

Université de Montréal

Méthodes entropiques appliquées au problème  
inverse en magnétoencéphalographie

par

Ervig Lapalme

Département de Physique  
Faculté des Arts et des Sciences

Thèse présentée à la Faculté des Études Supérieures  
en vue de l'obtention du grade de  
Philosophiæ Doctor (Ph.D.)  
en Physique

novembre 2004



QC

3

U54

2005

V.001

## AVIS

L'auteur a autorisé l'Université de Montréal à reproduire et diffuser, en totalité ou en partie, par quelque moyen que ce soit et sur quelque support que ce soit, et exclusivement à des fins non lucratives d'enseignement et de recherche, des copies de ce mémoire ou de cette thèse.

L'auteur et les coauteurs le cas échéant conservent la propriété du droit d'auteur et des droits moraux qui protègent ce document. Ni la thèse ou le mémoire, ni des extraits substantiels de ce document, ne doivent être imprimés ou autrement reproduits sans l'autorisation de l'auteur.

Afin de se conformer à la Loi canadienne sur la protection des renseignements personnels, quelques formulaires secondaires, coordonnées ou signatures intégrées au texte ont pu être enlevés de ce document. Bien que cela ait pu affecter la pagination, il n'y a aucun contenu manquant.

## NOTICE

The author of this thesis or dissertation has granted a nonexclusive license allowing Université de Montréal to reproduce and publish the document, in part or in whole, and in any format, solely for noncommercial educational and research purposes.

The author and co-authors if applicable retain copyright ownership and moral rights in this document. Neither the whole thesis or dissertation, nor substantial extracts from it, may be printed or otherwise reproduced without the author's permission.

In compliance with the Canadian Privacy Act some supporting forms, contact information or signatures may have been removed from the document. While this may affect the document page count, it does not represent any loss of content from the document.

Université de Montréal  
Faculté des Études Supérieures

Cette thèse intitulée

**Méthodes entropiques appliquées au problème  
inverse en magnétoencéphalographie**

présentée par

**Ervig Lapalme**

a été évaluée par un jury composé des personnes suivantes :

*Richard B. MacKenzie*  
\_\_\_\_\_  
(président-rapporteur)

*Bernard Goulard*  
\_\_\_\_\_  
(directeur de recherche)

*Jean-Marc Lina*  
\_\_\_\_\_  
(co-directeur)

*Alain Vincent*  
\_\_\_\_\_  
(membre du jury)

*Line Garnero*  
\_\_\_\_\_  
(examinatrice externe)

*Pierre Savard*  
\_\_\_\_\_  
(représentant du doyen)

Thèse acceptée le:  
*24 novembre 2004*  
\_\_\_\_\_

# SOMMAIRE

---

Cette thèse traite du problème de la localisation de sources biomagnétiques à l'aide de la magnétoencéphalographie. Ce problème est connu pour avoir une infinité de solutions ce qui signifie qu'il n'y a pas assez d'information sur la solution pour pouvoir en identifier une seule. Nous avons développé des méthodes pour introduire de l'information anatomique et fonctionnelle sur la solution. Nous utilisons la méthode du maximum d'entropie sur la moyenne pour résoudre ce problème. Cette méthode prend ses origines dans la mécanique statistique et la théorie de l'information.

Cette thèse se divise en deux parties constituées de trois chapitres chacune. La première introduit le problème inverse en magnétoencéphalographie. On y développe la théorie nécessaire pour en comprendre le contexte et les hypothèses servant à simplifier le problème. Dans le dernier chapitre de cette première partie, on présente la méthode du maximum d'entropie sur la moyenne. On y explique ses origines, son développement puis ses relations avec les autres méthodes de reconstruction.

La deuxième partie constitue le travail original de cette thèse. Elle est formée de trois articles dont le premier a déjà été publié et les deux autres ont été soumis pour publication<sup>1</sup>. Dans le premier, un modèle de sources biomagnétiques est développé et appliqué dans un contexte certes théorique mais démontrant tout de même son efficacité en comparant avec d'autres méthodes de reconstruction. Dans le deuxième article, on fait un pas de plus vers une modélisation réaliste de l'activité cérébrale. Les principaux *a priori* faits dans le premier article sont estimés à partir des données magnétoencéphalographiques. Cette méthode s'est avérée très efficace lors de simulations dans un contexte réaliste. Dans le troisième, on étend la méthode précédente à des signaux temporels. Ceci permet d'exploiter l'excellente résolution temporelle offerte par la magnétoencéphalographie. Contrairement aux travaux précédents, la méthode temporelle est appliquée

---

<sup>1</sup>En date du dépôt final de cette thèse.

à de véritables données magnétoencéphalographiques provenant d'une expérience de somatotopie. Il est montré que les résultats concordent avec les connaissances physiologiques que l'on a sur ce genre de procédés cognitifs.

**Mots Clés:** Magnétoencéphalographie, Problème inverse, Maximum d'entropie sur la moyenne, Régions d'activation cérébrale, Parcellisation du cortex, Régularisation spatio-temporelle, Imagerie cérébrale fonctionnelle.

## SUMMARY

---

This thesis is devoted to biomagnetic source localization using magnetoencephalography. This problem is known to have an infinite number of solutions. So methods are required to take into account anatomical and functional information on the solution. The work presented in this thesis uses the maximum entropy on the mean method to constrain the solution. This method originates from statistical mechanics and information theory.

