \log|F_s| + \log|1 + F_b/F_s|. \quad (3.94)$$

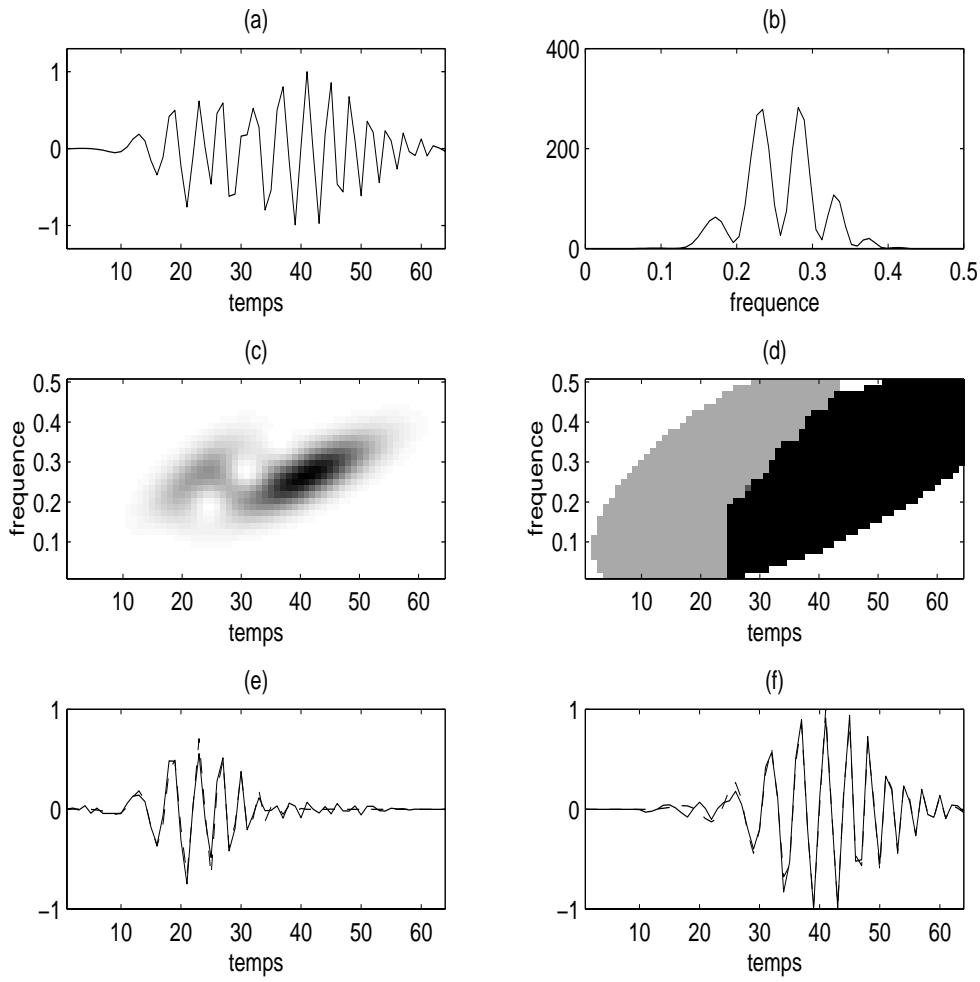


FIG. 3.11 – *Extraction–reconstruction de chirps à l'aide de l'algorithme de partition temps-fréquence.* (a) signal, (b) spectre d'énergie, (c) spectrogramme, (d) partition, (e) première composante (trait pointillé) et sa reconstruction (trait plein), (f) deuxième composante (trait pointillé) et sa reconstruction (trait plein).

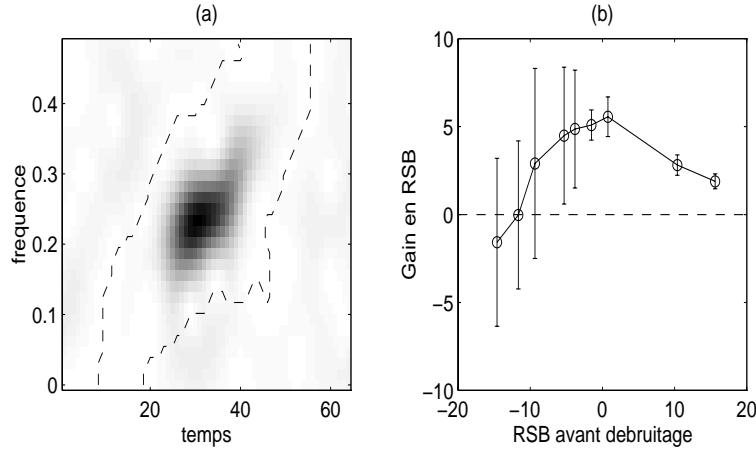


FIG. 3.12 – *Débruitage à l'aide de l'algorithme de partition temps-fréquence* (voir Sect. 3.3.4). (a) spectrogramme d'une réalisation bruitée ($RSB = 0\text{dB}$) et contour de la zone “signal” identifiée, (b) moyenne et écart-type du gain en RSB en fonction du RSB en entrée (10 réalisations de bruit indépendantes pour chaque point).

À fort rapport signal à bruit, i.e. lorsque $|F_b/F_s| \ll 1$, on développe au premier ordre le membre de gauche de l'équation précédente

$$\log|F_x| \approx \log|F_s| + \text{Re}\{F_b/F_s\}, \quad (3.95)$$

d'où l'on conclut que le potentiel associé au champ de vecteurs de réallocation du signal $x(t)$ bruité s'écrit comme la combinaison du potentiel obtenu avec le signal seul, à savoir $\log|F_s|$ et d'une perturbation du bruit $\text{Re}\{F_b/F_s\}$ qui causent l'apparition de maxima locaux et donc de composantes parasites. Nous avons envisagé plusieurs solutions pour accroître la robustesse de la partition au bruit, et plus généralement la robustesse à l'effet du battement de phase. Nous allons nous intéresser maintenant à la fusion de partition temps-fréquence.

Cette approche s'appuie sur la nécessité d'équivalence entre chaque élément de notre partition (chacune des régions sur la carte temps-fréquence) à un signal réel. Plus précisément, chaque restriction du spectrogramme à la zone correspondante donnée par l'algorithme de partition doit être lui-même un spectrogramme. Or, il s'avère que les partitions associées aux maxima parasites sont généralement de petites tailles. Par exemple, en Fig. 3.13, la composante n° 3 provient d'une coupure parasite due à un agencement particulier des phases des deux chirps. La zone qui lui est associée sur la carte de temps-fréquence est de faible surface : le spectrogramme réduit à cette zone ne peut pas correspondre au spectrogramme d'un signal qui existe réellement. Le but de cette section est de trouver un critère pour reconnaître les zones de la carte temps-fréquence non admissibles en termes de distribution temps-fréquence, puis de proposer un algorithme d'agrégation des partitions parasites (typiquement, on aimerait fondre les composantes n° 2 et n° 3 en une seule) pour le recouvrement des composantes qui leur sont effectivement associées.

