ÉCOLE DE TECHNOLOGIE SUPÉRIEURE
UNIVERSITÉ DU QUÉBEC

MÉMOIRE PRÉSENTÉ À
L'ÉCOLE DE TECHNOLOGIE SUPÉRIEURE

COMME EXIGENCE PARTIELLE
À L'OBTENTION DE LA
MAÎTRISE EN GÉNIE ÉLECTRIQUE
M.Ing.

PAR
Hicham MAHKOUM

TECHNIQUE DE PARCELLISATION ET DE LOCALISATION DES SOURCES CÉRÉBRALES À PARTIR DES SIGNAUX MEG

MONTRÉAL, LE 5 JANVIER 2012

PRÉSENTATION DU JURY

CE MÉMOIRE A ÉTÉ ÉVALUÉ

PAR UN JURY COMPOSÉ DE :

M. Jean-Marc Lina, directeur de mémoire
Département de génie électrique à l'École de technologie supérieure

M. Christian Gargour, président du jury
Département de génie électrique à l'École de technologie supérieure

M. Christophe Grova, membre du jury
Département de génie biomédical, Université de McGill

IL A FAIT L'OBJET D'UNE SOUTENANCE DEVANT JURY ET PUBLIC

LE 5 JANVIER 2012

À L'ÉCOLE DE TECHNOLOGIE SUPÉRIEURE

REMERCIEMENTS

J’aimerais dédier ce travail à l’ensemble des personnes qui ont contribué de près ou de loin à sa réalisation particulièrement à mon directeur de recherche Dr. Jean-Marc Lina pour son encadrement, sa patience et son assistance durant ce travail.

J’aimerais aussi remercier ma chère conjointe pour son support, sa présence et son soutien tout le long dans ce travail.

Je saisis aussi cette occasion pour dédier ce travail à mes chers parents et à tout le reste de ma famille.

TECHNIQUE DE PARCELLISATION ET DE LOCALISATION DES SOURCES CÉRÉBRALES À PARTIR DES SIGNAUX MEG

Hicham MAHKOUM

RÉSUMÉ

Ce travail consiste au développement d'une technique de résolution du problème inverse concernant la localisation de l'activité neuro-cérébrale à partir des données MEG (Magnétoencéphalographie). Cette technique se base sur la parcellisation sélective des sources du modèle corticale afin d'améliorer la régularisation du problème inverse, ensuite procéder à sa résolution.

La parcellisation sélective est basée sur deux principes; la pré-localisation des sources et la formation des parcelles à partir des sources sélectionnées. Pour la pré-localisation des sources on exploite la technique MSP (Multivariate Sources Prelocalisation) qui permet d'estimer le degré de l'activation des sources. Pour la sélection des sources on utilise la technique de test d'hypothèse FDR (False Discovery Rate) qui permet de définir un seuil de sélection sur les informations de pré-localisation des sources.

Après la parcellisation sélective, on procède à la résolution du problème inverse en exploitant la technique LCMV (Linearly Constrained Minimum Variance) adaptée pour un filtrage spatial sur des parcelles. Les résultats obtenus montrent que cette technique permet de localiser les sources avec une bonne sensibilité et spécificité.

Il aussi important de motionner que cette technique est modulaire; ainsi, on peut exploiter un module séparément des autres. Par exemple, on peut combiner la technique de parcellisation sélective avec la technique MEM (Maximum Entropy Method) (Amblard C 2004).

Mots clés : Localisation des sources, Problème inverse, Parcellisation, activité corticale et Technologie MEG.

TECHNIQUE DE PARCELLISATION ET DE LOCALISATION DES SOURCES CÉRÉBRALES À PARTIR DES SIGNAUX MEG

Hicham MAHKOUM

ABSTRACT

This work consists in the development of inverse problem resolution technique concerning the localisation of neuro-cerebral activity using MEG data. This technique, based on selective parcelling of cortical model sources, aims at providing a solution of the inverse problem after its regularisation.

The selective parcelling technique is based on two principles: sources pre-localisation and forming parcels using selected sources. For the sources pre-localisation, we used MSP (Multivariate Sources Prelocalisation) which helps estimating sources activity. Concerning sources selection we rely on hypothesis testing technique FDR (False Discovery Rate) which helps to define a selection threshold based on the sources pre-localisation data. ,

Based on results of selective parcelling technique, we then proceed to the inverse problem resolution using LCMV (Linearly Constrained Minimum Variance) technique which is adapted for Parcels spatial filtering. The results of using this technique of resolution show that we can achieve a good sensitivity and specify of activity sources localisation.

It is also important to mention that this technique is modular thus some parts of it can be re-used independently of the others. For example, the selective parcelling technique can be combined with MEM (Maximum Entropy Method) technique (Amblard C 2004).

Keywords : Sources localisation, Inverse problem, clustering, Cortical activity and MEG technology.

TABLE DES MATIÈRES

Page

LISTE DES FIGURES

Page

LISTE DES ABRÉVIATIONS, SIGLES ET ACRONYMES

MEG	Magnétoencéphalographie
EEG	Électroencéphalographie
IRM	Imagerie par résonance magnétique
IRMf	IRM fonctionnel
FDR	False Discovery Rate
MSP	Multivariate Sources Prelocalisation
LCMV	Linearly Constrained Minimum Variance
NMSE	Normalized Mean-Square Error
MEM	Maximum Entropy on the Mean
ROC	Receiver Operating Curve

INTRODUCTION

De nos jours, l'imagerie médicale est devenue essentielle pour obtenir une information visuelle de l'anatomie du corps humain voir même sa physiologie que les cliniciens sont capable d'interpréter. Cela permet aux cliniciens d'exprimer un diagnostique pour un problème médical. La neuroimagerie, par exemple, porte sur la reconstruction d'une image du système nerveux, en particulier le cerveau. On distingue deux classes de neuroimagerie cérébrale :

1- L'imagerie structurelle qui porte sur la reconstruction d'une image anatomique et l'identification de ses différentes structures cérébrales. Dans la pratique clinique médicale, l'image de l'anatomie cérébrale peut révéler des lésions et/ou des déformations d'une région cérébrale qui peuvent expliquer l'état médical du patient. Cette identification et localisation de lésions et de déformations permettent d'établir un diagnostique médical et parfois la recommandation d'une intervention chirurgicale.

2- L'imagerie fonctionnelle : ce type d'imagerie se base, en général, sur les résultats de la neuro-imagerie structurelle pour mesurer une activité physiologique dans une région cérébrale. Les images obtenues avec la neuro-imagerie fonctionnelle représentent une mesure d'activité cérébrale lue lorsque le sujet, sous observation, est entrain d'exécuter une tache déterminée (e.g. tache cognitive), ou activité spontanée.

Lorsqu'un sujet effectue une tâche comme, par exemple, la reconnaissance d'images, différents types d'activités neurophysiologiques sont observées. Pour mesurer ces activités neuronales, plusieurs technologies spécifiques ont été développées. Celles-ci ont pour objectifs de produire une image qui correspond à la mesure de l'activité cérébrale observée en indiquant sa localisation.

Les activités neurophysiologiques mesurables par différentes technologies d'imagerie fonctionnelles existantes sont les suivantes:

1- Le Signal BOLD (BLood-Oxygeneration-Level-Dependant) : il correspond à la consommation d'oxygène sanguin, dans le tissu neuronal, ce qui correspond aussi à une augmentation du flux sanguin. Pour mesurer cette activité, on utilise la technologie Imagerie par Résonance Magnétique fonctionnelle (IRMf) (Baillet 2001).

2- Le débit sanguin: il correspond au changement du débit sanguin dans une région en activité. Pour mesurer le débit sanguin, on utilise la technologie TEP (Tomographie par Émission de Positrons) qui consiste à injecter un traceur radioactif par voie intraveineuse et mesurer sa présence dans l'organe observé.

3- Potentiels électriques : contrairement aux activités précédentes, l'activité électrique cérébrale correspond aux potentiels électriques créés suite aux échanges électrochimiques entre les neurones de la surface corticale. La technologie utilisée pour mesurer cette activité électrique est l'EEG (Électro-Encéphalo Graphie).

4- Champs magnétiques : simultanément à l'activité électrique, le cerveau produit aussi une activité magnétique. La technologie de mesure de cette activité neuro-électromagnétique est la MEG (Magnéto-Encéphalo Graphie). Elle consiste à enregistrer à la surface de la tête d'un sujet les champs magnétiques, résultant d'une activité neuro-électrique. Les signaux magnétiques sont mesurés à travers des capteurs MEG (entre 75-250) qu'on appels Squids[1], placés au niveau de la tête.

Dans ce travail, la technologie de neuro-imagerie fonctionnelle qui nous intéresse est la MEG. Ce travail est une contribution au développement des techniques d'imagerie médicale et plus particulièrement, le développement d'une technique de localisation de l'activité neuronale au niveau cortical.

---

[1] **S**uperconducting **Q**uantum **I**nterference **D**evice **(SQUID)** se sont des supraconducteurs très sensibles exploitées comme magnétomètres ou capteurs MEG pour la mesure des champs magnétiques très faible.

Comme on peut le remarquer dans la figure 1.1 (http://www.4dneuroimaging.com/), la technologie MEG se distingue par sa très bonne résolution temporelle. Elle est aussi considérée parmi les technologies les moins invasives.

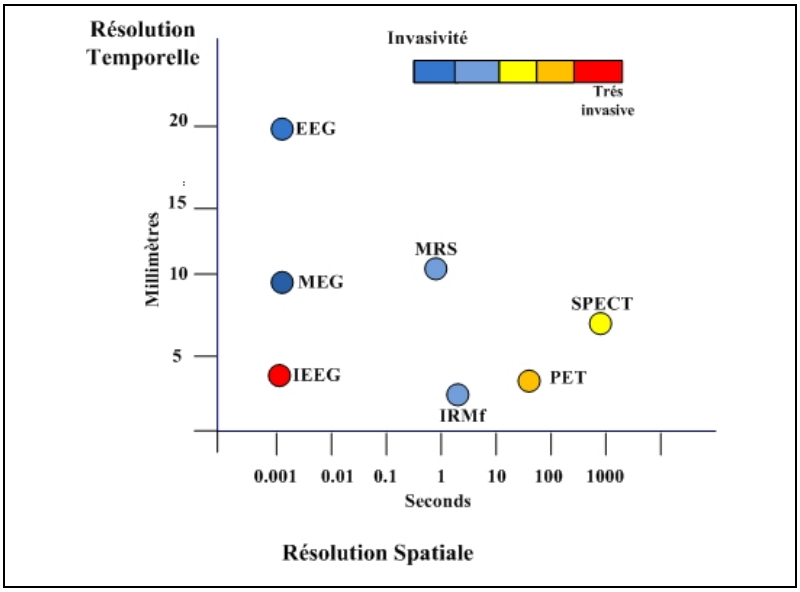


Figure 1.1 Résolution temporelle et spatiale des différentes technologies de neuro-imagerie

L'amplitude du signal MEG est considérablement faible, de l'ordre de 10 à 1.000 *fT* (femtoTesla), alors que le champ magnétique terrestre est de l'ordre de 10^7 *fT* et le bruit d'environnement (e.g. le bruit urbain) est de l'ordre de 10^8 *fT*. Pour mesurer le signal d'intérêt, la mesure doit être effectuée dans un environnement isolé de ce bruit.

L'activité neuronale mesurée par technologie MEG correspond principalement au courant électrique qui correspond à l'échange électrochimique entre les synapses (émettrices) d'une cellule nerveuse et les dendrites (réceptives) d'une autre cellule lors de la transmission d'un signal neuronal. Chaque synapse est capable de générer un courant d'environ ~20 *fAm*. A l'échelle cellulaire, ce courant est faible à mesurer. Toutefois, une région active d'une manière synchrone de 25 mm^2 peut générer un courant de l'ordre de ~10 *nAm*.

Comme illustré dans la figure 1.2 (Tosun, Rettmann et al. 2004), l'activité synchrone des neurones génère des courants électromagnétiques mesurables à la surface corticale. Ce signal est peu distordu par le scalp et la boite crânienne comparée au signal électrique mesuré par l'EEG.

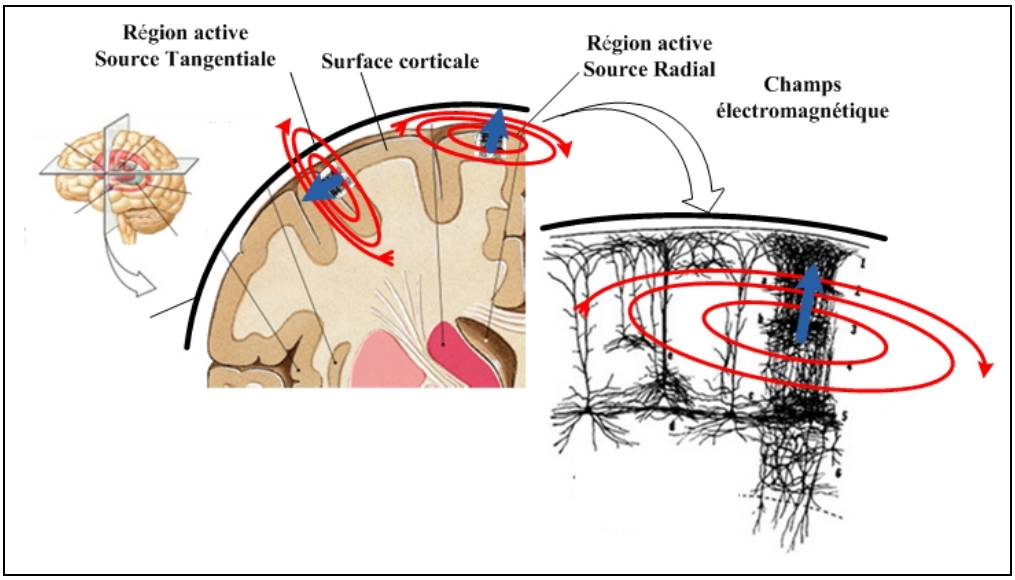


Figure 1.2 Région cérébrale et activité électromagnétique

Il est aussi important de noter qdue les cellules du cortex, en particulier les cellules dites pyramidales situées dans des régions perpendiculaires par rapport à la surface corticale, sont les principaux générateurs de cette activité électromagnétique. Comme on peut le remarquer sur la figure 1.2, elles peuvent être radiales ou tangentielles par rapport à la surface crânienne (ou scalp), précisément dans les régions de plis et sillons de la surface corticale. Ainsi, les champs provenant des colonnes tangentielles (au scalp) sont clairement mesurable à la surface du scalp, alors que les sources radiales sont difficilement "vues". En conséquence, l'activité dans les zones de plis (gyri) n'est pas facilement détectable par la MEG, ce qui est considéré comme une des principales limitations de cette technologie. Comme le montre cette même figure, un regroupement de neurones corticaux dans une région donnée peut être représenté par une source électromagnétique caractérisant la localisation, l'orientation et l'activité de cette région.

Dans le chapitre 1, on explique comment l'activité électromagnétique des neurones est modélisée en sources dipolaires. En fait, un regroupement de neurones est capable de générer un champ magnétique détectable, lorsqu'ils sont simultanément actifs. Ainsi, les neurones de la surface corticale peuvent être modélisés par des sources dipolaires caractérisées par leur localisation et leur orientation spatiales ainsi que leur amplitude.

En pratique, cette activité neuronale n'est mesurable qu'au niveau de la surface de la tête. A partir de ces mesures, on a besoin de déterminer la localisation cérébrale de la région active. Pour résoudre ce problème de localisation de l'activité corticale active à partir des mesures MEG, on formalise d'abord ce problème en modèle mathématique réaliste. Cette modélisation consiste essentiellement (1) à représenter l'activité neuronale sous forme de sources électromagnétiques caractérisées par leur positions, orientation et intensité et (2) à définir la relation entre les mesures (données MEG) et les sources à travers la définition d'une matrice de gain. Cette modélisation simplifiée se résume ainsi dans l'expression suivante :

$$\boldsymbol{M} = \boldsymbol{G}.\boldsymbol{J} + \boldsymbol{\eta}$$ Il s'agit du modèle linéaire avec contraintes anatomique (Dale A., Sereno M. 1993)

où $\boldsymbol{M}$ est la matrice des mesures (m x t), $\boldsymbol{J}$ correspond aux intensités des sources du modèle (n x t), $\boldsymbol{G}$ est la matrice de gain qui relie les sources et les mesures (m x n) et η représente le bruit (m x t). n est le nombre des sources, m est le nombre des capteurs et t est le nombre d'échantillons temporels.

A partir de cette modélisation on peut définir deux types de problème: le problème direct et le problème inverse. Le premier problème consiste à construire la matrice de gain $\boldsymbol{G}$ à partir de la modélisation de la surface corticale en sources (Baillet S. 2001). Cette matrice permet de calculer les mesures $\boldsymbol{M}$ étant donné une configuration d'intensités des sources $\boldsymbol{J}$. Le second correspond à la situation inverse : à partir des mesures MEG on veut retrouver les sources $\boldsymbol{J}$ actives. Dans ce cas, on considère que la matrice de gain $\boldsymbol{G}$ est connue.

Dans ce travail, on s'intéresse à la résolution du problème inverse qui est un problème mal-posé. Cela est dû, en particulier, à la grande différence entre le grand nombre de variables inconnues (environ 4000 variables exprimées dans $\boldsymbol{J}$) et le petit nombre de variables connues (environ 150 capteurs exprimées dans $\boldsymbol{M}$). Pour résoudre un problème inverse mal-posé on utilise des hypothèses ou des contraintes qui permettent d'améliorer son conditionnement. Ce formalisme s'appelle la régularisation.

Les techniques qui ne considèrent pas la régularisation du problème inverse mal-posé avant la résolution montrent des difficultés de précision et d'instabilité spatial. Par contre, lorsqu'on considère la régularisation les résultats peuvent être améliorés. Partant de là, dans ce travail, on considère une forme de régularisation du problème inverse avant la résolution proprement dite.

Notre approche de résolution se compose de deux parties. Dans la première partie de notre travail, on s'intéresse d'abord à une technique de régularisation. Celle-ci consiste à (1) estimer un degré (score) d'activation des sources et (2) la sélection qui permet de sélectionner les sources dont les scores ont assez importants par rapport à un seuil. L'objectif de cette étape est de définir un sous-ensemble de l'espace des sources du modèle. Cette réduction de l'espace des sources est fondée sur l'hypothèse suivante : l'activité neuronale d'intérêt se localise surtout dans une ou quelques régions limitées du cerveau (aires fonctionnelles) et ne concerne pas toute la surface corticale. Dans ce travail, on considère que cette information doit être exploitée, on obtient, ainsi, un problème inverse mieux posé et plus facile à aborder par une technique de résolution de problème inverse.

Dans la deuxième partie on s'intéresse (1) au regroupement des sources en parcelles à partir des sources considérées actives et (2) on ce qui concerne la détection de l'activité cérébrale. On considèrera une technique de résolution de type '*beamformer*' connue sous le nom de Linearly Constrained Minimum Variance (LCMV) (Barry Van-Veen 1997) où la localisation des sources actives qui repose sur un filtrage spatial des mesures.

Le principe de filtrage spatial consiste généralement à identifier la localisation des sources actives d'une façon individuelle à partir des mesures. Dans (Limpiti, Veen et al. 2006), les auteurs considère cette technique où l'ensemble des sources sont regroupées en parcelles. L'activité à localiser est alors représentée sous forme de régions, ce qui est une approche réaliste de la modélisation de l'activité neuronale. On s'intéresse à cette technique étant donné qu'elle peut être adaptée à notre approche de parcellisation.

Par contre, contrairement à la technique utilisée dans (Limpiti, Veen et al. 2006), notre technique de construction des parcelles corticales exploitent les données. Les parcelles formées par la technique dans (Limpiti, Veen et al. 2006) utilisent toutes les sources du modèle et elles couvrent totalement la surface corticale. Les parcelles construites dans notre approche de parcellisation utilisent seulement les sources estimées potentiellement actives. Ces groupements de sources constituent un sous-ensemble assez réduit de sources du modèle.

Notons qu'il existe des atlas (Lancaster JL 2000) qui identifient des régions corticales en fonction de leur fonction neuronale. En effet, on sait depuis les travaux de (Fischl, Salat et al. 2002; Fischl, VanDerKouwe et al. 2004) que l'activité neuronale normale respecte des gabarits assez universel par exemple celui présenté dans (Lancaster JL 2000). Bien qu'on ne respecte pas ces gabarits dans notre parcellisation, reste qu'on considère que l'activité neuronale se localise dans des régions fonctionnelles. C'est ce qui explique notre modélisation de la surface corticale en parcelles.

Il est important de noter que nous exploitons deux niveaux de modélisation : d'une part, la modélisation en sources ponctuelles et d'autre part, la modélisation en parcelles regroupant un ensemble de sources.

Le but de notre travail est d'étudier comment ces deux modélisations sont utilisables conjointement dans la localisation de l'activité cérébrale et l'application de cette combinaison dans une approche de résolution du problème inverse de type *beamformer.* Tel est l'objectif de ce mémoire.

CHAPITRE 1

ÉTAT DE L'ART ET CONTRIBUTION

La résolution d'un problème inverse est un sujet d'intérêt dans plusieurs domaines comme, par exemple, la géophysique, l'astronomie, l'océanographie, l'imagerie médicale etc.…. Lorsque le nombre des variables à déterminer dépasse considérablement celui des données dont on dispose, on parle alors d'un problème inverse mal-posé. La majorité des solutions proposées pour la résolution du problème inverse mal-posé prennent en compte certains *a priori* (des hypothèses). En général, ces *a priori* sont ajoutés au modèle sous forme de contraintes. En utilisant ces contraintes, il est possible d'améliorer le conditionnement du problème inverse mal-posé qu'on peut résoudre sans ambigüité. Généralement, ces *a priori* sont obtenus à partir du contexte du problème.

Par exemple, les auteurs dans (Lancaster JL 2000) exploitent la connaissance préalable de la cartographie anatomique cérébrale (Atlas de Talairach) pour déterminer automatiquement les régions d'activation fonctionnelles. Dans cet exemple, cette hypothèse peut être considérée assez forte, étant donné la grande variance des tailles et localisations des régions fonctionnelles entre individus (difficile à fixer au préalable). D'autres méthodes considèrent des hypothèses moins fortes comme, par exemple, le fait que l'activité cérébrale ne prend place que dans une région de taille finie regroupant un ensemble de sources synchrones et spatialement contigües. Dans ce mémoire on utilise les a priori d'une façon conservatrice.

Ce chapitre commence par une revue de littérature concernant les travaux proches et pertinents en regard de notre travail. Ensuite, on présente la problématique abordée dans cette maîtrise. On conclut par une présentation de nos contributions apportées dans ce travail.

1.1 État de l'art

Dans cette revue de littérature nous présentons les travaux ayant portés sur les trois principaux volets de notre travail : (1) la modélisation en parcelles de la surface corticale; (2) les techniques de résolution du problème inverse et (3) la technique LCMV du '*beamformer*'.

1.1.1 Travaux sur la modélisation en parcelles

Une parcellisation consiste à regrouper des "individus" qui partagent certains critères. Dans le cas de la modélisation corticale en sources dipolaires, ces dernières sont regroupées selon certains critères (ex. proximité spatiale) pour former une parcelle. Il existe plusieurs critères et processus pour réaliser ce regroupement. Ainsi, dépendamment des choix de ces critères et processus suivie par chaque méthode, on obtient une parcellisation différente.

Dans (Tzourio-Mazoyer, Landeau et al. 2002), les auteurs proposent 3 techniques d'automatisation de l'étiquetage anatomique des régions fonctionnelles à partir d'une parcellisation manuelle d'une image anatomique obtenue à partir d'une acquisition IRM. Pour la parcellisation manuelle les auteurs suivent les principaux sillons cérébraux. Ensuite, manuellement, ils effectuent le traçage des contours délimitant environ 90 parcelles volumiques. Concernant l'automatisation de l'étiquetage des régions fonctionnelles, les auteurs ont utilisés des données IRMf normalisées de 152 sujets (Gabarit anatomique de MNI) qui correspondent aux régions fonctionnelles de la parole. A partir de ces données d'IRMf, les auteurs proposent 3 techniques d'étiquetage suivantes. La première technique consiste à identifier le maxima local et lui assigne l'étiquètte de l'aire fonctionnelle (prédéfinie manuellement) où il se localise. La deuxième est similaire à la précédente, sauf que l'étiquetage concerne le volume sphérique de 10 *mm* autour du maxima local. La dernière technique considère l'étiquetage de toute la région identifiée par l'IRMf comme active.

Dans (Fischl, VanDerKouwe et al. 2004), les auteurs présentent une approche probabiliste pour l'assignement d'étiquettes neuro-anatomique à chaque position sur la surface corticale.

Leur approche repose sur deux types d'*a priori* provenant de l'expertise des anatomistes et des atlas : la localisation spatiale des parcelles et les relations de voisinage entre parcelles, indépendamment de leur localisation spatiale. À partir de ces informations, les auteurs estiment la probabilité de la localisation d'une parcelle en fonction de sa localisation par rapport aux autres parcelles.

Cette estimation probabiliste correspond à la variabilité de la structure anatomo-fonctionnelle de la surface corticale entre individus. Cela revient à estimer la probabilité d'avoir la parcellisation P connaissant le modèle de la surface corticale S; i.e. $p(P|S)$. Cette probabilité peut être exprimée en fonction de la probabilité *a priori* de la segmentation corticale $p(P)$ et de la probabilité de vraisemblance étant donnée les parcelles $p(S|P)$ comme suit :

$$p(P|S) \approx p(S|P) \cdot p(P)$$

Avec cette formulation bayésienne, on a incorporé l'information *a priori* $p(P)$ concernant la parcellisation P et on a également exploité la donnée de classification provenant des atlas $p(S|P)$ qui correspond à la probabilité d'avoir une surface corticale S sachant qu'on observe une parcellisation P. Vue que l'estimation de cette probabilité de parcellisation dépend de la classification utilisée (atlas), un paramètre (f) correspondant à cette classification est introduit pour raffiner le résultat :

$$p(P, f|S) \approx p(S|P, f) \cdot p(P|f) \cdot p(f)$$

Cette formulation, à partir d'une combinaison des *a priori* probabilistes, est en fait une représentation de la variabilité spatiale des régions anatomo-fonctionnelle entre sujets. Elle permet une estimation bayesienne de l'étiquetage d'une surface corticale donnée. Avec des images IRM volumiques de 36 patients, les résultats présentées dans (Fischl, VanDerKouwe et al. 2004) sont d'une grande précision par rapport aux résultats d'étiquetage manuel.

Dans (Thirion, Flandin et al. 2006), les auteurs présentent une technique de parcellisation corticale qui prend on compte des contraintes spatiales et fonctionnelles. Cette technique de parcellisation se compose de deux étapes: (1) parcellisation intra-sujet où les parcelles de la

surface corticale d'un sujet sont formées en considérant ses données anatomiques et fonctionnelles et (2) parcellisation inter-sujet où la formation des parcelles doit respecter une '*analogie spatiale et fonctionnelle*' par rapport à la parcellisation des autres sujets.

Pour la parcellisation intra-sujet, les voxels qui ont tendance d'avoir des profils fonctionnels très proches et leurs distances anatomiques ne dépassent pas un certain seuil sont groupées en parcelles. La surface corticale est présentée sous forme d'un graphe G dont les nœuds sont les voxels et les arêtes sont des liens entre elles (voxels voisins). On assigne au lien entre deux voxels voisins v et w une valeur qui correspond à la *'distance anatomo-fonctionnelle'* entre les deux voxels. Cette distance est définie comme suit :

$$\|\beta(v) - \beta(w)\|_{\Lambda} = \sqrt{\left(\beta(v) - \beta(w)\right) \Lambda^{-1} \left(\beta(v) - \beta(w)\right)^{T}}$$

où $\beta(v)$ estun paramètre fonctionnel du voxel v et Λ correspond à la matrice de covariance des données. A partir de ce graphe, les auteurs calculent la matrice Δ_G des plus cours chemins entre voxels utilisant l'algorithme de Dijkstra. En regroupant les voxels voisins d'ordre $k = 3$ selon la matrice Δ_G on obtient la parcellisation intra-sujet. On note que les parcelles dans ce cas respectent l'homogénéité fonctionnelle des signaux IRM, ainsi que spatiale étant donné qu'elles sont interconnectées.

Dans (Lapalme, Lina et al. 2006), les auteurs présentent une technique de parcellisation appelée '*Data-driven parceling*' à partir des données MEG. Cette technique exploite les données mesurées pour estimer un *'score*' pour chaque source du modèle qui représente le degré de possibilité que la source soit active. Pour le calcul des *scores*, ils utilisent la technique MSP (Multivariate Source Pre-localization) (Mattout J. 2005). Ensuite, ils procèdent à la sélection des sources dont les scores sont les plus élevés. La sélection des sources se fait en fixant un seuil d'activation minimal pour déclarer une source comme active. Dans (Mattout J. 2005), le seuil d'activation minimal est défini d'une façon arbitraire. A partir des scores des sources, les auteurs procèdent à la formation des parcelles de la façon suivante:

1- Établir une liste des scores des sources ordonnés en ordre décroissant.

2- Sélectionner la première source dans cette liste pour être le germe d'une parcelle.
3- Former une parcelle à partir des sources dans le voisinage du germe. Ces sources sont retirées de la liste ordonnée[2].
4- Répéter les deux étapes (2 et 3) jusqu'à l'épuisement de la liste des sources dont le score est supérieur au seuil de sélection.
5- Les sources qui ne sont dans aucune parcelle forment la "parcelle nulle".

La parcelle nulle peut être vide dépendamment de la valeur du seuil. C'est le cas si celui-ci est égal à zéro.

Les auteurs (Lapalme, Lina et al. 2006) notent que les résultats de résolution du problème inverse sont meilleurs lorsque la taille des parcelles construites est proche de la taille de la parcelle active.

1.1.2 Techniques de résolution

Les techniques de résolution du problème inverse parmi les plus utilisées sont les suivantes.