This thesis is divided into two main parts containing three chapters each. The first part reviews the magnetoencephalographic inverse problem: the theory needed to understand its context and the hypotheses for simplifying the problem. In the last chapter of this first part, the maximum entropy on the mean method is presented: its origins are explained and also how it is applied to our problem.

The second part is the original work of this thesis presenting three articles; one of them already published and two others submitted for publication. In the first article, a biomagnetic source model is developed and applied in a theoretical context but still demonstrating the efficiency of the method. In the second article, we go one step further towards a realistic modelization of the cerebral activation. The main priors are estimated using the magnetoencephalographic data. This method proved to be very efficient in realistic simulations. In the third article, the previous method is extended to deal with time signals thus exploiting the excellent time resolution offered by magnetoencephalography. Compared with our previous work, the temporal method is applied to real magnetoencephalographic data coming from a somatotopy experience and results agree with previous physiological knowledge about this kind of cognitive process.

**Keywords:** Magnetoencephalography, Inverse problem, Maximum entropy on the mean, Regions of activation, Cortex parceling, Spatio-temporal regularization, Functional brain imaging.

# Table des matières

---

Sommaire.....	iii
Summary.....	v
Table des matières.....	vi
Table des figures.....	ix
Liste des tableaux.....	xii
Remerciements.....	xiv
Introduction.....	1
<b>Chapitre 1. Physique de la Magnétoencéphalographie.....</b>	<b>7</b>
1.1. Le neurone et l'activité cérébrale.....	7
1.2. Calcul du champ magnétique induit par une source de courant....	10
1.3. Le dipôle de courant.....	14
<b>Chapitre 2. Problème inverse en Magnétoencéphalographie.....</b>	<b>18</b>
2.1. Modèles dipolaires.....	19
2.2. Modèles distribués.....	20
2.2.1. Méthode de norme minimale.....	22
2.2.2. LORETA.....	23
2.2.3. Méthodes bayésiennes.....	25
2.2.4. S-MAP.....	26
2.2.5. Présélection de sources.....	29
<b>Chapitre 3. Maximum d'entropie en moyenne.....</b>	<b>33</b>
3.1. Fondements de la mécanique statistique.....	34
3.1.1. Théorie ergodique.....	34
3.1.2. Entropie.....	36
3.2. Problèmes inverses en MEM.....	38
3.2.1. Expression duale du MEM.....	40
3.2.2. Critère primal du MEM.....	42
3.2.3. Bruit en MEM.....	43
3.3. Liens avec les autres méthodes de reconstruction.....	44



<b>Chapitre 4. Maximum d'Entropie et Magnétoencéphalographie</b>	
<b>Premier article</b> .....	47
4.1. Introduction .....	51
4.2. The Forward and Inverse Problems .....	53
4.3. Primal and Dual formulations of the MEM .....	55
4.3.1. The Maximum of Entropy on the Mean .....	56
4.3.2. Noisy observations .....	59
4.3.3. MEM and Maximum Entropy formalisms .....	59
4.4. Graphical Statistical Models and MEM .....	60
4.4.1. General Case .....	60
4.4.2. A degenerate case: the Finite Mixture Model .....	63
4.5. Simulations and Iterative MEM .....	66
4.5.1. Lattice, phantoms and simulated data .....	66
4.5.2. Reference measure on the dipole amplitude .....	68
4.5.3. MEM estimation .....	70
4.5.4. Learning the states: <i>Iterative MEM</i> algorithms .....	72
4.5.5. Results and Evaluation .....	73
4.6. Conclusion .....	78
Acknowledgements .....	80
Bibliography .....	85
<b>Chapitre 5. Estimation de la loi de référence</b>	
<b>Second article</b> .....	87
5.1. Introduction .....	90
5.1.1. The MEG inverse problem .....	90
5.1.2. Notations and classical inverse solutions .....	91
5.1.3. The entropic inference .....	93
5.2. Method .....	94
5.2.1. Maximum entropy on the mean .....	94
5.2.2. Data-driven cortex parcelling .....	100
5.3. Application .....	104
5.3.1. The MEG data simulation .....	104
5.3.2. The evaluation criteria .....	106
5.4. Results .....	107
5.5. Discussion .....	111
5.6. Acknowledgements .....	114
Bibliography .....	115

<b>Chapitre 6. Reconstruction spatio-temporelle</b>	
<b>Troisième article</b> .....	117
6.1. Introduction .....	120
6.2. Maximum entropy on the mean .....	122
6.2.1. Prior law .....	123
6.3. Data-driven parcelling .....	124
6.3.1. Source preselection .....	124
6.3.2. Parcelling .....	126
6.4. Hyperparameters estimation .....	127
6.5. Simulations .....	129
6.5.1. Framework .....	129
6.5.2. Results .....	130
6.6. Real MEG data .....	133
6.7. Conclusion .....	135
Acknowledgements .....	136
Bibliography .....	138
<b>Conclusion</b> .....	139
<b>Bibliographie</b> .....	142