Discrimination des partitions non-admissibles

On adopte ici une démarche heuristique. Nous faisons cinq propositions de critères que l'on sépare en deux catégories :

1. les critères qui font intervenir tous les points de chaque partition (approche composante),

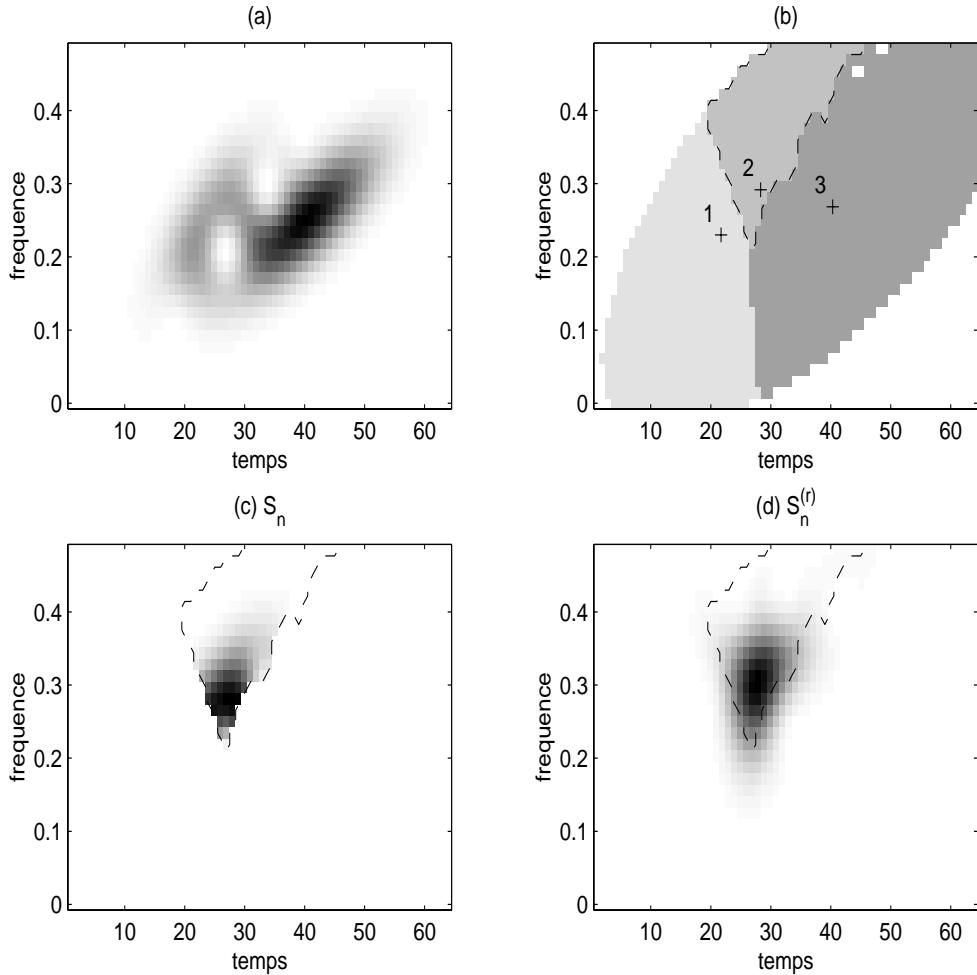


FIG. 3.13 – **Mise en évidence de coupure parasite dans la partition temps-fréquence due au battement de phase.** Cette figure illustre l'influence sur la partition, du battement de phase entre deux composantes (ici, deux modulations linéaires de fréquence d'enveloppe gaussienne ou chirps). L'algorithme de partition utilise le champ de vecteurs de réallocation du spectrogramme (**a**). Le battement de phase fait intervenir une coupure parasite qui scinde le chirp de plus petite amplitude en deux parties (composantes n°1 et 2 sur la carte temps-fréquence (**b**)). Il est clair que la composante n°2 n'est pas admissible. C'est ce que confirme les différences que l'on observe entre, d'une part, (**c**) la restriction du spectrogramme à la région qui lui est associée par l'algorithme et, d'autre part, (**d**) le spectrogramme de la reconstruction de la composante n°2 par la méthode décrite en Sect. 3.3.4.

2. et ceux qui ne prennent en compte que les points des frontières (approche frontière).

Nous nous limiterons à des mesures faites sur le spectrogramme, notre partition s'effectuant à partir de celui-ci.

• *Approche composante*

– *Mesure d'encombrement.*

Si, comme nous l'avons dit précédemment, les partitions non pertinentes ont une faible surface équivalente dans le plan temps-fréquence, elles ne respectent pas *a fortiori* les inégalités de concentration que l'on peut formuler sur des distributions d'énergie temps-fréquence en général. L'idée naturelle est alors de calculer, pour chacune des N composantes détectées pour le signal $x(t)$, une quantité que l'on comparera avec la borne donnée par l'inégalité.

Soit $S_n(t, \omega) = 1_{C_n}(t, \omega) S(t, \omega) / E_n$, la restriction du spectrogramme au domaine C_n du plan temps-fréquence associé à la $n^{\text{ième}}$ composante détectée par l'algorithme de partition. La fonction 1_{C_n} est la fonction indicatrice sur la région C_n . On a pris soin de normaliser S_n à 1 en norme L_1 , soit $E_n = \iint_{C_n} S(t, \omega) dt d\omega / (2\pi)$, pour que l'on puisse l'interpréter comme une densité de probabilité.

La première idée consiste à mesurer une surface équivalente par un moment du deuxième ordre conjoint en temps et en fréquence [60, 39] :

$$d_1(x, n) = \iint \left(\left(\frac{t - \bar{t}_n}{\Delta t_h} \right)^2 + \left(\frac{\omega - \bar{\omega}_n}{\Delta \omega_h} \right)^2 \right) S_n(t, \omega) \frac{dt d\omega}{2\pi}, \quad (3.96)$$

où $\bar{t}_n = \iint t S_n(t, \omega) dt d\omega / (2\pi)$ et $\bar{\omega}_n = \iint \omega S_n(t, \omega) dt d\omega / (2\pi)$ sont les coordonnées du point moyen de la distribution S_n . Si S_n est le spectrogramme d'un signal quelconque, alors nécessairement

$$d_1(x, n) \geq \frac{1}{2\pi}. \quad (3.97)$$

La borne de l'inéquation est atteinte [60] si et seulement si $h(t)$, la fenêtre d'observation est gaussienne (cf. éq. (1.40)) et le signal associé à C_n est gaussien.

On peut également profiter des inégalités faisant intervenir les mesures d'information [70, 61] sur les distributions temps-fréquence (déjà évoquées en Sect. 3.3.1). On note :

$$I_{SH}(S) = - \iint S(t, \omega) \log_2 S(t, \omega) \frac{dt d\omega}{2\pi}, \quad (3.98)$$

l'information de Shannon calculée sur la distribution S de norme L_1 unité. Si $h(t)$ est une fenêtre gaussienne et si S_n est le spectrogramme d'un signal arbitraire, on peut montrer que [61]

$$d_2(x, n) = 2^{I_{SH}(S_n) - I_0}, \quad (3.99)$$

(où I_0 est la quantité d'information minimum obtenue avec le spectrogramme d'un signal égal à la fenêtre) est toujours supérieure ou égale à 1. Cette quantité représente le nombre de composantes portées par $S_n(t, \omega)$, l'unité étant fixée par la fenêtre.

