



COORDINATEUR: Olivier DERRIEN

PARTENAIRES: LMA (F) – ARI (A)

Abstract: The goal of this project is on one hand to study new time-frequency representations for audio signals, with good mathematical properties and accurately adapted to auditory perception, on the other hand to study extensively the phenomenon of time-frequency auditory masking, and finally the application of this knowledge to the improvement of perceptual audio codecs (like MP3 or AAC).

► **OBJECTIFS DU PROJET**

Une **représentation optimale des signaux audio** doit permettre de décrire l'information utile avec un **minimum de données** tout en suivant au plus près la **perception auditive**. Les représentations temps-fréquence actuelles (Gammatone, MDCT ...) ne sont optimales que selon un seul aspect. Ce projet a pour but de répondre à :

- Peut-on obtenir une classe de représentations temps-fréquence bien adaptées à la perception auditive ?
- Peut-on rendre ces représentations inversible et non-redondante, ou a contrario fortement parcimonieuses ?
- Peut-on améliorer les algorithmes de codage audio perceptif (type MP3 ou AAC) ?

3. La mise au point de nouveaux **schémas de codage compressif** : Quantification non-linéaire, mise en forme spectrale, codage entropique. Un **schéma de codage complet pour décompositions parcimonieuses** a été développé, optimisé pour les bas débits.