## Table des figures

---

I.1	Exemple d'images IRM .....	3
I.2	Résolution spatiale et temporelle des différentes méthodes d'imagerie cérébrale.....	4
1.1	Illustration des différentes parties anatomiques d'un neurone .....	9
1.2	Schéma montrant les vecteurs de l'expression pour le champ magnétique $\mathbf{B}$ généré par un dipôle de courant $\mathbf{Q}$ situé à $\mathbf{r}_0$ à un point d'observation $\mathbf{r}$ (voir équation (1.13)). .....	16
2.1	Graphes des différents potentiels utilisé en reconstruction bayésienne.	28
2.2	Histogramme des $a_i$ pour un exemple pris du chapitre 6. La ligne verticale rouge indique la valeur du seuil $a = 0.65$ donné par une $p$ -value de $p = 0.05$ . .....	31
4.1	Schematic representation of the source brain model. Dipoles are located at sites “ $i$ ”. Position ( $d_i$ ) and orientation of the dipoles are fixed; the dipole intensities (or amplitudes) $r_i$ are random quantities to be estimated from the magnetic field $\mathbf{b}$ . The present approach consists in introducing <i>latent</i> (or <i>hidden</i> ) state variables $S_i$ that describe locally the activation state of the cortical strip.....	53
4.2	MEM: the solution $\tilde{p}$ is the “closest” density to the reference law ( $\mu$ ) that explains the data (the “true solution” is also in $\mathcal{C}_m$ ).....	57
4.3	Graphical model: nodes, links and neighborhood. The set $\gamma(2)$ is the statistical neighborhood of node 2.....	64
4.4	Mixture model for the activation sources and associated graphical model. A white node represents a region and its state variable. It “emits” the intensities of the $n_k$ dipoles (black nodes) that belong to the corresponding region $P_k$ .....	64
4.5	$G_u(x) = \frac{x}{x+(1-x)e^{-u}}$ for $u > 0$ (dashed-dot), $u = 0$ (plain) and $u < 0$ (dashed).....	67
4.6	“Cortical mesh” used in simulations and an example of phantom. <b>Upper:</b> the “cortical surface” is modeled as an assembly of 16 patches with 64 nodes each. <b>Lower:</b> <i>Phantom0</i> . Intensities $\mathbf{r}$ of the 1024 dipoles are coded in grayscale. The patch 3 is active (with outgoing currents) and patch 13 is partially active (upper half dipoles are ingoing	

	currents; lower half dipoles are “silent” with vanishing currents). Notice that this phantom mixes different scales of activation.....	69
4.7	$F_{\mu_g}^*(\xi)$ for Normal laws $N(m_+ > 0, \sigma_+)$ and $N(m_- < 0, \sigma_-)$ (line and dashed line respectively).....	70
4.8	MEM reconstruction of <i>Phantom0</i> with various activation scales: $\tilde{r}_{K=16}$ (a), $\tilde{r}_{K=64}$ (b), $\tilde{r}_{K=256}$ (c) and $\tilde{r}_{K=1024}$ (d). The scale of the region is indicated with a box in the upper left corner of each display. Data are noise free.....	71
4.9	MEM reconstruction $\tilde{r}^*$ without privileged scale. $\rho_K = 1/4$ and $K = 16, 64, 256$ and $1024$ in (4.46).....	72
4.10	MEMiter: the reference law “lives” in a space of prior models.....	73
4.11	Progression of the iterative MEM algorithm applied to <i>Phantom2</i> reconstruction.....	74
4.12	Phantom 1 (top,left) with 4 reconstructions from <b>non-noisy</b> observations: MEM, MEMiter, LORETA and S-MAP. (a): $K = 16$ (64 dipoles per region); (b): $K = 64$ (16 dipoles per region); (c): $K = 256$ (4 dipoles per region); (d): $K = 1024$ (1 dipole per region). (e) is the reconstruction of the form (4.46) with $K = 16, 64$ and $256$ . .....	81
4.13	Phantom 2 (top,left) with 4 reconstructions from <b>non-noisy</b> observations: MEM, MEMiter, LORETA and S-MAP. (a): $K = 16$ (64 dipoles per region); (b): $K = 64$ (16 dipoles per region); (c): $K = 256$ (4 dipoles per region); (d): $K = 1024$ (1 dipole per region). (e) is the reconstruction of the form (4.46) with $K = 16, 64$ and $256$ . .....	82
4.14	Phantom 1 (top,left) with 4 reconstructions from <b>noisy</b> observations (30 dB): MEM, MEMiter, LORETA and S-MAP. (a): $K = 16$ (64 dipoles per region); (b): $K = 64$ (16 dipoles per region); (c): $K = 256$ (4 dipoles per region); (d): $K = 1024$ (1 dipole per region). (e) is the reconstruction of the form (4.46) with $K = 16, 64$ and $256$ . .....	83
4.15	Phantom 2 (top,left) with 4 reconstructions from <b>noisy</b> observations (30 dB): MEM, MEMiter, LORETA and S-MAP. (a): $K = 16$ (64 dipoles per region); (b): $K = 64$ (16 dipoles per region); (c): $K = 256$ (4 dipoles per region); (d): $K = 1024$ (1 dipole per region). (e) is the reconstruction of the form (4.46) with $K = 16, 64$ and $256$ . .....	84
5.1	Illustration of the clustering process.....	103
5.2	Clustering example.....	108
5.3	Hyperparameters estimation example.....	109
5.4	Results of the reconstruction over different activation scales using noisy data with non-iterative MEM.....	110
5.5	Results of the reconstruction for the methods with 20 dB noisy data.....	112