Notons bien que ces deux conditions d'admissibilité sont nécessaires, et que des distributions non admissibles peuvent aussi les vérifier.

– *Comparaison au signal reconstruit.*

Le deuxième point de vue consiste à confronter la restriction S_n du spectrogramme, avec le spectrogramme $S_n^{(r)}$ de la composante reconstruite $x_n^{(r)}$ par la méthode décrite en Sect. 3.3.4. Ceci revient

approximativement à faire la comparaison de S_n avec son projeté sur l'ensemble des spectrogrammes. Si cette comparaison met en évidence deux distributions très différentes alors $S_n(t, \omega)$ n'est pas admissible en tant que spectrogramme ; il faut rejeter la partition correspondante. Pour quantifier les dissemblances entre S_n et $S_n^{(r)}$, on préfère ignorer les différences qui peuvent apparaître à l'intérieur de C_n et mesurer la proportion d'énergie de $S_n^{(r)}$ à l'extérieur de C_n :

$$d_3(x, n) = \frac{1}{E_n^{(r)}} \iint S_n^{(r)}(t, \omega) (1 - \mathbf{1}_{C_n}(t, \omega)) \frac{dt d\omega}{2\pi}, \quad (3.100)$$

où $E_n^{(r)} = \|x_n^r\|^2$ est l'énergie du signal reconstruit. On estime que, si cette proportion dépasse un seuil arbitraire que l'on placera pour fixer les idées à 10%, la partition n'est pas pertinente. La raison de ce choix est que, si la partition C_n est trop “étroite”, alors $S_n^{(r)}(t, \omega)$, étant contraint aux inégalités d'encombrement dans le plan temps-fréquence, va devoir s'étaler hors des frontières fixées par C_n . C'est, par exemple, ce que l'on constate en Fig. 3.13.

On peut adapter la même philosophie dans un contexte informationnel et mesurer la différence des informations portées par S_n et $S_n^{(r)}$:

$$d_4(x, n) = 2^{I_{SH}(S_n^{(r)}) - I_{SH}(S_n) - I_0}. \quad (3.101)$$

Ceci revient à compter le nombre de composantes qu'il faudrait ajouter au signal virtuel de spectrogramme S_n pour obtenir $x_n^{(r)}$. S'il faut ajouter plus d'une composante, il est alors probable que la partition C_n ne soit alors pas pertinente.

- *Approche frontière : énergie le long de la coupure*

Pour tout signal d'énergie finie, le spectrogramme finit par décroître dans toutes les directions du plan temps-fréquence pour s'annuler à l'infini. Si on désire que chacune de nos partitions soit admissible en tant que spectrogramme, il nous faut donc exiger que le spectrogramme ait une valeur négligeable sur chacune des frontières. L'idée n'est donc plus de vérifier *directement* si chaque partition représente ou non un signal réel, mais plutôt de valider chacune des coupures faites par l'algorithme de partition.

Pour chaque couple C_n et C_m de composantes adjacentes, on note l'intersection $(\overset{\circ}{t}_{n,m}, \overset{\circ}{\omega}_{n,m})$ entre le segment qui relie leur centres respectifs (t_n, ω_n) et (t_m, ω_m) et leur frontière. Le rapport de l'énergie mesurée en $(\overset{\circ}{t}_{n,m}, \overset{\circ}{\omega}_{n,m})$ et de la moyenne géométrique des maxima du spectrogramme dans C_n et C_m

$$d_5(x, n, m) = \frac{S(\overset{\circ}{t}_{n,m}, \overset{\circ}{\omega}_{n,m})}{\sqrt{S(t_n, \omega_n)S(t_m, \omega_m)}}, \quad (3.102)$$

nous indique, lorsqu'il est petit, i.e., inférieur à un seuil que l'on a fixé arbitrairement pour l'exemple à 0.5, que le spectrogramme décroît suffisamment vers 0 à la frontière, et valide par conséquent la frontière considérée. Dans le cas contraire, la coupure entre C_n et C_m est à remettre en question.

Algorithme de fusion

Notre but est ici d'agréger au mieux les composantes non pertinentes dans une carte temps-fréquence. Plutôt que d'agir sur la carte temps-fréquence elle-même, nous allons, à partir d'elle,

construire un graphe sur lequel l'algorithme d'agrégation agira. La construction du graphe et l'algorithme de fusion diffère selon l'approche que l'on considère.

- **Approche composante : critères d_1 à d_4 .**

Le graphe est constitué de *nœuds* qui correspondent chacun à un centre temps-fréquence (et donc à une composante de la partition). A chacun de ces noeuds, on associe la mesure d'admissibilité de la composante idoine. La structure de ce graphe est fixée par l'information de voisinage (ou de connexité) dans le plan temps-fréquence : deux centres sont connectés si les régions temps-fréquence qui leur sont associées sont limitrophes. Enfin, la distance entre chaque nœud est simplement la distance euclidienne (en coordonnées normalisées par les durée et bande de la fenêtre, resp. en temps et en fréquence) dans le plan temps-fréquence.

L'algorithme va modifier ce graphe pour en obtenir un nouveau (du même type) où tous les nœuds seront associés à des composantes admissibles. La fusion est effective nécessairement sur des composantes connexes, et de préférence sur des composantes proches.

Le déroulement des opérations suit la trame suivante :

Algorithme de fusion (Approche composante) :

- Calcul du diagramme de connexité

Pour toutes les branches du diagramme, et dans l'ordre croissant de leur longueur,

faire

calculer le critère d'admissibilité pour les deux composantes de la branche courante,

si l'une des deux composantes, au moins, de la branche courante n'est pas admissible :

- procéder à la fusion des deux composantes de la branche courante,
- réévaluer la mesure d'admissibilité de la nouvelle composante,
- fixer son centre au milieu des deux centres de la branche courante,
- réévaluer le graphe de connexité et sélectionner la branche de plus petite longueur,

sinon

- passer à la branche suivante.

fin si

fin faire

- **Approche frontière : critère d_5 .**

Pour cette approche, le graphe est similaire au précédent à la différence que l'on n'associe plus une valeur du critère d'admissibilité à chaque noeud mais à chaque branche. Le traitement suit alors le déroulement suivant :

Algorithme de fusion (Approche frontière) :

- Calcul du diagramme de connexité

Pour toutes les branches du diagramme,

```

faire
    calculer (par dichotomie) la position du point ( $\overset{\circ}{t}_{n,m}$ ,  $\overset{\circ}{\omega}_{n,m}$ ),
    évaluer le critère d'admissibilité de la branche courante,
    si la branche courante n'est pas admissible :
        – procéder à la fusion des deux composantes,
        – fixer le centre de la nouvelle composante au milieu des deux centres de la branche
            courante,
    sinon
        – passer à la branche suivante.
fin si
fin faire

```

Résultats, signaux synthétiques

Nous avons testé nos algorithmes sur deux types de signaux. Le premier, en Fig. 3.13 met en évidence les problèmes liés au battement de phase entre deux composantes, en l'occurrence deux modulations linéaires de fréquence (ou chirp) d'enveloppe gaussienne. L'interaction entre ces deux chirps provoque l'apparition d'une coupure parasite qui vient scinder un des deux chirps (celui de plus faible amplitude) en deux. Nous donnons les résultats pour les critères d_1 à d_4 uniquement. Une analyse rapide du tableau 3.1 montre que le critère d_1 n'est pas suffisamment contraignant pour détecter les mauvaises partitions. Cependant, les trois autres critères donnent des résultats satisfaisants. En effet, leur utilisation dans l'algorithme de fusion conduit à l'agrégation des composantes n°1 et n°2 et laisse la composante n°3 intacte. La partition obtenue en fin de traitement est donc le résultat espéré.