A. Techniques de régularisation spatiale :

En général, ces techniques de régularisation consistent dans l'usage de contraintes spatiales dans la résolution du problème inverse. On distingue trois grandes approches:

Norme Minimale ($\boldsymbol{l_{1-Norm}}$)*:* connue sous le nom de '*minimum norm*' (Matsuura and Okabe 1995), cette technique garantit une solution unique; $\tilde{J} = G^{Pinv} \cdot M$, où G^{Pinv} est la pseudo-inverse de la matrice de gain G. Cette technique consiste à calculer une distribution de sources qui expliquent les données dont leur nombre est proche de celui des données et leur norme (la somme des valeurs absolues) est minimale et cela avec une

[2] L'ensemble des sources formant cette parcelle ne participent plus dans les prochaines itérations de la formation de d'autres parcelles. Cette condition assure que les parcelles formées ne se chevauchent pas

tolérance d'erreur marginale. Cette solution correspond à la résolution de l'expression suivante:

$$\mathrm{J}^* = \underset{\mathrm{J}}{\mathrm{Argmin}}\left(\sum_{k=1}^{\mathrm{n}}|q_k|\right) \quad \text{tel que} \quad \|G.J^* - M\| \leq d \tag{1.1}$$

avec $J^* = \{q_1, q_k, \dots, q_n\}$ est l'ensemble des sources vérifiant les contraintes définie dans l'expression (1.1). En pratique, pour calculer J^* on fait appel à la méthode simplex de la programmation linéaire.

Norme quadratique minimale (l_{2-Norm}) : cette technique est très semblable à l_{1-Norm}. La différence entre les deux techniques est que cette dernière cherche à minimiser la norme euclidienne des courants des sources définie comme suit (Uutelaa, Hämäläinena et al. 1999);

$$\mathrm{J}^* = \underset{\mathrm{J}}{\mathrm{Argmin}}\left(\sum_{k=1}^{n}(q_k)^2\right) \quad \text{tel que} \quad \|G.J^* - M\|^2 \leq d \tag{1.2}$$

Les résultats obtenus avec la régularisation l_{1-Norm} sont focaux; cela est dû au faite que l'expression de minimisation correspond à la somme des valeurs absolues des sources. Par contre, ceux obtenus avec la l_{2-Norm} sont beaucoup plus lisses spatialement vue que l'expression de minimisation est une somme des carrées des amplitudes des sources calculées en utilisant la pseudo-inverse de $\boldsymbol{G}$.

Il aussi important de remarquer que les méthodes sont sensibles à la localisation en profondeur des sources. Autrement, si on a deux sources; une profonde avec une grande intensité et l'autre superficielle (proche des capteurs) et elles génèrent le même champ magnétique au niveau des capteurs, les deux techniques avantages la deuxième source par rapport à la première. Cela est dû au fait que ces techniques favorisent une solution lisse ou la variation d'intensité est petite ce qui est le cas de la deuxième source comparée à la

première (Silva, Maltez et al. 2004). Pour remédier à ce problème on introduit des poids attribués à chaque source (s_i) selon leur localisation. Les variantes de cette technique de minimum norme sont connues sous le nom de 'minimum norme pondérée' (*Weighted minimum norm*) (Silva, Maltez et al. 2004).

Loreta (LOw Resolution Electromagnetic TomogrAphy)(Pascual-Marqui RD 1994): Cette technique est parmi les techniques les plus populaires (tout autant que) dans la localisation des sources actives. Elle est aussi similaire aux techniques 'minimum norme' dans le sens où les sources de la solution doivent aussi satisfaire une contrainte de minimisation et elles doivent aussi expliquer les données avec une tolérance d'erreur prédéterminée. La solution obtenue par *Loreta* est une solution spatialement *'lisse'* obtenue à partir de la résolution de problème suivant :

$$\mathrm{J}^* = \operatorname{Arg}\min_{\mathrm{J}} ||\nabla \mathrm{J}||^2 \qquad \text{tel que} \quad \|G.J^* - M\|^2 \leq d \tag{1.3}$$

où ∇ est l'operateur Laplacien et d est la variance résiduelle dans les données estimer à partir du signal de *baseline*. La solution J^* doit aussi expliquer les données avec une tolérance $\boldsymbol{d}$. Comme les méthodes dites minimum norme, LORETA elle aussi favorise une solution avec des sources superficielles que celles profondes.

B. **Techniques probabilistes :** Contrairement aux techniques précédentes qui calculent *la solution* en appliquant certaines contraintes, les techniques probabilistes, par contre, calculent les solutions possibles en attribuant une valeur probabiliste à chacune de ces solutions selon les contraintes considérées (*a priori* ajoutés dans la modélisation du problème).

Méthode bayesienne *:* La formulation du problème inverse, selon l'approche bayesienne (Baillet and Garnero 1997), consiste à déterminer la distribution des probabilités des sources J sachant les mesures, i.e. $p(J|M)$. Pour le calcul de cette probabilité, on utilise la

formule de Bayes : $p(J|M) = p(M|J) \cdot p(J)/p(M)$ où $p(J)$ correspond à l'*a priori* sur les sources et $p(M|J)$ représente la solution du problème directe.

Cette approche se base sur des informations spatio-temporelles comme *a priori* (contraintes) pour la régularisation du problème inverse. Les contraintes exploitées dans (Baillet and Garnero 1997) sont de deux types: (1) contraintes spatiales où les sources sont organisées dans des régions *planaire* localisées sur la surface corticale et (2) contraintes physiologiques qui stipulent que les amplitudes des sources de la même région *planaire* doivent évoluer avec une homogénéité temporelle.

La modélisation bayesienne consiste alors à trouver un estimateur des amplitudes des sources $(\tilde{J})$ qui maximise la probabilité d'expliquer les mesures considérant des contraintes prédéfinies. Cet estimateur est défini par:

$$\tilde{\boldsymbol{J}} = Argmax_{\boldsymbol{J}} \, [\boldsymbol{p(J|M)}] \tag{1.4}$$

où $p(J|M)$ est la probabilité d'avoir $\boldsymbol{J}$ sachant qu'on observe les mesures M. L'estimateur a posteriori utilisé dans cette modélisation, au sens MAP (Maximum A Posteriori), peut s'exprimer comme une distribution de probabilité en fonction de J comme suit :

$$p(M|J) = \frac{1}{Z} \cdot e^{-U(J)}.$$

Z est une constante de normalisation et $U(J) = U_1(J) + \lambda U_2(J)$ où $U_1(J)$ et $U_2(J)$ sont des fonctions d'énergie associées respectivement à $p(M|J)$ et $p(J)$; λ est un paramètre de mélange qui "équilibre" les contributions de la vraisemblance $p(M|J)$ et de l'*a priori* $p(J)$. Ainsi la résolution de l'expression (1.4) correspond à la résolution de l'expression suivante: $\tilde{J} = \text{Argmin}_J \, U_1(J) + \lambda U_2(J)$.

Cette approche bayesienne décompose le problème inverse en une minimisation d'une expression qui se compose de deux parties: (1) $U_1(J) = \|M - G \cdot J\|^2$ qui correspond à la

contrainte d'expliquer les données et (2) $U_2(J) = U_s(J) + U_t(J)$ où $U_s(J)$ et $U_t(J)$ sont des fonctions qui représentent respectivement à l'a priori spatiale et temporelle.

MEM (Maximum Entropy on the Mean) : La méthode MEM permet d'inférer de l'information d'un système à partir des données. Cette méthode est exploitée dans plusieurs domaines notamment la résolution du problème inverse en imagerie médicale (Amblard C 2004). D'après (Lapalme, Lina et al. 2006), le problème direct en M/EEG peut être formulé de la façon suivante :

$$G\ E_p(J) = E(M) \tag{1.5}$$

où p est une loi de probabilité, $E_p(J)$ est l'espérance des intensités des sources J selon cette loi et E(M) l'espérance des mesures M. Si on considère J comme une variable aléatoire et $d\mathcal{P}(J)$ sa densité de probabilité, la méthode MEM répond à la question suivante : Quelle est la densité de probabilité optimale $d\mathcal{P}$ qui satisfait l'expression **(1.5)**. L'espace solution C_m à considérer pour répondre à cette question est l'ensemble des densités de probabilités $d\mathcal{P}$ qui satisfont l'expression (1.5). Ainsi,

$$C_m = \{ d\mathcal{P}(q) = f(q)\ d\mathcal{P}_{ref}(q) \mid G\ E_p(J) = M \}$$ (Si on néglige le bruit)

où la probabilité des sources $d\mathcal{P}$ est exprimée en fonction de la probabilité de référence $d\mathcal{P}_{ref}$ comme *a priori* et la fonction de densité $f(q)$.

La solution MEM consiste à calculer la probabilité des sources $d\mathcal{P}^*$ qui maximise l'entropie de Shannon pour laquelle l'*a priori* $d\mathcal{P}_{ref}$ est égale à zéro :

$$d\mathcal{P}^* = Argmax_{d\mathcal{P} \in C_m} S_{ref}(d\mathcal{P}) \qquad \text{où} \quad S_{ref}(d\mathcal{P}_{ref}) = 0$$

$$\text{et} \quad S_{ref}(d\mathcal{P}) = -\int f(q) \cdot \log f(q) \cdot d\mathcal{P}_{ref}(q)$$

Ainsi la solution retenue permet de maximiser l'information exprimée dans la densité de probabilité $d\mathcal{P}$ et elle explique les données puisqu'on cherche la solution dans C_m.

C. Techniques de filtrage spatial

Beamformer and LCMV (Linear Constraint Minimal Variance)

Issue du domaine du traitement des signaux radar et sonar, la technique du '*beamformer*' a été aussi utilisée dans le cadre des signaux MEG et EEG (Lin, Witzel et al. 2008).

La méthode de filtrage spatial '*beamfomer*' est une méthode de 'balayage'. Elle consiste à définir, à partir des données, un filtre spatial W_k pour chaque position spatiale k. En appliquant ce filtre sur les mesures, on peut estimer la contribution J_k provenant de la source s_k à la position k. L'expression suivante résume cette relation :

$$J_k = W_k M \tag{1.6}$$

Pour déterminer le filtre W_k, la technique *beamformer* le formule sous forme d'un problème de minimisation suivant :

$$\mathbf{min}_{\boldsymbol{w_k}} W_k^T R_x W_k \text{ sujet à } W_k^T G_k = 1$$

où la première partie consiste à déterminer $\boldsymbol{W_k}$ qui minimise la puissance total du signal des données considérant la matrice de covariance des données $\boldsymbol{R_x}$ et la deuxième partie est la contrainte qui garantie que $\boldsymbol{W_k}$ favorise le signal provenant de la source désirée dans la position k. Dans ce travail, on utilise une variante du beamformer, notament *beamformer LCMV* (Barry D. Van Veen 1997) qui permet un filtrage sur les parcelles.

1.1.3 Parcellisation et approche LCMV

Dans (Tulaya Limpiti 2006), les auteurs utilisent une modélisation en parcelles dans une approche LCMV pour la résolution du problème inverse. Leur technique se compose de deux étapes :

- L'espace des sources est modélisé par un ensemble de K parcelles qui se chevauchent à 50% et dont les tailles varient entre 70 mm^2, 270 mm^2 et 590 mm^2. A chaque parcelle k

correspond une sous-matrice de gain G_k dont les colonnes ou *'leadfields'* correspondent à celles de l'ensemble des sources de la parcelle k. En fait, la matrice de gain décrivant la parcelle k utilisée se limite seulement aux sources les plus représentatives. Pour définir cette matrice de gain $\tilde{G}_k$, les auteurs utilisent la technique NMSE (Normalized Mean-Square Error). Ce seuil permet de sélectionner un sous-ensemble de sources de la parcelle k représentatifs de la parcelle avec une tolérance $E_{NMSE} = G_k - \tilde{G}_k$. Cette approximation des sources par un modèle de parcelles permet de réduire la dimension de l'espace des sources dans une perspective de régularisation du problème inverse.

- Calcul d'un filtre spatial w_k pour chaque parcelle k (Tulaya Limpiti 2006). Pour calculer les amplitudes de chaque parcelle, on applique un filtre de type *beamformer* sur les mesures; l'amplitude de la parcelle sera l'amplitude des sources qui la composent.

D'après les auteurs (Tulaya Limpiti 2006), les résultats de simulations montrent qu'avec une tolérance E_{NMSE} de 85% et de des parcelles dont la taille est environ 270 mm^2, il est possible d'identifier correctement une parcelle active de ~253 mm^2 (dans une région somato-sensorielle du cortex) ce qui correspond à une résolution spatiale de 1.5 cm.

1.2 Problématiques et contributions

Dans cette section, on présente la problématique de la résolution du problème inverse en utilisant une modélisation en parcelles et nos contributions.

Les questions abordées dans ce mémoire sont les suivantes:

- Étant donné les données fonctionnelles (MEG) et anatomiques (IRM), comment peut-on construire un ensemble de parcelles à partir des sources du modèle de sorte qu'on a un problème inverse mieux-posé ?
- Quelle parcellisation doit-on choisir qui représente au mieux l'espace des sources et en même temps assure la construction d'un espace de parcelles à partir des sources potentiellement actives ?

- Considérant cette modélisation en parcelles, quelle approche doit-on utiliser pour la résolution du problème inverse et produire une image de l'activité cérébrale ?

Les principales contributions de notre travail portent sur la modélisation de la surface corticale en parcelles. Une nouvelle technique de parcellisation est développée; elle permet la pre-localisation des régions actives considérant les sources sélectionnées par une technique connue sous le nom de "Multivariate Source Pre-localisation" (MSP) (Mattout J. 2005). Dans ce travail, on procède à la validation empirique de l'hypothèse suivante : les scores MSP des sources au repos suivant une loi beta. Cette hypothèse est avancée dans certains travaux (Lapalme, Lina et al. 2006) mais elle n'a jamais était validée. Il faut noter que les parcelles qu'on construit se caractérisent par leur localisation anatomique et le degré d'activation qui correspond à celles des sources qui les composes. On présente ensuite une technique d'estimation de l'activité des sources étant donné leur appartenance aux parcelles.

1.2.1 Problématiques

La principale problématique de ce travail est la localisation de l'activité cérébrale à partir d'informations fonctionnelles et anatomiques. Ici, les données fonctionnelles sont de deux types : (1) les mesures (par exemple, le sujet effectue une tache donnée) et (2) les données de repos (*'baseline'* ou mesures prises quand le sujet est au repos). Les données anatomiques proviennent essentiellement du maillage cortical obtenu à partir de la segmentation d'une acquisition IRM. À partir de ces données, on établit l'espace des sources et la matrice de gain qui représente la contribution de chacune des sources du modèle.

Pour la réduction de l'espace des sources et un meilleur conditionnement du problème inverse (voir chapitre 2), les parcelles doivent respecter deux critères importants : elles doivent (1) inclure la totalité des régions actives et (2) n'inclure relativement que peu de régions inactives. Ainsi, avec ces deux critères on peut garantir un meilleur conditionnement du problème inverse sans pertes dans la fiabilité du résultat. Pour satisfaire le premier critère, le processus de formation des parcelles se base sur le calcul pre-localisation des sources

actives (Mattout J. 2005). Pour satisfaire le deuxième critère, seules les sources potentiellement actives sont considérées dans le processus de parcellisation. Pour la sélection des sources, on définit un seuil de sélection en utilisant test d'hypothèse de type FDR (False Discovery Rate) (Benjamini and Hochberg 1995).

Étant donné que l'activité corticale se localise, en général, dans des régions et non en des points isolés, notre processus de parcellisation doit tenir compte de cette contrainte. Ainsi, les parcelles seront formées à partir de sources sélectionnées en utilisant une technique de diffusion spatial. Le degré de voisinage de la diffusion est une variable à définir. Cette technique respecte aussi nos deux critères de formation des parcelles. En effet, elle garantit la réduction de l'espace des sources à un sous-ensemble qui satisfont les contraintes fonctionnelles (scores MSP) et anatomiques (diffusion autour des germes).

1.2.2 Contributions

Les principales contributions de ce travail sont les suivantes :

A. Développement d'un algorithme de modalisation en parcelles qui permet une pré-localisation des régions actives. Cet algorithme se compose des étapes suivantes :
- Estimation de l'activation des sources (scores MSP).
- Définition d'un seuil de sélection des scores (données de *baseline*).
- Sélection des sources potentiellement actives utilisant un seuil adaptatif.
- Formation de parcelles à partir des sources sélectionnées (parcellisation).

Cet algorithme se distingue par son exploitation des données dans la régularisation du problème inverse. Cela à travers une sélection contrôlée des sources à partir de laquelle on établit une parcellisation autour des régions où l'activité neuronale est estimée potentiellement présente. En plus de la régularisation, il est important de mentionner que à la sortie de cet algorithme on dispose de données d'inférence (concernant le degré de participation de sources dans la parcellisation) qui peuvent être exploitées dans la partie de résolution du problème inverse.

Il est à noter que bien qu'on ait choisi la technique LCMV pour la résolution du problème inverse dans ce travail, cette technique de modélisation en parcelles reste valide pour d'autres méthodes de résolution du problème inverse telle que le MEM (Lapalme 2004).

B. Notre modélisation en parcelles se base sur une sélection de sources estimées potentiellement actives. Cette liste de sources est déterminée à partir d'un processus de sélection des scores des sources. On utilise un calcul statistique pour déterminer un seul seuil de sélection pour l'ensemble des sources du modèle. On exploite une technique de contrôle statistique ou plus exactement test multi-hypothèses qui permet une comparaison multiple des scores des sources d'une façon simultané.

C. Pour la définition de l'hypothèse nulle (nécessaire pour le test d'hypothèse multiple), on calcule les scores des sources à partir des données de "*baseline*". Cette distribution des scores est supposée suivre une loi beta ce qu'on a démontré dans un test empirique. Ce test, qu'on appel test de *'beta-Fit*', révèle aussi d'autres informations concernant la stationnarité de l'activité physiologique des sources au repos.

D. Une des contributions dans ce travail est le développent de la technique de parcellisation qui repose sur une technique de diffusion autour des sources sélectionnées. Le résultat du groupement des sources en parcelles (chevauchantes) détermine les régions où se localise fort probablement une activité neuronale.

E. En utilisant le résultat de la parcellisation on utilise la technique de filtrage spatial de type *beamformer* (LCMV) pour la résolution du problème inverse. Cela se fait en deux étapes; (1) l'estimation du degré d'activation des parcelles à partir du résultat de filtrage spatial sur les parcelles et (2) le calcul des contributions des sources des parcelles à partir des données de filtrage sur les parcelles.

CHAPITRE 2

MODÉLISATION EN PARCELLES A PARTIR DU MODÈLE DISTRIBUÉ

Ce chapitre présente les étapes pour aboutir à une modélisation de la surface corticale en parcelles. On présente le modèle en sources distribuées utilisé pour la définition du problème direct qui permet d'exprimer les données à partir du modèle de sources.

2.1 Description de la surface corticale

Cette section se compose de trois parties présentant les différents types de ségmentation; anatomiques, fonctionnels et anatomo-fonctionnels de la surface corticale. On commence par présenter une description anatomique de la surface corticale comme étant une région cérébrale caractérisée par une densité neuronale et sa forme anatomique. Ensuite, on présente la surface corticale comme étant une région cérébrale fonctionnelle où se localise l'activité neuro-cérébrale. La surface corticale peut aussi être repartie en parcelles considérant les données anatomiques et fonctionnelles (MEG/EEG) observées, c'est ce qu'on introduit dans cette dernière partie.

2.1.1 Répartition anatomique de la surface corticale

Le cerveau est l'organe principal du système nerveux. La plupart des organes du corps humain sont anatomiquement connectés au cerveau à travers un système nerveux (central et périphérique). Le cerveau compte environ 100 milliards de neurones dans un volume d'environ 700 cm^3 et pèse environ 1,3 Kg. Le cerveau, comme le reste du système nerveux, se compose de deux substances; la substance blanche qui correspond aux ramifications des neurones de conduction et la substance grise qui correspond à la concentration des neurones. Cette dernière se répartit en une couche superficielle épaisse (entre 2 à 4 mm) et elle correspond essentiellement au cortex cérébral formé principalement de matière grise où se localise la plus grande densité de neurones du corps humain. Le cortex se compose lui-même

de plusieurs couches renfermant différentes classes de neurones, d'inter-neurones et de cellules gliales (Scheibel, Davies et al. 1974).

Du point de vue macroscopique, l'anatomie du cortex cérébral est formée de deux hémisphères cérébraux, chacun divisé par 3 sillons principaux en quatre lobes (frontal, temporal, pariétal, occipital) (voir figure 2.1 (http://www.psycho-neuro-feedback.com/)). Sur la surface de ces 5 lobes, on distingue des circonvolutions causées par des plis et d'autres sillons secondaires que les anatomistes utilisent pour identifier des sous-régions anatomiques à l'intérieur de chaque lobe. Grace aux plis et aux sillons, la surface du cortex peut loger dans un volume réduit, de sorte que la surface réelle du cortex est beaucoup plus importante que la surface crânienne. Aussi, il est à noter que certaines régions corticales qui sont exposées à la surface alors que d'autres sont plus profondes dans le cerveau.

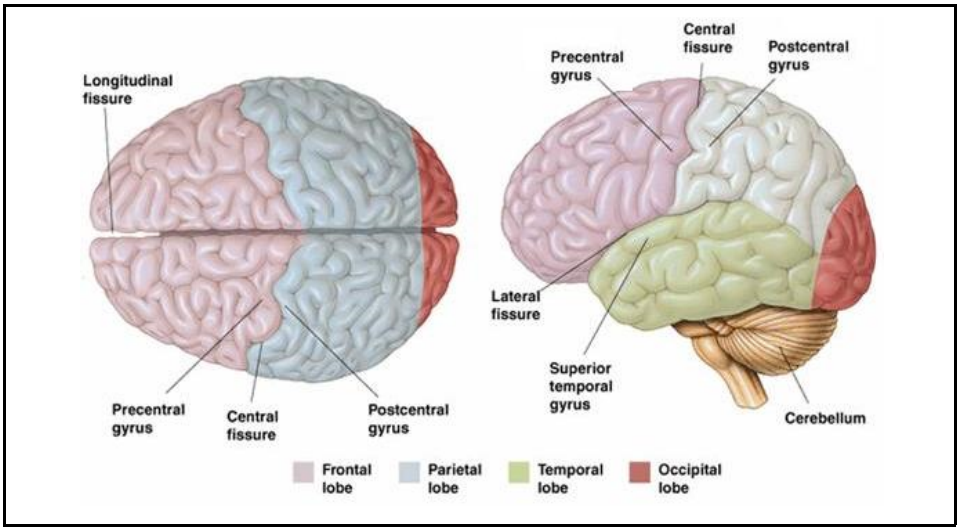


Figure 2.1 Montre les deux hémisphères et les 5 lobes principaux

D'un point de vue microscopique, le cortex cérébral peut être segmenté en différentes couches (voir figure 2.2 (Scheibel, Davies et al. 1974)) chacune se caractérise par un type de neurones. Par exemple, dans la couche interne (couche II), les neurones ont des dendrites courtes et orientées perpendiculairement à la surface du cortex et les axones sont longs et orientés parallèlement. Au niveau de la couche IV (couche pyramidale), on trouve

principalement de cellules pyramidales qui sont perpendiculaires à la surface du cortex. Les neurones de la sous-couche pyramidale interne reçoivent des signaux des autres régions corticales, par contre celles de la sous-couche pyramidale interne y envoient des signaux.

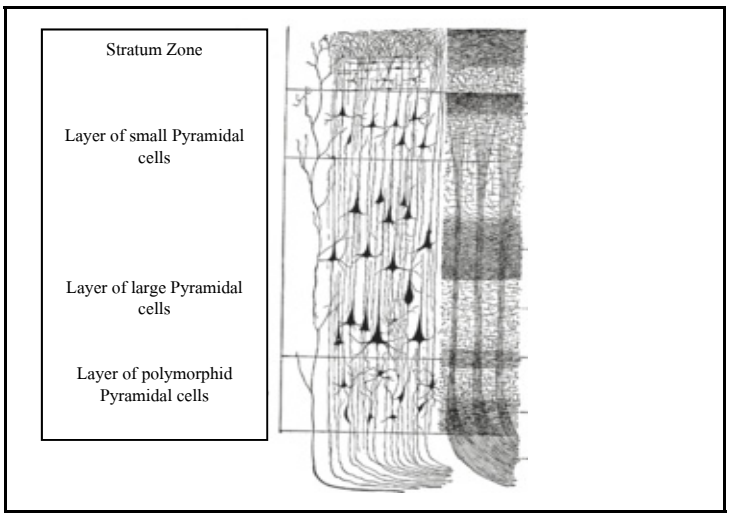


Figure 2.2 Montre l'organisation histologique corticale

2.1.2 Description de l'anatomie fonctionnelle corticale

Le cerveau communique avec le reste du corps à travers le système nerveux. Différents messages sont reçus et envoyés au cerveau concernant les fonctions cérébrales primaires. Celles-ci se résument essentiellement au système sensoriel qui gère les 5 sens classiques (la vue, l'odorat, l'ouïe, le toucher et le goût) ainsi que d'autres données comme l'état d'équilibre du corps ou l'état du sang, le système moteur responsable de générer des signaux de contrôle des mouvements volontaires du corps à travers des stimulations envoyées aux muscles et le système limbique qui joue un rôle dans la mémoire et les émotions (comme la peur, le plaisir et l'agressivité).

A part la segmentation anatomique du cortex en lobes (frontal, temporal, pariétal, occipital) définis par leur position, plis et les sillons qui les séparent, on a aussi une répartition du cortex en aires fonctionnelles. Elles sont déterminées selon la fonction neurologique

(moteur/sensoriel) qui leur correspond. La figure 2.3 (Maurieés) indique des aires sensoriels-motrices.

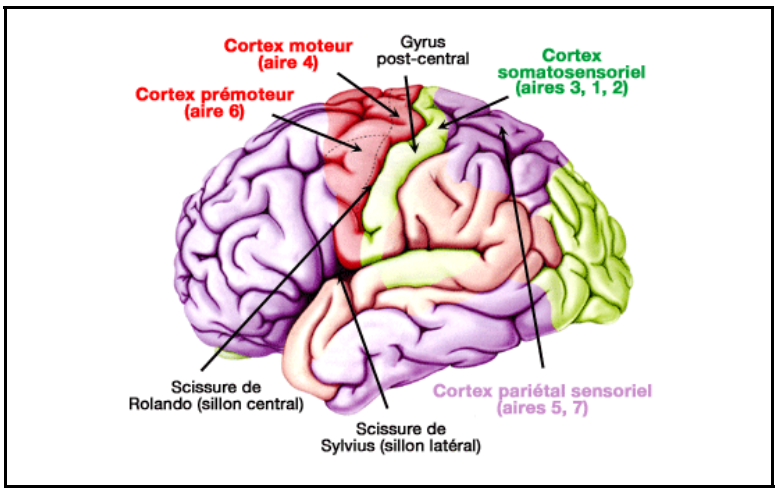


Figure 2.3 Exemples d'aires fonctionnels identifiés

2.1.3 Notion de parcelles et hypothèse de modélisation

Malgré les progrès en neuroscience dans la segmentation anatomique du cortex, selon des critères anatomiques et/ou fonctionnels, la variabilité entre individu fait qu'il est difficile d'établir un modèle anatono-fonctionel universel. En fait, vue la plasticité corticale, la cartographie fonctionnelle chez le même individu varie aussi avec le temps, selon plusieurs critères (l'âge, l'apprentissage, etc…).

Étant donné qu'on s'intéresse, dans ce travail, à la localisation de l'activité fonctionnelle, il est important de considérer les deux notions suivantes: (1) la répartition corticale en aires fonctionnelles et (2) la variance de cette répartition entre individus.

Ces hypothèses vont définir notre stratégie à suivre dans la localisation de l'activité neuronale dans le reste de ce travail. Ainsi, notre algorithme de résolution sera-t-il basé sur la définition et la formation de parcelles comme régions anatomo-fonctionnelles à partir des données anatomiques et fonctionnelles. Ensuite, on effectue un filtrage spatial des parcelles pour déterminer les régions qui expliquent mieux les données.

2.2 Modèle distribué

Le modèle distribué repose sur des données d'acquisition IRM (image anatomique) cérébrale à partir desquelles on extrait le maillage de la surface corticale. Ces données sont utilisées pour définir un ensemble de dipôles ou source qui représentent l'activité électromagnétique neuronale. Ces sources sont reparties uniformément pour former le *modèle distribué*. Dans cette section on présente cette modélisation de la surface corticale en modèle distribué.

2.2.1 Modélisation de la surface corticale et modèle distribué

Pour la modélisation de la surface corticale en sources MEG reparties uniformément, on fait appel à la technologie IRM anatomique. Cette dernière permet d'obtenir une description 3D du volume cérébrale en voxels (maillage cubique). A partir de cette acquisition IRM cérébrale pondérée en T1, on extrait la surface corticale du cerveau (ou de la tête) (voir figure 2.4). Pour isoler et extraire les différentes couches corticales formant le cerveau, on fait appel aux techniques de segmentation (Cointepas 1999; Tosun D. 2004). En fait, la segmentation se fait en plusieurs étapes de reconnaissance et d'extraction des constituants cérébraux qui se base sur l'identification des voxels appartenant seulement au cortex. La surface corticale est finalement représentée par un maillage d'environ 10.000 voxels qu'on exploite pour définir les sources du modèle distribué.

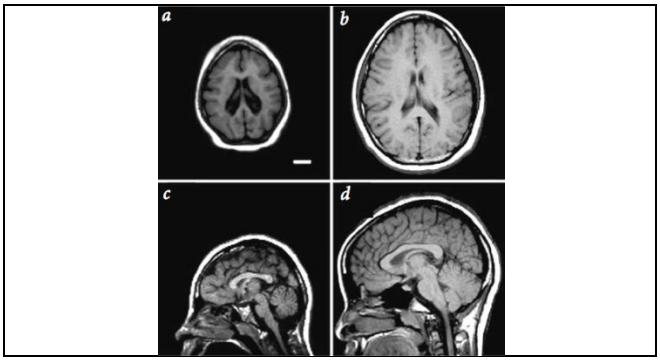


Figure 2.4 Images IRM d'une tête humaine en 3D ((a,b) est une vue transversale et (c ,d) est une vue longitudinale)

2.2.2 Sources

Le champ magnétique généré par un neurone n'est pas suffisamment fort pour être mesurable par la technologie MEG/EEG. Par contre, quand un groupe de neurones est actif de façon synchrone, ils génèrent un courant électrique suffisamment intense qui est mesurable; de l'ordre de 10 nA/mm^2 l'équivalent à 10^{-13} *Teslas* (Baillet S. 2001).

