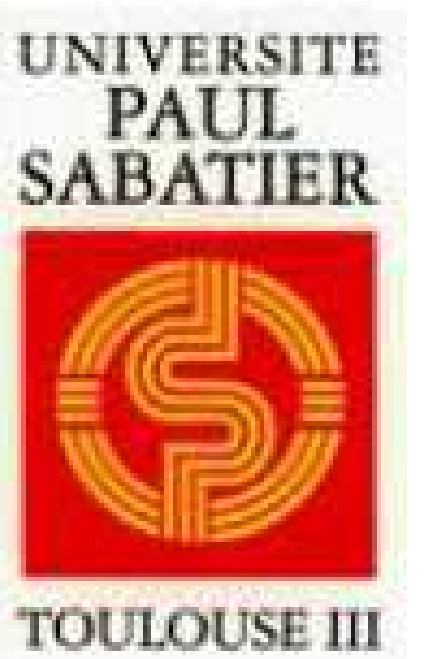


UTILISATION DE LA PHASE POUR L'AMELIORATION DE LA LOCALISATION TEMPORELLE ET FREQUENTIELLE DE L'ANALYSE SPECTROGRAPHIQUE

Le **Spectrogramme d'Amplitude et de Fréquence Instantanée Réalloué (SAFIR)** est un outil de représentation des signaux audio reprenant le principe du spectrogramme. Plusieurs opérations de **réallocation de la Transformée de Fourier à Court Terme (TFCT)** rendent la représentation temps-fréquence plus fine, plus souple et plus intuitive.



V. Emiya^{*(1)(2)}, B. David⁽¹⁾, V. Gibiat⁽²⁾
⁽¹⁾Ecole Nationale Supérieure des Télécommunications, Paris
⁽²⁾Université Paul Sabatier, Toulouse
 E-mail : {valentin.emiya, bertrand.david}@enst.fr, gibiat@cict.fr