Le deuxième exemple en Fig. 3.14 illustre l'action du bruit sur la partition temps-fréquence. Le signal est un chirp linéaire d'enveloppe gaussienne. Il est ajouté à un bruit blanc gaussien avec un rapport signal à bruit (quotient des énergies) égal à +2dB. La partition détecte six composantes. Il est clair que la présence de bruit provoque trois coupures parasites qui viennent découper le signal en tranches (composantes n°s 1,2,3 et 5). Les tableaux 3.2 et 3.3 regroupent les valeurs des cinq critères pour les six composantes détectées. On tire la même conclusion que précédemment au sujet de la surface équivalente, à savoir son inefficacité. Pour les trois critères de l'approche région, l'algorithme de fusion (cf. Fig. 3.15) se trouve piégé par le fait que les partitions dues au bruit sont à la fois non admissibles en tant que spectrogramme, et proches des partitions du signal en distance euclidienne normalisée (c'est par exemple vrai pour la configuration des nœuds 3, 4 et 5. La distance entre 3 et 4 est inférieure à celle entre 3 et 5). L'algorithme va donc avoir tendance à assembler les partitions du signal avec celles du bruit. Le critère d_5 , au contraire, fait bien la distinction entre frontières pertinente et non pertinente et conduit au résultat que l'on attend.

Résultats, signaux réels

Dans le cadre d'une collaboration avec Marianne Nardin (LIS, INPG, Grenoble), nous avons appliqué la partition temps-fréquence au problème de la caractérisation de modes de propagation d'une onde acoustique dans un guide d'ondes.

On désire caractériser la propagation dans un guide d'ondes qui est, en pratique, un canal qui conduit de l'eau avec, au fond, une couche sédimentaire. Pour ce faire, on provoque une explosion

Numéro de composante	Surface équivalente éq. (3.96)	Information de Shannon éq. (3.99)	Énergie éq. (3.100)	Information de Shannon éq. (3.101)
Composante 1	0.34	0.86	0.12	0.54
Composante 2	0.21	0.41	0.36	1.49
Composante 3	0.92	1.66	0.03	0.41
Borne théorique ou seuil acceptable	≥ 0.16	≥ 1	≤ 0.1	≤ 1

TAB. 3.1 – Ce tableau regroupe les valeurs des quatre critères d’admissibilité d’une partition (introduits en Sect. 3.3.5) pour chacune des trois composantes du signal présenté en Fig. 3.13.

Numéro de composante	Surface équivalente éq. (3.96)	Information de Shannon éq. (3.99)	Énergie éq. (3.100)	Information de Shannon éq. (3.101)
Composante 1	0.61	1.31	0.04	0.49
Composante 2	0.21	0.56	0.16	0.86
Composante 3	0.40	0.90	0.09	0.60
Composante 4	0.19	0.70	0.17	1.50
Composante 5	0.51	1.14	0.05	0.52
Composante 6	0.22	0.72	0.10	1.34
Borne théorique ou seuil acceptable	≥ 0.16	≥ 1	≤ 0.1	≤ 1

TAB. 3.2 – Ce tableau regroupe les valeurs des quatre premiers critères d’admissibilité d’une partition (introduits en Sect. 3.3.5) pour chacune des six composantes du signal présenté en Fig. 3.14.

Numéro de branche	Énergie à la coupure, éq. (3.102)
Branche 1 : Comp. 1 → Comp. 2	0.94
Branche 2 : Comp. 2 → Comp. 3	0.90
Branche 3 : Comp. 3 → Comp. 4	0.20
Branche 4 : Comp. 3 → Comp. 5	0.74
Branche 5 : Comp. 3 → Comp. 6	0.09
Branche 6 : Comp. 4 → Comp. 5	0.11
Branche 7 : Comp. 5 → Comp. 6	0.15
Seuil acceptable	≤ 0.5

TAB. 3.3 – Ce tableau regroupe les valeurs du critère d’admissibilité d_5 (approche frontière) d’une partition (introduit en Sect. 3.3.5) présentée en Fig. 3.14.

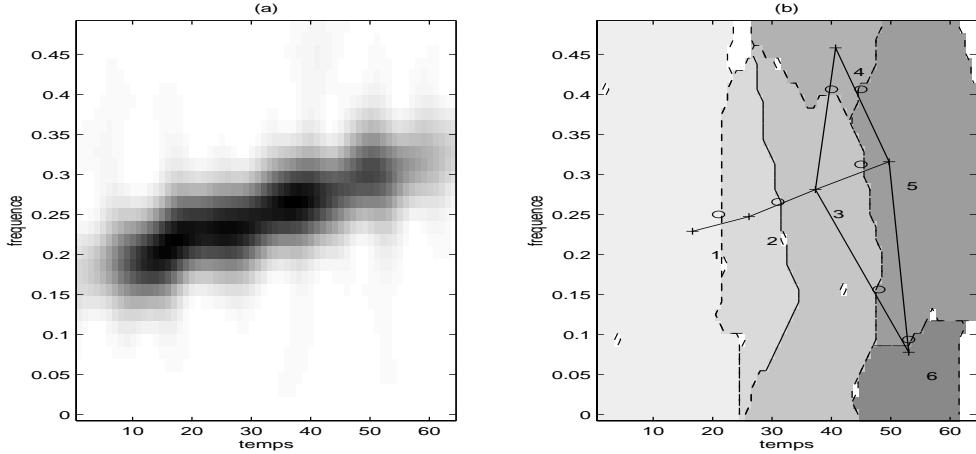


FIG. 3.14 – *Mise en évidence de coupure parasite dans la partition temps-fréquence due à la présence de bruit.* Le signal, dont le spectrogramme est donné en (a), est formé d'un chip linéaire d'enveloppe gaussienne ajouté à un bruit blanc gaussien ($RSB = 2\text{dB}$). (b) La présence de bruit cause l'apparition de coupures parasites qui sectionnent en tranches la partition que l'on aimerait associer au signal (composantes n° 1, 2, 3 et 5). Les centres de chacune des composantes sont marqués par des croix et sont reliés par le graphe de connexité (voir Sect. 3.3.5). Les points indiqués d'un cercle sont les (points de la grille temps-fréquence les plus proches de) intersections entre le segment qui relie deux centres connexes et la frontière entre les deux composantes qui leur sont associées.

à la surface de l'eau i.e., on émet une impulsion en entrée du guide. On observe ensuite le signal reçu après propagation. La nature du milieu et les différentes réflexions vont conduire à ce que l'onde de départ emprunte différents trajets et soit modifiée pendant ce trajet. Le résultat est que le signal observé est composé de différentes composantes, chacune due à un mode de propagation particulier.