Partant delà, il est acceptable de modéliser un groupe de neurones voisins dans une région d'environ quelques mm^2, comme une source génératrice de courant électrique (dipôle) d'où la notion de dipôle de courant ou ECD (Equivalent Current Dipole). D'après (Dale A., Sereno M. 1993) et (Fischl, Sereno et al. 1999), la modélisation en ECD est une modélisation valable pour définir un modèle en sources. En général, l'activité cérébrale est modélisée en sources (dipôles) de courant qui génèrent un champ électromagnétique lorsqu'ils sont actifs.

2.2.3 Modélisation en sources distribuées

Il y a deux types de famille de modèles de sources utilisées pour la formulation et la résolution du problème inverse; (1) modèle en sources non-distribuées et (2) modèle en sources distribuées (M. Wagner Th. Köhler 2000), dont le nombre des sources à localiser et leurs caractéristiques sont pré-déterminées. C'est ce dernier qu'on exploite dans ce travail.

En effet, la modélisation de l'activité (bioélectrique) des neurones localisées dans la région corticale considère que la plupart des neurones sont distribués d'une façon perpendiculaire par rapport à la surface corticale.

Par contre, il faut mentionner que le nombre des sources, obtenues avec cette modélisation, est de millier de fois plus important que le nombre des mesures (ou capteurs). Cette disproportion, entre l'espace des sources et celui des mesures, fait qu'on se retrouve avec un très grand degré de liberté lors du calcul d'amplitude des sources sachant les mesures.

La modélisation distribuée est la plus fréquente, même si le nombre des paramètres à déterminer est très grand, vu qu'elle permet, en générale une meilleure résolution spatiale. Il faut aussi reconnaitre que cette modélisation doit sa popularité à la technologie IRM qui permet d'avoir une description précise du cortex humain et permet ainsi une modélisation proche du problème réel. En générale, la localisation de l'activité neuronale, utilisant cette modélisation, fait toujours appel aux contraintes anatomiques et fonctionnelles afin d'améliorer la régulation et faciliter la résolution du problème inverse.

2.3 Dipôles et matrice de gain

Dans la section précédente, on a introduit le modèle distribué qui permet une représentation du cortex en sources reparties uniformément. Dans cette section, on explique en détails la modélisation en dipôles (ou sources) et on présente la matrice qui modélise le liens entre les sources et les données qu'on appel *'Matrice de gain'*.

2.3.1 Dipôles

Un dipôle sans contrainte est caractérisé par trois paramètres : sa position dans un référentiel 3D (x, y, z), son orientation (θ, φ) et son intensité (ou moment dipolaire). Un dipôle actif génère un champ magnétique reçu à la surface de la tête, comme le montre la figure 2.5. Comme on peut le constater dans cette figure, si on fait varier l'intensité du dipôle, on va certainement observer une variation des intensités des champs magnétiques à la surface de la tête. Pour mieux définir cette relation, on utilise les *'lois de Maxwell'* (Baillet S. 2001) qui permettent de calculer le champ magnétique créer par une source de courant à l'intérieur d'un milieu donné (Voir Annexe 1). Il est aussi important de noter que dans cette même figure (de gauche) le champ magnétique produit par le même dipôle est vu dans deux points de la surface avec la même intensité mais de signes opposés.

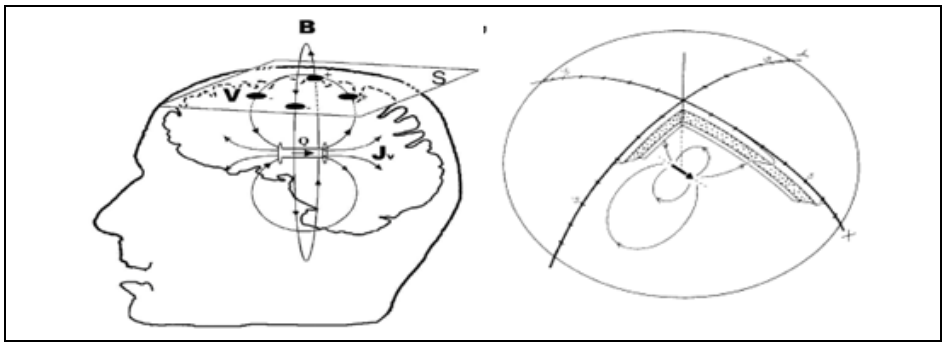
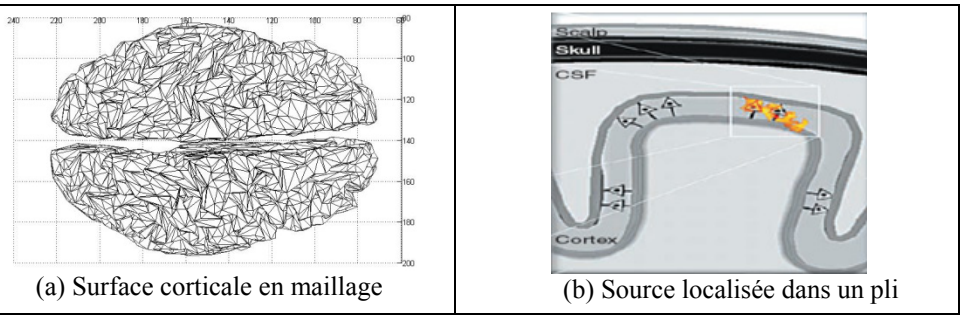


Figure 2.5 Champ électromagnétique génère par un dipôle actif au niveau de la surface de la tête

Il est aussi important de rappeler qu'à cause de la structure anatomique de la surface corticale et plus exactement l'orientation des sources par rapport à la surface crânienne (voir figure 2.6), le champ magnétique provenant des sources dans les plis (sources longitudinales) est mieux perçu par les capteurs MEG que celui par les sources à la surface (sources radiales).



(a) Surface corticale en maillage

(b) Source localisée dans un pli

Figure 2.6 Source de la surface corticale localisées dans les plis sont mieux perçu par la MEG que celles à la surface

Comme introduit précédemment, les dipôles sont construits à partir des données anatomiques (IRM) et sont présentés sous forme d'un maillage (figure 2.6 (a)). Ce maillage est en fait une représentation numérique de la surface corticale où les dipôles sont des nœuds organisés sous forme de faces triangulaires.

2.3.2 Matrice de Gain

Nous venons de présenter la modélisation des neurones en dipôles ayant les paramètres suivants : la position et l'orientation spatiale. Ces deux paramètres (la position et l'orientation) peuvent être considérés comme des données statiques et connues à partir d'un maillage issu de l'anatomie. Par contre, un troisième paramètre, l'intensité des sources, peut être qualifié comme une donnée dynamique dans le temps, étant donné que l'intensité change dans le temps. On a aussi introduit l'influence des sources actives sur les capteurs. Dans cette section, on s'intéresse à la relation entre les sources et les données mesurées par les capteurs qui se modélise par une matrice qu'on appelle '*Matrice de Gain*'.

Le champ magnétique $\mathbf{B}_i$ généré par un dipôle dépend linéairement de son moment dipolaire. Ce dernier est caractérisé son courant $q \equiv |q|$ et son orientation $\Theta = {}^{q}/_{|q|}$ où $\Theta = (\theta, \varphi)$ exprimé en coordonnés sphériques. Ainsi, la mesure sur le scalp à une position r causée par le dipôle placé à la position r' peut être exprimée par l'expression suivante :

$$m(r) = g(r, r', \Theta)q$$

avec $g(r, r', \Theta)$ est le vecteur associé au dipôle à l'emplacement r' ayant l'orientation Θ par rapport à un point d'observation r. En pratique, la surface corticale est modélisée par un ensemble de sources $\{s_1, \dots, s_i\}$ localisées aux points $\{r_1', \dots, r_i'\}$. La mesure lue correspond à la somme linéaire des contributions de chacune de ces sources:

$$m(r) = \sum_i g(r_i, r_i', \Theta_i)q_i$$

Finalement, la relation algébrique qui lie l'ensemble des sources et les mesures est donnée dans l'expression suivante :

$$M = \begin{pmatrix} m_1 \\ m_2 \\ \vdots \\ m_m \end{pmatrix} = \begin{pmatrix} g_{1,1} & g_{1,2} & \cdots & g_{1,n} \\ g_{2,1} & g_{2,2} & \cdots & g_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ g_{m,1} & g_{m,2} & \cdots & g_{m,n} \end{pmatrix} \begin{pmatrix} q_1 \\ q_2 \\ \vdots \\ q_n \end{pmatrix}$$

c'est-à-dire $M_m = G_{m,n} J_n$ où

m: correspond aux nombres des capteurs.

n: correspond aux nombres des sources.

On remarque que $m << n$ et M_m et J_n sont des vecteurs respectivement de m et n éléments. Cette différence importante entre m et n correspond à la différence entre les degrés de résolution spatiale de la représentation de la surface corticale et de celle des mesures.

On note que la matrice de Gain $G_{m,n}$ ne varie pas avec le temps. Les vecteurs 'ligne' de la matrice de gain constituent ce qu'on appelle le *Forward-Field* et les vecteurs 'colonne' sont les *Lead-Fields* (voir figure 2.7).

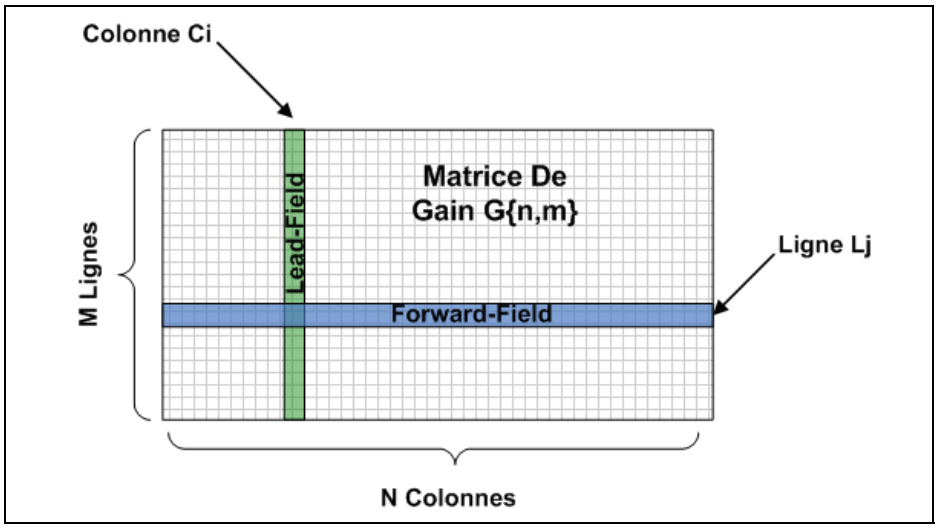


Figure 2.7 Lead-Fields et de Forward-Fields d'une Matrice du Gain

Si on considère une série de mesures temporelles de *t* échantillons de mesures, la relation entre les sources et les mesures devient:

$$M = \begin{pmatrix} m_{1,1} & \cdots & m_{1,t} \\ m_{2,1} & \cdots & m_{2,t} \\ \vdots & \ddots & \vdots \\ m_{m,1} & \cdots & m_{m,t} \end{pmatrix} = \begin{pmatrix} g_{1,1} & g_{1,2} & \cdots & g_{1,n} \\ g_{2,1} & g_{2,2} & \cdots & g_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ g_{m,1} & g_{m,2} & \cdots & g_{m,n} \end{pmatrix} \begin{pmatrix} q_{1,1} & \cdots & q_{1,t} \\ q_{2,1} & \cdots & q_{2,t} \\ \vdots & \ddots & \vdots \\ q_{n,1} & \cdots & q_{n,t} \end{pmatrix}$$

équivalent à $M_{m,t} = G_{m,n}\, J_{n,t}$ où t : correspond au nombre de mesures acquises dans un temps entre 1 et t.

Il est important de remarquer que chaque *Forward-Field* correspond aux contributions des dipôles du modèle sur un capteur et que chaque *Lead-Field* correspond à la contribution d'un dipôle donné sur les capteurs. Ainsi, les *Forward-Fields* ce sont des coefficients attribués aux dipôles par rapport aux mesures lues sur un capteur. Il est aussi à noter que les sources profondes ont des coefficients de contribution faibles.

2.4 Problème inverse

La relation entre les sources et les mesures peut être formulée sous deux formes: (1) le problème direct qui consiste à exprimer les observations (i.e. les mesures) en termes des sources (voir figure 2.8) et (2) le problème inverse qui correspond à la situation où on dispose de mesures à partir desquelles on détermine les intensités des sources correspondantes.

Dans cette section on présente le problème inverse dont la résolution consiste à trouver une solution à la formulation suivante : $J = G^{-1}.m$

G est une matrice considérablement rectangulaire, le calcul de son inverse n'est pas évident et cela présente un problème de la résolution de cette expression $J = G^{-1}.m$

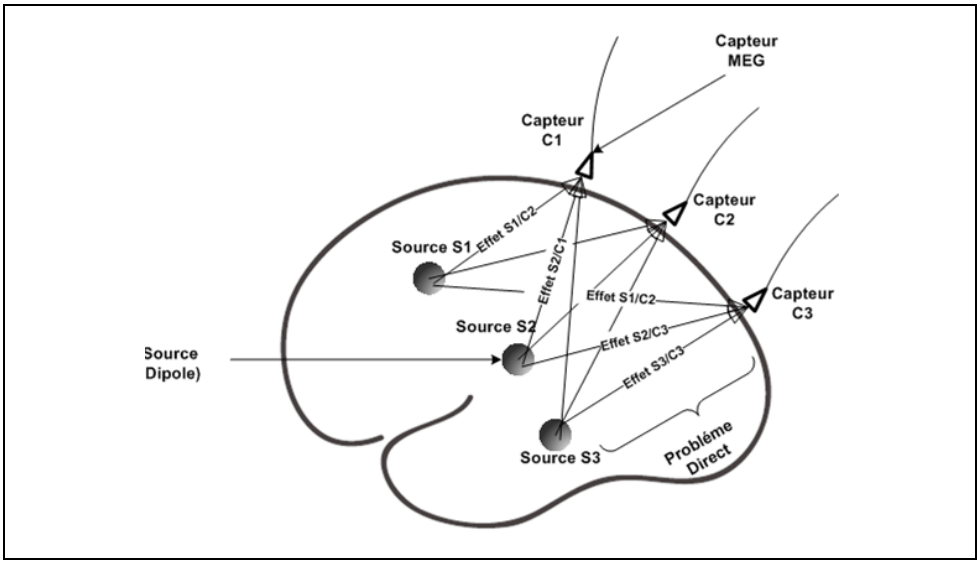


Figure 2.8 Problème direct : Modèle simplifié à 3 sources et 3 capteurs

Le problème inverse consiste dans l'estimation de sources à partir des mesures obtenues sur les capteurs. Cette estimation doit être unique, stable et correspondre au mieux à la réalité du processus qui explique les données. Avant de présenter de la résolution du problème inverse, on commence d'abord par sa formulation algébrique.

L'expression algébrique suivante permet d'exprimer la configuration des intensités des sources (exprimées en $J(t)$) en fonction des mesures $M(t)$ et de la matrice de gain inversée (d'où le nom problème inverse) : $J(t) = G^{-1} \cdot M(t)$.

L'inversion de la matrice de gain n'est pas immédiate simplement à cause du fait que $\boldsymbol{G}$ est une matrice rectangulaire. Autrement dit, plusieurs configurations des sources peuvent produire la même observation. Étant donnée qu'il est impossible de calculer G^{-1}, plusieurs techniques sont proposées pour estimer *une* configuration des sources qui explique le mieux les données. Généralement ces techniques s'appuient sur des a priori pour déterminer cette configuration.

Une difficulté rencontrée par ces techniques est le grand nombre de possibilités de configurations des sources qui expliquent les données si on applique le problème direct; c'est-à-dire un espace solutions très large. Par contre, on sait à l'avance qu'il n'y a qu'*une seule* configuration (celle désirée) qui représente l'activité cérébrale observée. Face à cette situation ces techniques appliquent des contraintes sur l'espace solution pour ne favoriser que les solutions qui satisfont ces contraintes. Il faut souligner que la formulation du problème inverse considérant ces contraintes rend le problème plus complexe.

Il est important de noter que le choix des contraintes à considérer est assez grand et c'est ce qui fait la différence entre ces techniques. Il faut aussi admettre que ces contraintes sont en faite des biais introduits dans le processus de la résolution. Par conséquence, les résultats de ces techniques dépendent de ces biais; ils peuvent être biaisés ou proches de la réalité. Cela dépend du degré fidélité de la représentation de la relation entre les sources et les données. Des biais ou contraintes qui ajoutent une information réelle concernant la relation entre les sources et les données va aider certainement à obtenir un résultat plus réaliste. Si par contre, l'information introduite s'éloigne de la relation entre les sources et les données alors le résultat sera certainement baisé.

On a, en général deux types de contraintes : contraintes spatiales et temporelles. Les contraintes spatiales sont liées à l'anatomie de la surface corticale et la distribution spatiale des sources; leurs positions et orientations. Ce sont des données statiques prédéterminées avant la résolution. Par contre, les contraintes temporelles sont des informations dynamiques qui décrivent l'évolution de l'activité électromagnétique dans le temps. Ces deux types de contraintes peuvent être combinées pour obtenir des résultats fiables.

Dans notre approche la définition des contraintes provient des observations de l'activité cérébrale. La contrainte spatiale qu'on considère ici est la modélisation de l'activité cérébrale en parcelles. Ainsi, une seule source ne peut pas être considérée active sans que ses voisines soient aussi engagées. En effet, l'activité cérébrale n'est perceptible que lorsqu'un

groupement de sources dans la même région sont engagées ensemble avec une certaines synchronie (Baillet S. 2001).

Notre modélisation en parcelles repose essentiellement sur deux types d'information : anatomique (proximité spatiale) et dynamique (données MEG). On note que cette parcellisation n'est pas prédéterminée au départ, elle sera obtenue à partir des données. Autrement dit, une parcellisation qui consiste à la répartition de la surface corticale en parcelles prédéterminées comme contrainte et que la résolution du problème inverse doit simplement déclaré parmi ces parcelles celles qui sont actives, est une parcellisation dont avec une contrainte spatiale très forte, étant donné qu'on ne dispose pas d'une carte anatomo-fonctionnelle valide pour tout sujet en observation.

2.5 Qualification des sources avec la MSP

Dans (J. Mattout 2005), les auteurs proposent une technique de régularisation appelée (MSP) *Multivariate Source Pre-localisation*. Ils proposent de réduire l'espace des sources du modèle. Cela permet d'améliorer le conditionnement du problème inverse. Cette technique permet la pré-localisation des sources qui expliquent le plus les données. Lors de la présence d'une activité on a juste quelques régions cérébrales qui sont engagées, ainsi on peut supposer qu'il y aura peu de sources qui seront pré-localisées. Ainsi, après cette pré-localisation, l'espace des sources est réduit et le conditionnement du problème inverse est considérablement amélioré.

La technique MSP repose sur les données suivantes: la matrice de gain et les données MEG pour la pré-localisation des sources. Après l'application de la MSP sur le modèle, on obtient des coefficients associés aux sources qu'on appelle *'coefficients d'activation'* ou plus simplement *'scores'*. La valeur de ces scores varie dans [0,1]. Un coefficient d'activation attribué à une source permet d'avoir une information relative sur son rôle possible d'expliquer les données. Ce coefficient correspond au degré (ou possibilité) que la source soit active ou non à un instant donné ou durant une période donnée.

Ces scores permettent non seulement d'évaluer le degré d'activation des sources mais ils sont aussi utilisés comme base de comparaison entre les sources. Pour assurer une comparaison équitable entre les sources indépendamment de leur localisation par rapport aux capteurs, on normalise les *lead-fields* de la matrice de gain $\boldsymbol{G}$. Rappelons nous que les *lead-fields* sont les vecteurs colonnes de la matrice $\boldsymbol{G}$ qui représentent la relation entre les sources et les capteurs tenant compte de leurs distances. Il important de souligner que l'influence de la variation de la distance entre une source et un capteur par rapport à une mesure peut être compensée par une variation de l'intensité sur la source. Autrement, deux sources à des distance différentes par rapport à un capteur peuvent générer les mêmes mesures si la différence des distances est ajustée par une différence proportionnelle des intensités des sources. Ainsi, il serait difficile de différencier le degré d'influence des sources si on considère leur position par rapport aux capteurs. Pour limiter cette influence on a besoin de normaliser les lead-fields. La normalisation d'un lead-field g_i de $\boldsymbol{G}$ soit $\widetilde{g_i}$ qui correspond à sa division sur sa norme $||g_i||$, ainsi on a $\widetilde{g_i} = \frac{g_i}{||g_i||}$. Après la normalisation de la matrice de gain $\boldsymbol{G}$, la comparaison de leur contribution par rapport aux données observés n'est plus influencée par leur position par rapport aux capteurs mais seulement leur capacité d'expliquer les mesures. L'algorithme de pré-localisation se compose des trois étapes qui suivent.

2.5.1 Décomposition en valeurs singulières

Cette étape consiste dans la décomposition en valeurs singulières (SVD) de la matrice de gain normalisée. Le résultat de cette décomposition est une factorisation de la forme suivante:

$$\tilde{G} = U\Lambda V^T \tag{2.1}$$

où

$\tilde{G}$: est la matrice de gain normalisée.

U: est la matrice composée des vecteurs colonnes u_i orthogonaux entre eux. Cette matrice est de dimension m x m.

Λ: est la matrice diagonale composée des valeurs propres σ_i.

V: est la matrice composée des vecteurs colonnes v_i orthogonaux entre eux. Cette matrice est de dimension n x n.
avec m: est le nombre des mesures dans le modèle et n: est le nombre des sources dans le modèle.

A partir du produit de cette décomposition la matrice de gain $\widetilde{\boldsymbol{G}}$ peut aussi être présentée comme suit (voir figure 2.9): $\widetilde{\boldsymbol{G}} = \sum_{i=1}^{m} \sigma_i \,.\, u_i \,.\, v_i$.

Dans le domaine du traitement d'images, on exploite la SVD pour reconstruire une image à partir de sa décomposition en sous images. L'image est reconstruite à partir des sous images qui correspondent aux vecteurs propres σ_i les plus élevées.

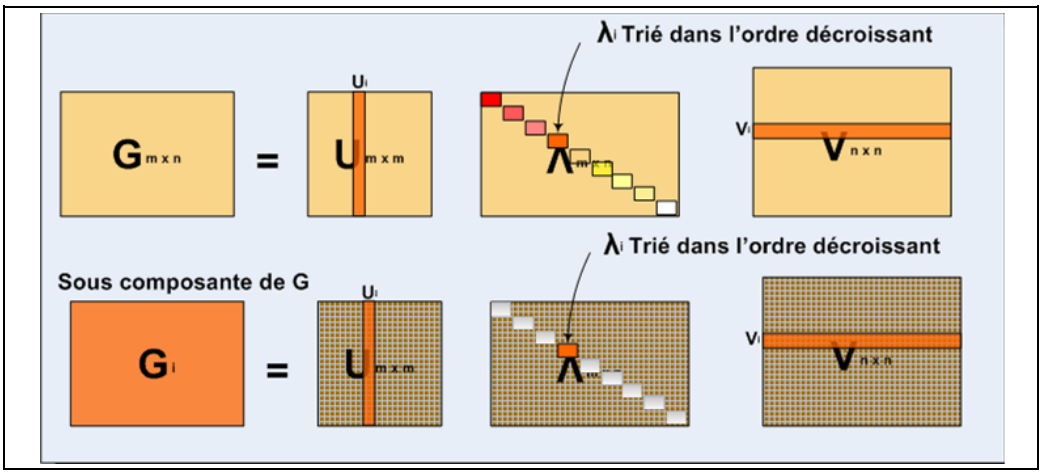


Figure 2.9 Décomposition en valeurs singulière de la matrice de gain et reconstruction de ses sous-composantes en utilisant la SVD

Les valeurs propres σ_i de $\boldsymbol{\Lambda}$ sont triées en ordre décroissant. Autrement, les sous composantes $\widetilde{\boldsymbol{G}}_i = \sigma_i \,.\, u_i \,.\, v_i$ de $\widetilde{\boldsymbol{G}}$, dont les σ_i sont les plus élevées; les premières dans $\boldsymbol{\Lambda}$, ont une contribution plus importantes que les dernières composantes. Si on considère que la matrice $\widetilde{\boldsymbol{G}}$ représente la relation entre l'ensemble des sources et les capteurs, par analogie, on peut aussi dire que $\widetilde{\boldsymbol{G}}_i$ représente la relation entre une source (virtuelle) et un capteur (virtuel). En fait, cette source virtuelle n'est qu'une combinaison donnée de sources et le capteur virtuel

n'est qu'une combinaison de capteurs. Partant de là, on peut conclure que les combinaisons des sources et des -capteurs les plus explicable sont celles dont les σ_i sont les plus importantes.

2.5.2 Quantification des Lead-Fields

Cette étape consiste essentiellement à la quantification des vecteurs colonnes de $\boldsymbol{U}$ (l*ead-fields virtuels*). On a considéré la normalisation de la matrice de gain pour permettre une comparaison équitable entre les vecteurs colonnes, dans cette étape on considère la même chose pour les mesures. Ainsi, on normalise la matrice des mesures pour permettre une comparaison équitable entre les capteurs sans considérer leurs distances par rapport aux sources.

Après normalisation de la matrice des mesures on a $\widetilde{\boldsymbol{M}} = \widetilde{\boldsymbol{G}} \cdot \boldsymbol{J}$ ou encore

$$\widetilde{M} = U\Lambda V^T J \tag{2.2}$$

Puisque V est une matrice orthogonale alors $U = U^T$ et on peut dire que :

$$\Gamma = U^T \widetilde{M} \qquad \text{avec} \qquad \Gamma = \Lambda V^T \cdot J \tag{2.3}$$

On remarque qu'on peut calculer $\boldsymbol{\Gamma}$, à partir de $\boldsymbol{U}$ et $\boldsymbol{M}$ même si $\boldsymbol{J}$ est inconnue. Ainsi, pour chaque '*lead-field*' $\boldsymbol{u_i}$ de $\boldsymbol{U}$ lui correspond $\boldsymbol{\gamma_i}$ calculé comme suit :

$$\gamma_i = u_i^T \widetilde{M}$$

Comme le montre la figure 2.10, γ_i est le produit de la projection de u_i sur les mesures M et c'est aussi un coefficient qui caractérise u_i.

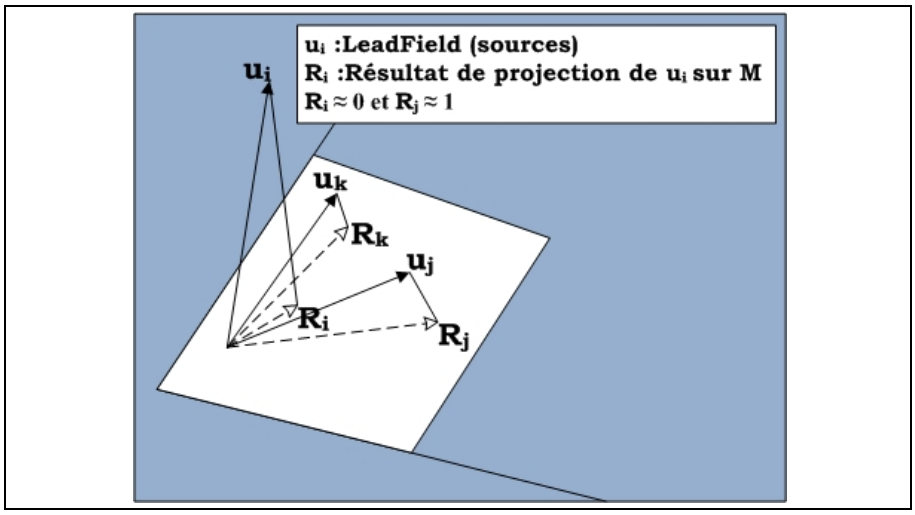


Figure 2.10 : Projection des vecteurs $\boldsymbol{v_j}$ sur l'espace des mesures M

D'après (J. Mattout 2005; Mattout J. 2005), $\boldsymbol{\Gamma}$ représente la corrélation entre les vecteurs propres de $\boldsymbol{U}$ et les données décrites dans $\widetilde{\boldsymbol{M}}$. En effet, dans (MARDIA, KENT et al. 1979), on définit un coefficient $\mathbf{R_i^2}$, qu'on appelle *'coefficient de corrélation multiple'* qui correspond à :

$$\mathrm{R}_\mathrm{i}^2 = \frac{u_\mathrm{i}^\mathrm{t}\ \widetilde{M}\left(\widetilde{M}^T\widetilde{M}\right)^{-1}\widetilde{M}^T u_\mathrm{i}}{u_\mathrm{i}^\mathrm{t}\ u_\mathrm{i}}$$

Après simplification on obtient : $\mathrm{R}_\mathrm{i}^2 = \gamma_\mathrm{i}(\Gamma^T\,\Gamma)\gamma_\mathrm{i}^\mathrm{t}$

Ce coefficient de corrélation multiple R_i^2 permet de comparer la contribution des vecteurs propres u_i par rapport aux mesures. A la sortie de cette étape, on obtient deux valeurs caractérisant u_i; (1) la valeur propre σ_i qui caractérise le degré de contribution d'une combinaison de sources et (2) un coefficient de corrélation R_i^2 qui caractérise la relation entre les vecteurs colonnes de $\boldsymbol{U}$ et les mesures. En utilisant ces deux coefficients, on peut sélectionner les u_i et réduire ainsi l'espace de la base $\boldsymbol{U}$ pour obtenir une nouvelle matrice $\boldsymbol{U_s}$ formé seulement des u_i considérés expliquant le mieux les données.

2.5.3 Scores MSP

Dans l'étape précédente, l'intérêt était de réduire l'espace des vecteurs de la base V pour ne garder que ceux qui expliquent le mieux les données observées. Dans cette dernière étape de la technique MSP, l'intérêt est de calculer un coefficient d'activation pour chaque source à partir de la base U_{s}. Ce coefficient indique la possibilité ou le degré qu'une source soit active.