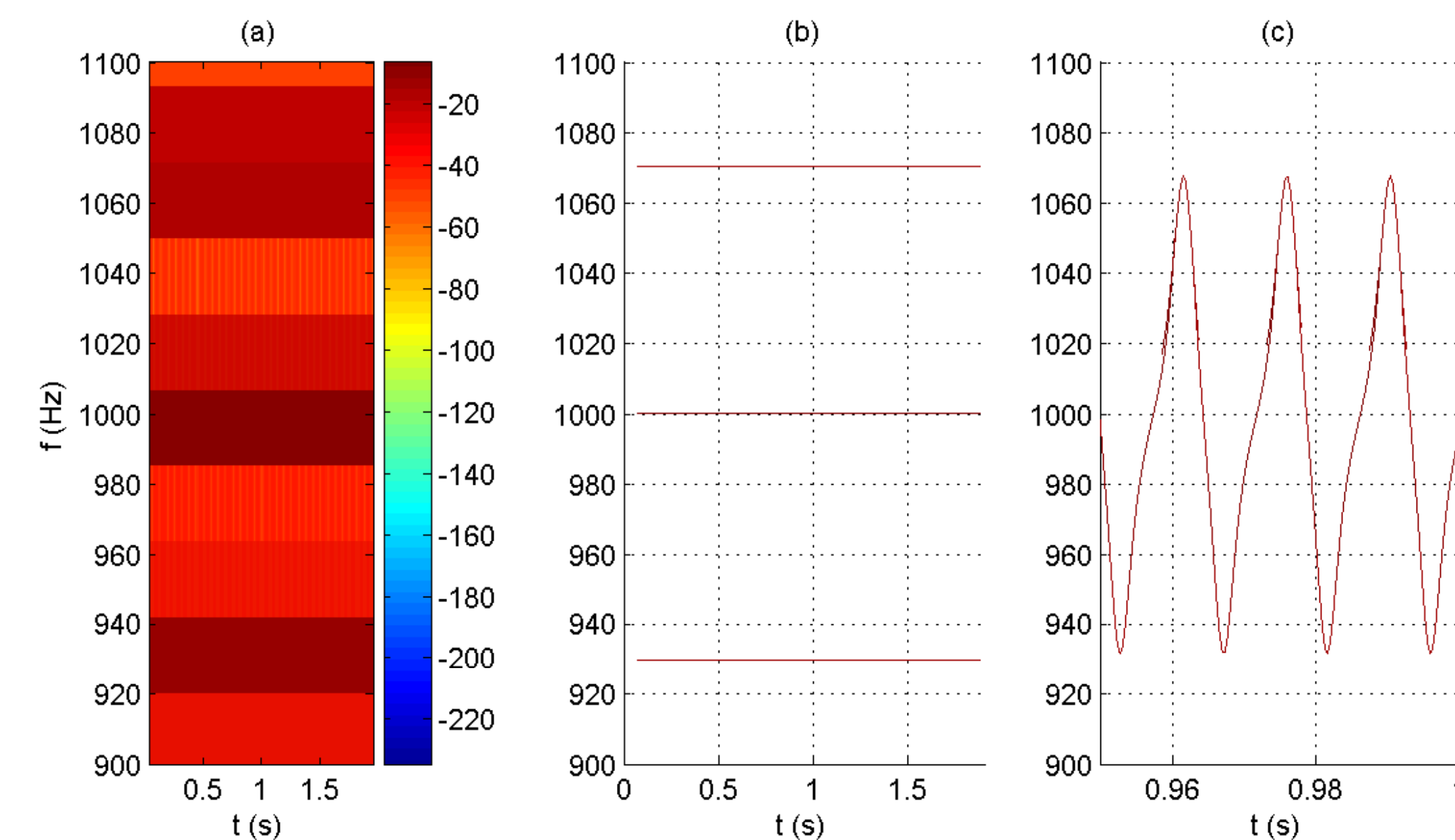
Principe du SAFIR

– Tout comme le spectrogramme, le SAFIR repose sur la représentation de l'énergie (module au carré) de la TFCT $X(n, k)$ calculée sur N_{fft} points à partir de N échantillons du signal x pondérés par une fenêtre w :

$$X(n, k) = \frac{1}{N} \sum_{l=k-\frac{N-1}{2}}^{k+\frac{N-1}{2}} x(l)w(l-k+\frac{N+1}{2})e^{-2i\pi\frac{nl}{N_{fft}}}$$