L'utilisation d'une distribution temps-fréquence pour ce problème est naturelle : les signaux observés sont non stationnaires et la représentation temps-fréquence permet une visualisation immédiate de leur structure. Le signal est, en fait, une superposition de chirps comme l'indique clairement le spectrogramme en figure 3.16-(a). Chacun de ces chirps correspond à un mode de propagation.

Faire la partition temps-fréquence pour isoler un de ces chirps est donc équivalent à faire l'analyse de la propagation mode par mode. Ceci est utile si l'on veut faire l'estimation des paramètres (comme par exemple, la largeur de bande moyenne) du chirp dont on possède, par ailleurs, une description théorique précise. On en déduit ainsi une caractérisation des modes de propagation puis, avec l'ensemble des estimations obtenues pour chaque mode, l'extraction des paramètres du guide d'ondes (comme la profondeur du guide, le gradient de vitesse et de densité du milieu, . . .).

Nous montrons en Fig. 3.16 l'extraction du quatrième harmonique par la partition temps-fréquence suivie d'une fusion. Il est important de signaler que la fusion a été effectuée “à la main” i.e., sans l'aide des algorithmes présentés en Sect. 3.3.5 (nous ne disposons pas d'algorithme pour l'évaluation automatique du graphe de connexité).

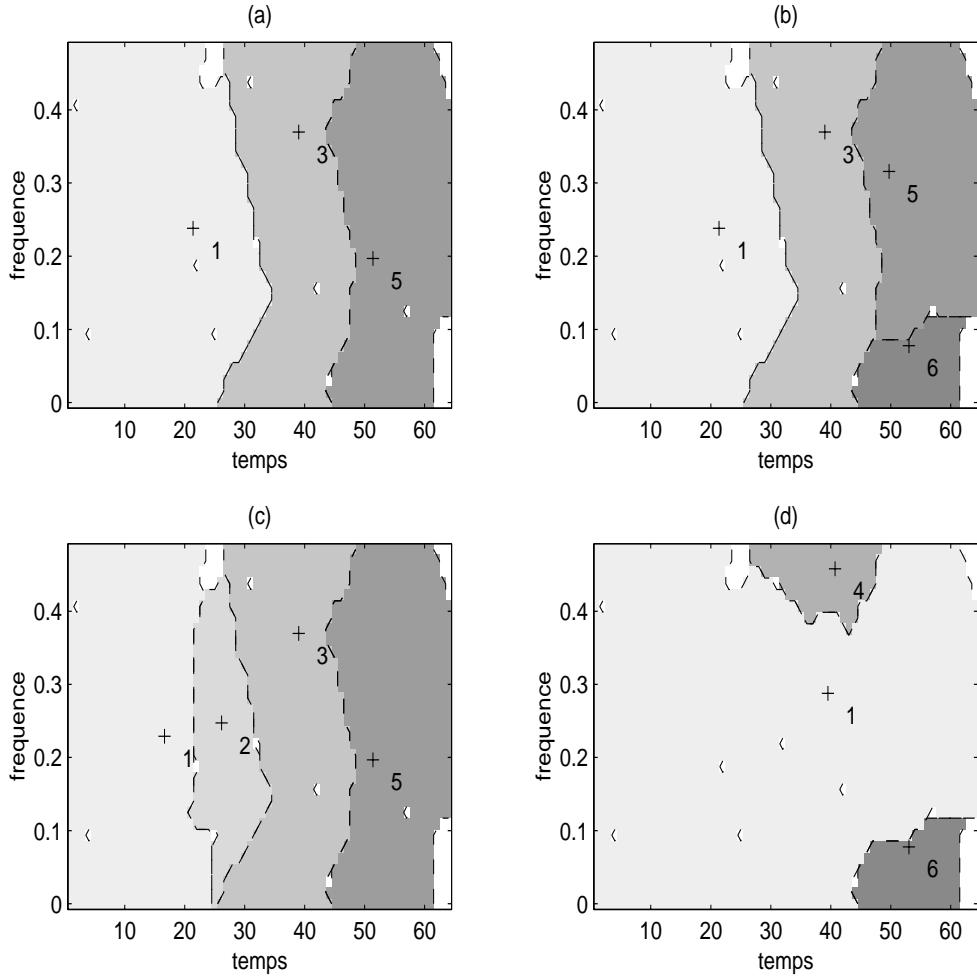


FIG. 3.15 – **Fusion de partition temps-fréquence.** Cette figure rassemble les partitions obtenues après traitement de la partition en Fig. 3.14 par les algorithmes de fusion proposés. (a) Algorithme 1, critère d_2 , (b) Algorithme 1, critère d_3 , (c) Algorithme 1, critère d_4 , (d) Algorithme 2, critère d_5 . Ce dernier seulement donne une solution acceptable en regroupant les composantes 1, 2, 3 et 5, toutes clairement associées au signal.

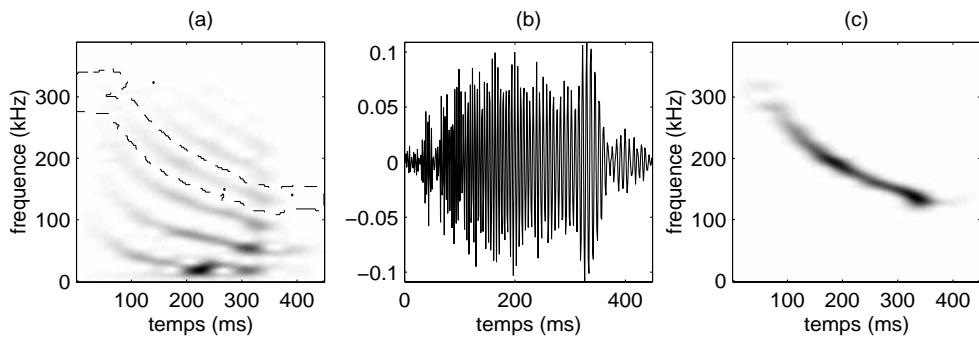


FIG. 3.16 – *Application de la partition temps-fréquence – Extraction de modes de propagation d'une onde acoustique dans un guide d'onde.* On observe le signal reçu après propagation d'une onde de choc (explosion de surface) dans un guide d'ondes (canal aquatique) pendant une distance de 14 km. (a) spectrogramme du signal observé. Ce signal est clairement composé d'une superposition de chirps qui correspondent, chacun, à un mode de propagation différent. À l'aide de l'algorithme de partition, on extrait un de ces chirps pour le caractériser précisément et le comparer à des attentes théoriques. (b) Quatrième harmonique extrait à partir de la région indiquée en pointillé en (a). (c) Spectrogramme du signal extrait.

3.4 Conclusions sur la réallocation différentielle et la partition temps-fréquence

Une analogie aux systèmes physiques supportée par une description géométrique du champ de réallocation nous a amené à introduire une nouvelle version de la réallocation : *la réallocation différentielle*.