A partir de la base $\boldsymbol{U_{\mathrm{s}}}$ on définit un projecteur $P_{\mathrm{s}} = U_{\mathrm{s}}\, U_{\mathrm{s}}^{T}$. Les mesures sont à leur tour projetées sur P_{s} pour sélectionner les mesures $\widetilde{M}_{\mathrm{s}}$ en corrélation avec l'espace des sources de la base U_{s}. La projection des mesures sur le projecteur P_{s} est exprimée comme suit:

$$\widetilde{M}_{\mathrm{s}} = U_{\mathrm{s}}\, U_{\mathrm{s}}^{T}\, \widetilde{\mathrm{M}} \tag{2.4}$$

où $P_{\mathrm{s}} = U_{\mathrm{s}}\, U_{\mathrm{s}}^{T}$ est un projecteur dans l'espace des capteurs.
Rappelons que les conditions que doit satisfaire un projecteur sont :

1- P_{s} doit être symétrique : $P_{\mathrm{s}}^{T} = P_{\mathrm{s}}$

On a; $P_{\mathrm{s}} = U_{\mathrm{s}}\, U_{\mathrm{s}}^{T} \iff P_{\mathrm{s}}^{T} = (U_{\mathrm{s}}\, U_{\mathrm{s}}^{T})^{T}$

$\iff P_{\mathrm{s}}^{T} = (U_{\mathrm{s}})^{T}\, (U_{\mathrm{s}}^{T})^{T} \iff P_{\mathrm{s}}^{T} = U_{\mathrm{s}}\, U_{\mathrm{s}}^{T}$

Donc: $P_{\mathrm{s}}^{T} = P_{\mathrm{s}}$

2- P_{s} doit aussi respecter la règle suivante :$P_{\mathrm{s}}^{2} = P_{\mathrm{s}}$

On a; $P_{\mathrm{s}}^{2} = P_{\mathrm{s}} P_{\mathrm{s}} \iff P_{\mathrm{s}}^{2} = U_{\mathrm{s}}\, U_{\mathrm{s}}^{T}\, U_{\mathrm{s}}\, U_{\mathrm{s}}^{T}$

$\iff P_{\mathrm{s}}^{2} = V_{\mathrm{s}}\, (V_{\mathrm{s}}^{T}\, V_{\mathrm{s}}) V_{\mathrm{s}}^{T} \iff P_{\mathrm{s}}^{2} = U_{\mathrm{s}}\, U_{\mathrm{s}}^{T}$ (vue que $U_{\mathrm{s}}^{T}\, U_{\mathrm{s}} = \mathrm{I}$)

Donc: $P_{\mathrm{s}}^{2} = P_{\mathrm{s}}$

A partir de la $\widetilde{M}_{\mathrm{s}}$ on définie le projecteur P_{m} tel que :

$$P_{\mathrm{m}} = \widetilde{M}_{\mathrm{s}} \left(\widetilde{M}_{\mathrm{s}}^{T}\, \widetilde{M}_{\mathrm{s}}\right)^{-} \widetilde{M}_{\mathrm{s}}^{T} \tag{2.5}$$

P_{m} vérifie aussi les conditions suivantes: $P_{\mathrm{m}}^{T} = P_{\mathrm{m}}$ et $P_{\mathrm{m}}^{2} = P_{\mathrm{m}}$.

P_{m} est un projecteur qui représente l'espace des mesures qui peuvent être expliquées par (en corrélation avec) l'espace des sources sélectionnées. Autrement, si on a une matrice donnée $\boldsymbol{X}$ qu'on projette sur P_{m}, on aura $P_{\mathrm{m}} \,.\, \boldsymbol{X} = \widetilde{M}_{\mathrm{s}} \left(\widetilde{M}_{\mathrm{s}}^{T}\, \widetilde{M}_{\mathrm{s}}\right)^{-} \widetilde{M}_{\mathrm{s}}^{T} \,.\, \boldsymbol{X}$. Si $\boldsymbol{X} = \widetilde{M}_{\mathrm{s}}$ alors le résultat de la projection $P_{\mathrm{m}} \,.\, \boldsymbol{X} = \widetilde{M}_{\mathrm{s}} = \boldsymbol{X}$. De même le produit de la projection suivante $\boldsymbol{X^T} \,.\, P_{\mathrm{m}} \,.\, \boldsymbol{X} = \boldsymbol{I}$ si $\boldsymbol{X} = \widetilde{M}_{\mathrm{s}}$. Par contre, si $\boldsymbol{X}$ est une matrice orthogonal par rapport à $\widetilde{M}_{\mathrm{s}}$ alors $\boldsymbol{X^T} \,.\, P_{\mathrm{m}} \,.\, \boldsymbol{X} = \boldsymbol{0}$. Ainsi, le projecteur P_{m} permet de mesurer le degré de corrélation entre l'espace des mesures et les données exprimées dans $\boldsymbol{X}$.

Considérant le même principe, la projection de chaque *lead-field* de G (l'espace des sources) sur P_{m} (dans l'espace des mesures) permet d'obtenir une valeur indiquant le degré d'activation de chaque source du modèle; c'est le *score*. Ce coefficient est en fait un indice de corrélation entre les sources et les mesures. Ainsi, le résultat de cette projection est représenté dans l'expression suivante:

$$A_{\mathrm{s}} = \tilde{G}^{T}\, P_{\mathrm{m}} \tilde{G} \tag{2.6}$$

où $a_i = \tilde{g}_i^t \cdot P_m \cdot \tilde{g}_i$ et $a_i = \|P_m \cdot \tilde{g}_i\|^2 = \tilde{g}_i^t \cdot P_m^T \cdot P_m \cdot \tilde{g}_i$

On peut lire sur la diagonale de la matrice A_{s} les coefficients d'activation pour chaque source. Comme illustré dans la figure 2.11, ce coefficient peut être comparé au cosinus de l'angle qui sépare deux vecteurs (ici sources et mesures). Lorsque cet angle est proche de '0', la probabilité que la source soit active est forte, c'est le contraire dans le cas où cet angle est droit. Il est important de remarquer que pour les sources où cet angle tend vers 0.5, il est difficile de décider l'état de la source.

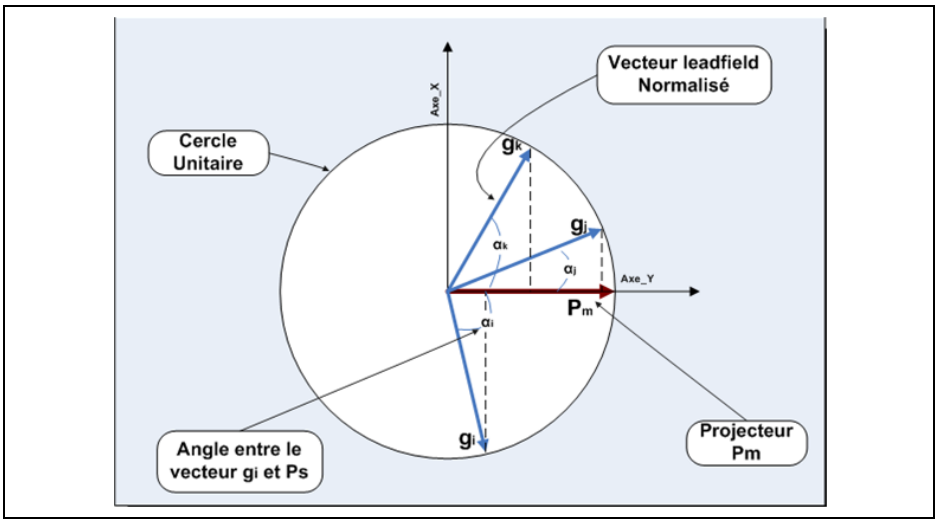


Figure 2.11 Calcul des coefficients d'activation en projetant les leadfields $\widetilde{\boldsymbol{g}}_{\boldsymbol{i}}$ de $\widetilde{\boldsymbol{G}}$ sur le projecteur$\boldsymbol{P}_{\boldsymbol{m}}$

Ainsi, avec la MSP on a une classification des sources selon leur score qu'on exploite dans la prochaine étape de notre algorithme de sélection.

A l'issue de cette classification des scores, on n'a pas encore décidé les états (actif ou non) des sources du modèle. En faite, on a besoin d'un critère de sélection des scores qui permet de discriminer les sources et identifier celles qui sont impliquées dans les données observées. Étant donnée que les scores varient entre 0 et 1, on a besoin de fixer un seuil entre 0 et 1 qui permet de départager les sources entre celles qu'on considère 'implicables' dans les données et celles qui ne le sont pas.

Une façon simple et 'naïve' de déterminer ce seuil de sélection c'est de dire que les sources ayant un score au-delà de 0.5 sont impliquées et celles en dessous ne le sont pas. Bien que cette façon à l'avantage d'être simple, il faut reconnaitre que la distribution des scores varie selon l'activité observée et par conséquent, le seuil à fixer doit tenir compte des données observées.

Dans ce travail, on a choisi une technique plus élaborée pour fixer ce seuil de sélection qui considère les données observées. Cette technique, même si elle est un peu complexe, est aussi élaborée dans le sens qu'elle permet de fixer un seuil de sélection considérant les données et cela avec un contrôle d'erreur des décisions de sélection. On permet ainsi d'augmenter l'efficacité du processus de discrimination entre les sources. Dans le prochain chapitre, on présente cette technique et son exploitation.

CHAPITRE 3

RÉGULARISATION DU PROBLÈME INVERSE

Ce chapitre porte sur le seuillage des scores et la sélection des sources. Pour déterminer un seuil de sélection des scores, on utilise le test d'hypothèse FDR (False Discovery Rate) (Benjamini and Hochberg 1995) suivi d'une parcellisation. La sélection des sources est une réduction de l'espace des sources, ce qui permet la régularisation du problème inverse. A partir de cette sélection des sources, on forme des parcelles en utilisant notre technique de parcellisation. Ces deux techniques de sélection et de parcellisation sont détaillées dans ce chapitre.

3.1 Contrôle des faux positifs

La FDR est une technique de test d'hypothèses multiples qu'on exploite pour tester une distribution de scores par rapport une hypothèse nulle. Pour ce test d'hypothèse, l'hypothèse nulle H_0 est 'il n'y a pas d'activité spécifique à cette position'. La spécifité sera calculée par rapport à un modèle obtenue à partir des données de *baseline* (segment de données sans activité spécifique) qu'on décrira plus loin. La FDR permet de tester cette hypothèse nulle pour l'ensemble des sources en un seul test tout en assurant un contrôle des faux positifs, c'est à dire l'erreur de type I.

Avec un test d'hypothèse, on a toujours des erreurs de décisions qui l'accompagnent. Les 4 cas possibles avec un test d'hypothèse, incluant les erreurs de décisions, sont les suivants:

- S_{aa} : l'ensemble des sources déclarées actives (H_0 est rejetée) et elles sont réellement actives.
- S_{ai} : l'ensemble des sources déclarées actives et elles ne sont pas réellement actives. Ce sont les faux positifs.
- S_{ii} : l'ensemble des sources déclarées inactives (H_0 n'est pas rejetée) et elles sont réellement inactives.

- S_{ia} : l'ensemble des sources déclarées inactives et elles sont réellement actives. Ce sont les faux négatifs.

A partir de ces données, on peut déduire l'ensemble des sources réellement actives, réellement inactives, considérées actives et considérées inactives comme suit :

- R_a est l'ensemble des sources réellement actives et ca correspond à:

$$R_a = S_{aa} + S_{ai}$$

- R_i est l'ensemble des sources réellement inactives et ca correspond à:

$$R_i = S_{ia} + S_{ii}$$

- C_a est l'ensemble des sources considérées actives et ca correspond à:

$$C_a = S_{aa} + S_{ia}$$

- C_i est l'ensemble des sources considérées inactives et ca correspond à:

$$C_i = S_{ii} + S_{ai}$$

Tableau 1.3 Test d'hypothèse avec les erreurs de décisions

Tableau du test d'hypothèse		**La décision**		**Sources réellement actives vs. inactives**
		H_0 rejeté	H_0 n'est pas rejeté	
La réalité	H_0 est vraie	Faux Positifs (S_{ia}) Erreur type I (α)	Vraie Positifs (S_{aa})	Ensemble des sources réellement actives (R_a)
	H_0 est fausse	Vraie Négatifs (S_{ii})	Faux Négatifs (S_{ai}) Erreur type II (β)	Ensemble des sources réellement inactives (R_i)
Sources considérées actives vs. inactives		Ensemble des sources considérées actives (C_a)	Ensemble des sources considérées inactives (C_i)	Ensemble des sources du modèle (N)

A partir des valeurs C_a et S_{ia}, on peut déduire le taux de faux positifs:

$$FDR = \frac{S_{ia}}{C_a} = \frac{S_{ia}}{S_{aa} + S_{ia}}$$

En fait, en pratique, on ne dispose que des valeurs C_a et C_i.

Dans la figure 3.1, on présente un exemple de distribution de scores des sources d'une activité observée où on fixe un seuil de sélection arbitraire $\lambda^* = 0.75$. Dans cet exemple, on a une région en dessous de la courbe entre 0.75 et 1. En faite, cette région correspond au taux des *faux positifs* S_{ia}; les hypothèses nulles faussement rejetées.

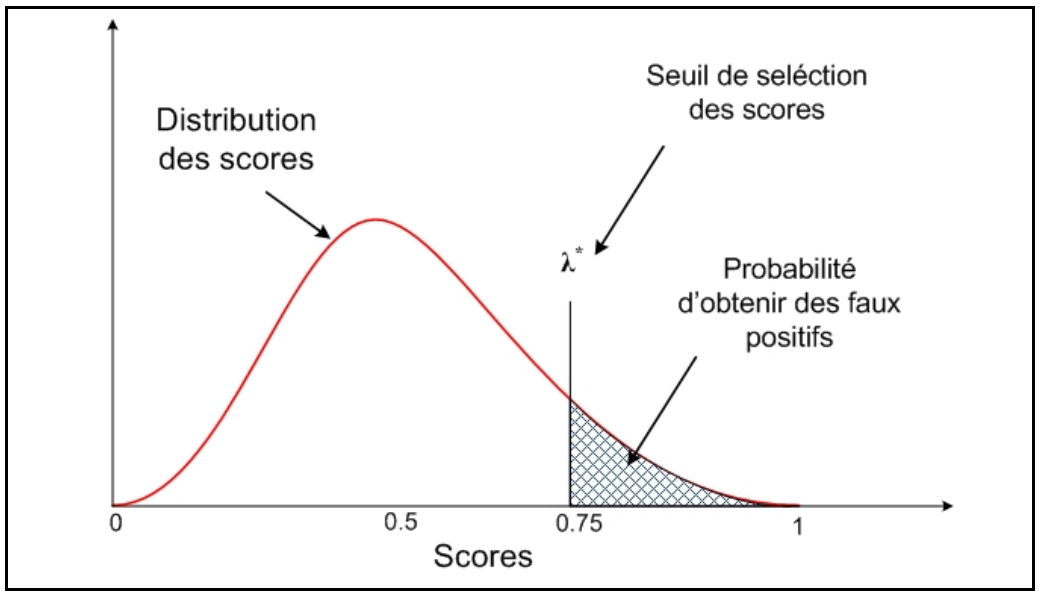


Figure 3.1 Seuillage et sélection d'une distribution de scores

Pour mieux expliquer comment on peut contrôler ce taux d'erreur, dans la figure 3.2 on présente deux distributions de scores; l'hypothèse nulle H_0 et les scores des données observées. On fixant un seuil de sélection, ici $\lambda^* = 0.75$, on a deux groupes des scores, ceux en dessous au seuil et ceux au-delà, dans la zone hachurée. Ce sont les positifs. Parmi les positifs C_a de la zones hachurée, on a les vrais positifs S_{aa} et les faux positifs S_{ia} (voir figure 3.2). Ainsi, étant donné H_0, les données observées et le taux des faux positifs toléré, la

technique FDR permet de définir un seuil de sélection des scores qu'on explique dans la prochaine section.

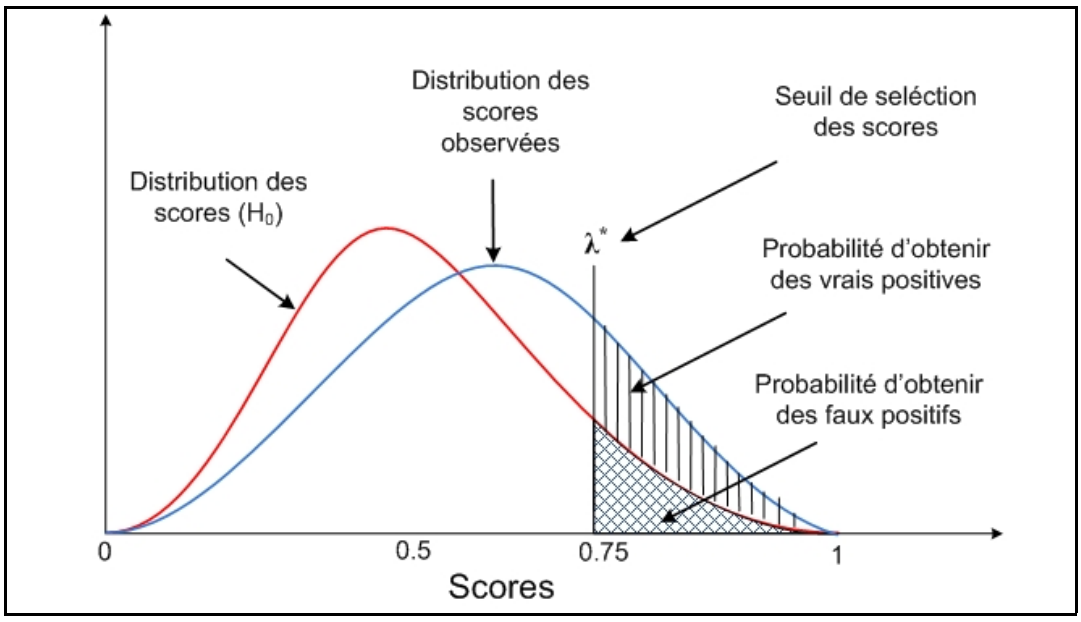


Figure 3.2 Contrôle du taux des faux positifs basées sur l'hypothèse nulle et un seuil de sélection FDR

3.2 Sélection des sources utilisant la FDR - False Discovery Rate

En général, une façon simple de sélectionner les scores est de fixer un seuil arbitraire (ex. un pourcentage) qu'on applique sur l'ensemble des données observées. Une façon un peu plus élaborée est de définir un seuil de sélection est d'utiliser un test d'hypothèse pour chaque élément des données observées. Une technique de seuillage qui permet de définir un seuil pour l'ensemble des données observées est la FDR. C'est une technique de test d'hypothèses multiples qu'on utilise pour calculer un seul seuil de sélection pour l'ensemble des hypothèses.

Dans cette section, on présente comment on exploite la technique FDR (False discovery Rate) (Benjamini and Hochberg 1995) pour calculer un seuil de sélection sur les scores qui

garantit un contrôle du taux des faux positifs. Cette sélection permet de discriminer entre les sources selon leurs scores pour décider celles qu'on considère actives et celles qui ne le sont pas.

Rappelons que l'objectif de cette étape est de réduire la dimension du modèle à un sous-ensemble qui peuvent expliquer le mieux les données observées dans une fenêtre temporelle donnée. Après sélection/réduction des sources, le conditionnement du problème inverse est amélioré.

On considère la technique FDR pour la sélection des sources pour les raisons suivantes: (1) elle permet d'effectuer des tests d'hypothèses multiples en un seul test. Dans notre cas, on a besoin d'effectuer un test d'hypothèse sur l'ensemble des sources du modèle; (2) elle permet de contrôler le taux des faux positifs des sources sélectionnées. Cela permet le control du taux d'erreurs de la sélection et (3) elle permet de définir un seuil adapté aux données observées considérant le taux d'erreur qu'on veut tolérer.

3.2.1 Présentation

La technique FDR est introduite pour la première fois par Benjamini et Hochberg en 1995 (Benjamini and Hochberg 1995). Cette technique fut exploitée dans plusieurs domaines (i.e. médical, pharmaceutique, marketing etc.). Elle est toujours un sujet de recherche dans des domaines, notamment dans le domaine de l'imagerie cérébrale. La FDR est une technique de test d'hypothèses multiples qui permet de prévoir/contrôler le taux d'erreur des faux positifs (des hypothèses faussement rejetées) sur l'ensemble des hypothèses rejetées. La FDR est définie comme suit :

$$FDR = E(Q) \qquad Q = \begin{cases} V/R, & si\ \ R > 0. \\ 0, & si\ \ R = 0. \end{cases} \tag{3.1}$$

où R est le nombre d'hypothèses nulles rejetées et V est le nombre des faux positifs. V et Q sont des variables aléatoires inconnues. Cependant, il est possible de prévoir en *moyenne* le taux des hypothèses rejetées par erreur parmi l'ensemble des hypothèses rejetées; cela correspond à $\mathrm{E}(\mathrm{Q})$.

Si on fixe une valeur α entre 0 et 1 on a toujours cette inégalité:

$$V/R \cdot \alpha \leq \alpha \qquad \text{étant donnée qu'on a toujours} \quad V \leq R$$

Puisque que V et R sont toujours positifs, on peut dire que

$$E(Q) = V/R \leq V/R \cdot \alpha \leq \alpha$$

La technique FDR permet de déterminer un seuil de sélection P^* tel que :

$$P^* \leq \frac{i}{n} \alpha \tag{3.2}$$

Généralement, α correspond à la tolérance des faux positifs admissibles dans un test d'hypothèse, sa valeur est choisie entre 1 et 5%. i correspond à l'indice d'une hypothèse nulle et n est le nombre total des hypothèses nulles

3.2.2 Seuillage et sélection des scores par la FDR

Comme pour tout test d'hypothèse, on doit d'abord établir l'hypothèse nulle H_0 . Dans ce travail, l'hypothèse nulle H_0 de la FDR correspond à la distribution des scores des données observées lorsque les sources sont non-actives. Le choix de cette hypothèse nulle repose sur le fait qu'il est plus facile, en théorie, d'obtenir une distribution des scores où les sources sont non-actives. Autrement dit, les scores correspondent aux données où on a une absence d'activité spécifique. L'activité des sources observée correspond à de l'activité physiologique. On remarque que le cas contraire où l'ensemble des sources sont actives comme hypothèse nulle, ce n'est pas pratique.

Le test FDR est basé sur la comparaison de deux distributions des scores : (1) la distribution pour la quelle l'ensemble des sources sont au repos et (2) la distribution où les sources sont engagées dans une activité particulière.

Distribution des sources au repos

En pratique, les états des sources de la deuxième distribution (sources actives) sont à déterminer. Par contre, les scores de la première distribution sont considérés comme données statistiques qui suivent une loi statistique donnée. En fait, plusieurs travaux (Lapalme, Lina et al. 2006) ont supposé que ces scores suivent une loi beta sans toutefois valider cette hypothèse. Dans ce travail, on effectue un test de validation empirique de cette supposition qui s'avère valide (voir la prochaine section).

Il faut aussi ajouter que pour nos tests de simulation, on a besoin d'un modèle analytique qui représente cette distribution. Ainsi, considérant le score de chaque source (non-active) comme une variable aléatoire X dont la valeur varie entre 0 et 1, on suppose qu'il suit une loi Beta définit comme suit:

- X est une variable aléatoire ayant une valeur entre [0,1], suit une loi de probabilité beta, ayant les paramètres α et β tel que ($\alpha > 0$ et $\beta > 0$), noté ($X \hookrightarrow \boldsymbol{B}(\alpha, \beta)$)
- $\boldsymbol{B}(\alpha, \beta)$ est continue et admet une fonction de densité de probabilité sur cette intervalle défini comme suit:

$$f(x) = \begin{cases} \frac{1}{B(\alpha,\beta)}\, x^{\alpha-1}\,(1-x)^{\beta-1} & si\ x \in [0,1]; \\ 0 & sinon. \end{cases} \tag{3.3}$$

où $\boldsymbol{B}(\alpha, \beta)$ est une loi beta. L'espérance de la loi Beta est :

$$E(X) = \frac{\alpha}{\alpha + \beta}$$

et la variance

$$V(x) = \frac{\alpha\,\beta}{(\alpha+\beta)^2\,(\alpha+\beta+1)}$$

Dans le cas de ce travail, la variable *X* correspond aux scores a. En faisant l'hypothèse que les scores des sources non-actives suivent une loi Beta, pour le calcul de leurs p-values P_i, on utilise la fonction de distribution cumulative beta suivante:

$$P_i = \int_{a_i}^{1} f(s)ds$$

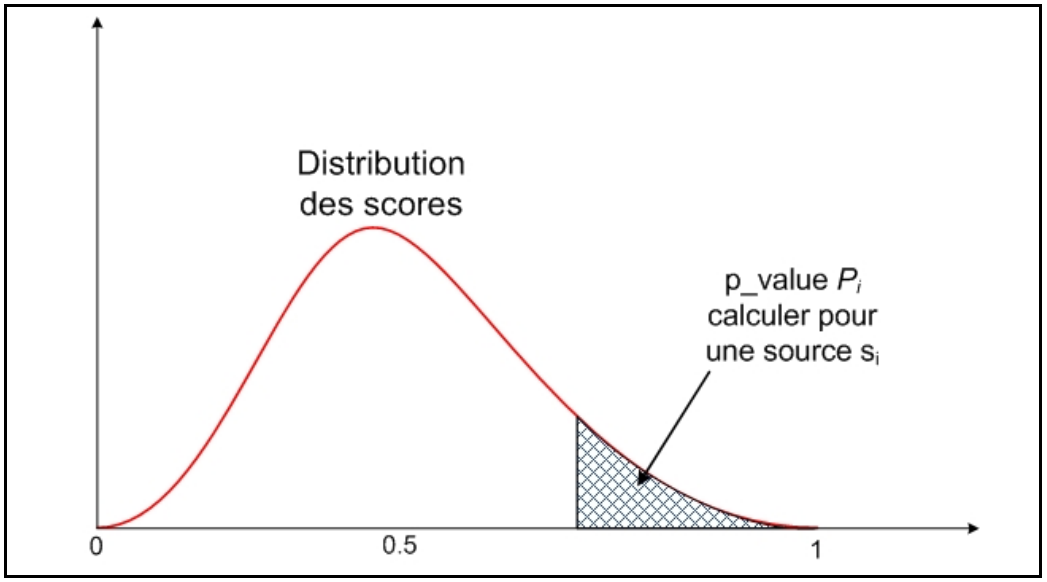


Figure 3.3 P-value d'une source pour une distribution des scores d'une donnée

pour tout score a_i lui correspond une p-value P_i avec $i = 1,\ 2, 3, \ldots, n$ où n est le nombre de sources.

Il est à noter que les p-values P_i des scores a_i sont calculées pour une prise de mesure M_j durant une fenêtre temporelle. On remarque aussi que pour définir notre hypothèse nulle H_0 et obtenir une distribution des scores qui la définit, on a plusieurs scenarios possibles. Par exemple, effectuer une prise de mesure où le sujet n'effectue aucune tâche

particulière. Ces mesures seront utilisées comme *baseline* pour calculer les scores des sources lorsqu'elles sont au repos.

Calcul du seuil de sélection FDR

La figure 3.4 illustre comment la FDR permet de déterminer le seuil de sélection. D'une part, on a la droite (discontinue en rouge) qui correspond à l'expression $FDR_i = \frac{i}{n}\,\alpha_{FDR}$ où α_{FDR} est un seuil de tolérance. D'autres part, on a la courbe (en continue en bleu) qui correspond aux p-values des scores triés en ordre croissant. Comme on peut le voir sur la figure 3.4, le premier point de croisement de la droite $FDR_i = \frac{i}{n}\,\alpha_{FDR}$ et la courbe des p-values correspond à un point (c, P_c). En fait, on a $i = r$ où $P_r \leq \frac{r}{n} \cdot \alpha_{FDR}$ correspond au seuil de sélection des scores ($\lambda^* = a_r$).

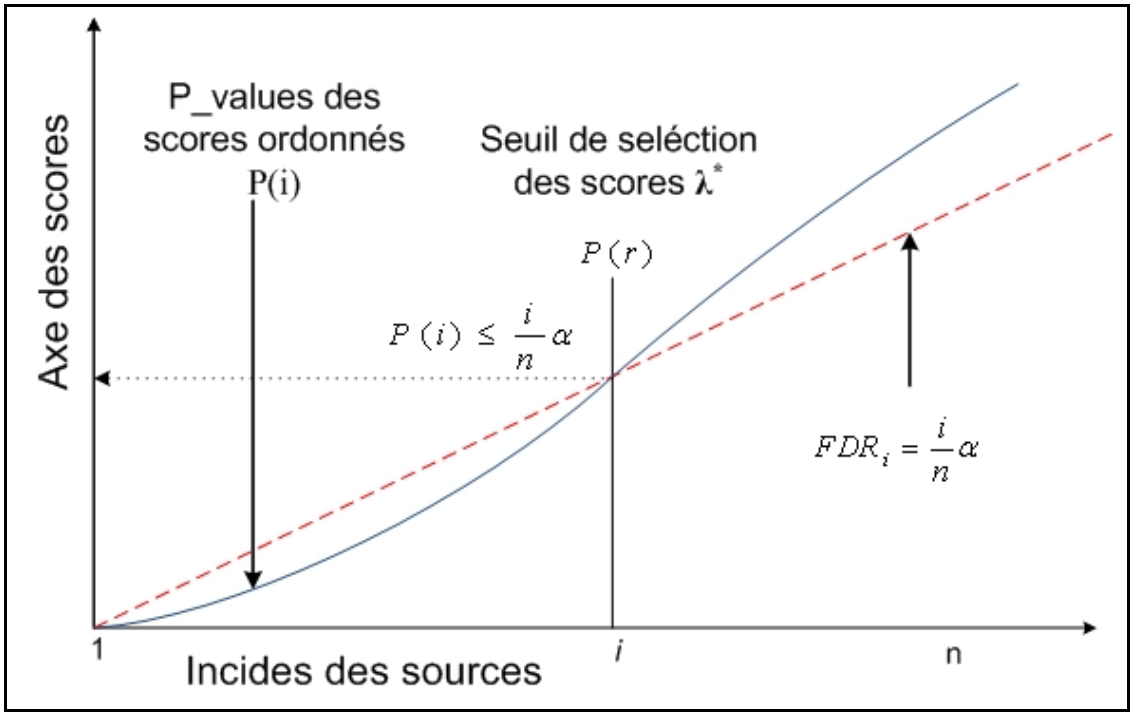


Figure 3.4 Calcul du seuil de sélection utilisant la technique FDR

Ainsi, après ce seuillage et sélection des scores, on définit un ensemble de sources sélectionnées $\mathcal{P}_s$ considérées actives, telles que : $\mathcal{P}_s = \{ s_i \; source \; active \; | \lambda^* \leq a_i \}$. Avec $\mathcal{P}_s$ est l'ensemble des sources dont le score est supérieur au seuil de sélection λ^*.