6.1	Histograms of the $a_{j;l}$ . Left, $a_{j;l}$ with the data from $t = 1, \dots, 10$ ( $l = 1$ ) for the phantom in Figure 6.3. Center, $a_{j;l}$ with $t = 51, \dots, 60$ ( $l = 6$ ). Right, $a_j = \max_{1 \leq l \leq L} a_{j;l}$ valid for all time steps, $t = 1, \dots, 60$ . The threshold $a$ , which is different for each histogram, is indicated by the vertical red line. ....	126
6.2	Example of spline smoothing of the $\alpha_k$ 's. The dashed line is the discontinuous $\alpha_{k;l}$ . We see the steps at $t = 10, 20, 30, \dots$ . The solid line is the $\alpha_{k;l}$ smoothed by a cubic spline. ....	128
6.3	These figures present one of the phantoms used in the simulations. The left image shows the spatial distribution of the activation at $t = 40$ . The right image contains the time evolution of the dipole intensity in the colored patches A and B. ....	129
6.4	Parcelling with $D = 9$ mm. On the left side, the cortex displays the reconstructed region. On the right, each image shows the phantom mean activation of each region which contains at least one activated dipole. The number on top of each box corresponds to the region in the parcelling while the number ( $n_k$ ) of dipoles in each region is in parentheses. Axes are the same as in Figure 6.3. This image can be used to observe the reconstruction performance of each method. ....	131
6.5	Cortex views of the reconstruction at $t = 40$ . The phantom activation is in Figure 6.3. (a): " <i>MEM+time+smooth</i> ". (b): " <i>MEM+time-smooth</i> ". (c): " <i>MEM-time</i> ". ....	132
6.6	Temporal reconstructed activation ....	134
6.7	The images (a) and (b) show the real MEG data. In (a), it is the complete set for 256 milliseconds. In (b), we extracted 80 times samples around the signal peak to do the reconstruction and it is displayed with the time windows used in the parcelling process. The images (c) and (d) show the reconstruction made with " <i>MEM+time+smooth</i> ". Time course of some dipoles is shown at left. The number on the top of each box in is the number of dipoles displayed. In (d), we show the spatial organization of the solution at $t = 40$ . The dipoles displayed in the image (c) are the ones circled in image (d). The color bar at the very bottom is in units of $10 \text{ nA} \cdot \text{m}$ and refers to the image (d). ....	135

## Liste des tableaux

---

4.1	Performance of MEM, MEMiter, LORETA and S-MAP. No Noise. . . .	76
4.2	Performance MEM, MEMiter, LORETA and S-MAP (scale average). No Noise. . . . .	77
4.3	Performance MEM, MEMiter, LORETA and S-MAP. Noisy data. . . .	79
5.1	Results of the reconstruction for non-iterative (noniter), iterative MEM (iter <sub>1</sub> , iter <sub>2</sub> ) and LORETA for non-noisy and noisy (20 dB) MEG data. The acronym DDC stands for data-driven clustering. . . . .	113
6.1	Objective comparison between the reconstruction methods. Each box displays the average over the three phantoms together with the standard deviation. . . . .	133

La vie c'est court, mais c'est long des p'tits bouts.

André Fortin (Les Colocs)  
dans la chanson "Le Répondeur"  
sur Dehors Novembre

## REMERCIEMENTS

---

Je remercie tout d'abord mes directeurs de recherche Bernard Goulard et Jean-Marc Lina. Vous avez toujours respecté mes décisions qui n'étaient pas toujours pour le bien du groupe de recherche. Vous avez aussi enduré mes sautes d'humeur lorsque les relations devenaient plus tendues. Vous avez respecté les courtes échéances que je vous ai imposées lors des derniers mois de rédaction. Votre disponibilité durant cette période a été grandement appréciée.

Karine est la personne à qui je dois le plus. Depuis que nous sommes ensemble, elle m'a toujours poussé à être meilleur pour moi et m'a toujours épaulé lorsque venaient les décisions difficiles. Et je lui lève mon chapeau pour avoir enduré mon absence durant les derniers mois de rédaction. Le boulot en plus de la thèse me ne laissait pas beaucoup de temps à lui consacrer. Elle s'est toujours débrouillé pour je ne manque de rien durant cette période. Karine, si je termine cette thèse, c'est grâce à toi et seulement toi. Karine, merci du fond du cœur. Je t'aimerai toujours.

Mes parents, Diane et Guy, ont toujours été là pour m'encourager. Même lorsque j'ai pris des décisions qui mettaient en péril cette thèse, ils ont toujours crû en moi et même parfois plus que moi-même. Je vous en remercie. Un merci spécial à mon père pour avoir lu cette thèse à deux reprises en un temps record. Si cette thèse n'est pas une catastrophe, c'est un peu à cause de lui.

Qu'auraient été ces années d'études sans quelques collègues avec qui on chiâle et on refait le monde. Chaint, Flon et Miguel<sup>2</sup>, merci pour ce bon temps passé ensemble et pour ces heures de rires gravées dans ma mémoire. Un merci spécial à Philippe St-Jean pour m'avoir permis d'emprunter un ordinateur lequel j'ai surutilisé pour faire les calculs numériques nécessaires à la production des résultats contenus dans cette thèse.

---

<sup>2</sup>Vous n'êtes pas en ordre!