À partir de cette nouvelle méthode, nous avons proposé un algorithme qui permet d'obtenir un découpage du plan temps-fréquence dans lequel chaque partie contient ce que nous appelons une *composante*. Pour ce faire, nous utilisons une information directement issue du champ des vecteurs de réallocation du spectrogramme, lui-même lié à un certain potentiel. Chaque composante peut alors être décrite par un domaine temps-fréquence (un puits de potentiel) et un attracteur (minimum local dans ce puits). Nous détectons donc autant de composantes qu'il existe de minima locaux dans le potentiel prescrit. Nous montrons que cet algorithme peut s'appliquer à l'extraction de chacune des composantes d'un signal, ainsi qu'à des problèmes de débruitage.

Nous positionnons notre algorithme par rapport à d'autre solution existentes : nous montrons que

- dans certaines conditions particulières, la méthode proposée devient équivalente à une partition en polygônes de Voronoï du plan temps-fréquence,
- la partition obtenue est cohérente avec les résultats donnés par les méthodes basées sur des mesures d'information dans le plan temps-fréquence.

Nous nous attachons ensuite aux difficultés liées à la présence de plusieurs minima locaux (causée par exemple par la présence de bruit) dans une zone que l'intuition aurait marquée comme une seule et même composante. Une partie du travail qui a été accomplie sur la régularisation du champ des

vecteurs de réallocation a été passée sous silence. L'idée est d'appliquer un traitement (une opération comparable à un lissage) sur le champ de réallocation destiné à atténuer les perturbations dues au bruit (en augmentant la régularité) et par conséquent à limiter l'influence des minima locaux parasites dans la partition. Le caractère marginal de l'espace des champs de vecteurs de réallocation a fait que cette approche s'est pour l'instant soldé par des échecs. Il reste cependant encore des voies de recherche à approfondir pour y donner solution. Néanmoins, pour le même problème, nous proposons une stratégie alternative qui consiste à fusionner les composantes non pertinentes de la partition. Nous concluons finalement qu'entre tous les critères de détection des mauvaises composantes envisagés celui qui procède à l'examen du spectrogramme le long de la frontière entre deux partitions conduit aux meilleurs résultats.

3.5 Réallocation supervisée

Dans les situations bruitées (bruit large bande), la réallocation telle qu'elle a été introduite par Kondera *et al.* [67] présente des inconvénients. En effet, elle crée des paquets d'énergie concentrés dans les régions associées au bruit uniquement (régions "bruit seul") alors qu'on y aurait espéré une distribution lisse. De plus, alors qu'il peut être prouvé que le spectrogramme d'un chirp est indépendant de la fenêtre d'observation, cela n'est plus vrai pour les mélanges chirp+bruit. Dans cette section, nous proposons une amélioration de la réallocation du spectrogramme qui

1. préserve une bonne concentration temps-fréquence pour les signaux modulés en fréquence,
2. n'applique pas le principe de réallocation dans les régions "bruit seul",
3. réduit la dépendance de l'analyse à l'égard de la fenêtre d'analyse.

Nous présenterons dans un premier temps le principe de cette amélioration. Nous en discuterons ensuite sa mise en œuvre et ses performances.

3.5.1 Superviser la réallocation

La méthode proposée se divise en deux étapes :

1. La première consiste à discriminer les régions "signal+bruit" des régions "bruit seul" en extraignant de l'information du champ des vecteurs de réallocation. Le principe de détection repose sur l'observation suivante : lorsque la longueur de la fenêtre d'analyse change, les vecteurs de réallocation dans le voisinage d'un signal déterministe obéissent à une certaine loi d'évolution tandis que, dans les régions "bruit seul" leur évolution devient erratique. Plus précisément, pour faire la discrimination, nous allons nous appuyer sur l'observation de la répartition des angles des vecteurs de réallocation pour différentes tailles de fenêtre. Pour donner une autre interprétation à la démarche présentée ici et en parlant abusivement, on pourrait dire que l'on compense le fait que l'on dispose d'une seule réalisation de bruit en observant le contenu énergétique dans le plan temps-fréquence du signal de différentes façons, i.e. à travers plusieurs fenêtres. On pourrait considérer les différents vecteurs de réallocation obtenus pour chacune des fenêtres utilisées comme autant d'épreuves d'une même variable aléatoire. Notons bien que cet argument a pour seule utilité de donner un nouvel éclairage aux idées données ici.
2. La seconde étape exploite complètement le contexte d'analyse multi-fenêtre. Elle consiste à mélanger la collection des spectrogrammes et des champs de vecteurs de réallocation associés,

qui ont été calculés lors de l'étape précédente. Le but est de réduire la dépendance de l'analyse vis-à-vis de la longueur de la fenêtre et aussi de stabiliser le processus de réallocation en faisant la moyenne des vecteurs de réallocation.

En conséquence, la distribution temps-fréquence qui résulte de cette nouvelle méthode est un mélange de spectrogrammes, modifié par la réallocation sélectivement appliquée dans les régions étiquetées en tant que régions “signal+bruit”. C'est ce que nous appellerons “supervision”.

Étape 1. Détection : le chirp linéaire comme modèle.

- *Fonctions de densité*

Soit la modulation linéaire de fréquence $s(t) = A \exp(i\beta t^2/2)$, nous avons déjà montré que le vecteur de réallocation adimensionné $r(t, \omega)$ du spectrogramme avec une fenêtre gaussienne s'écrit (cf. (2.13))

$$r(t, \omega) = \left(\frac{\hat{t} - t}{\Delta t_h}, \frac{\hat{\omega} - \omega}{\Delta \omega_h} \right)^t = \frac{\sqrt{2}(\omega - \beta t)}{\lambda(1/\lambda^4 + \beta^2)} \left(\beta, -\frac{1}{\lambda^2} \right)^t. \quad (3.103)$$

L'expression de l'angle θ du vecteur de réallocation se déduit directement par :

$$\tan \theta = \frac{\Delta t_h}{\Delta \omega_h} \frac{\hat{\omega} - \omega}{\hat{t} - t} = -\frac{1}{\beta \lambda^2}. \quad (3.104)$$

Si l'on considère θ comme une fonction de λ et en supposant que λ est uniformément distribuée entre λ_{min} et λ_{max} i.e., a une densité constante dans cet intervalle, nous obtenons, pour θ dans l'intervalle image $[\theta_{min}, \theta_{max}]$, la densité

$$\rho_\Theta(\theta) = \frac{C}{2\sqrt{|\beta|}} \left| \frac{1 + \tan^2 \theta}{\tan^{3/2} \theta} \right|, \quad (3.105)$$

où C est une constante de normalisation. Nous proposons d'utiliser (3.105) comme référence pour l'identification de la proximité d'un signal dans l'observation.

- *Critère de détection*

Nous effectuons en tout point (t, ω) du plan temps-fréquence, une estimation empirique (histogramme) de $\rho_\Theta(\theta)$ à partir d'une collection de champs de vecteurs de réallocation associés à des fenêtres de différentes longueurs. Une mesure de Kullback-Leibler est alors utilisé pour comparer la densité estimée $\hat{\rho}_\Theta(\theta)$ au modèle en éq. (3.105) :

$$d(\hat{\rho}_\Theta, \rho_\Theta) = \int \hat{\rho}_\Theta(\theta) \log \frac{\hat{\rho}_\Theta}{\rho_\Theta} d\theta. \quad (3.106)$$

Si cette distance dépasse un certain seuil (choisi sur des observations de bruit), on décide d'inhiber la réallocation.