Algorithme de sélection par FDR :

L'algorithme de sélection des scores par le test FDR se compose des 4 étapes suivantes :

Ces étapes sont précédées par une étape préliminaire où on utilise la loi beta de l'hypothèse nulle paramétrée par (α_0 et β_0) qui sont déterminées à partir de la moyenne et de la variance des scores obtenues avec le signal de *baseline* (voir prochaine section).

1- On fixe un seuil de tolérance α_{FDR}= 5%, par exemple.
   A partir d'une fenêtre de mesures MEG, on calcule les scores a_i des sources et leurs p-values P_i correspondantes.

2- On trie les p-values P_i en ordre croissant
$$P_1 \leq P_2 \leq P_3 \leq \leq \cdots P_i \leq \cdots \leq P_r \leq \cdots P_n$$
   où $i = 1, 2, \ldots, n$ et n est le nombre total des sources.

3- Recherche de l'indice r le plus élevé tel que P_r vérifie la condition suivante:
$$P_r \leq P^* = \frac{r}{n} . \alpha_{FDR}$$
   On remarque qu'il est possible qu'aucune source ne soit active, c'est le cas auquel on a $r = 1$.

4- Finalement pour toutes, les sources dont les p-values sont inferieures au seuil P^*, l'hypothèse nulle H_0 est rejetée. Ainsi, ces sources sont considérées comme actives.

Il faut noter que la technique FDR a l'avantage de déterminer un seuil de sélection variable en fonction des données observées. Comme le montre la figure 3.5 pour des p-values de deux distributions de scores différentes avec le même seuil de tolérance aux faux positifs, la FDR détermine deux seuils de sélections différents (λ^1et λ^2).

Pour conclure, la technique FDR permet d'avoir un seuil de sélection des scores considérant les données observées (dans le signal). Ainsi, elle est considérée comme une technique de *seuillage adaptatif*.

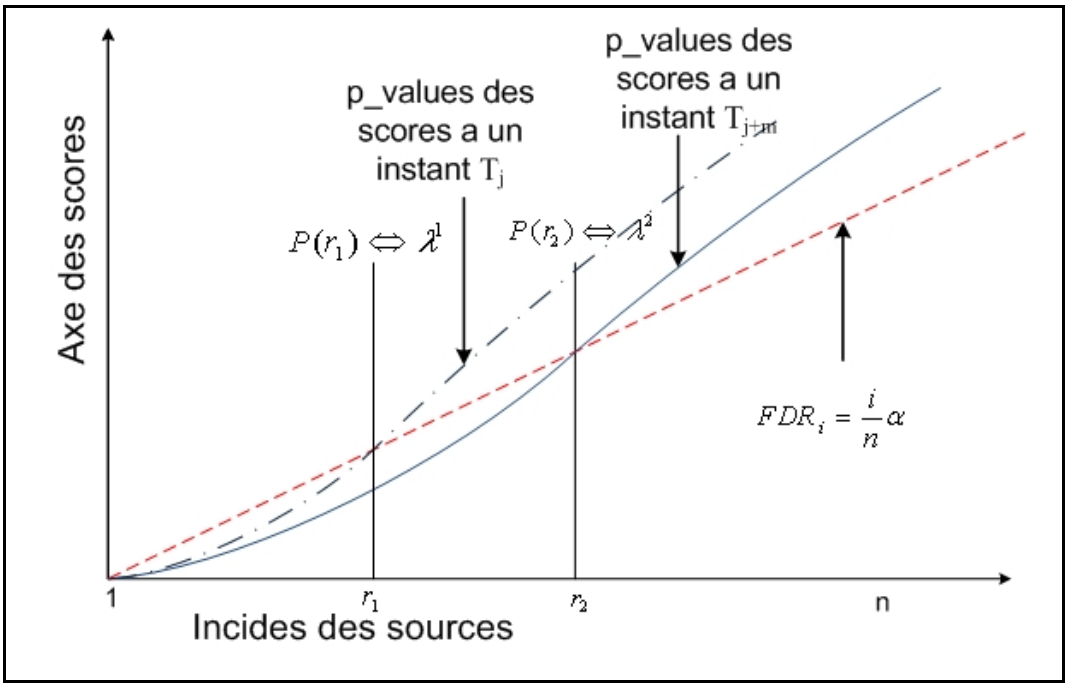


Figure 3.5 Sélection par FDR qui est une technique de seuillage adaptative

3.3 Validation et choix des paramètres

Dans la section précédente, on a présenté la technique de sélection des sources considérant les données et l'hypothèse nulle. Pour définir l'hypothèse nulle, on utilise les scores des sources 'au repos' modélisées par une loi beta dont les paramètres (α_0, β_0) sont estimées a partir de ces scores. A notre connaissance cette hypothèse n'a jamais été validée.

Dans cette section, on présente une validation empirique de cette hypothèse. En fait, cette section se compose de deux parties : (1) validation empirique de cette hypothèse; les scores des sources au repos suivent une loi beta et (2) estimation des paramètres de la loi beta (α_0, β_0) à choisir comme modèle pour l'hypothèse nulle à partir de la *'baseline'* (données au repos).

3.3.1 Validation du modèle pour la loi '*Beta* '

Considérons des mesures au repos où le sujet n'exerce aucune activité particulière. En supposant que les scores des sources suivent une loi beta, on calcule les paramètres α_0 et β_0 de cette loi. Ces paramètres sont calculés à partir du calcul de la moyenne μ_0 et la variance ν_0 des scores comme suit :

$$\alpha_0 = \frac{\mu_0(\mu_0 - \mu_0^2 - \nu_0)}{\nu_0}$$

Et

$$\beta_0 = \frac{\alpha_0}{(\mu_0 - \alpha_0)}$$

Une fois les paramètres α_0 et β_0 obtenus, on peut déterminer la loi beta décrivant ces scores en appliquant la formule suivante :

$$\boldsymbol{f}(a_i) = \frac{1}{B(\alpha_0, \beta_0)}\, a_i^{(\alpha_0 - 1)}\, (1 - a_i)^{(\beta_0 - 1)}$$

où a_i est le score de la source s_i.

Il reste, maintenant, à déterminer si les scores correspondent à la loi beta dont les paramètres (α_0, β_0) ont été estimés à partir de la moyenne et la variance des scores. Pour cela, on propose un algorithme qu'on appelle ***BetaFit*** sous-forme d'un test statistique de validation empirique sur des échantillons de mesures. Ces derniers sont des fenêtres temporelles d'une prise de mesure, localisées à différentes positions (voir figure 3.6).

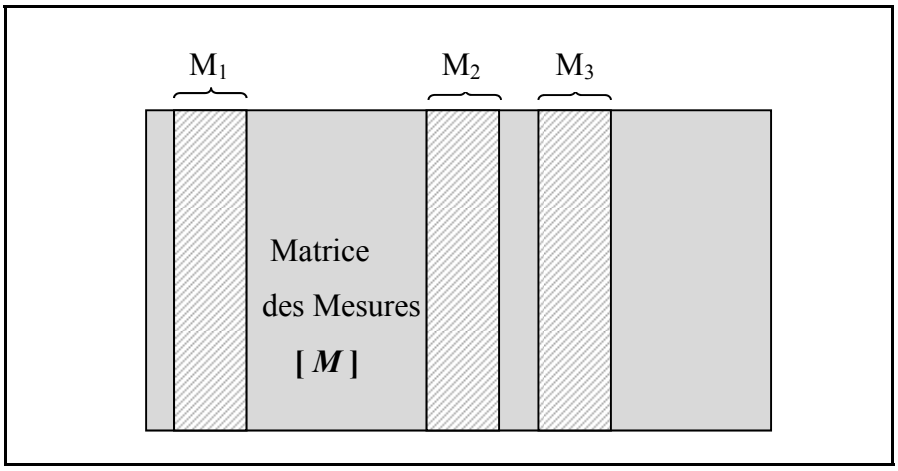


Figure 3.6 Sous-matrices de la matrice des mesures choisies aléatoirement

Algorithme *BetaFit*

Cet algorithme se compose de 5 étapes :

Étape 1:

Définir un ensemble de fenêtres (données de mesures) dont la position temporelle est prise aléatoirement. Ainsi, dans cette étape on a R groupes de N fenêtres temporelles. Les fenêtres du même groupe sont de même taille w.

Étape 2 :

On calcule les scores des sources pour chaque échantillon des données choisit dans l'étape précédente.

Ainsi, à chaque fenêtre w_j correspond un vecteur des scores $a_{i,j}$ pour estimer l'état d'activation des sources.

Étape 3 :

Dans cette étape, on établit une donnée statistique en utilisant les scores calculés dans l'étape précédente pour chaque fenêtre w_j. Les scores sont groupés dans des classes b_x (x est l'indice des plages). Pour chaque classe b_x, on calcule le nombre des scores qui tombent dans cette plage. Ainsi, on obtient une statistique de la distribution des scores a_i.

La figure 3.7 ci-dessous illustre l'exemple d'une distribution de scores comme statistiques des scores a_i.

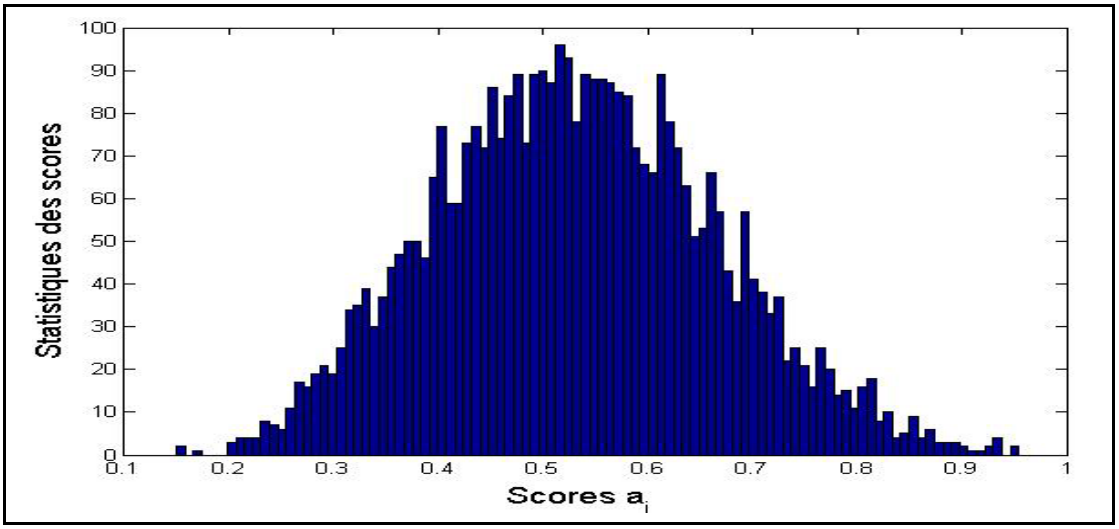


Figure 3.7 Histogrammes des scores MSP

A partir de cette statistique, on estime la densité de probabilité P_i de chaque score en utilisant la relation suivante :

$$P_i = \frac{a_i}{\sum_i a_i \cdot (d)} \qquad \text{où } d = b_{x+1} - b_x$$

Étape 4 :

Dans cette étape, on calcule les paramètres (α_0, β_0) de la distribution de loi beta, à partir de la moyenne et la variance (μ_0 ,ν_0) de la distribution de probabilité des scores calculés dans l'étape précédente. Ainsi, on calcule les paramètres α_0 et β_0, de cette loi beta, à partir de la distribution des scores utilisant les formules suivantes :

$$\alpha_0 = \frac{\mu_0(\mu_0 - \mu_0^2 - \nu_0)}{\nu_0} \qquad \beta_0 = \frac{\alpha_0}{(\mu_0 - \alpha_0)}$$

Étape 5 :

Dans cette étape, on calcule l'écart entre la distribution de probabilité des scores et la loi beta décrite avec les paramètres (α_0, β_0). Cela permet de vérifier si la fonction de densité de probabilité, décrite par les probabilités P_i (présentée par la courbe en rouge dans la figure 3.8 correspond à la loi beta avec paramètres α_0 et β_0 (présentée dans la même figure par la courbe en bleu). La superposition des deux courbes, pour une fenêtre de mesures donnée, montre que la fonction de densité de probabilité peut très bien modélisée par une loi beta.

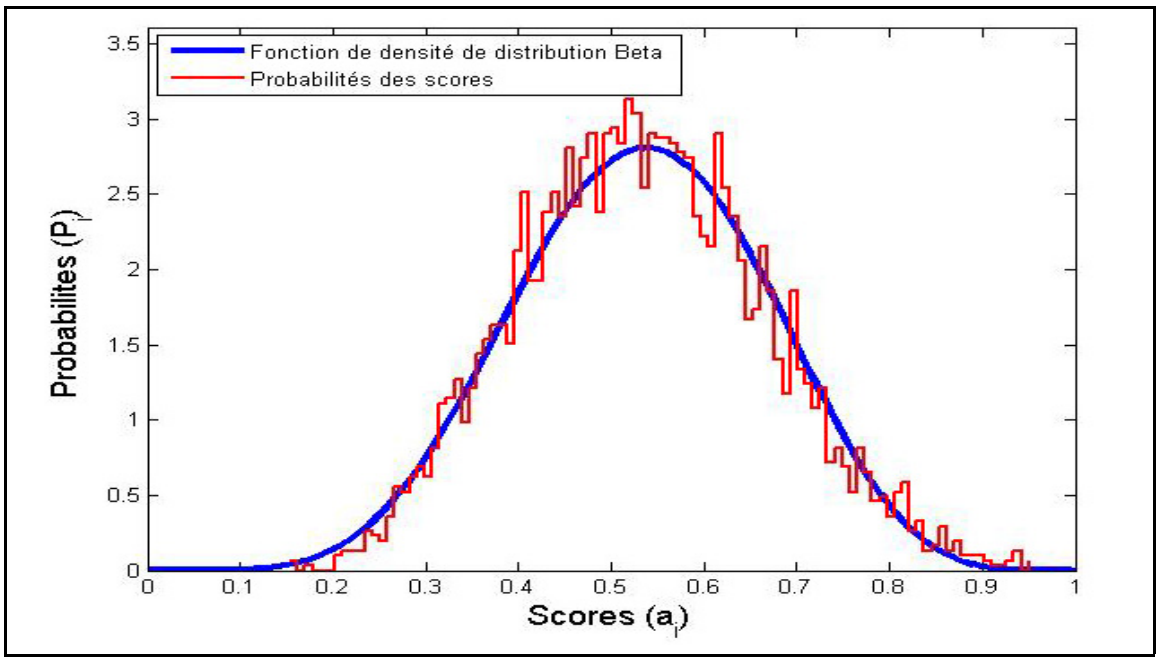


Figure 3.8 Superposition de la loi beta et la fonction des probabilités des scores

Pour avoir une meilleure idée concernant cette superposition, on calcule l'erreur entre la loi beta avec les paramètres α_0 et β_0 et la fonction de probabilité des scores. Les résultats de ce calcul seront présentés plus tard.

Étape 6 :

Reprendre les étapes 1 à 5 pour chaque groupe R d'échantillons avec des tailles des fenêtres différentes. Ces fenêtres varient entre 3 à 11 échantillons par fenêtre.

Résultats

Dans la pratique, on utilise la fonction '*betafit*' de l'outil ***Matlab*** qui prend comme paramètre d'entrée la distribution des P_i et on calcul les paramètres α_0 et β_0 ainsi que leurs intervalles d'erreurs.

La figure 3.9 montre le calcul d'écarts de ces paramètres pour 100 fenêtres dont la taille varie entre 3 à 11 d'enchantions/fenêtre. Cette figure d'erreurs montre des écarts très marginaux pour les deux paramètres (α_0 et β_0).

Pour pouvoir exploiter ce résultat dans le processus de seuillage et sélection des sources on a besoin de fixer des valeurs pour ces paramètres (α_0 et β_0). Dans la prochaine section, on présente comment l'hypothèse de la stationnarité permet de choisir ces valeurs.

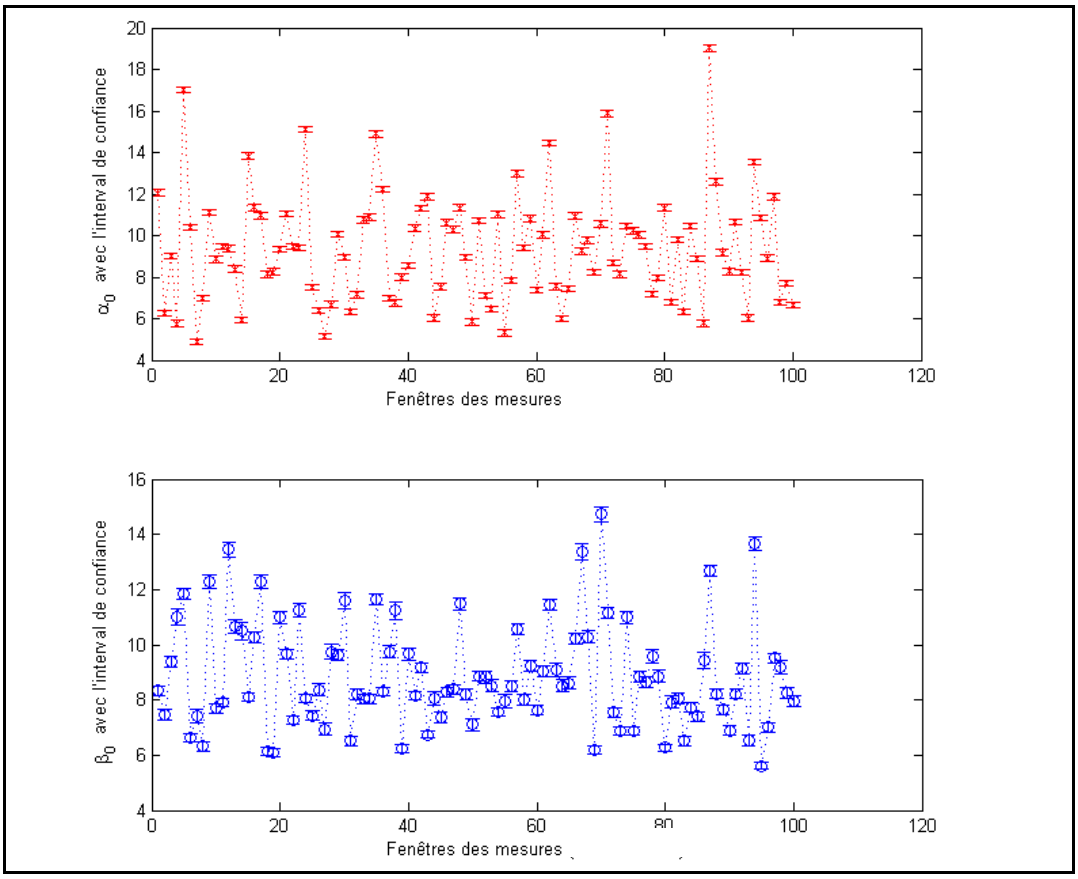


Figure 3.9 Erreurs d'estimation des paramètres α_0 et β_0 pour des fenêtres de 3 à 11 échantillons

3.3.2 Choix des paramètres (α_0 et β_0)

Le test de validation présenté dans la section précédente, a permis de vérifier que les scores des sources lorsqu'elles sont au repos suivent une distribution beta. Ainsi pour modéliser une *baseline*, il suffit de fixer des valeurs les paramètres (α_0 et β_0) d'une la loi beta et ensuite calculer les scores de l'hypothèse nulle. La question concernant le choix de ces paramètres (α_0 et β_0) porte sur leur variance en fonction de la largeur de la fenêtre.

Le but de ce test 'Betafit' est de déterminer la taille qui reflètent le plus l'activité des sources au repos. Autrement dit, la problématique est de déterminer la taille de fenêtre qui correspond

à la stationnarité temporelle de l'activité observée. Ainsi, avec une petite fenêtre, on n'observe qu'une partie de l'activité, par contre, avec une grande fenêtre, on observe plus que l'activité.

Dans le reste de cette section, on présente l'algorithme qui permet d'étudier la variance des paramètres beta en fonction la taille de la fenêtre observée. A la fin de cette étude on détermine la taille de fenêtre à choisir pour déterminer les valeurs des paramètres beta qui décrivent le mieux l'activité physiologique et la distribution des scores de '*baseline*'.

Algorithme du choix des paramètres beta

Pour établir des statistiques concernant la variance des paramètres beta, on utilise des mesures (*baseline*) dont la taille globale est de 501 échantillons. Cet algorithme se compose des étapes présentées ci-dessous:

Étape 1:
Dans cette étape, on établit un jeu de fenêtres de mesures utilisant la *baseline* (501 échantillons). Les tailles de fenêtres considérées varient entre 3 à 51 échantillons par fenêtre. Pour chaque taille donnée, on a 100 groupes de 40 fenêtres. La position de chaque fenêtre, dans la matrice de mesures (*baseline*), est choisie aléatoirement.

Étape 2:
Dans cette étape, on calcule les scores des sources pour chaque fenêtre et cela pour chaque groupe. Ensuite, on calcule la moyenne des scores des sources pour les 40 fenêtres.

Étape 3:
A partir des scores moyennés dans l'étape précédente, on calcule la moyenne et la variance (μ_0 ,ν_0) de ces scores. Ensuite, on calcul les paramètres beta (α_0 et β_0), tel qu'il est présenté dans la section précédente.

Étape 4:

On répète les étapes 1 à 3, pour chaque groupe de fenêtres de même taille. Ainsi, chaque groupe de 40 fenêtres de même taille on a un couple de valeurs (α_0 , β_0).

Étape 5:

Dans cette étape, on calcule les variances V_α et V_β des variables α et β des groupes de fenêtres de même taille. On calcule aussi la variance jointe des deux variances comme suit :

$$V_{Jointe} = V_\alpha + V_\beta$$

Résultats et analyses

Dans cette section, on présence les résultats obtenues en appliquant cet algorithme, avec un jeu de fenêtres dont la largeur varie entre 3 à 51 échantillons/fenêtre. Les résultats obtenus et leurs analyses se résument dans les points suivants :

A. Comme on peut le voir sur la figure 3.10 la variance des paramètres beta dépend de la largeur des fenêtres de mesures. On remarque que plus la largeur des fenêtres des mesures augmente, plus les valeurs du paramètre α augmentent et ceux de β diminuent.

Ce résultat montre que le choix de la taille des fenêtres de mesures influence le résultat de la modélisation de l'hypothèse nulle. Le résultat sur cette même figure nous amène à poser les questions suivantes : (1) quelles valeurs des paramètres beta à considérer pour modéliser notre hypothèse nulle? et (2) quelle taille de fenêtre doit-on considérer pour déterminer ces paramètres?

Dans le point suivant, on répond à ces questions tout en s'intéressant à la question de stationnarité de l'activité neuronale au repos.

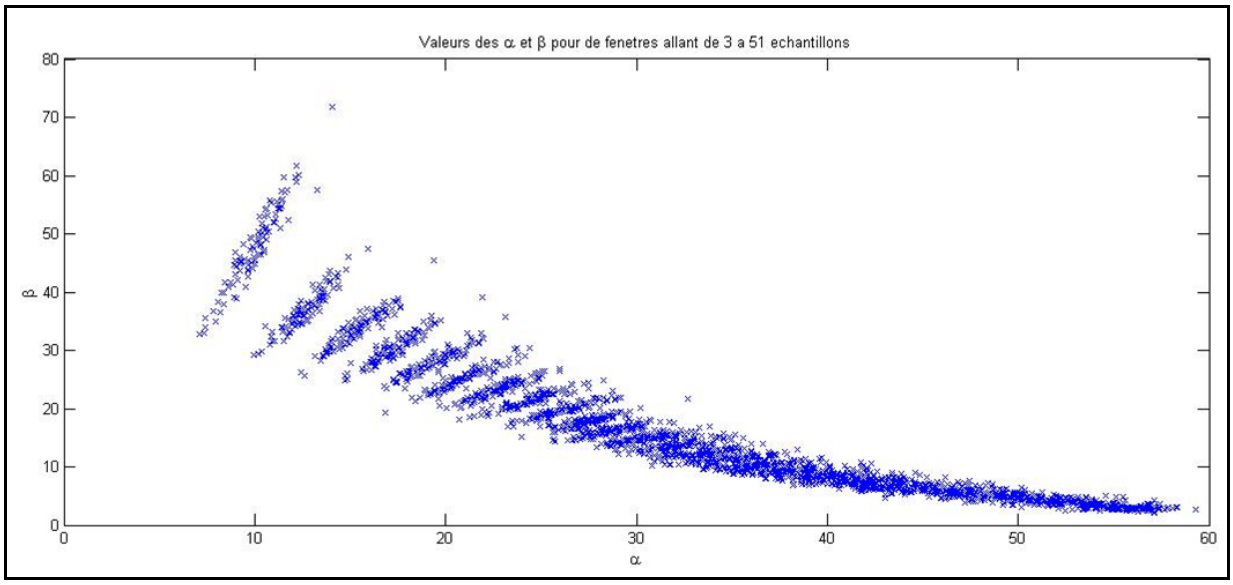


Figure 3.10 Valeurs des paramètres α et β des fenêtres dont la largeur (W) varie entre 3 à 51 échantillons

B. Concernant la question de la stationnarité soulevée précédemment, on étude la variance des paramètres beta, et leur variance jointe. Dans la figure 3.11, on présente un exemple de l'évolution logarithmique des variances des paramètres alpha, beta et leur variance jointe en fonction de la taille des fenêtres d'observation.

   Comme on peut le voir sur cette figure, la variance jointe logarithmique diminue progressivement quand la largeur des fenêtres augmente. Ensuite, elle augmente légèrement lorsque la taille des fenêtres dépasse 20 échantillons/fenêtres.

   On remarque aussi que ce comportement est dû au degré de contribution de chacune des ces variances (α et β) dans la variance jointe. Ainsi, pour des petites fenêtres la contribution de la variance β est plus importante que celle de α et inversement pour des fenêtres assez large.

Pour déterminer, la largeur qui correspond le mieux à la stationnarité de cette activité, il faut choisir celle dont les contributions des deux variances (α et β) est proportionnellement les plus proches (équivalentes).

Partant de cette hypothèse et considérant le résultat présenté dans la figure 3.11, on remarque que les courbes des variances logarithmiques tendent à se rejoindre autour des tailles de fenêtres comprissent entre w=15 et w=20.

Ainsi, pour notre processus de sélection des sources par la FDR et le choix de la définition de l'hypothèse nulle, on a adopté des fenêtres de largeur comprise entre 15 et 17 échantillons/fenêtre. Pour conclure, on peut dire que la stationnarité de l'activité neuronale au repos peut être représentée dans des fenêtres (de la *baseline*) de cette largeur.

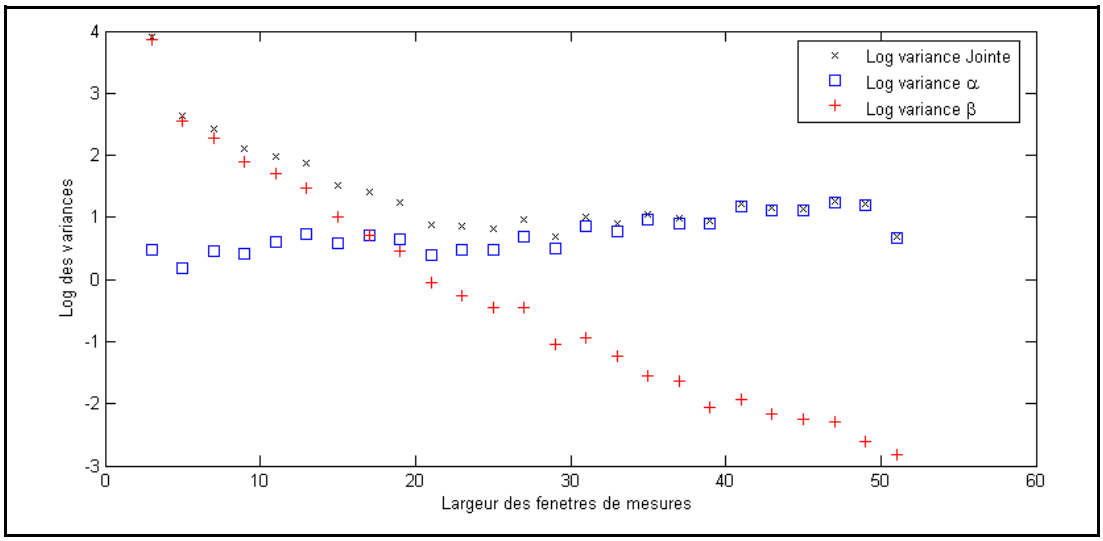


Figure 3.11 Variance jointe des paramètres α et β en fonctions de la taille de la fenêtre des mesures

CHAPITRE 4

RÉSOLUTION DU PROBLÈME INVERSE : PARCELLISATION ET FILTRAGE SPATIAL

Dans le chapitre précédent, on a présenté la modélisation de la surface corticale en sources électromagnétiques. Pour la régularisation du problème inverse, on a d'abord calculé une estimation d'activation de chaque source par rapport aux données sous forme de scores. Ensuite, on a calculé le seuil de sélection des scores pour déterminer un sous-ensemble de sources considérées potentiellement actives. Ainsi, on a une réduction du nombre des sources à considérer dans l'étape de résolution.

En faite, cette régularisation du problème inverse est la première partie de notre approche de résolution. La deuxième partie consiste dans la parcellisation des sources et la résolution par filtrage spatial en utilisant la technique LCMV (Limpiti, Veen et al. 2006).