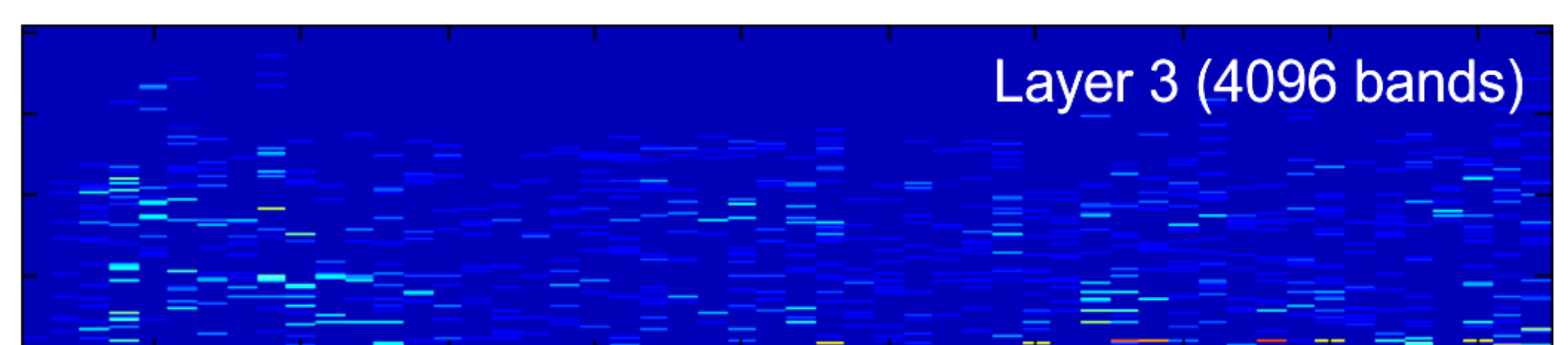
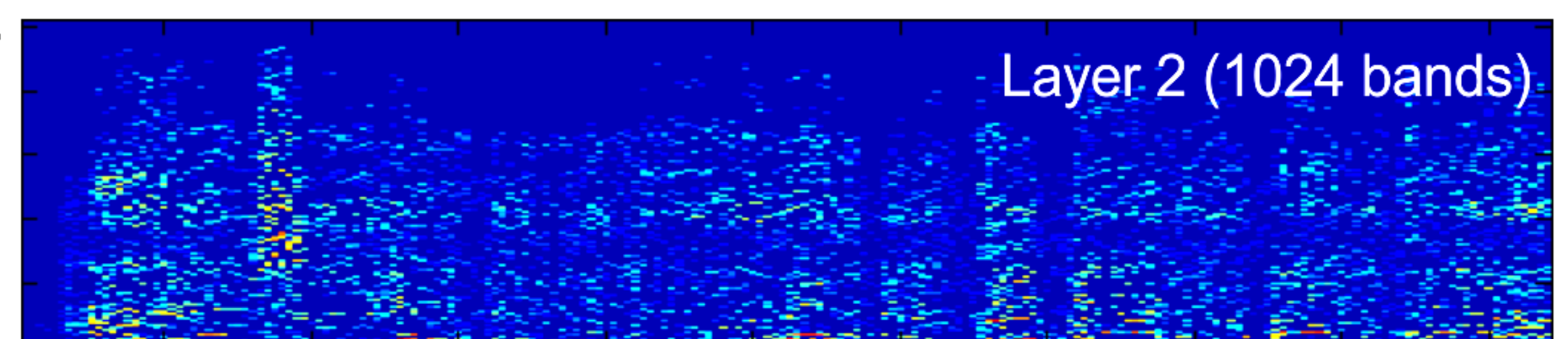
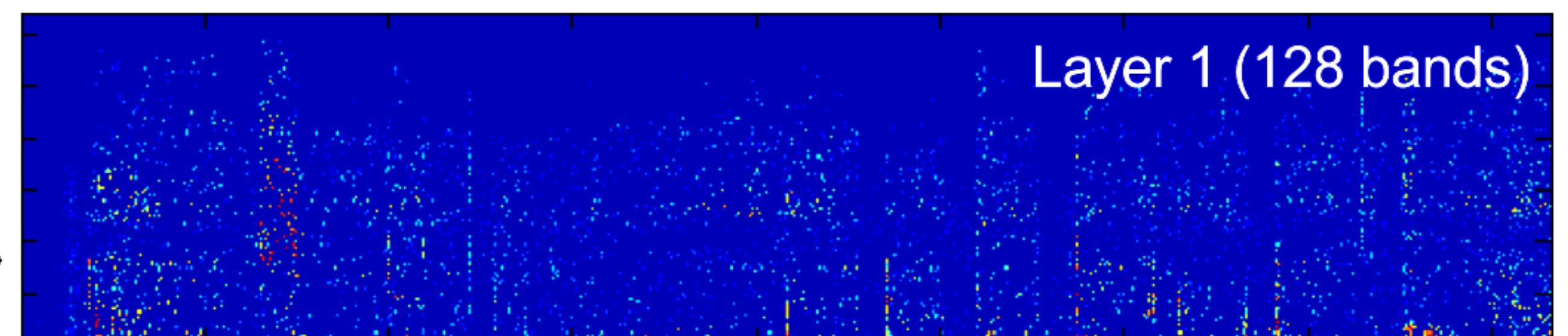
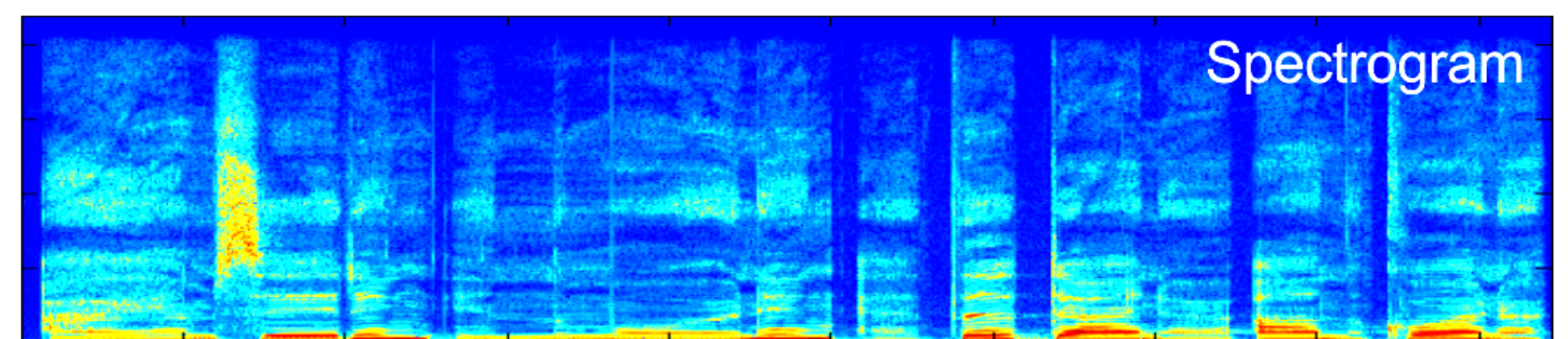
► **MÉTHODOLOGIE ET RÉSULTATS**

Le projet a été structuré en trois tâches :

1. L'étude de **nouvelles transformées temps-fréquence** inversibles et perceptives : **ERBLet** et **ERB-MDCT**, ainsi que de **nouvelles décompositions parcimonieuses** : **Unions de MDCT avec optimisation d'une mesure de distorsion perceptive**.
2. La mise au point d'un **modèle d'audition temps-fréquence** le plus réaliste possible : Un **noyau de masquage temps-fréquence** partiel a été mesuré.