Étape 2. Moyenne

Nous avons à notre disposition une batterie de spectrogrammes associés à des fenêtres de longueurs variées. Le problème est alors de savoir comment les combiner pour en renforcer les caractéristiques communes. Ceci revient à “moyenner” les différentes distributions, ce qui appelle la question

du choix arbitraire de la moyenne. Cette question est en fait assez générale et il peut être montré que la sélection d'une règle de moyenne spécifique est reliée avec le choix *a priori* d'une mesure de distance entre distributions temps-fréquence (voir sur ce point [71]). Dans le contexte qui nous intéresse, la solution exige quelques approfondissements qui n'ont pas été fait. L'approche que nous suivons ici est plus pragmatique et consiste à simplement faire la moyenne arithmétique (qui correspond à une distance L^2 entre distributions) des spectrogrammes et à faire la même chose avec les vecteurs de réallocation. On ne peut marier de manière cohérente et avec une simple moyenne arithmétique les spectrogrammes et les vecteurs de réallocation calculés avec les très grandes et très petites fenêtres. On choisit de garder la simplicité de la moyenne choisie, mais de ne faire la moyenne que dans un sous-ensemble des spectrogrammes et des vecteurs disponibles (en excluant les très grandes et les très petites fenêtres). Rappelons ici que le spectrogramme et les opérateurs de réallocation peuvent être vu respectivement comme une moyenne et les coordonnées d'un centre de masse évalués sur la distribution de Wigner Ville du signal dans un voisinage du plan temps-fréquence. Nous nous restreignons donc à faire la moyenne de quantités évaluées sur des voisnages semblables du plan temps-fréquence.

3.5.2 Mise en œuvre et résultats

Algorithm

La méthode que nous venons de décrire conduit à l'algorithme suivant :

1. calculer les N spectrogrammes à fenêtre gaussienne avec N longueurs de fenêtres distribuées uniformément, ainsi que les N champs de vecteurs de réallocation qui leur sont associés,
2. faire la moyenne des N spectrogrammes et champs de vecteurs de réallocation comme spécifié en Sect. 3.5.1,
3. Pour chaque point du plan temps-fréquence,
 - (a) estimer la densité empirique de l'angle des vecteurs de réallocation,
 - (b) calculer la distance de Kullback-Leibler (3.106) entre la densité de référence (3.105) et son estimée,
 - (c) si cette distance est plus petite qu'un certain seuil, réallouer la valeur du spectrogramme moyen avec le vecteur de réallocation moyen. Dans le cas contraire, laisser sur place la valeur du spectrogramme moyen.

Quelques images

La figure 3.17 illustre le fonctionnement de l'algorithme proposé dans le cas du mélange d'un chirp linéaire avec du bruit, le rapport signal à bruit (RSB) étant fixé à 10dB. La famille de fenêtres utilisée est une suite de 30 fenêtres gaussiennes (éq. (1.40)) avec une taille λ uniformément répartie entre $\lambda_{min} = 1.8$ et $\lambda_{max} = 7$. Les trois fenêtres de longueur les plus proches de la valeur centrale de l'intervalle $[\lambda_{min}, \lambda_{max}]$ sont sélectionnées pour calculer les spectrogramme et vecteurs de réallocation moyens. Il est clair sur la carte de supervision (seuil de décision fixé à 0.5) que la proximité d'un chirp est détectée dans une bande entourant la ligne de fréquence instantanée. À l'extérieur de cette bande, i.e., dans le bruit, on décide majoritairement de ne pas réallouer. Cela se traduit sur la distribution finale par une bonne représentation du signal (l'énergie est bien concentrée autour du chemin de fréquence instantanée que sur la moyenne des spectrogrammes réalloués) et du bruit (dans les

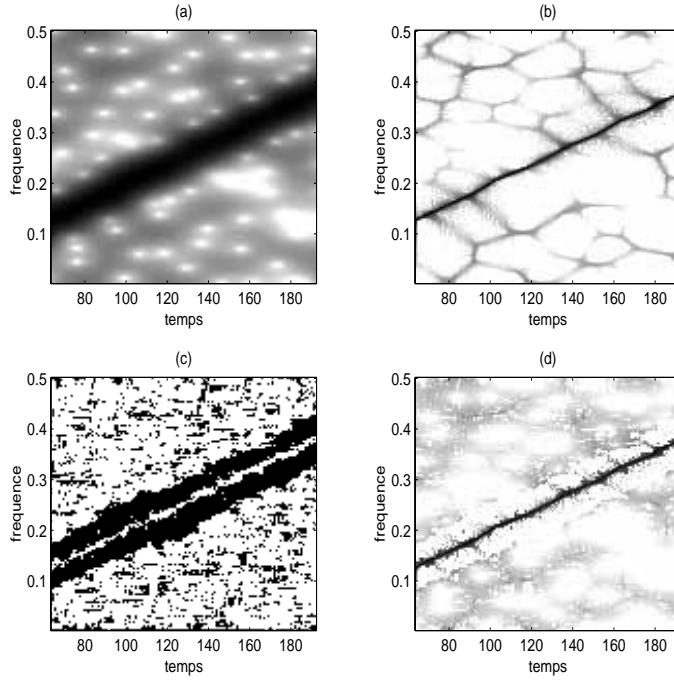


FIG. 3.17 – **Réallocation supervisée d’un chirp linéaire.** Chirp linéaire bruité, RSB = 10dB (voir le texte pour les détails), **(a)** Moyenne arithmétique des spectrogrammes, **(b)** Moyenne arithmétique des spectrogrammes réalloués, **(c)** Carte des distances de Kullback–Leibler : la région noire (resp. blanche) correspond aux points où la supervision autorise (resp. inhibe) la réallocation, **(d)** La distribution d’énergie temps-fréquence résultante (la dynamique est donnée en échelle logarithmique).

régions “bruit seul”, la distribution finale est une fonction lisse) relativement aux spectrogrammes et spectrogrammes réalloués moyens.

Cet exemple est basé sur un modèle simple de signal monocomposante modulé en fréquence, mais la procédure peut être généralisée à des situations plus compliquées (signaux multicomposantes modulés non linéairement en fréquence) étant admis que le modèle reste valide localement à l’échelle de chacune des fenêtres. Par exemple, la figure 3.18 montre la robustesse du processus de supervision dans le cas où le signal analysé est un chirp en loi de puissance correspondant au modèle newtonien des ondes gravitationnelles issues de la coalescence d’un système de deux objets astrophysiques massifs [93] (sur ce point, on se reportera au chapitre 4). Les longueurs des 47 fenêtres utilisées sont, cette fois, réparties entre $\lambda_{min} = 0.2$ et $\lambda_{max} = 6$. On sélectionne de la même façon, trois fenêtres pour les spectrogramme et vecteurs de réallocation moyens. Pour compenser l’écart au modèle, on doit augmenter légèrement le seuil de détection (fixé maintenant à 0.8). La détection donne alors de bons résultats mais ses performances sont diminuées lorsque β tend vers 0 ou l’infini : lorsque $\beta = 0$, fréquence pure, (resp. $\beta \rightarrow \infty$, impulsion, l’angle $|\theta|$ passe brutalement de $\pi/2$ (resp. 0) à 0 (resp. $\pi/2$) quand la taille de fenêtre augmente. Un échantillonnage très fin de la gamme de longueurs de fenêtre est alors nécessaire pour que l’angle θ décrivent toutes les valeurs entre 0 et $\pi/2$. Il serait donc important d’ajouter une prise en compte spécifique de ses cas dégénérés par le détecteur.