$$= |X(n, k)|e^{i\Phi(n, k)}$$

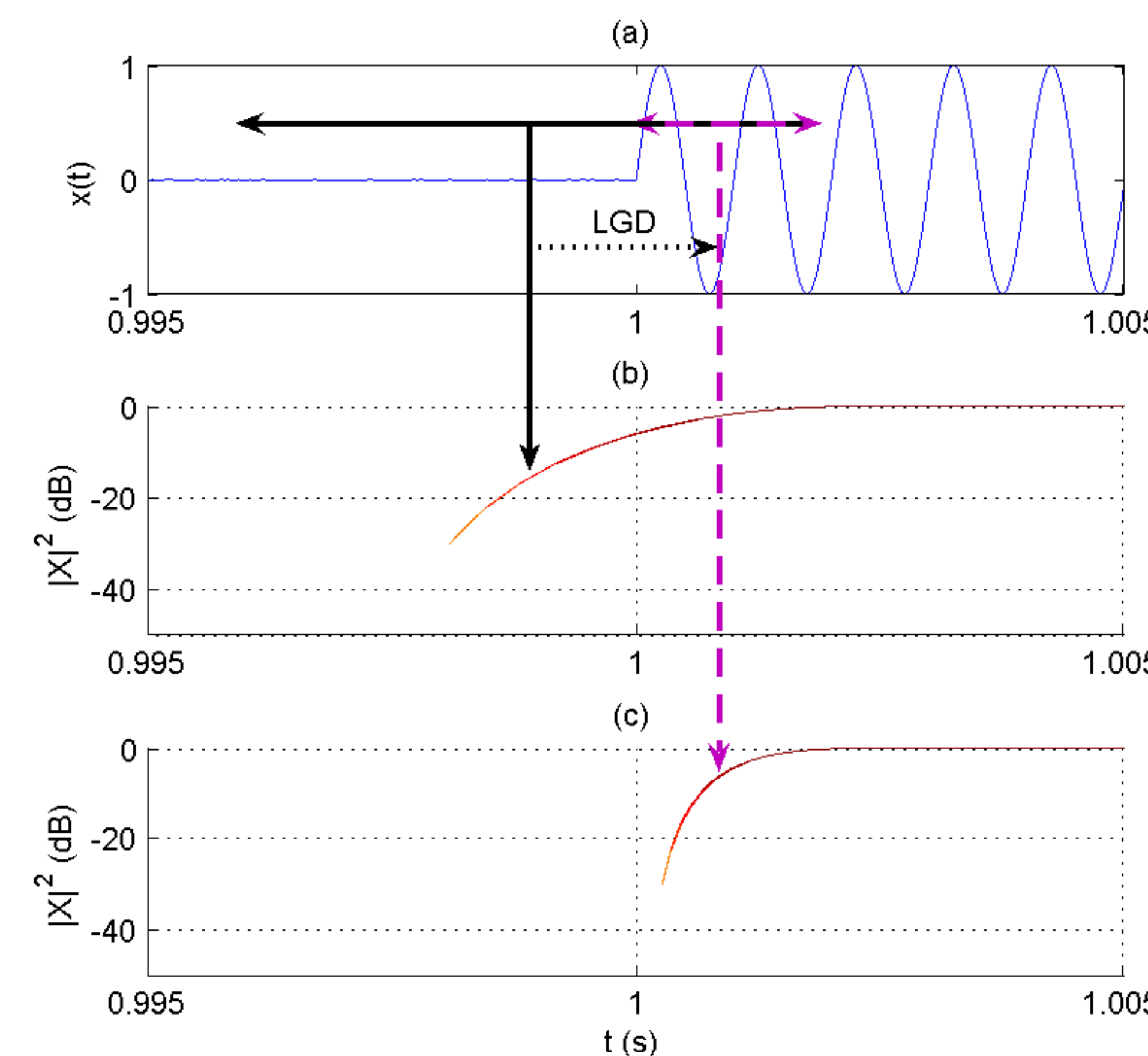
– **Réallocation fréquentielle** : dans chaque canal fréquentiel n , la TFCT est le résultat d'un filtrage passe-bande du signal initial. La **fréquence instantanée** $f_i(n, k) = \frac{1}{2\pi} \frac{d\Phi}{dk}$ de la TFCT extrait la fréquence (ou la combinaison des fréquences) éventuellement présente dans cette bande. Si les bandes sont suffisamment étroites (i.e. pour N suffisamment grand), toutes les fréquences peuvent être isolées.



Réallocation fréquentielle.

Représentations du signal $x(t) = \sin(2\pi 930t) + \sin(2\pi 1000t + \phi_0) + \sin(2\pi 1070t)$ échantillonné à 44.1kHz. Le spectrogramme (a) est réalloué en fréquence en conservant des bandes étroites afin de séparer les composantes (b) ($N = 4096$) ou en prenant des bandes larges pour observer la modulation résultante (c) ($N = 64$).