Quelques rencontres au début de nos études universitaires parfois sont déterminantes. C'est le cas avec Thomas, Hugo, Philippe et J-F<sup>3</sup>. Vous avez changé ma vie et je vous en remercie.

Je souligne aussi la motivation que m'a procurée Marie lorsqu'elle a elle-même terminé sa thèse. Elle l'a obtenue alors que mes espoirs étaient enterrés ce qui m'a fait comprendre qu'il était possible d'achever une thèse dans un seul morceau et ce, même dans des conditions pas toujours idéales.

Je voudrais remercier Pr. Line Garnero pour nous avoir procuré des données expérimentales de magnétoencéphalographie qui ont permis de tester les méthodes présentées dans cette thèse. Je la remercie aussi d'avoir organisé un atelier de formation auquel j'ai participé en juin 2001 à Paris. Cet atelier a été déterminant pour approfondir mes connaissances du domaine. Je la remercie aussi d'avoir si gentiment accepté de faire partie du jury de cette thèse comme examinatrice externe.

La rencontre avec Cécile Amblard au début de ma thèse a été déterminante puisqu'elle a su insuffler un momentum au projet. Son expertise avec la méthode du maximum d'entropie sur la moyenne fut d'une grande utilité. Aussi, elle a toujours pris le temps de me conseiller. De plus, elle a gentiment offert de relire ma thèse ce qui fut grandement apprécié.

J'aimerais souligner les discussions enrichissantes que j'ai eu avec Jean Daunizeau. Même si nous ne nous sommes pas vu très souvent, il a toujours été disponible pour m'écouter.

Je souligne la gentillesse de Richard Hogue, mon patron chez Environnement Canada, de m'avoir fait comprendre qu'il fallait que je prenne une décision pour finir cette thèse. Il a généré en moi une réflexion qui m'a enfin fait bouger (pour le plus grand plaisir de Karine). Il m'a aussi permis d'assouplir mes heures de travail durant les derniers milles de ma thèse.

Que feraient les étudiants en physique sans Louise Lafortune? Louise m'a toujours conseillé le mieux possible et toujours pris soin de mon dossier même lorsque je n'étais plus incrit. Sa patience est précieuse et sa bonne humeur, d'or.

---

<sup>3</sup>Vous n'êtes pas en ordre!

Durant mes études à temps plein à l'Université de Montréal, j'ai reçu le soutien financier des organismes CRSNG et FCAR. Grâce à eux, j'ai pu payer mon loyer, faire mon épicerie : donc vivre!

Je voudrais aussi lever mon chapeau à mon sous-sol qui m'a toujours été fidèle malgré le nombre incalculable d'heures passé en sa compagnie et ses fissures si charmantes.

# INTRODUCTION

---

L'étude du siège de la pensée est l'un des plus grands défis scientifiques auxquels l'humain s'est attaqué. Un long chemin tortueux a mené aux découvertes constituant les fondements de la neurologie. Il y a plus de deux cents ans, Galvani mettait en évidence l'existence de l'électricité animale, par hasard, en faisant des expériences sur des grenouilles. De fil en aiguille, nous avons découvert que le transport d'information dans le corps est réalisé via un courant électrique. Ces courants permettent aux différentes parties du corps de communiquer entre elles. L'élément de base du cerveau (et de tout le système nerveux) est le neurone. Les neurones sont, comme l'avait remarqué Galvani, les porteurs des courants électriques, ces courants étant un signe d'activation neuronale. Cette activité neuronale s'accompagne de phénomènes autres qu'électromagnétiques. Pour générer un courant, un neurone a besoin d'énergie. Cette énergie est puisée dans le sang qui est apporté par les réseaux sanguins. Puis pour que le courant électrique soit engendré, un échange d'ions entre les membranes a lieu ce qui donne une signature chimique de l'activité cérébrale.

Cependant, avant le vingtième siècle, il était impossible d'observer le cerveau en action. Les scientifiques se limitaient à des études sur des cadavres ou des animaux. Il a fallu attendre la fin des années 20 du siècle dernier pour avoir un instrument de mesure de l'activité cérébrale. En effet, en 1929, H. Berger a inventé l'électroencéphalographie (EEG). Quoique les équipements ont depuis beaucoup évolué, surtout avec l'avènement de l'informatique, le principe de mesure est resté le même. L'EEG consiste à mesurer les différences de potentiel électrique, sur la surface du crâne, engendrées par les courants neuronaux. L'EEG permet donc d'observer l'activité cérébrale en temps réel ou presque.

Le complément naturel de l'EEG, la magnétoencéphalographie (MEG), a été inventé beaucoup plus tard. En effet, la MEG consiste à mesurer les très faibles champs magnétiques créés par l'activation cérébrale. Ils sont un milliard de fois plus petits que le champ magnétique terrestre. Le physicien D. Cohen [9] fut le

premier à observer ces champs magnétiques en utilisant des dispositifs supraconducteurs, des SQUID<sup>4</sup>, qui n'ont été mis au point qu'à la fin des années soixante. Il n'est pas étonnant que la MEG ait été inventée par un physicien. En effet, les SQUID utilisent des effets quantiques pour détecter des champs magnétiques d'une extrême faiblesse. La MEG ayant été découverte par des physiciens, la reconstruction des foyers d'activation cérébrale avec les données MEG fut pour eux un problème qu'il était naturel d'attaquer. Ce n'est qu'un autre problème d'électromagnétisme, qui a ses difficultés propres bien sûr, mais nous essaierons de le résoudre dans cette thèse. L'EEG, qui pourtant fut inventée quarante ans auparavant, n'avait jamais encore été utilisée à cette fin. Ayant été inventée par des neurophysiologistes, ces derniers n'utilisaient l'EEG que pour son information temporelle. Or, le problème de la localisation des sources en MEG (et aussi en EEG) est difficile. Il a fallu attendre plus d'une dizaine d'années après l'invention de la MEG pour que des techniques mathématiques, développées pour la plupart dans d'autres contextes, par exemple en traitement d'images, viennent poser les fondements de l'imagerie cérébrale en MEG.