Dans ce chapitre, on présente notre approche qui se compose essentiellement de deux grandes parties; (1) la modélisation et la régularisation du problème inverse qui permet d'améliorer son conditionnement pour faciliter sa résolution et (2) la modalisation en parcelles (parcellisation) et localisation et estimation des sources actives.

4.1 Approche et algorithme de localisation

4.1.1 Approche

Les techniques de résolution du problème inverse, en général, sont de deux types. Le premier type d'approche considère la localisation de l'activité cérébrale au niveau des sources ponctuelles. Le deuxième type considère la localisation au niveau de régions cérébrales, c'est-a-dire un groupe de sources contigües en référence à la notion d'aires cérébrales.

Certes, la modélisation de l'activité cérébrale en sources (dipôles) électromagnétiques permet d'avoir une bonne résolution spatiale selon la première approche. Par contre, il y a aussi un

inconvénient qui vient avec cette approche, c'est le mauvais conditionnent du problème inverse qui le rend difficile à résoudre. Ainsi, le gain de résolution spatiale espéré est compromis par la qualité de la résolution (localisation) obtenue.

Pour remédier à cette difficulté, dans ce travail on a choisie de passer par la réduction du degré de liberté pour améliorer le conditionnement du problème inverse. C'est ce que propose le deuxième type d'approche. En effet, la solution de cette approche souffre d'une faible résolution spatiale et cela affecte aussi la qualité de la résolution.

On considère que les deux types d'approches sont valides et elles sont complémentaires. Ainsi, notre approche de résolution est une combinaison de ces deux types d'approches, tirant profit des avantages des deux approches ce qui permet d'améliorer les résultats de résolution du problème inverse.

4.1.2 Algorithme

Notre approche pour établir l'algorithme présenté ci dessous consiste d'abord (1) à formuler le problème inverse en utilisant une modélisation en sources (dipôles) pour assurer une bonne résolution spatiale; (2) à réduire l'espace des sources à travers une modélisation en parcelles, ce qui améliore le conditionnement du problème inverse (régularisé) et finalement (3) à résoudre le problème sans compromettre (perdre) la résolution spatiale initial.

Comme le montre la figure 4.1 cette approche se compose essentiellement de 5 étapes regroupées en deux parties qui considèrent au départ les informations concernant la surface corticale modélisée en sources prédéterminées, les données observées et la *baseline*.

La première partie de cet algorithme se compose de deux étapes : (1) l'estimation des scores des sources à partir des données et (2) la sélection des sources ce qui permet la réduction de l'espace des sources et la régularisation du problème inverse.

La deuxième partie qui fait l'objet de ce chapitre se compose de 3 étapes : (1) la parcellisation qui permet de former des parcelles à partir des sources sélectionnées, (2) l'estimation des contributions des parcelles utilisant une technique de filtrage spatial et (3) l'estimation des contributions des sources et la localisation des sources actives en utilisant les données de la parcellisation et du filtrage spatial.

Il est important de rappeler que les données exploitées dans cette approche (régularisation/résolution), se limitent essentiellement aux données MEG et l'hypothèse nulle, aucune information externe n'a était utilisée.

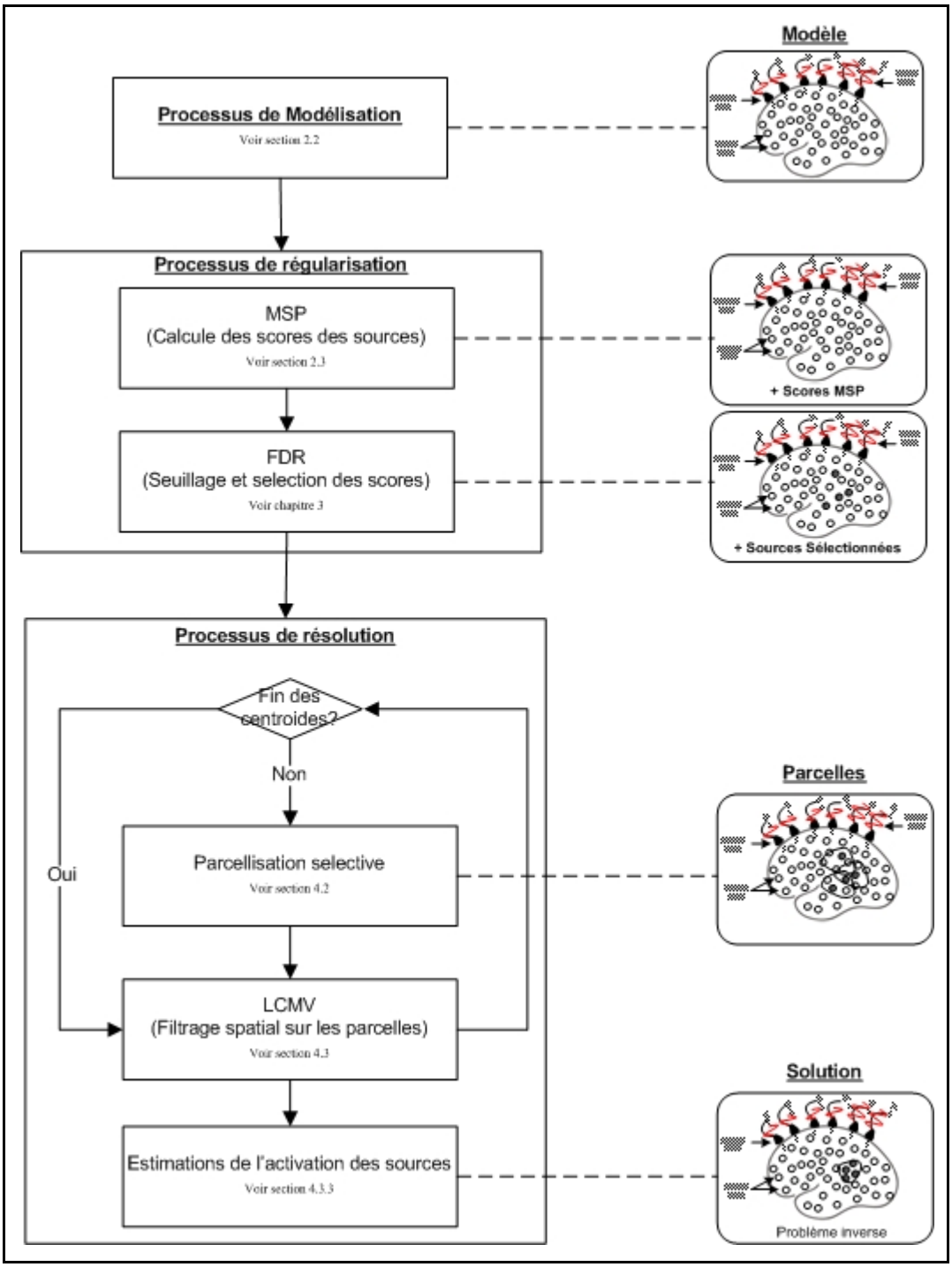


Figure 4.1 Les différentes étapes de notre approche de résolution du problème inverse

4.2 Technique de Parcellisation

Après sélection et réduction de l'espace des sources, on a un jeu de sources sélectionnées qui ne sont pas forcement localisées en régions contigües. En conséquence, les sources sélectionnées à date sont dispersées en points isolées dans la surface corticale.

Cependant, on rappelle que l'activité cérébrale ne se manifeste pas dans des points ponctuels mais plutôt dans des régions neuro-fonctionnelles. Autrement dit, l'activité cérébrale des sources est homothétique et elle respecte la caractéristique de proximité régionale : si une source est considérée active, il y a de forte chance que celles avoisinantes, appartenant à la même région anatomique que cette source, le soient aussi.

Partant de cette hypothèse, on définie un espace de proximité autour d'une source sélectionnée qu'on appelle *Germe* ou *Noyau*. Cet espace correspond aux sources adjacentes à un degré de voisinage 'D' par rapport au germe. Plusieurs critères peuvent être considérés pour définir cette distance. Dans ce travail, cette distance géodésique des sources voisines est comprise entre [3-5] voisines du maillage en face triangulaire.

Cette technique de parcellisation présentée ci-dessous permet de former, à partir des sources sélectionnées, des parcelles ou régions considérées comme potentiellement actives.

4.2.1 Stratégies de modélisation en parcelle

Comme présenté précédemment, avec la méthode distribuée, l'activité cérébrale est modélisée par des dipôles bipolaires (sources) caractérisées par leurs positions, orientations et l'intensité. Après parcellisation, on assigne un autre label aux sources permettant d'étiqueter les parcelles auxquels elles appartiennent. Ces groupements de sources en parcelles sont caractérisées par (1) le fait qu'elles sont localisées spatialement dans le même voisinage et (2) leurs synchronie fonctionnelle pendant une fenêtre temporelle (scores MSP).

On rappelle qu'une des difficultés de la localisation de sources est le mauvais conditionnement du problème inverse. Pour y remédier, on considère deux niveaux de modélisations : (1) la modélisation en sources et (2) celles en parcelles. La première modélisation assure une bonne résolution spatiale et la deuxième assure le conditionnement du problème inverse. La modélisation en parcelles est basée sur la modélisation en sources sélectionnées; ce qui permet de préserver les acquis de la première (bonne résolution spatiale) tout en améliorant la localisation des sources.

Il y a plusieurs stratégies possibles pour former des parcelles à partir des sources sélectionnées. On présente 2 stratégies dont les parcelles se chevauchent:

Stratégie 1: Parcelle sans sélection

Cette stratégie, ne nécessite pas la sélection des sources. On peut aussi considérer que le seuil de sélection est égal à 0. Autrement dit, toutes les sources sont utilisées pour former des parcelles. Ainsi, pour chaque source, on détermine ses sources voisines à un degré donné qui correspond à une parcelle. A la fin de cette stratégie on a autant de parcelles que de sources. Cette stratégie est un cas généralisé de la prochaine stratégie où on utilise seulement les sources sélectionnées.

L'inconvénient avec cette stratégie est qu'elle n'est pas efficace à cause du dégrée de liberté qui reste important. Plus précisément, le nombre des parcelles générées par cette stratégie est égale exactement au nombre des sources du modèle. Par conséquence, on n'a pas de réduction de l'espace des sources. C'est la stratégie utilisée dans le travail (Limpiti, Veen et al. 2006).

Stratégie 2 : Parcelle avec sélection

Dans cette stratégie, on utilise le résultat de sélection des sources présentées dans une liste S_s qu'on exploite pour former des parcelles.

avec $S_s = \{s_k^{Sel}$ est une source sélectionnée $|\ k = 1, 2, \ldots, K$ et K est le nombre des sources sélectionnées$\}$

Chaque source s_k^{Sel} de S_s est utilisée pour former une parcelle C_k. A la sortie de ce processus de parcellisation, on a un ensemble de parcelles $\mathcal{C}$ tel que :

$$\mathcal{C} = \{C_k \text{ avec } k = 1, \ldots, K \text{ est } K \text{ est le nombre total des clusters}\}$$

avec $C_k = \{s_{k,i} \text{ est l'ensemble des sources formant le cluster } k\}$

et s_k^{Sel} est le germe de la parcelle k.

Étant donnée la liste des sources S_s, on calcule pour chaque source s_k^{Sel} de S_s ses sources voisines a une distance géodésique *'D'*. Ces sources voisines a s_k^{Sel} forment la parcelle C_k. Ce processus est appliqué de la même façon pour l'ensemble des sources de S_s sans répétition. L'ordre du choix des sources de S_s n'est pas important.

Il faut noter que les sources qui ont participé pour former une parcelle C_k peuvent être aussi choisies pour former une autre parcelle $C_{k'}$. Autrement dit, une source peut appartenir à plusieurs parcelles qui se chevauchent.

On remarque aussi que des sources non-sélectionnées sont choisies par cette parcellisation pour participer dans la résolution du problème inverse. Les sources qui ne sont pas choisies pour former aucune parcelle forment la parcelle C_0. Les sources de la 'parcelle nulle' C_0 ne sont plus considérées dans le processus de résolution; elles sont considérées comme définitivement non-actives.

Algorithme

A l'entrée de cet algorithme l'ensemble des sources du modèle par défaut appartiennent à la parcelle C_0. Les étapes de cet algorithme de parcellisation sont les suivantes:

1. On choisit un élément s_k^{Sel} de la liste des germes S_s; sources sélectionnées.
2. On détermine les voisins a s_k^{Sel} dont la distance géodésique est inferieur ou égale a *D*.

3. L'ensemble des sources voisines, déterminées dans l'étape 2, forment la parcelle C_k incluant le germe s_k^{Sel}.
4. On répète les étapes 1, 2 et 3 jusqu'à épuisement de la liste des germes S_s.

A la sortie de cet algorithme on a une liste de parcelles $\mathcal{C}$ de K parcelles C_k.

La figure 4.2 ci-dessous est une présentation schématique d'une parcellisation utilisant cet algorithme. Comme le montre cette figure, la majorité des sources de la surface corticale ne font partie d'aucune parcelle autre que la parcelle par défaut, la parcelle nulle C_0. On a aussi 3 autres configurations possibles : on a (1) des sources groupées dans des parcelles isolées; (2) des sources groupées dans des parcelles qui se chevauchent en faible densité et (3) finalement des sources groupées dans des parcelles qui se chevauchent en grande densité.

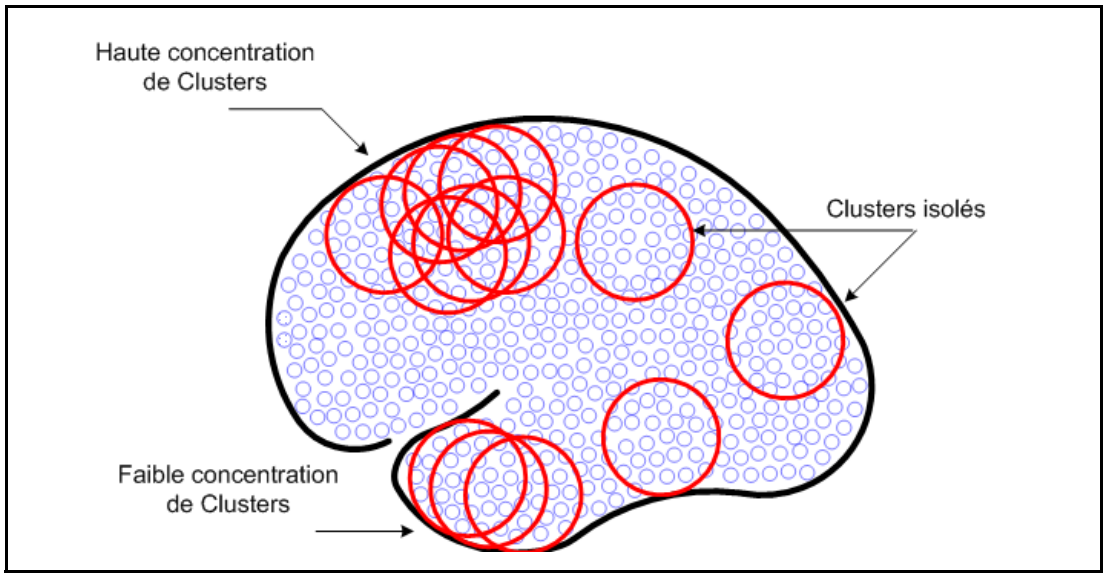


Figure 4.2 : Formation des clusters chevauchant autour des germes (ou centroides)

Étant donné cette différence des répartitions des sources dans des parcelles, il est important d'évaluer le degré d'appartenance des sources aux parcelles. Pour calculer ce degré d'appartenance, on utilise le résultat de la parcellisation pour établir une matrice dont les

colonnes correspondent à l'ensemble des sources du modèle et les lignes correspondent aux parcelles C_k. Les cellules de cette matrice contient le digit 1 si la source fait partie de la parcelle et 0 sinon.

A partir de cette matrice, pour chaque source on calcule le nombre de parcelles auxquelles elle appartient. Autrement dit, pour chaque colonne on calcule la somme des '1'. Ce résultat des degrés d'apparence des sources aux parcelles est une donnée que peuvent exploiter les techniques de résolutions existantes telles que 'Maximum d'Entropie' (Lapalmea, Lina et al. 2006) pour la localisation et l'estimation de l'activité des sources.

4.3 Résolution du problème inverse régularisé

L'étape précédente a permis de former des parcelles à partir des sources sélectionnées en utilisant la stratégie de parcellisation. A ce stade de notre approche, le nombre des sources retenues pour participer à la résolution du problème inverse est assez réduit, ainsi, on a une réduction de l'espace des sources. Les sources non-sélectionnées jusqu'à ce stade sont définitivement considérées comme non-actives et elles sont écartées de la résolution du problème inverse.

Dans cette étape, on utilise une technique de filtrage spatial LCMV pour la résolution du problème inverse régularisé. Avant de passer à la présentation de notre stratégie de résolution, on commence par une brève présentation de cette technique.

4.3.1 Présentation de la LCMV

La technique LCMV (Linearly Constrained Minimum Variance) (Barry D. Van Veen 1997), aussi appelée LCMV beamforming, est une technique de filtrage spatial. Cette technique a été développée initialement pour des applications radar, détection et classification des sonars. Plus tard, elle fut exploitée dans des applications de communication acoustiques et d'imageries biomédicales (B. Van Veen Sep. 1997.).

L'origine du beamforming (Veen and Buckley 1988) ou 'Forming beams' (formation de faisceaux) signifie radiation d'énergie; toutefois, elle s'applique aussi bien pour l'émission d'énergie qu'à sa réception. Les systèmes de réception radio utilisés pour recevoir un signal à propagation spatial d'intérêt donné risque de recevoir un autre signal d'interférence. Plus la fréquence du signal d'interférence est proche de la bande fréquentielle (temporelle) du signal désiré, plus il devient difficile d'exploiter le signal d'intérêt. Pour récupérer le signal d'intérêt, dans ce cas, on ne peut pas utiliser un filtre fréquentiel, on a besoin d'un filtre spatial. Généralement, le signal d'interférence est à une position différente du signal d'intérêt, cette séparation spatiale peut être exploitée pour filtrer le signal d'intérêt. Cela nécessite la conception du filtre spatial (Limpiti, Veen et al. 2006).

En fait, pour la conception d'un filtre temporel/fréquentiel on considère le traitement des données temporelles (reçues dans des fenêtres temporelles), par analogie, pour la conception d'un filtre spatial on considère le traitement des données spatiales (reçus dans des fenêtres spatiales).

4.3.2 Filtre spatial et localisation des sources

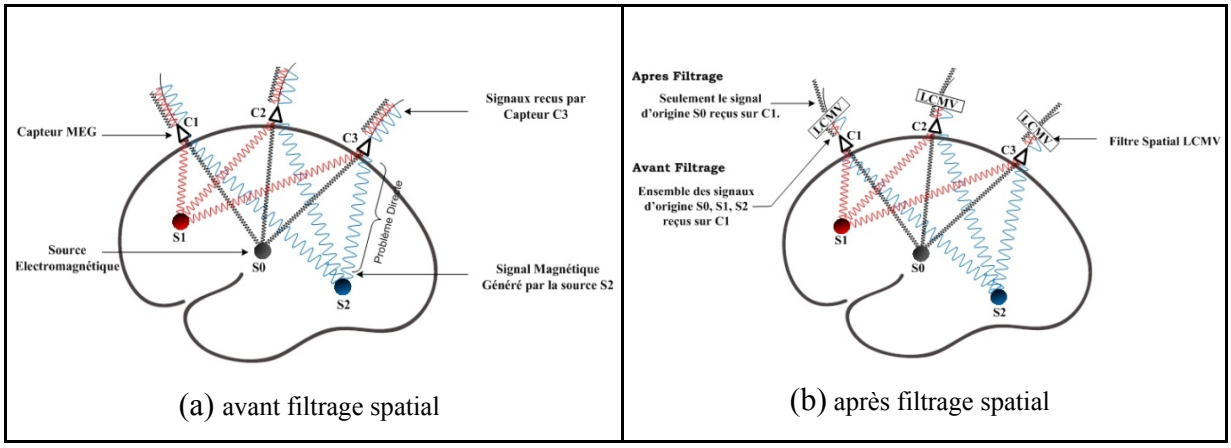
A. Localisation des sources par filtrage spatial

La localisation des sources par filtrage spatial repose sur l'utilisation des filtres conçus de façon à ce qu'ils laissent passer l'activité cérébrale provenant d'une région cérébrale spécifique et qu'ils atténuent ou ne laissent pas passer les signaux générés par des sources ailleurs.

Le principe du filtre spatial LCMV est de minimiser la variance de la puissance tout en assurant une réponse linéaire (comme contrainte) à sa sortie. Ainsi, un signal provenant d'une région d'intérêt passe sans atténuation à travers ce filtre. Par contre, tout autre signal est atténué.

Dans notre cas, comme l'illustre la figure 4.3 (a) en dessous, les dipôles sont des sources de radiation et les capteurs forment le système de réception. Les signaux provenant des différentes sources à différentes localisation s'interférent au niveau des capteurs. Ainsi, le signal reçu sur chaque capteur est le résultat d'une contribution de des signaux générés par des sources à différentes localisation.

Pour retrouver un signal provenant d'une source à une position donnée, on utilise un filtre spatial qui est conçu pour filtrer tous les signaux provenant des autres sources sauf celui de la source d'intérêt. En appliquant ce filtre au niveau de chaque capteur sur les données tel que montrée dans la figure 4.3(a), à la sortie on obtient seulement les contributions de cette source au niveau des capteurs (figure 4.3(b)).



(a) avant filtrage spatial

(b) après filtrage spatial

Figure 4.3 Filtrage du signal provenant d'une source à une localisation spatial donnée appliquant la LCMV-Beamformer

Dans la section qui suit on présente les étapes de la conception d'un filtre spatial pour une parcelle donnée.

B. Calcul du filtre spatial

Considérons une région corticale $\boldsymbol{k}$ dont l'activité de source électromagnétique est $\boldsymbol{J_{k,j}(t)}$ pendant une période $\boldsymbol{j}$ et le signal de mesure résultant de cette activité est $\boldsymbol{M_j(t)}$. En appliquant le filtre spatial $\boldsymbol{W_k^T}$ à ce signal de mesure, on extrait la contribution du signal généré par la région $\boldsymbol{k}$.

L'expression ci-dessous exprime le filtrage spatial $\boldsymbol{W_k^T}$ pour extraire la contribution aux mesures qui correspond à la région $\boldsymbol{k}$.

$$\boldsymbol{J_{k,j}(t) = W_k^T M_j(t)}$$

$\boldsymbol{W_k^T}$: Filtre spatial

$\boldsymbol{M_j(t)}$: Données mesurées

Le calcul du filtre spatial LCMV pour une région corticale donnée se résume dans la résolution problème d'optimisation présentée ci-dessous. Les étapes du calcul de la solution de ce problème d'optimisation pour déterminer le filtre $\boldsymbol{W_k^T}$sont les suivantes :

1. Formulation du problème d'optimisation suivant :

$$\min_{\boldsymbol{w_k}} \boldsymbol{W_k^T R_m w_x} \; sujet \text{ à } \boldsymbol{W_k^T U_k V_k = 1}$$

avec

$\boldsymbol{R_x}$ est la matrice de corrélation de mesures,

$\boldsymbol{U_k}$ est la matrice des vecteurs colonnes de la décomposition en valeurs singulières de la sous-matrice de gain $\boldsymbol{G_k}$, et

$\boldsymbol{V_k}$ est un coefficient d'activation de la parcelle $\boldsymbol{k}$.

2. Définition du filtre :

$$\boldsymbol{W_k = [V_k^T U_k^T R_x^{-1} U_k V]^{-1} R_x^{-1} U_k V_k}$$

3. Calcul de la puissance de sortie a minimisé :

$$\boldsymbol{P_{out}(k) = \frac{1}{V_k^T U_k^T R_x^{-1} U_k V_k}}$$

avec $\boldsymbol{V_k}$ est la matrice des vecteurs colonnes de la décomposition de la sous-matrice de gain $\boldsymbol{G_k}$.

4. Calcul du vecteur propre qui assure la minimisation de la puissance de sortie $\boldsymbol{P_{out}}$. En fait, ce vecteur propre correspond à la valeur propre la plus petite de cette expression.

$$\boldsymbol{U_k^T R_x^{-1} U_k}$$

Avec

$$\boldsymbol{G_k = U_k B_k^T}$$

Les détails à propos de ces 4 étapes de formation du filtre spatial concernant une parcelle, sont présentées dans (B. Van Veen Sep. 1997.) où on exploite aussi la LCMV pour la localisation des sources.

4.3.3 Résolution basée sur une modélisation en parcelles

Dans ce travail, on utilise la technique LCMV comme technique de résolution du problème inverse dont on se sert pour améliorer le résultat de la pré-localisation et de la parcellisation des sources. En fait, on l'utilise précisément pour calculer les amplitudes des parcelles et à partir de ce résultat on estime les amplitudes des sources.

Problème direct

Contrairement aux techniques de résolution classiques où on estime l'activité au niveau des sources, dans ce travail, la résolution est appliquée au niveau des parcelles. Dans cette

section on présente le principe de la résolution du problème inverse à l'échelle de régions corticales.

Comme le montre la figure 4.4, la surface corticale se présente sous forme de régions d'activations où parcelles formées d'un groupement de sources. Ainsi, selon ce modèle, les sources sont plutôt des parcelles qui sont responsable des données. Autrement dit, dans le modèle classique les techniques de résolution s'intéresse à estimer l'état d'une source étant donnée les mesures observées. Dans le modèle en parcelles présenté ci-dessous, on s'intéresse plutôt à estimer et évaluer l'état des régions étant donne une distribution d'une parcellisation donnée. Par exemple, les parcelles présentées dans la figure 4.4 se chevauchent et elles ne sont pas toutes de même taille.

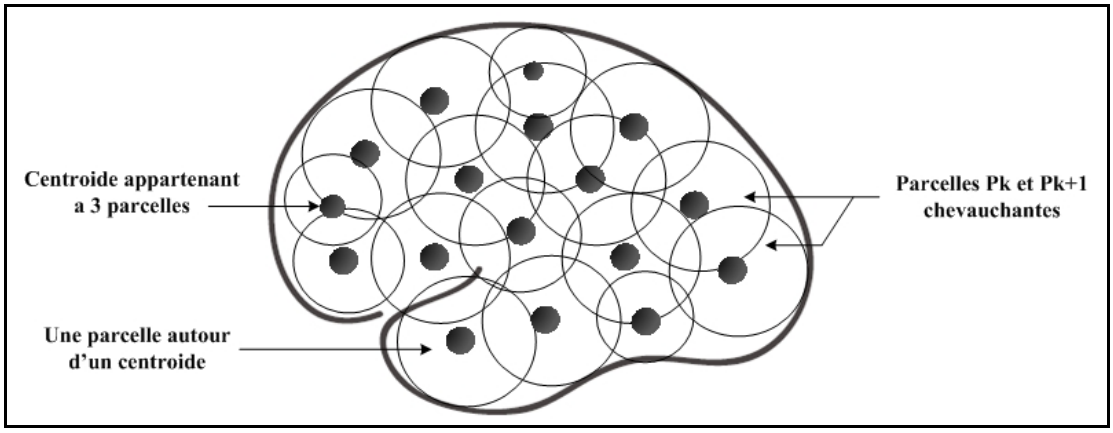


Figure 4.4 Formation de parcelles cortical de différentes taille et chevauchantes autour des centroides s_k

Considérant cette modélisation en parcelles, on active une parcelle donnée pour laquelle on obtient une mesure qui lui correspond (voir figure 4.5). Etant donné que les autres parcelles ne sont pas actives, la matrice de gain $\boldsymbol{G}$ du modèle en sources peut se résumer dans la sous-matrice $\boldsymbol{G_k}$ qui correspond à l'ensemble des vecteurs colonnes des sources de la parcelle $\boldsymbol{k}$. La même chose pour la matrice des intensités $\boldsymbol{J}$ elle aussi peut être résumée a $\boldsymbol{J_k}$**.** Ainsi, le problème direct dans ce cas est le suivant : $\boldsymbol{M} = \boldsymbol{G_k} . \boldsymbol{J_k}$.

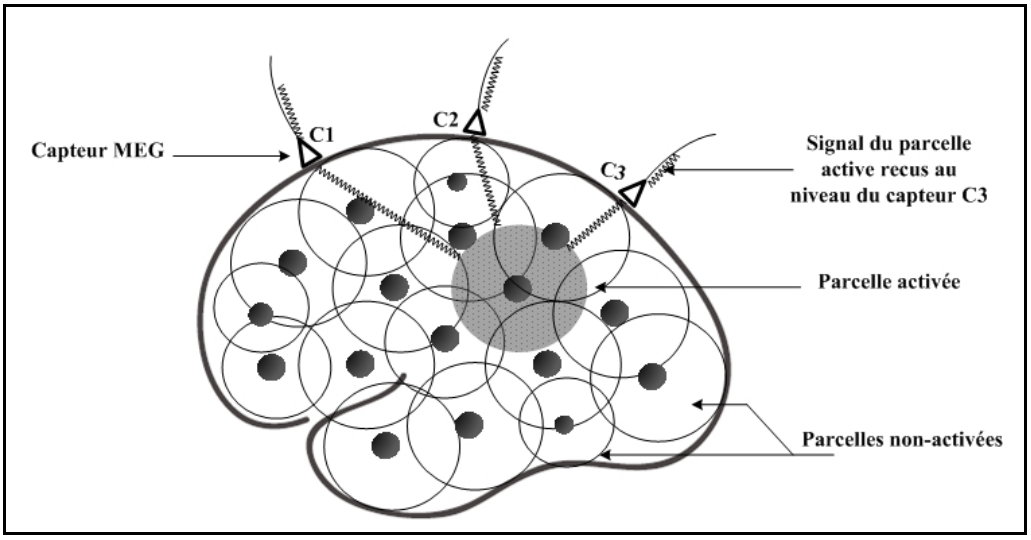


Figure 4.5 Signal reçu par les capteurs MEG provenant de la parcelle activée

Problème inverse

La résolution du problème inverse, dans le cas d'une modélisation en parcelles correspond à : étant donné une distribution de parcelles, quelle est la configuration des parcelles qui explique le mieux les données observées ?