Performance

Nous proposons deux tests pour évaluer l’efficacité de la réallocation supervisée.

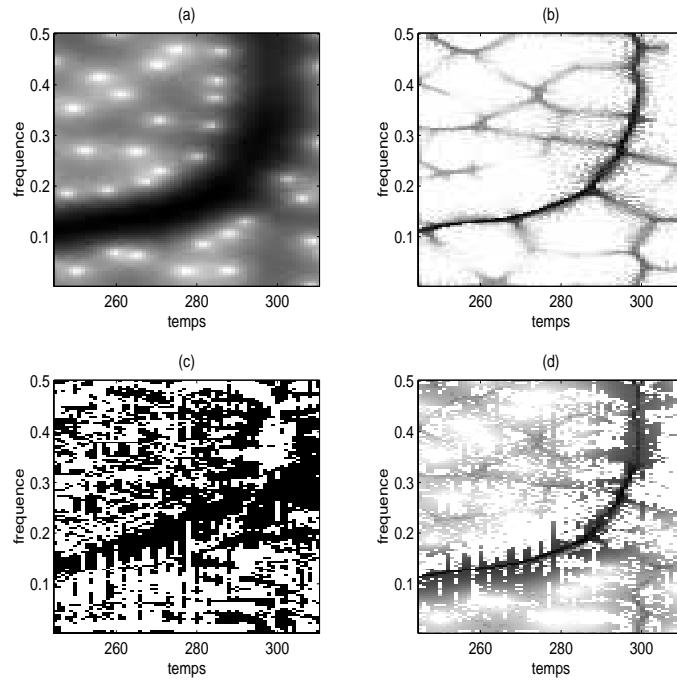


FIG. 3.18 – **Réallocation supervisée d'un chirp en loi de puissance.** Chirp en loi de puissance bruité, RSB = 10dB (voir le texte pour les détails) **(a)** Moyenne arithmétique des spectrogrammes, **(b)** Moyenne arithmétique des spectrogrammes réalloués, **(c)** Carte des distances de Kullback–Leibler: la région noire (resp. blanche) correspond aux points où la supervision autorise (resp. inhibe) la réallocation, **(d)** La distribution d'énergie temps-fréquence résultante (la dynamique est donnée en échelle logarithmique).

- *Courbes COR*

Les Caractéristiques Opérationnelles du Récepteur (COR) [96] de notre règle de décision ont été évaluées et sont présentées en Fig. 3.19. On peut remarquer que pour un RSB inférieur à -7dB , la règle de détection proposée n'est pas meilleure qu'un tirage aléatoire. Chose plus importante, il s'avère que la probabilité de fausse alarme est approximativement liée à une valeur du seuil de décision.

- *Critères de qualité de la représentation*

Pour donner des informations supplémentaires au sujet de la qualité de la représentation finale, la Fig. 3.20 exhibe deux critères :

1. Critère 1. La différence entre l'énergie totale du signal sans bruit et l'intégration de la distribution temps-fréquence (du signal bruité) le long de la ligne de fréquence instantanée quantifie comment le chirp initial est représenté.
2. Critère 2. L'entropie de Shannon de la distribution d'énergie 2D (normalisée) quantifie la platitude de la distribution dans les régions "bruit seul".

Ce qui est mis en évidence par ces deux critères, est que l'utilisation de la supervision conduit à un résultat intermédiaire entre celui obtenu par les spectrogrammes conventionnel ou totalement réalloué. La supervision assure donc un compromis entre ces deux situations extrêmes.

3.5.3 Conclusions sur la supervision

Nous présentons un prolongement de la méthode de réallocation dont l'objectif principal est l'obtention d'une distribution temps-fréquence qui donne à la fois une bonne représentation du signal et du bruit. Cette distribution est obtenue par un audit de la situation en chaque point temps-fréquence, chargé de décider si oui ou non il est opportun de réallouer. La décision s'effectue à partir d'information extraite des variations des vecteurs de réallocation du spectrogramme lorsque la taille de la fenêtre d'analyse change. On profite pleinement du cadre multi-fenêtre en faisant agir cette réallocation *supervisée* sur une moyenne de spectrogrammes associés à des fenêtres de tailles différentes.

Il serait important de comparer les résultats obtenus ici avec ceux obtenus par des méthodes concurrentes comme, par exemple, les distributions à noyau adaptatif radiallement gaussien [63] ou plus récemment les distributions modifiées par des équations de diffusion dans le plan temps-fréquence [49]. Le positionnement par rapport à l'optimal du détecteur de chirps que nous mettons en œuvre lors de l'étape de supervision serait également nécessaire.

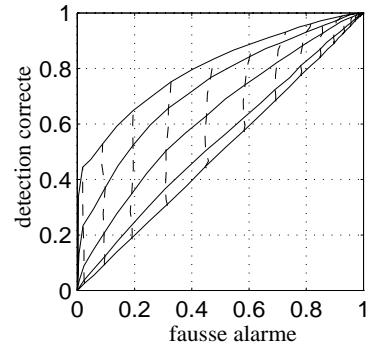


FIG. 3.19 – **Performances de la supervision : Courbes COR.** Ces courbes COR sont associées aux RSB = −7, 0, 7, 13 et 20dB (de bas en haut). Les lignes en pointillé relient les points obtenus avec le même seuil.

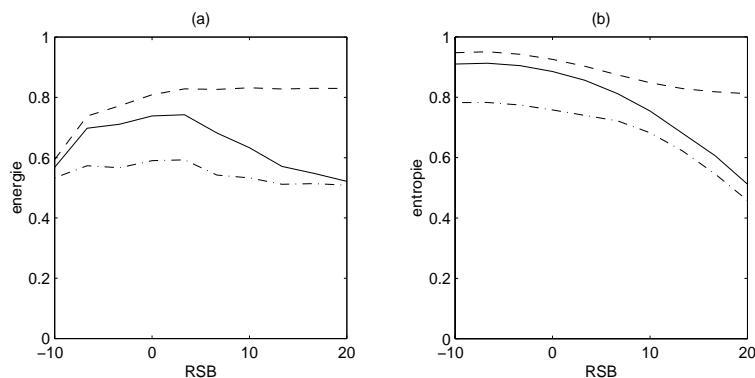


FIG. 3.20 – **Performance de la supervision : critères de qualité de la représentation.** (a) Critère 1 (voir 3.5.2), (b) Critère 2 (voir 3.5.2). Ligne pleine : spectrogramme réalloué avec supervision, ligne pointillée : spectrogramme moyen, ligne mixte : spectrogramme réalloué moyen