Depuis, l'imagerie cérébrale est devenue un domaine scientifique à part entière. Ce domaine comprend des techniques et méthodes multidisciplinaires dont l'objectif est de mettre en images le cerveau et, ultimement, de le voir au travail. Ces images permettront de développer une meilleure compréhension du fonctionnement du cerveau. Depuis les quinze dernières années, il y a eu des développements majeurs dans ce domaine avec, entre autres, la venue de nouvelles méthodes d'imagerie. Nous nous limiterons ici à deux exemples. La tomographie par émission de positrons (TEP), développée à partir de 1975, consiste à injecter un traceur radioactif dans le sang du sujet. En s'écoulant dans le système sanguin, le sang diffuse le traceur. En se désintégrant, l'élément radioactif du traceur émet un positron qui va s'annihiler avec un électron avoisinant. Cette réaction émet un photon qui est détecté par des instruments extérieurs. Des méthodes de reconstruction tomographique permettent ensuite de trouver l'origine du photon donnant ainsi une image de l'activité sanguine. On observe alors l'activité sanguine qui est signe d'activation cérébrale. Comme deuxième exemple, mentionnons l'imagerie par résonance magnétique fonctionnelle (IRMf). C'est une méthode beaucoup plus récente, développée à partir de 1993, se basant sur l'imagerie par résonance

---

<sup>4</sup>Superconducting QUantum Interference Device

magnétique (IRM) anatomique introduite une dizaine d'années auparavant. Elle utilise un intense champ magnétique pour identifier les tissus qui composent le cerveau à l'aide de leurs différentes propriétés magnétiques. Cette technique permet d'obtenir des images d'une précision impressionnante de l'ordre du millimètre. Un exemple est présenté dans la figure I.1. La variante fonctionnelle consiste à identi-

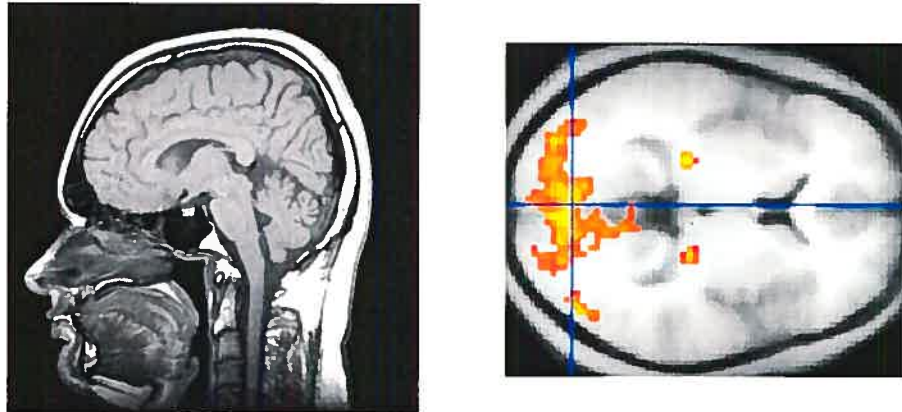


FIG. I.1. Exemple d'images IRM. À gauche, IRM anatomique. À droite, IRM fonctionnelle. Source : [en.wikipedia.org/wiki/{MRI,Functional\\_MRI}](http://en.wikipedia.org/wiki/{MRI,Functional_MRI})

fier les concentrations de la désoxyhémoglobine (protéine paramagnétique). Cette protéine est émise par les neurones en activation permettant à l'IRMf d'identifier les zones du cerveau activées. C'est une signature chimique de l'activité cérébrale. Le point commun entre l'IRMf et la TEP est qu'elles dépendent toutes deux de conséquences métaboliques de l'activité cérébrale (activation sanguine pour la TEP et émission d'une protéine pour l'IRMf). Quoique ces méthodes atteignent des niveaux de résolution spatiale assez élevés à cause du nombre de mesures disponibles, elles ne permettent pas d'obtenir une résolution temporelle supérieure à la seconde. Or, il s'avère que les phénomènes cérébraux se déroulent à une échelle temporelle beaucoup plus petite.