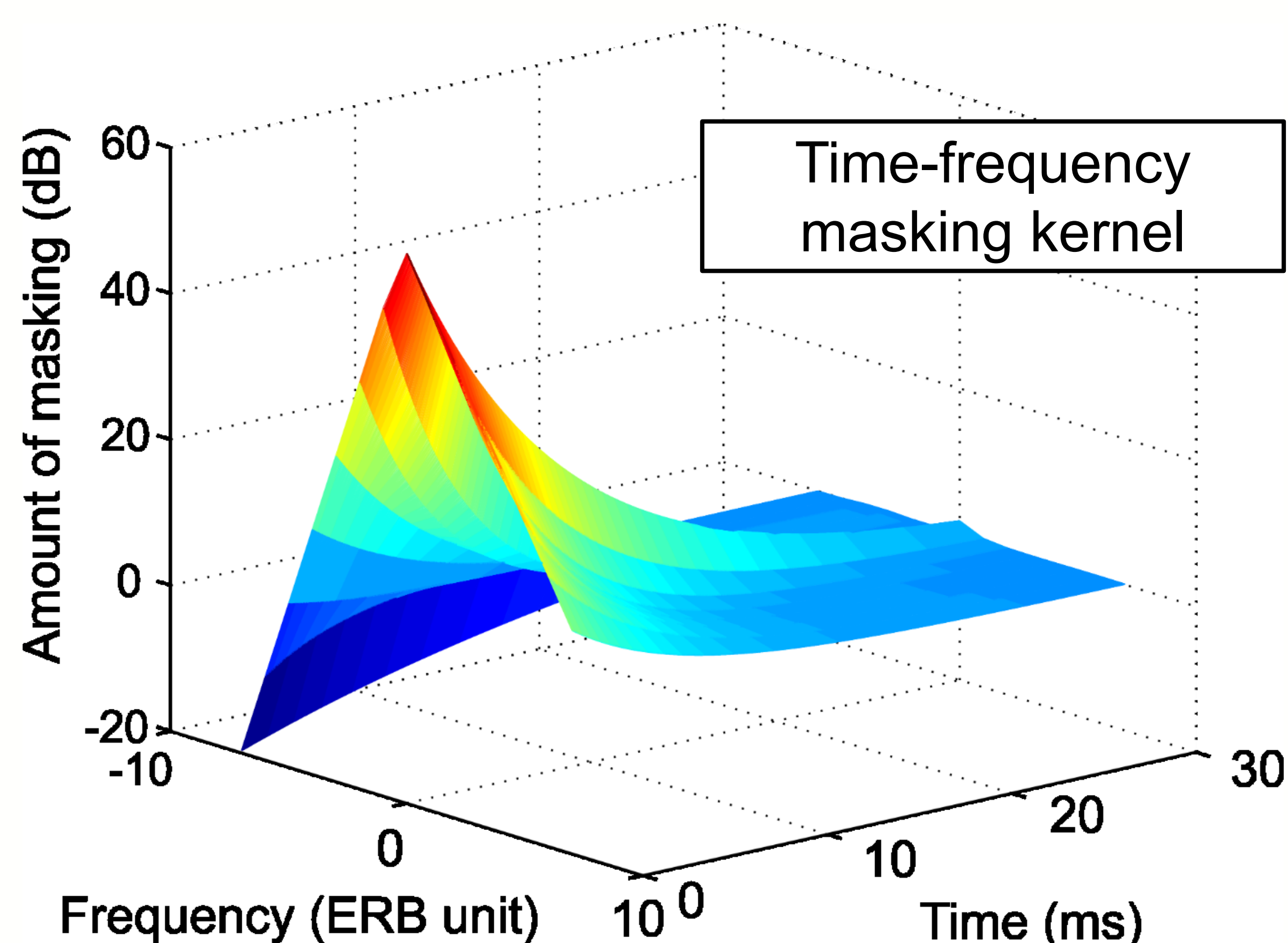
Sparse decomposition on 3 MDCT bases (layers) using a perceptual time-frequency model

Total sparsity = 30%



Frequency

Time



► **CONCLUSION ET PERSPECTIVES**

- Les nouvelles décompositions temps-fréquence apportent de **nouveaux outils** dont les applications vont être progressivement développées.
- La mise au point du **modèle d'audition** est en cours.
- Les **schémas de codage** testés ont montré qu'ils sont **plus performants que l'état de l'art (AAC HEv2 / AAC+)** à bas débit : **16 - 32 kbps (pour 1 canal)**.