Ainsi, on calcul le filtre spatial utilisant les étapes présentées dans la section précédente pour chaque parcelle. En appliquant ces filtres sur les données observées on obtient plusieurs distribution des intensités qui expliquent ces données selon le filtre utilisées, tel que : $\hat{\boldsymbol{J}}_{\boldsymbol{k}} = \boldsymbol{W}_{\boldsymbol{k}}^{\boldsymbol{T}} \boldsymbol{M}$.

On remarque qu'on peut bien avoir une différence entre $\hat{\boldsymbol{J}}$ et $\hat{\boldsymbol{J}}_{\boldsymbol{k}}$ cela dû aux interférences possible parvenant des parcelles voisines.

Pour diminuer l'effet de cette inférence, dans la conception du filtre spatial d'une parcelle plus particulièrement, dans la matrice de gain qui représente cette parcelle, on ne considère pas l'ensemble des vecteurs colonnes des sources de cette matrice. Les colonnes qu'on

considère sont celles qui contribuent le plus dans la représentation de la parcelle. En effet, on utilise la décomposition en valeurs singulières de la matrice de gain $\boldsymbol{G_k}$. Ensuite, on calcule les valeurs propres telles que:

$$\boldsymbol{E_n} = \frac{\sum_{i=1}^{n} \boldsymbol{\sigma_i}^2}{\sum_{i=1}^{N} \boldsymbol{\sigma_i}^2}$$

Avec N est le nombre total des sources de la parcelle $\boldsymbol{k}$, $\boldsymbol{\sigma_i}$ est la valeur propre qui correspond à la source $\boldsymbol{i}$ de la parcelle et $\boldsymbol{E_n}$ est l'erreur tolérée qui dépend de l'indice de la source $\boldsymbol{n}$. La valeur de $\boldsymbol{E_n}$est comprise entre 0 et 1. Il faut noter que les sources sont ordonnées selon l'ordre décroissant des valeurs propres. Ainsi, en fixant la valeur de $\boldsymbol{E_n}$ on fixe le sous-ensemble des sources qui représente la parcelle $\boldsymbol{k}$ avec une tolérance $\boldsymbol{E_n}$. Partant de là, la sous-matrice de la parcelle $\boldsymbol{k}$ dans ce cas devient : $\boldsymbol{G_k = G_{k'} + \mathcal{E}_n}$.

Cette approximation de la matrice de gain permet de définir un filtre spatial peu sensible à l'interférence des parcelles voisines en acceptant une erreur de représentation de la parcelle d'intérêt.

A la sortie du filtrage spatial des parcelles, on a une amplitude pour chaque parcelle. Les amplitudes des parcelles actives sont plus importantes par rapports aux autres.

Parcellisation sélective

Dans le cas de notre parcellisation sélective, les parcelles sont en faite de même taille et elles sont limitées aux régions autours des 'sources germes'. Ainsi, leurs nombres est plus réduit et elles sont concentrées dans des régions où l'activité est pré-localisée.

Cela permet d'améliorer (1) le conditionnement du problème inverse, (2) de réduire le degré d'interférence possible et (3) d'accélérer le temps du calcul. Il faut noter que les deux premiers points contribuent à une meilleure localisation des sources dans régions actives.

Estimation des amplitudes des sources

Jusqu'à présent, le filtrage spatial des parcelles a permet de déterminer les niveaux d'activation des parcelles. Dans cette section, on explique comment on estime les amplitudes d'activation des sources à partir des résultats du filtrage et de la parcellisation.

D'abord, on assigne à chaque source les amplitudes des parcelles correspondantes. Ainsi pour chaque source lui correspond un jeu d'amplitudes de ses parcelles. Par défaut, on attribue aux sources de la parcelle nulle C_0 une valeur nulle.

Ensuite, on calcule le degré d'appartenance d'une source aux parcelles. Ainsi, pour chaque source on attribue un facteur d''*occurrence*' qui correspond aux nombres de parcelles auxquelles elle appartient. Par défaut, ce nombre est initialisé à 1, puisque toutes les sources appartiennent à la parcelle nulle.

Il faut remarquer que le calcul de la moyenne des amplitudes pour chaque source n'est pas une référence fiable pour différentier entre le degré d'activation des sources. Il important de considérer l'information présentée dans le facteur '*occurrence'*.

CHAPITRE 5

RÉSULTATS ET ANALYSES DE SIMULATION

Dans ce chapitre, on présente les résultats de localisation et résolution du problème inverse en appliquant notre technique de parcellisation sélective. D'abord, on présente les données de simulation qui correspondent à l'activation d'une région prédéterminée, en utilisant des données de '*baseline*' fournies par L. Garnero (CNRS, Fance) en 2003. Ensuite, on présente les résultats obtenus à chaque étape du processus de résolution. Finalement, on effectue une évaluation des performances de notre approche de résolution du problème inverse. Ces résultats de localisation sont sommairement évalués en utilisant une analyse ROC (Receiver Operating Characteristic).

5.1 Simulation

5.1.1 Présentation de la simulation

Pour l'évaluation de notre approche on utilise un modèle cortical dont on dispose dans le laboratoire 'LATIS'(L. Garnero, 2003). Les données de repos sur 5 minutes à une fréquence de 600 Hz filtrées entre 0-70 Hz (avec 4000 nœuds et 150 capteurs CTF Vancouver). Ce modèle se compose d'un ensemble d'arêtes (*faces*) et de nœuds (*vertices*) qui décrivent la surface corticale d'un sujet. On utilise aussi des mesures effectuées où le sujet est au repos (*baseline)*. On utilise aussi une matrice de gain MEG (modèle sphérique à une couche) qui décrit la relation entre les sources et les mesures.

On utilise ces données pour construire les simulations qui se composent des éléments suivants:

- Région d'activation: pour former cette région, on choisit une source de façon aléatoire, ensuite on calcule ses sources voisines autour d'une distance géodésique donnée. Sur la figure 5.1 on a un exemple d'une région choisie pour être active.

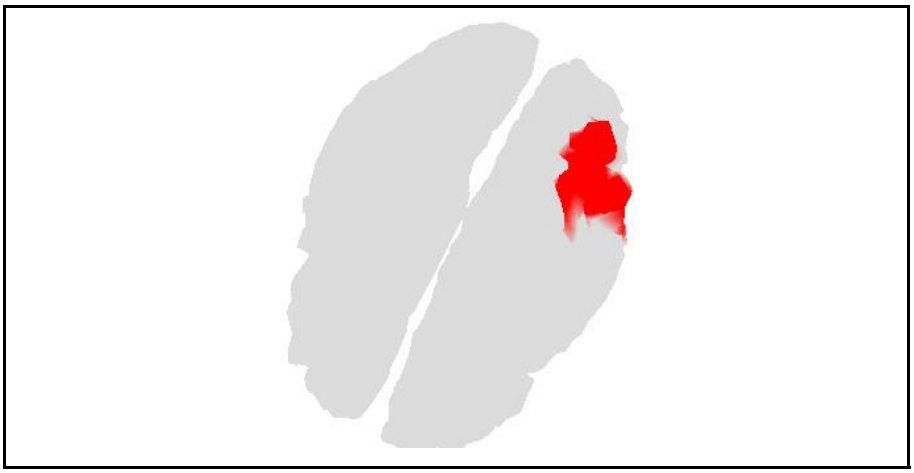


Figure 5.1 Région de simulation de l'activité électromagnétique

- Activation neuro-électrique : On applique une intensité dipolaire à chaque source de la parcelle d'activation pendant une durée donnée. Le signal de simulation est une gaussienne dont les paramètres sont: la moyenne $\mu = 0$ et la variance $\sigma = 0.2$. En fait, dans cette simulation on a considéré deux signaux identiques à celui sur la figure 5.2 avec une différence d'intensités.

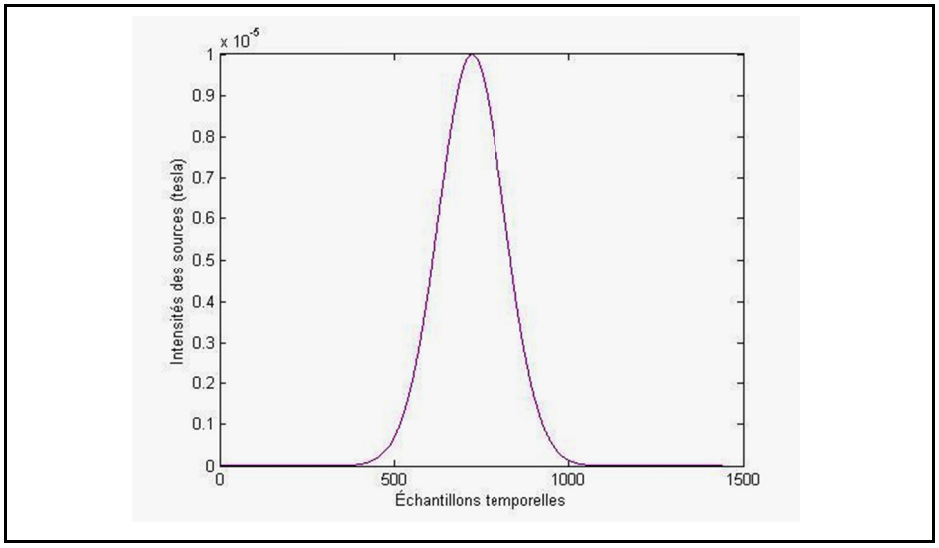


Figure 5.2 L'intensité de l'activité de simulation

- Matrice des amplitudes: à partir du signal d'activation on détermine les intensités des sources présentées dans la matrice $\boldsymbol{J}$ des intensités des sources actives.

- *Baseline*: La figure 5.3 présente des données de *baseline* utilisées dans cette simulation. Ce sont des données réelles fournies par L. Garnero (CNRS, Fance) en 2003.

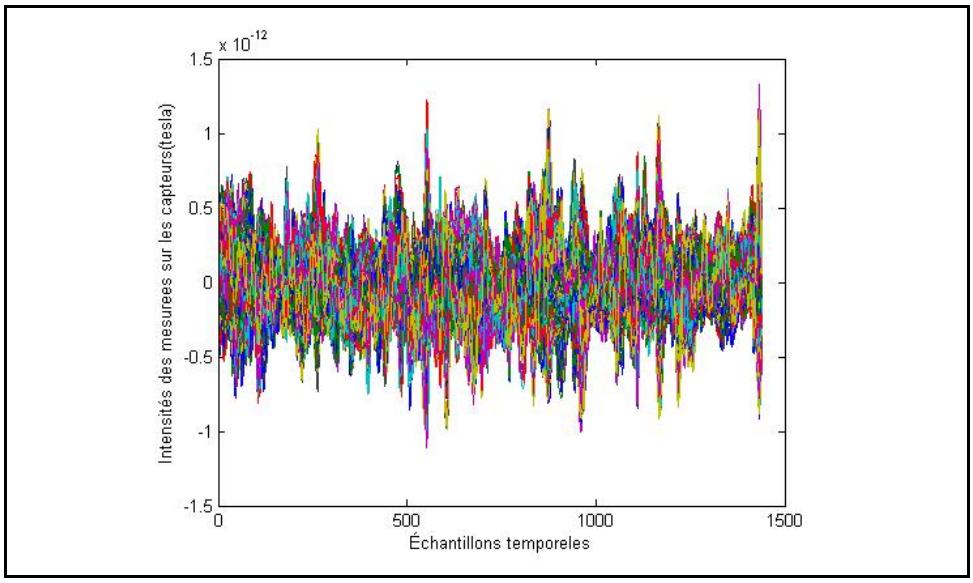


Figure 5.3 Mesures effectuées quand le sujet et repos

- Données simulées : en utilisant la matrice de gain $\boldsymbol{G}$ et les intensités de simulation de l'activation des sources $\boldsymbol{J}$, on calcule les mesures qui seront obtenus si on a une activité électromagnétique des sources de la parcelle d'activation avec une intensité telle que présenté dans la figure 5.2. Ensuite, on somme les données obtenues aux données de *baseline* présentées dans la figure 5.3 pour établir les données simulées. Les résultats de ces calculs permettent de définir deux signaux de données de simulation finales qu'on utilise dans cette simulation (voir figure5.3).

A la fin de ce stade de préparation des données de simulation, on a le signal de mesures *M* qui correspond aux mesures de l'activité de la parcelle de simulation en combinaison avec le signal de *Baseline* (bruit) comme le montre la figure 5.4.

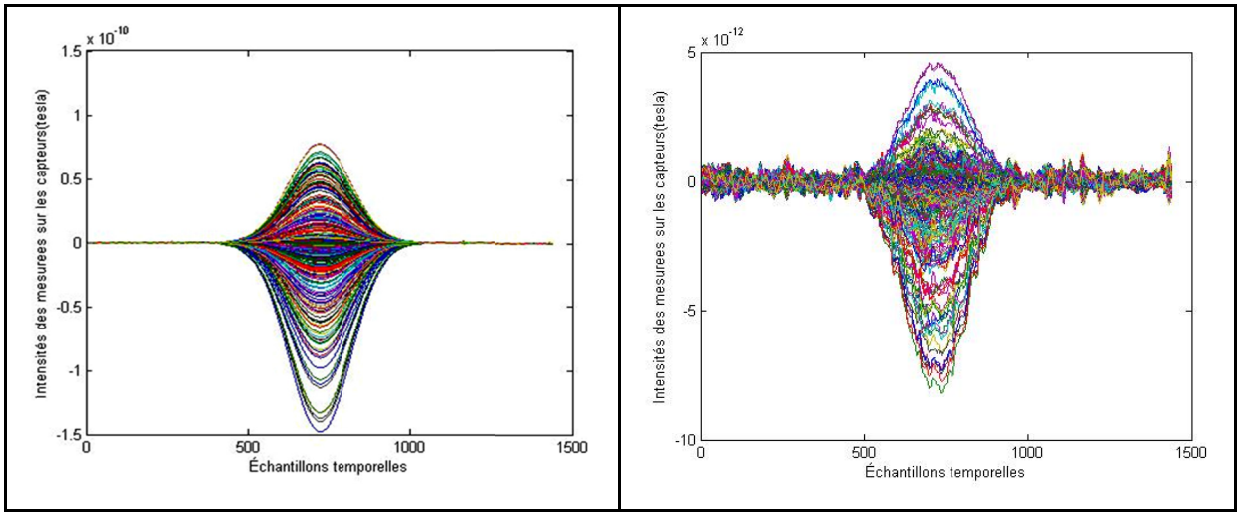


Figure 5.4 Mesures qui correspondent à l'activité de simulation

5.2 Résultats

Dans cette section, on présente les résultats obtenus à chaque étape de notre algorithme de résolution du problème inverse.

5.2.1 Scores MSP/ Pré-localisation

Cette étape première de notre algorithme de résolution du problème inverse consiste à calculer les scores des sources du modèle. Pour le calcul des scores, on utilise une fenêtre de 15 échantillons autour du pic des mesures de simulation ce qui est optimal par rapport à la stationnarité du signal de *baseline* (section 3.3). Ensuite, on attribue un score à chaque source du modèle selon la technique MSP avec un facteur de réduction des mesures, lors de la deuxième étape de la MSP, de 85%.

La figure 5.5, ci-dessous, représente la carte des scores MSP des sources. Comme on peut bien le voir sur cette figure, les sources dont les scores sont les plus élevées sont localisées autour de la région du patch actif (voir figure 5.1). On remarque aussi, que plus on s'éloigne de la région du patch actif, plus les scores sont faibles.

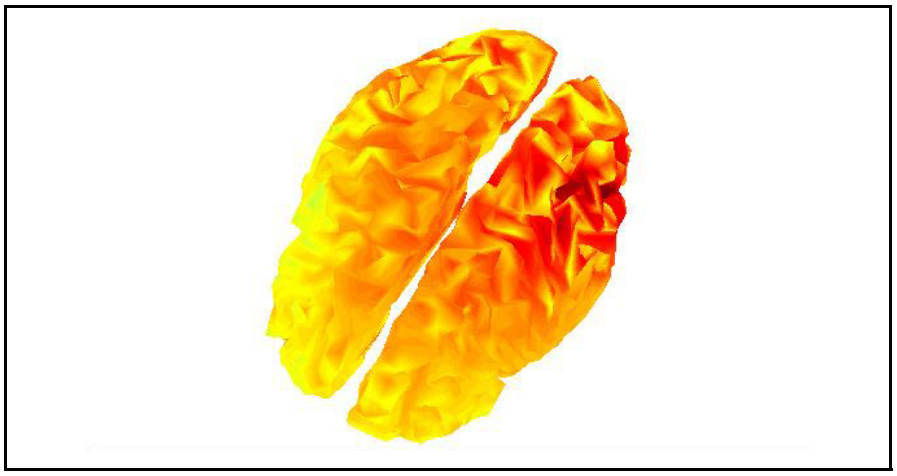


Figure 5.5 Carte des scores MSP de l'ensemble des sources

Bien que les scores permettent de pré-localiser la région où les sources sont actives, la localisation qu'ils donnent n'est pas assez précise. On doit alors appliquer une sélection en utilisant un seuil de sélection calculé par la technique FDR telle que présentée dans la prochaine section.

5.2.2 Sélection et seuillage avec la FDR

Dans cette étape, on procède à la sélection des sources selon leur score en utilisant la technique FDR.

Avant de présenter les résultats de sélection, on rappel qu'on a besoin de définir l'hypothèse nulle H0, de la technique FDR pour cette simulation. Pour calculer les paramètres α_0 et β_0 de la loi beta qui modélise l'hypothèse nulle, on calcul la moyenne des variances ϑ_0 et des

moyennes μ_0 des scores pour des fenêtres de mesures dont la taille varient entre 13 et 17 échantillons par fenêtre.

Une fois H0 déterminée, on fixe le taux de contrôle des faux positifs α. Pour cette simulation on a utilisé deux valeurs 1 et 5%.

Après application de la technique FDR, avec ces paramètres d'hypothèse nulle, comme le montre la figure 5.6, on remarque que la FDR permet de sélectionner des sources qui appartiennent à la région active. Dans cet exemple $\alpha = 1\%$.

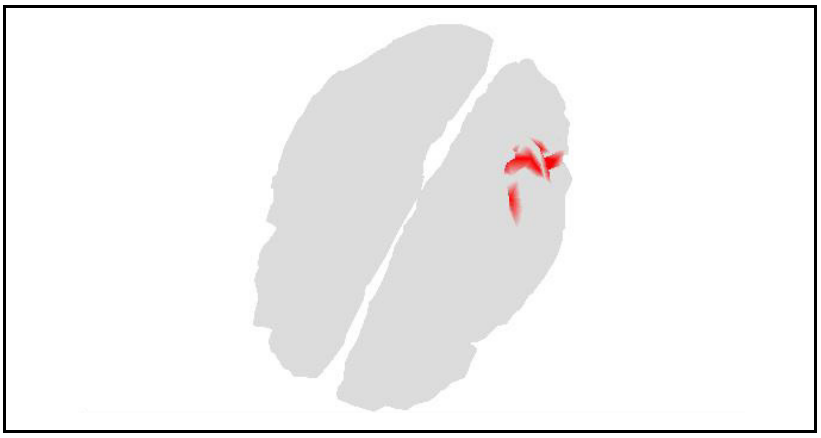


Figure 5.6 Sources sélectionnées selon leurs scores MSP
après application de la FDR

Il est important de remarquer que les sources de la région activée ne sont pas toutes forcement sélectionnées par la FDR. En fait, ce n'est pas l'objectif de cette sélection étant donne que dans la prochaine étape d'autres sources seront annexées à celles sélectionnées dans cette étape.

5.2.3 Parcellisation

Dans cette étape, on procède à la formation des parcelles à partir des sources sélectionnées. La formation des parcelles se fait autour de chaque source sélectionnée avec un degré de voisinage géodésique qui varie entre 2 à 4. Par conséquence, comme le montre la figure 5.7, à la fin du processus de parcellisation, le nombre des sources à considérer dans la résolution du problème inverse est relativement 20 fois plus grand comparé au nombre de sources sélectionnées par la FDR.

Notons aussi que l'intensité des sources sur cette figure correspond au nombre de fois qu'elles sont visitées pour former des parcelles. Rappelons-nous qu'on a autant de parcelles que de sources sélectionnées, mais une source sélectionnée peut faire partie de plusieurs parcelles. Ainsi, les régions qui apparaissent sur la figure 5.7 avec une grande intensité sont visitées plus de fois que celles qui apparaissent avec une faible intensité. A cette étape, on peut aussi dire que les sources qui ne sont pas choisies pour former des parcelles sont considérées non actives et ne feront partie de l'espace solutions.

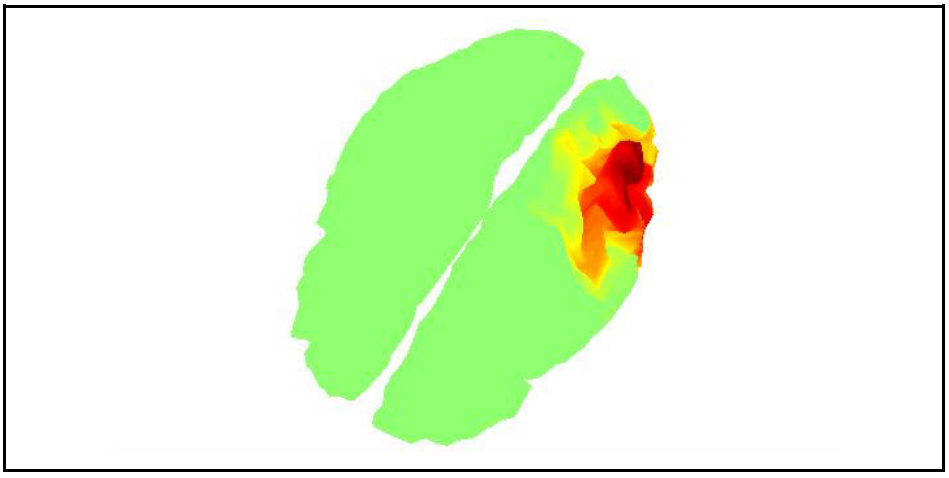


Figure 5.7 Sources sélectionnées à la fin du processus de parcellisation

Pour le même exemple, l'histogramme de la figure 5.8 montre qu'on a un grand nombre de sources qui sont visitées peu de fois, par contre, peu de sources sont visitées plusieurs fois. Effectivement, les sources qui sont activent se retrouvent dans des régions avec une grande densité de parcelles et inversement dans le cas des sources qui sont proches de la région d'activation sans forcement faire partie de cette activation.

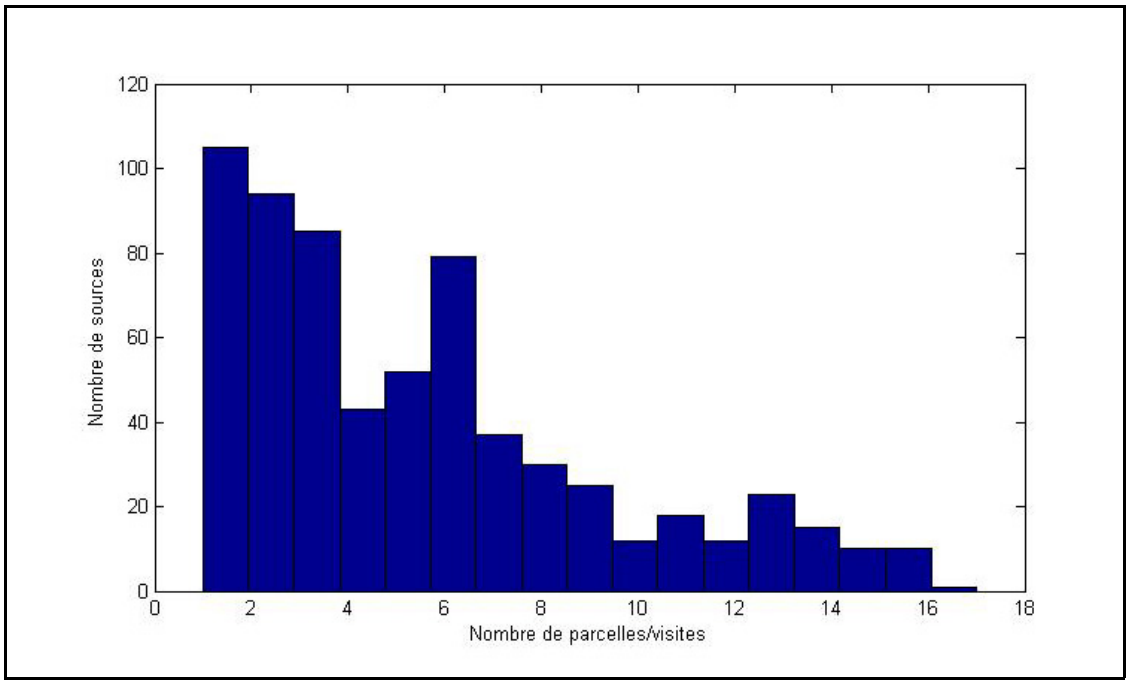


Figure 5.8 Histogramme d'occurrence de visites des sources pour former des parcelles

D'après ce résultat, on peut dire que plus le nombre de visite par source augmente, plus la probabilité que cette source appartient à l'espace solution. Par contre, à cette étape il reste difficile de mettre un seuil de sélection sur le nombre de visites pour décider, à partir de cette information seulement, qu'une source soit active ou non.

5.2.4 Calcul des amplitudes des parcelles par la LCMV

Dans cette étape on utilise la technique LCMV pour calculer les amplitudes des parcelles, ainsi, pour chaque parcelle et ses sources on a une amplitude qui leur correspond.

En sortie, on a une matrice des amplitudes dont les lignes correspondent aux parcelles et les colonnes correspondent aux sources sélectionnées. A partir de cette matrice on calcule la moyenne des amplitudes LCMV de chaque colonne (source). La figure 5.9 présente le résultat du calcul des moyennes des amplitudes des sources.

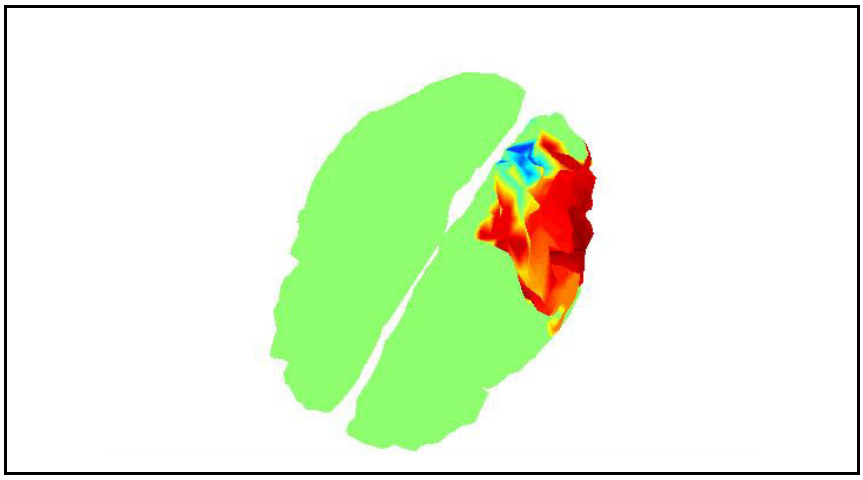


Figure 5.9 Amplitudes LCMV pour l'ensemble des sources sélectionnées par parcellisation

Comme on peut bien le constater sur cette figure, les sources ont presque le même niveau d'amplitude. Cela peut s'expliquer de deux façons : (1) les amplitudes des sources originalement dans la parcelle active se trouvent dans la même parcelle après la parcellisation et (2) le fait de calculer la moyenne simple des amplitudes LCMV sans tenir compte du nombre de visites pour former des parcelles, cela défavorise les sources visitées plusieurs fois et avantage les sources qui sont visitées très peu de fois.

Pour remédier à cette situation, dans la prochaine étape on présente une solution qui permet de tenir compte des résultats obtenus après la parcellisation.

5.2.5 Estimation de l'activation des sources

Dans cette dernière étape, on combine les résultats obtenus par les deux étapes précédentes pour estimer l'activation des sources tenant compte de leur occurrence dans des parcelles. Il y à plusieurs façons de combiner ces résultats, dans cette simulation, on a utilisé la façon suivante. Elle consiste principalement à calculer la moyenne des amplitudes LCMV pour chaque source. On multiplie chacune des amplitudes moyennées par des coefficients d'occurrences; nombres de visites des sources pour former des parcelles.

La figure 5.10 présente la carte des amplitudes moyennées pondérées avec les coefficients d'occurrence dans la formation des parcelles.

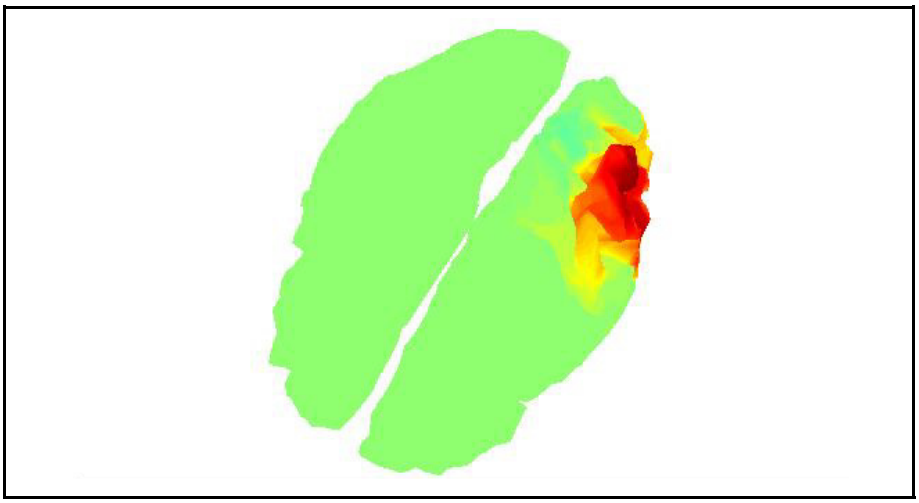


Figure 5.10 Carte des moyennes des amplitudes LCMV des sources pondérées avec leurs coefficients d'occurrences

En comparant les cartes d'activation des sources présentés dans les figure 5.7 et 5.9 avec le résultat présenté dans la figure 5.10, il est évident qu'il y a une amélioration de la localisation de l'activité.