L'EEG et la MEG ne présentent pas ces difficultés. Les champs électromagnétiques détectés sont créés par les courants neuronaux qui sont la véritable signature de l'activité cérébrale. Ces champs sont mesurés à l'extérieur du crâne sans qu'aucune altération des tissus ne soit nécessaire. L'EEG et la MEG sont donc deux méthodes non-invasives. De plus, les champs voyageant à la vitesse de la lumière, on obtient une lecture quasi-instantanée de l'activité cérébrale. La résolution temporelle n'est donc limitée que par la fréquence d'échantillonnage

donnant, un avantage supplémentaire aux méthodes EEG et MEG. Comme on le voit dans la figure I.2 qui compare les résolutions spatiale et temporelle des méthodes d'imagerie cérébrale, l'excellente résolution temporelle de la MEG se fait au détriment de sa résolution spatiale. Quoique leurs résolutions temporelles

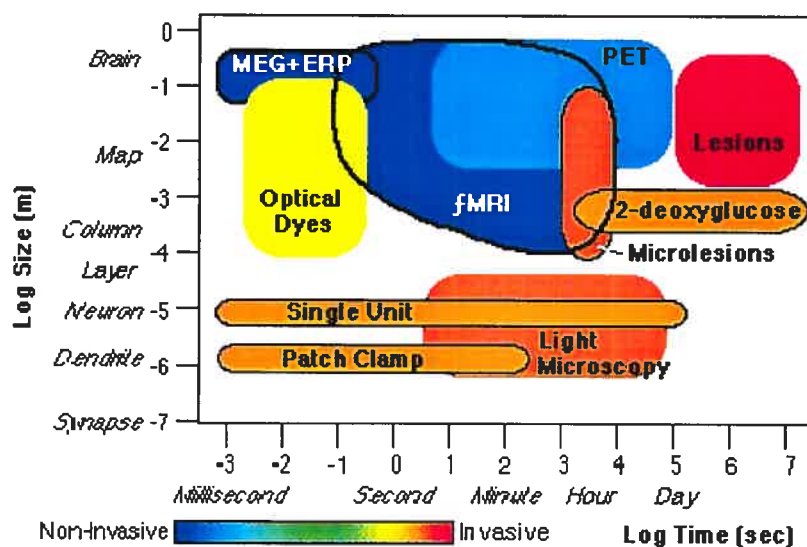


FIG. I.2. Résolution spatiale et temporelle des différentes méthodes d'imagerie cérébrale [10]

soient excellentes, le point faible de la MEG et de l'EEG est la faible quantité de données spatiales (de l'ordre de  $10^2$  comparativement à plus de  $10^5$  pour la TEP et l'IRMf). Ce petit nombre complique le problème de la localisation des sources et suscite d'intenses recherches auxquelles collaborent neurologues, neuropsychologues, physiciens et spécialistes en traitement de signal et d'image. Les travaux présentés ici se situent dans ce contexte multi-disciplinaire.

Dans cette thèse, nous traitons du problème de localisation de sources biomagnétiques sous l'angle du physicien en nous limitant à l'étude de la MEG préférable à l'EEG à cause de sa faible sensibilité aux propriétés conductrices des différents tissus de la tête. Cependant, la MEG ne pourra jamais donner une connaissance de l'activation cérébrale qu'au millimètre près. En effet, pour véritablement connaître les processus cérébraux en jeu, il faudrait pouvoir étudier les phénomènes à l'échelle du neurone. Cette échelle n'est accessible aujourd'hui à aucune technique d'imagerie non-invasive. Cependant, il y a beaucoup à apprendre sur le fonctionnement du cerveau à l'échelle macroscopique du millimètre

et la MEG est une technique idéale pour identifier les processus cognitifs. Plusieurs études psycho-cognitives ont été faites à l'aide de la MEG<sup>5</sup> qui constitue le contexte de cette thèse. Nous développons une méthode d'identification des sources biomagnétiques cérébrales basée sur le principe du maximum d'entropie en moyenne (MEM). Cette méthode est présentée dans le chapitre 3 de façon générale. Puis, l'application du MEM à la MEG est présentée dans les 3 derniers chapitres de cette thèse.

Le chapitre 1 commence par un exposé sur la physiologie des phénomènes cérébraux pour familiariser le lecteur avec les processus à l'origine de l'activité cérébrale. Nous poursuivons avec la modélisation du problème électromagnétique sous-jacent à la localisation de sources en MEG, ce qu'on appelle le problème direct. Le problème direct consiste à trouver le champ magnétique émis par une configuration de sources modélisant l'activité cérébrale. Par opposition, l'identification des sources de courants à partir du champ magnétique mesuré est appelé problème inverse. Pour le physicien, les problèmes inverses sont monnaie courante. En effet, en termes généraux, un problème inverse se présente lorsqu'on a des mesures sur un système dont on veut connaître les propriétés, ce qui est le cas en MEG. La particularité principale de ce problème inverse est qu'il n'admet pas de solution unique. Différentes configurations de courants à l'intérieur de la tête peuvent donner le même champ magnétique à l'extérieur de la tête. Des méthodes spécifiques ont été développées pour identifier une solution unique à l'aide d'information supplémentaire sur la solution.

Dans le chapitre 2, on présente les principales méthodes de résolution du problème inverse ayant en commun l'intégration d'information pour contraindre la solution. Ces méthodes sont variées et ont suscité d'intenses recherches au cours des quinze dernières années.

Dans le chapitre 3, le principe du maximum d'entropie en moyenne (MEM) est présenté. Cette technique de résolution de problèmes inverses prend ses origines dans une interprétation différente de la mécanique statistique qui la considère comme de l'inférence statistique. Cette méthode d'inférence n'utilise nullement les équations du mouvement de la physique. Cela permet alors d'appliquer ce

---

<sup>5</sup>Dans le recueil de l'atelier de formation à la MEG et l'EEG donné à l'Hôpital Pitié-Salpêtrière, Paris, du 4 au 8 juin 2001, on retrouve beaucoup d'exemples de telles études.

principe à des problèmes distincts de la mécanique statistique, comme c'est le cas dans cette thèse.