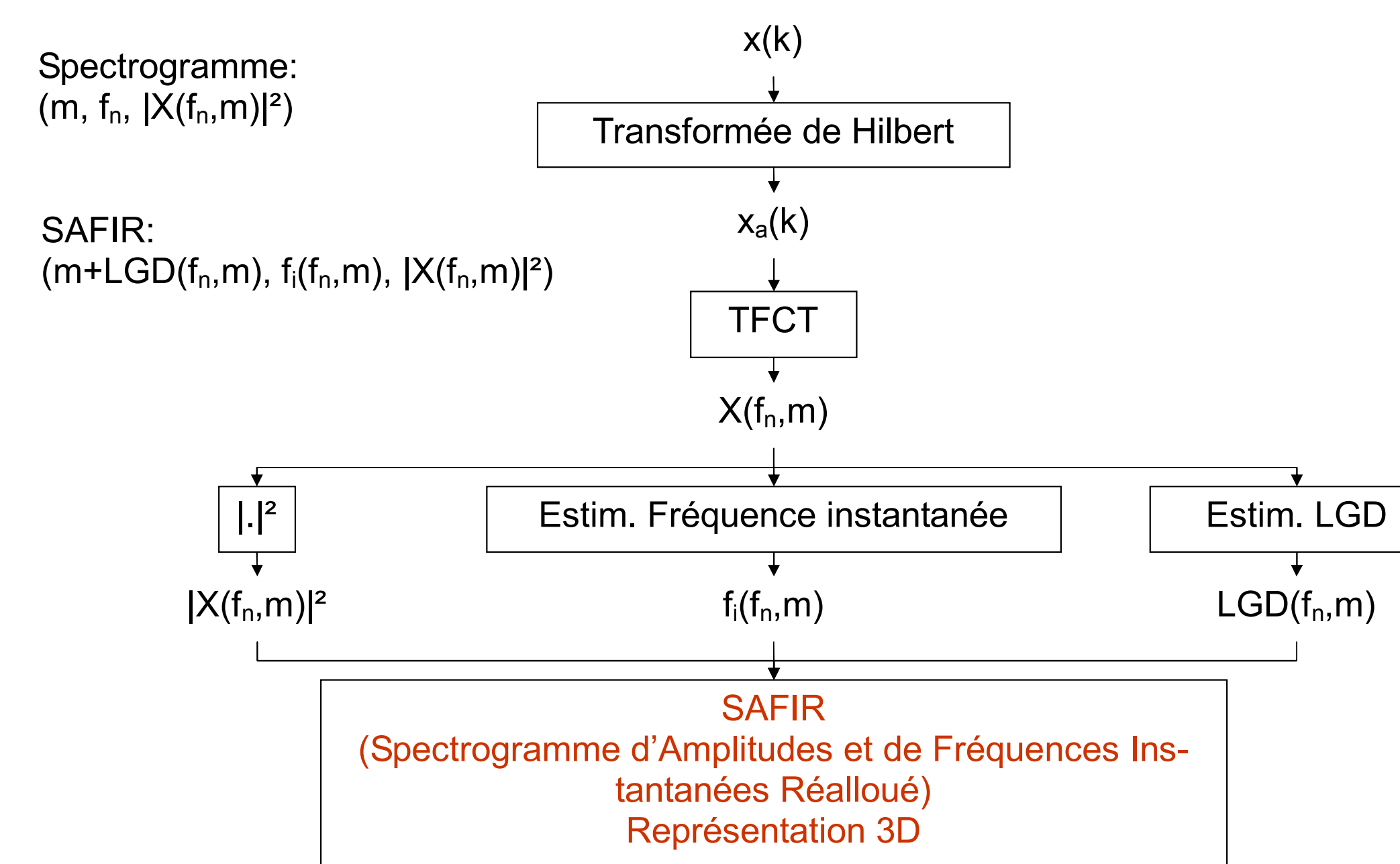
– **Réallocation temporelle** : un coefficient de la TFCT est issu de la contribution de N échantillons successifs. Dans le cas du spectrogramme, l'énergie calculée est attribuée au milieu de la fenêtre d'analyse. Cette approximation a tendance à étaler l'énergie dans le temps et introduit des effets gênants comme le pré-écho. L'instant choisi dans le SAFIR correspond à l'instant moyen de localisation de l'énergie dans la fenêtre d'analyse. Il est donné par le **retard de groupe** $LGD(n, k) = -\frac{1}{2\pi} \frac{d\Phi}{dn}$.