5.3 Analyse ROC

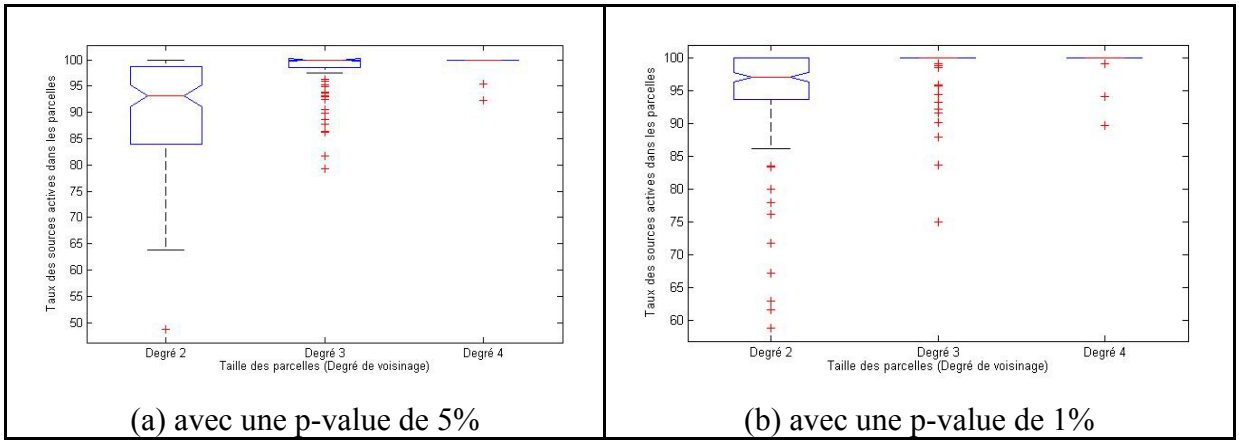
Dans cette section, on présente les résultats de l'analyse ROC concernant la localisation des sources avec des patchs activés à différentes localisation corticale. La taille des patchs activés est centrée autour d'une source germe avec un degré de voisinage géodésique de 3.

Pour cette évaluation ROC, on a considéré les données suivantes concernant :

- La sélection par FDR : on considéré deux seuils de tolérances des p-values; 1% et 5%.
- La parcellisation : on a considéré trois cas éventuels où la taille des parcelles formées soit de degré de voisinage géodésique 2, 3 et 4.

5.3.1 Taux des sources actives dans les parcelles

Dans cette section, on présente les résultats de la localisation des sources actives après parcellisation sélective dans les parcelles. Comme on peut le voir sur la figure 5.11, le taux des sources retrouvées s'améliore lorsque le degré de la parcellisation augmente et lorsque la p-value augmente aussi.



(a) avec une p-value de 5% (b) avec une p-value de 1%

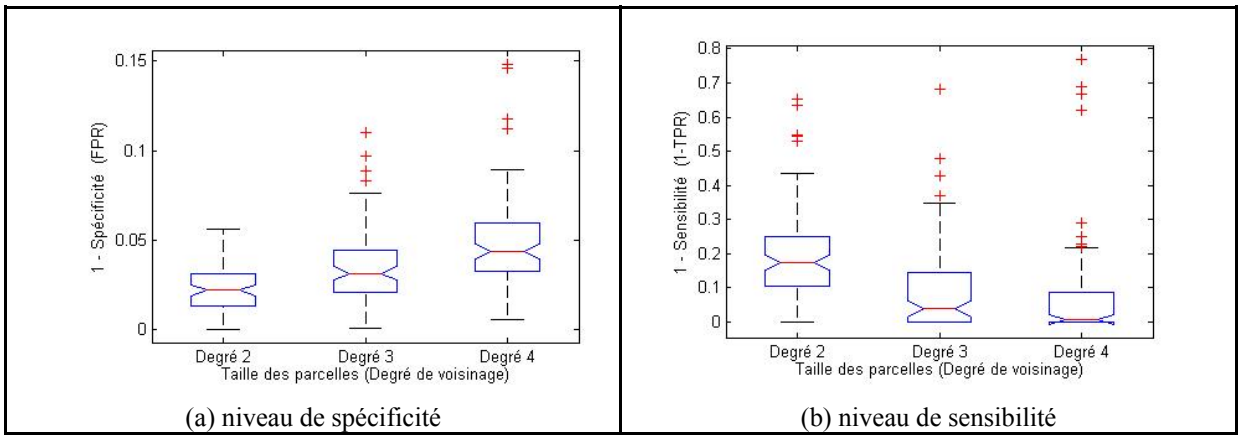
Figure 5.11 Taux des sources actives retrouvées après la parcellisation sélective

Selon ces résultats, on peut déduire que la région d'activation est déjà bien pré-localisée dans les parcelles. Cela montre que la parcellisation sélective permet d'améliorer la régularisation

du problème inverse et peut même être considérée comme une étape préliminaire dans la localisation des sources.

5.3.2 Localisation des sources avec un seuil FDR de 1% et faible bruit

Les résultats de l'analyse ROC présentés dans la figure 5.11 correspondent respectivement au cas où la taille des parcelles sont des degrés 2, 3 et 4 avec une tolérance aux faux positifs de 1%. D'après ces résultats, on peut dire que cet algorithme permet une bonne localisation des régions actives. On peut aussi remarquer que ces résultats varient selon la taille des parcelles formées durant la parcellisation. Plus précisément, on remarque que plus la taille des parcelles augmente plus la sensibilité augmente et plus la spécificité diminue.

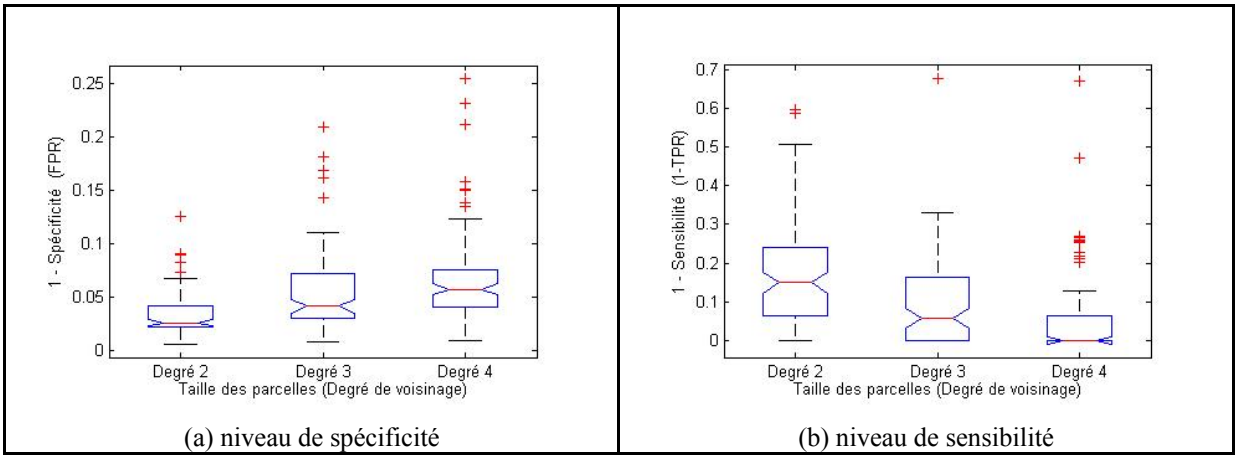


(a) niveau de spécificité (b) niveau de sensibilité

Figure 5.12 Analyse ROC de la localisation des sources avec une tolérance FP de 1%

5.3.3 Localisation des sources avec un seuil FDR de 5% et faible bruit

Les résultats de l'analyse ROC présentés dans la figure 5.13 correspondent respectivement au cas où la taille des parcelles sont de degrés 2, 3 et 4 avec une tolérance aux faux positifs (FP) de 5%. Les résultats obtenus avec une tolérance de 5% sont relativement comparables avec les résultats précédents avec une différence. Cette différence concerne l'amélioration de la sensibilité au compte de la spécificité qui diminue lorsqu'on augmente la tolérance aux faux positifs à 5%.

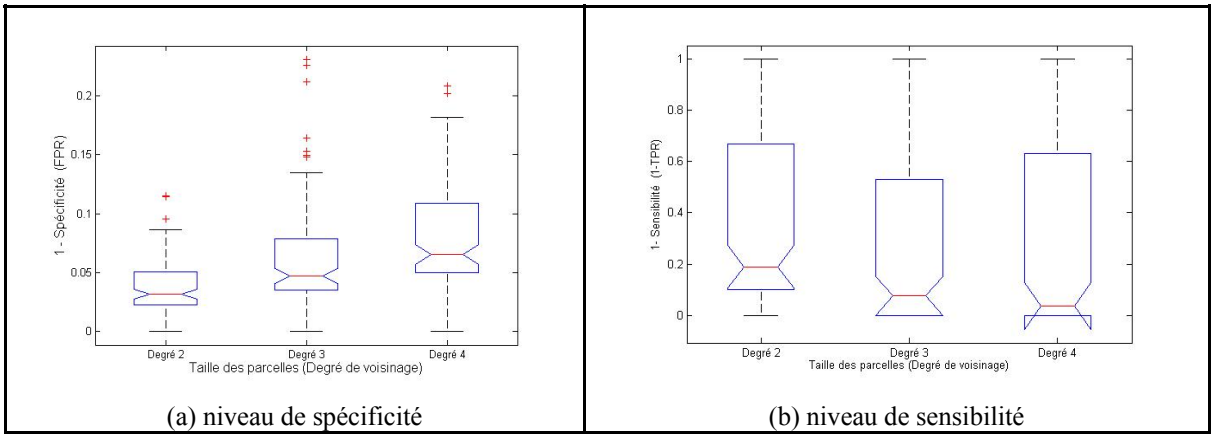


(a) niveau de spécificité (b) niveau de sensibilité

Figure 5.13 Analyse ROC de la localisation des sources avec une tolérance FP de 5%

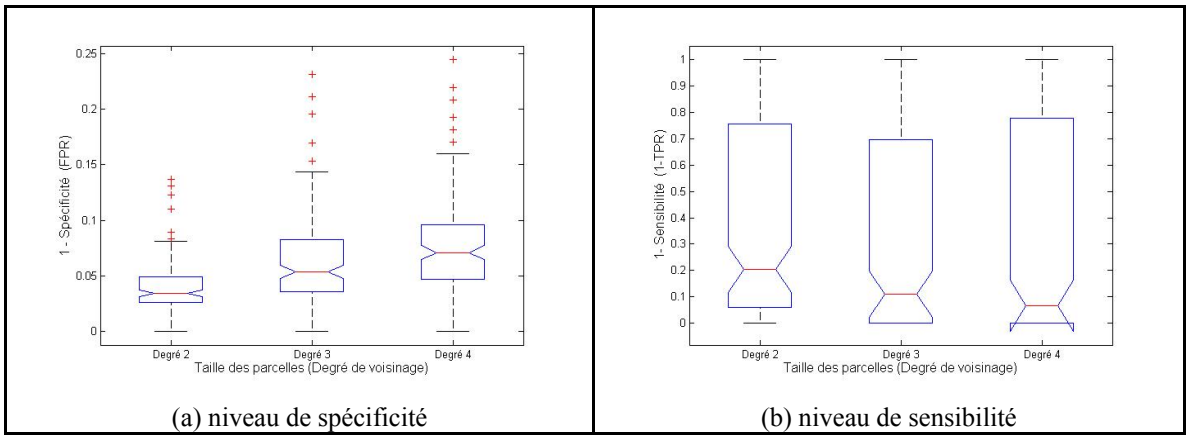
5.3.4 Localisation des sources avec un seuil FDR de 1% et 5% avec un bruit important

Les résultats de l'analyse ROC présentés dans les figures 5.14 et 5.15 sont comparables aux résultats présentées dans les figures 5.11 et 5.12 sauf qu'ici les mesures sont assez bruitées. On remarque une légère dégradation de la sensibilité ainsi que la spécificité au même degré presque lorsque le niveau de bruit dans le signal est devient important.



(a) niveau de spécificité (b) niveau de sensibilité

Figure 5.14 Analyse ROC de la localisation des sources avec une tolérance FP de 1% et des données avec un rapport signal/bruit de 10^{-4}



(a) niveau de spécificité (b) niveau de sensibilité

Figure 5.15 Analyse ROC de la localisation des sources avec une tolérance FP de 5% et des données avec un rapport signal/bruit de 10^{-4}

CONCLUSION

Ce travail est une contribution du développement d'un outil d'imagerie de détection de l'activité neuro-fonctionnelle. Il permet la localisation de l'activité cérébrale en utilisant une technique basée sur la modélisation en parcelles sélectives. L'ensemble des algorithmes de sélection des sources et de formation des parcelles sont implémentés et intégrés dans cet outil d'imagerie et de localisation corticale (Braintropy). L'implémentation de l'algorithme de filtrage spatial sur les parcelles et l'estimation de l'activation des sources sert beaucoup plus comme test comparatif.

En fait, l'originalité de cette technique est qu'elle utilise deux niveaux de la modélisation de la surface corticale. Le premier c'est la modélisation en sources qui permet d'assurer une bonne résolution spatiale. Le deuxième c'est la modélisation en parcelles sélectives qui permet la régularisation et l'amélioration de l'efficacité du résultat de la localisation des régions actives.

Dans le développement de cette technique on a proposé des principes qui peuvent être réutilisées et exploitées par d'autres techniques de résolution du problème inverse. Notamment, la sélection contrôlée par les données observées pour discriminer entre les sources potentiellement actives et ceux peu probablement actives. On a aussi proposé le concept de parcelles sélectives pour améliorer le conditionnement du problème inverse. Ce concept qui repose sur la sélection par la FDR des scores MSP permet d'élargir le spectre de la solution pour passer de la localisation de l'activité dans des sources et aller vers la notion de sa localisation dans des régions cérébrales. Même si on a développé ce concept pour notre technique de résolution basé sur la technique LCMV, elles peuvent facilement être adaptées pour d'autres techniques.

On a aussi proposé, dans ce travail, une technique de validation empirique concernant l'activité physiologique qui montre que les scores des sources au repos peuvent être

modélisés par une loi beta dont les paramètres peuvent être calculés à partir des fenêtres dont la taille varie entre 13 et 20 échantillons par fenêtre.

Les résultats de notre technique, selon les courbes ROC, montrent qu'elle permet la résolution du problème inverse avec un rapport de sensibilité et d'efficacité très intéressantes. Ces résultats varient essentiellement selon la localisation de la région active et le degré de la parcellisation choisie.

Malgré les bons résultats obtenues avec cette technique, on pense on peut les améliorer en considérant les points suivants :

1. Pour le calcul de la LCMV on utilise la matrice de corrélation des mesures estimé à partir d'une fenêtre de mesures donnée. Le résultat de résolution peu être amélioré si le choix de cette fenêtre est optimale.
2. La taille de la fenêtre lue sur les mesures de la *baseline* pour modéliser la distribution des scores au repos peut aussi influencer le résultat. Pour un résultat optimal, on a besoin de plus de tests pour estimer l'influence de cette taille sur le résultat.

On recommande de tester cette technique en utilisant d'autres techniques de résolution du problème inverse tel que la méthode de Maximum d'Entropie sur la Moyenne. On pense que les résultats peuvent être améliorés si plusieurs méthodes de résolution sont utilisées conjointement (i.e. MEM, LCMV etc…).

ANNEXE I

MODÉLISATION DE L'ACTIVITÉ NEURONALE

Les lois physiques, plus exactement celle qu'on appelle *'loi de MaxWell'* (Baillet S. 2001) permet de calculer le champ magnétique créer par une source de courant à l'intérieur d'un milieu comme cervicale humaine par exemple. Considérant une source avec un courant de densité J_i, due à une activité biologique ayant un facteur de conductivité σ et un facteur de perméabilité magnétique $\mu = \mu_0$ le même dans tout le tissu (pour simplification). Pour calculer le champ électrique **E** et le champ magnétique **B**, qui correspondent à J_i, on utilise une approximation quasi-statique des équations de MaxWell qui sont:

$$\nabla E = -\nabla V$$
$$\nabla \times B = \mu_0 J \qquad \nabla \cdot B = 0$$
$$J = J_i + \sigma E \tag{AI.1}$$

Il est à noter que l'équation (AI.1) est composée de deux sources:

1- J_i: un courant engendrer par le mouvement des charges électriques, c'est ce qu'on appelle courant primaire.
2- σE: un courant de conduction du tissue biologique, qu'on appelle courant secondaire. Il ne contribue pas dans **B** si on considère σ comme facteur de conductivité du tissu, qu'on suppose que c'est un milieu homogène.

avec **V** est le potentiel électrique, J est la densité totale du courant et σE est le courant Ohmic.

Il faut aussi noter que $\nabla \cdot J = 0$ (voir équation (AI.1)) et que le vecteur d'identité $\nabla \cdot \nabla \times B = 0$. Étant donné que J est le courant total, **B** peut être exprimé par la loi de Biot-Savart (théorie de conduction dans des volumes) dans l'équation suivante:

$$B(r) = \frac{\mu_0}{4\pi} \int_{\Omega} J(r') \frac{r - r'}{|r - r'|^3} dv' \qquad \text{(AI.2)}$$

Ω représente le volume de la tête.

Cette équation (après quelques calculs) est la solution de l'équation avec $B(r) \to 0$ quand $r \to \infty$.

Pour avoir **E** et J on doit trouver **V**. Si on considère que σ est une constante alors on peut écrire $J = J_i - \sigma \nabla V$. Puisque $\nabla \cdot J = 0$ on obtient $\nabla \cdot (\sigma \nabla V) = \nabla \cdot J_i$ et ainsi on a; $\Delta V = \frac{\nabla \cdot J_i}{\sigma}$

D'après (Sarvas 1987), si on considère **V** et **B** comme produit d'une source dipolaire ayant une densité de courant $J_i(r) = Q\ \delta(r - r_0)$, où Q contient l'intensité et r_0 est la position du dipôle alors ils peuvent être exprimés par les équations suivantes:

$$V(r) = \frac{1}{4\pi\sigma}\ Q \cdot \frac{r - r'}{|r - r'|^3} \qquad \text{(AI.3)}$$

$$B(r) = \frac{\mu}{4\pi\sigma}\ Q \times \frac{r - r'}{|r - r'|^3} \qquad \text{(AI.4)}$$

Il est à remarquer de l'équation (AI.4) que la mesure du champ magnétique **B** est linéaire par rapport au moment Q du dipôle et non linéaire vise à vis de la distance r'.

ANNEXE II

TEST D'HYPOTHÈSE

Dans la section précédente on a présenté la technique FDR comme technique d'inférence d'information statistique sur les coefficients d'activation des sources basée sur un test d'hypothèse multiple. Cette section est une présentation générale d'un test d'hypothèse et des notions suivantes : (1) hypothèse nulle; (2) test d'hypothèse; (3) une p-value et (4) erreur type (I & II).

I. Hypothèse nulle :

En statistiques, l'hypothèse nulle, généralement notée comme H_0, correspond à une hypothèse qu'on assume en avance concernant une donnée en observation. Cette hypothèse sera soit rejetée (en faveur d'une hypothèse alternatif H_1) ou non, cela dépend du résultat du test de cette hypothèse. Quand on utilise l'hypothèse nulle, elle est toujours considérée vraie, sauf si le test d'hypothèse (sous-forme d'une évidence statistique) indique autrement. Dans les sciences médicales, par exemple, l'hypothèse nulle est utilisée pour tester la différence entre les traitements médicaux et leurs effets sur diffèrent types de groupes (ayant des caractéristiques différentes).

En générale l'objectif de définir l'hypothèse (nulle) n'est pas de la confirmer mais plutôt de la rejeter. Par exemple, par convention (ca dépend des pays) "tout le monde est innocent jusqu'à preuve du contraire". C'est facile d'accuser quelqu'un sans preuve. Dans la vraie vie devant un accusé on le considère innocent (hypothèse nulle) mais si on l'accuse c'est qu'on cherche généralement à avoir une preuve suffisamment convaincante pour le considérer coupable ou non (rejeter l'hypothèse nulle ou non). Cette preuve est livrée suite à un test d'hypothèse.

i. Test d'hypothèse :

Un test d'hypothèse permet d'inférer une information à partir des données (observation) statistiques. Comme son nom l'indique un test d'hypothèse se base sur une hypothèse prédéfinie, l'hypothèse nulle (voir plus loin), qu'on considère vraie mais qu'elle n'est pas encore prouvée. Le test d'hypothèse est donc une démarche qui consiste à rejeter ou à accepter une hypothèse statistique concernant une donnée (un jeu de données ou échantillon). En générale, on distingue deux types de test d'hypothèse:

a. Test d'homogénéité : Le test d'homogénéité compare deux échantillons entre eux. Dans ce cas, l'hypothèse nulle H_0 considère que l'homogénéité entre les deux échantillons est vraie (ex. comparaison des moyennes).
b. Test de conformité : Par contre le test de conformité détermine si un échantillon suit une loi statistique particulière (ex. l'hypothèse nulle H_0 suppose que l'échantillon sous test suit une loi Normale).

En générale le test d'hypothèse suit les étapes suivantes : (1) définir l'hypothèse nulle H_0; (2) fixer ou calculer la valeur de la variable de décision qui correspond soit à la distance entre les deux échantillons, dans le cas d'un test d'homogénéité, ou à la conformité entre l'échantillon et la loi statistique, dans le cas d'un test conformité; (3) calculer la valeur pour la variable de décision, qu'on appelle p-value (voir en bas), en supposant que H_0 est vraie; ce qui correspond à calculer la probabilité de rejeter H_0 et (4) définir un seuil (α), qu'on appel seuil de tolérance, en dessous duquel H_0 est rejetée. Généralement α est plus que 1% et ne dépasse pas 5%, ce qui correspond au risque de rejeter H_0 alors qu'elle est vraie (plus d'explication s'envient).

ii. P_value

En générale, un test statistique lui est toujours associé une hypothèse nulle considérée comme vraie et sert comme référence pour le test. La p-value peut être définie comme la probabilité qu'un échantillon statistique observé (d'une même population sous-test) soit extrêmement possible sachant que l'hypothèse nulle est vraie. Il est important de motionner qu'en générale, lorsque le p-value est petit rejeter ou non l'hypothèse nulle n'est pas une décision fiable.

Dans l'industrie, on utilise un seuil auquel on compare la p-value pour décider si l'hypothèse nulle est à rejeter (ou non) en faveur d'une hypothèse alternative H_1. Ce seuil, c'est ce qu'on appel *seuil de rejet* ou *seuil de signification* et il est aussi noté α. Ce seuil est souvent autour de 5%, il est considéré sévère lorsqu'il est de l'ordre de 1% (ex. domaines critiques, santé, etc. Dans le cas où le seuil est fixé à 5% par exemple, si le p-value est inférieur à ce seuil alors on a un argument statistique qui permet de rejeter l'hypothèse nulle.

iii. Erreur de type (I & II)

Il faut admettre, qu'il y a toujours une possibilité d'erreur lorsqu'on rejette une hypothèse nulle; on risque de rejeter une hypothèse nulle alors qu'elle est vraie. La probabilité d'avoir cette erreur s'appelle *erreur de type I*; elle correspond au seuil de signification α. Ce seuil est préétabli, par conséquence, il n'est pas affecté par la taille des échantillons statistiques.

L'erreur de type I (α) correspond au faux positif et son complément est (1- α) qui correspond à la situation où l'hypothèse nulle est acceptée et cette décision est correcte; c'est ce qu'on appel un *vraie négatif*. On parle aussi de *l'erreur de type II* noté β, elle correspond à la probabilité d'accepter l'hypothèse nulle alors qu'elle ne doit pas l'être. Le complément de cette probabilité est (1- β) qui correspond à rejeter l'hypothèse nulle alors que cette décision est correcte. La figure (en haut voir section FDR) illustre les quartes possibilité considérant les deux types d'erreur (I \& II).

Pour conclure cette section, le test d'hypothèse est utilisé pour vérifier l'hypothèse nulle (la source est inactive), par rapport au seuil de signification, pour enfin décider la sélection ou non de chaque source du modèle. Ainsi, on utilisant un test d'hypothèse et un seuil de signification (tolérance) α, on assure le contrôle du taux de l'erreur (faux positifs) après sélection des sources.

BIBLIOGRAPHIE

Amblard C, L. E., Lina JM (2004). "Biomagnetic source detection by maximum entropy and graphical models." IEEE Trans Biomed Eng. : 427-442.

Barry D.V., Wim v.D., Moshe Y., and Akifumi. S., (1997). "“Localization of brain electrical activity via linearly constrained minimum variance spatial filtering,” " IEEE Trans. Biomed. Eng. (9): 867–880.

Baillet, S. and L. Garnero (1997). "A Bayesian Approach to Introducing Anatomo-Functional Priors in the EEG/MEG Inverse Problem." IEEE Transactions on Biomedical Engineering (5).

Baillet S., M. J. C., and Leahy R.M. (2001). "Electromagnetic Brain Mapping,." IEEE, Signal Processing Magazine (6): 14, 30.

Benjamini, Y. and Y. Hochberg (1995). "Controlling the False Discovery Rate: a Practical and Powerful Approach to Multiple Testing." Journal of the Royal Statistical Society (B): 289, 300.

Cointepas, Y. (1999). "Modélisation Homotopique et Segmentation 3D du Cortex Cérébral à partir d'IRM pour la Résolution des Problèmes Directs et Inverse en EEG et en MEG". Paris, France, ENST Paris.

Dale A., Sereno M. (1993). "Improved localization of cortical activity by combining EEG and MEG with MRI cortical surface reconstruction: A linear approach." Journal of Cognitive Neuroscience 5:162-176.

Fischl, B., David. H. Salat, et al. (2002). "Whole Brain Segmentation: Neurotechnique Automated Labeling of Neuroanatomical Structures in the Human Brain." Neuron: 341–355.

Fischl, B., Sereno M. I., et al. (1999). "Cortical Surface-based Analysis II: Inflation, Flattening, and a Surface-based Coordinate System,." NeuroImage : 195, 207.

Fischl, B., VanDerKouwe A., et al. (2004). "Automatically Parcellating the Human Cerebral Cortex." Cerebral Cortex: 11–22.

http://www.4dneuroimaging.com/WhatIsMEG. "Properties of MEG" from http://web.mit.edu /kitmitmeg/whatis.html.

http://www.psycho-neuro-feedback.com/, http://www.psycho-neuro-feedback.com/Articles/

BasicintroductiontopIRHEG.aspx. image001.jpg,

Lancaster JL, Woldorff M., Parsons LM, Liotti M, Freitas CS, Rainey L, Kochunov PV, Nickerson D, Mikiten SA, Fox PT, (2000). "Automated Talairach Atlas labels for functional brain mapping." Human Brain Mapping: 120-131.

Lapalme E., (2004). Méthodes Entropiques au Problème Inverse en Magnétoencéphalographie. Montreal, Canada Université de Montréal, Département de physique. PhD thesis.

Lapalme, E., Lina J.-M., et al. (2006). "Data-driven Parceling and Entropic Inference in MEG." NeuroImage (1): 160-171

Limpiti, T., Barry D. Van-Veen, et al. (2006). "Cortical Patch Basis Model for Spatially Extended Neural Activity." IEEE Transaction on Biomedical Engineering (9): 1740-1754.

Lin, F.-H., Witzel T., et al. (2008). "Linear Constraint Minimum Variance Beamformer Functional Magnetic Resonance Inverse Imaging." NeuroImage 297–311.

Mardia K., KENT J. T., et al. (1979). "Multivariate Analysis. ." Academic Press, Biometrical Journal (5): 502, 518.

Matsuura, K. and Okabe Y. (1995). "Selective Minimum-norm Solution of the Biomagnetic Inverse Problem." Biomedical Engineering, IEEE Transactions (6): 608 - 615.

Mattout J., Peligrini-Issac M., Garnero L. and Benali H. (2005). "Multivariate source prelocalization (MSP): use of functionally informed basis functions for better conditioning the MEG inverse problem." NeuroImage (2005) ((2)): 356–373.

Maurieés, J. P. Anatomie du Cortex Somatosensoriel Primaire. http://www.vetopsy.fr/sens/soma/soma_cortex1.php.

Pascual-Marqui RD, Michel C.M., Lehmann D. (1994). "Low resolution electromagnetic tomography: a new method for localizing electrical activity in the brain." Int Journal of Psychophysiologie : 49-65.

Sarvas, J. (1987). "Basic Mathematical and Electromagnetic concepts of the Biomagnetic Inverse Problem." Low Temperature Laboratory, Helsinki University of Technology n^{o} 1: 11,22.

Scheibel M.E., Davies T. L., et al. (1974). "Basilar Dendrite Bundles of Giant Pyramidal Cells." Exp Neurol.: 307-319

Silva C., Maltez J. C., et al. (2004). "Evaluation of L1 and L2 Minimum Norm Performances on EEG Localizations." Clinical Neurophysiology 1657–1668.

Thirion B., Flandin G., et al. (2006). "Dealing With the Shortcomings of Spatial Normalization: Multi-Subject Parcellation of fMRI Datasets." Human Brain Mapping: 678–693.

Tosun D., Rettman. M. E., Hana X., Tao X., Xu C., Resnick S. M., Pham D. L. and Prince J. L. (2004). "Cortical surface segmentation and mapping." Elsevier NeuroImage 108 - 118.

Tulaya Limpiti, B. D. V. V. a. R. T. W. (2006). "Cortical Patch Basis Model for Spatially Extended Neural Activity." IEEE Transactions on Biomedical Engineering **53**(9): 1740-1754.

Tzourio-Mazoyer, N., B. Landeau, et al. (2002). "Automated Anatomical Labeling of Activations in SPM using a Macroscopic Anatomical Parcellation of the MNI MRI Single-subject Brain." NeuroImage: 273–289.

Uutelaa, K., Hämäläinen M., et al. (1999). "Visualization of Magnetoencephalographic Data Using Minimum Current Estimates." NeuroImage (2): 173-180

Van-Veen B.D. and K. Buckley (1988). "Beamforming: a versatile approach to spatial filtering." IEEE ASSP Magazine: 4–24.

Wagner M. Köhler Th.,fuchs M., and J. Kastner (2000). "An extended source model for current density reconstructions." Biomag2000, Proc. 12th Int. Conf. on Biomagnetism.