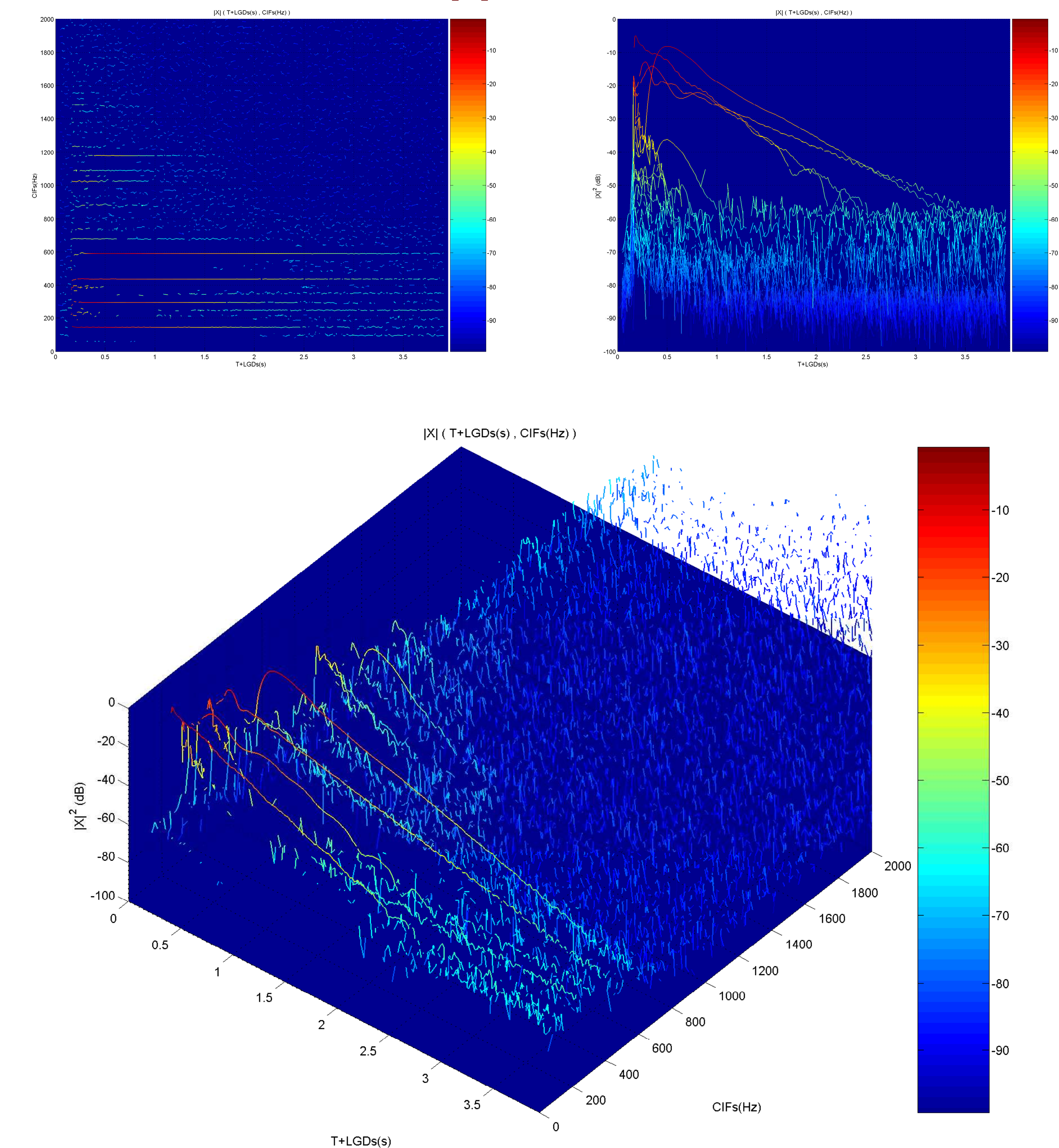
Réallocation temporelle.

Le pré-écho (b) obtenu par l'analyse d'une partie non-stationnaire du signal (a) est évité en localisant l'énergie des coefficients de la TFCT grâce au retard de groupe (c).

Schéma du SAFIR



Application



SAFIR d'un son de steel-drum (instrument de percussion caribéen aux notes accordées) : vue en 3 dimensions (en bas), projection dans le plan {fréquence, temps} (à gauche) et projection dans le plan {amplitude, temps} (à droite). Les différentes composantes sont localisées avec précision, en fréquence, en amplitude et en temps. On observe des comportements caractéristiques de cet instrument en termes de fréquence des composantes, de transitoire d'attaque, de retard d'apparition et de décroissance de l'énergie.

Conclusion et perspectives

- SAFIR offre une représentation améliorée de la TFCT avec un faible temps de calcul.
- La réallocation repousse certaines approximations du spectrogramme.
- Le choix de la longueur de la fenêtre d'analyse, principal paramètre de SAFIR, permet d'envisager des analyses en bandes étroites (séparation des composantes) ou en bandes larges (mise en évidence de modulations d'amplitude et de fréquence).