Les trois chapitres subséquents contiennent les articles qui constituent le travail original de cette thèse. Chacun de ces chapitres est constitué d'un court avant-propos, expliquant le contexte et les liens avec les résultats précédents, puis d'un article (publié ou soumis) qui présente la méthodologie développée et les résultats obtenus.

Dans le chapitre 4, une étude de principe démontre la pertinence de la méthode pour la résolution du problème inverse en MEG à temps fixé. L'ingrédient principal du principe du maximum d'entropie est la loi de probabilité de référence qui sert d'état d'entropie nulle. Ici, cette loi englobe toute l'information, anatomique ou fonctionnelle, que l'on peut avoir sur les courants neuronaux et, donc, sur l'activité cérébrale sous-jacente. Une loi de référence est proposée en faisant l'hypothèse que l'activation est regroupée en régions fonctionnellement cohérentes. La méthode est comparée aux méthodes de reconstruction standard avec des simulations.

Le cinquième chapitre fait un pas de plus vers l'application du MEM à la MEG dans un contexte réaliste. Une méthode de construction des régions d'activation est présentée. Des simulations sont faites sur des cortex réalistes et notre méthode est comparée à d'autres méthodes classiques d'inversion.

Alors que les travaux précédents n'utilisent que l'information spatiale présente dans les données MEG, dans le dernier chapitre, l'aspect temporel est intégré au modèle spatial utilisé jusqu'ici. Ce dernier volet permet d'exploiter l'information temporelle de la MEG. Dans ce cas, la méthode a été appliquée sur des données provenant d'une expérience cognitive réelle. Il est montré que l'information temporelle est indispensable pour résoudre le problème inverse et apporte de l'information pertinente sur l'évolution temporelle des signaux cérébraux.

Notre contribution originale est d'avoir appliqué la méthode de maximisation de l'entropie en moyenne à la MEG en développant un modèle intuitif de sources biomagnétiques. Cette modélisation a permis de produire une méthode générale avec des hypothèses faibles sur la forme de la solution. Nous espérons que ces travaux permettront de sonder les plus fines propriétés fonctionnelles du cerveau.



# Chapitre 1

---

## PHYSIQUE DE LA MAGNÉTOENCÉPHALOGRAPHIE

Dans ce chapitre, les principes physiques de base en MEG sont présentés dans le but d'établir une relation entre les courants neuronaux présents à l'intérieur du crâne et le champ magnétique. Les arguments présentés dans ce chapitre ont été extraits des différentes références données tout au long du texte. Ce problème, dit direct, consiste à calculer le champ magnétique étant donné des sources de courant connues. Puis, on désigne par problème inverse, le fait d'identifier les sources ayant donné un champ magnétique connu. En faisant certaines hypothèses sur la géométrie et les propriétés électromagnétiques des tissus cérébraux, il sera possible de réduire le problème direct à une équation linéaire simplifiant du même coup le problème inverse que nous exposerons dans le prochain chapitre. Auparavant, il faut partir de l'origine de l'activité cérébrale en présentant d'abord le métabolisme qui est à l'origine des courants électriques cérébraux.

### 1.1. LE NEURONE ET L'ACTIVITÉ CÉRÉBRALE

Dans cette section, nous nous penchons sur l'unité de base du traitement de l'information dont est constitué le cerveau, le neurone, pour mettre en évidence l'origine des champs électromagnétiques. Une connaissance de l'anatomie du cerveau a trois implications majeures [17] :

- au niveau cellulaire, les structures déterminent comment les courants électriques neuronaux produisent des courants macroscopiques observables,
- la conductivité électrique est nécessaire pour avoir une solution adéquate pour le problème direct et
- l'anatomie peut être utilisée dans la solution du problème inverse. En particulier, si on suppose que les sources de courants sont confinées à la surface

du cortex, le problème inverse tri-dimensionnel devient essentiellement bi-dimensionnel.

Le neurone est constitué de 4 parties principales (voir figure 1.1) :

- le corps cellulaire qui contient le noyau et le métabolisme cellulaire,
- les dendrites qui reçoivent les stimuli des autres neurones environnants,
- l'axone qui est une longue fibre conductrice qui porte le courant vers les autres neurones et
- l'extrémité synaptique qui contient les terminaisons neuronales et qui émet des signaux pouvant être captés par les autres neurones.

Le neurone est entouré, comme les autres cellules, d'une membrane qui divise la cellule en compartiments intra-cellulaires et extra-cellulaires avec différentes concentrations d'ions chargés. Cette différence est maintenue par des protéines sur la membrane qui pompent certains ions contre le gradient de concentration. Les principaux ions en jeu sont le  $\text{Na}^+$ , le  $\text{K}^+$  et  $\text{Cl}^-$ . C'est le mouvement de ces ions qui crée les courants électriques. Lorsqu'une impulsion électrique arrive par l'axone, des molécules transmettrices (neurotransmetteurs) sont libérées par les synapses dans l'espace libre autour de celles-ci. Les molécules se positionnent ensuite à la surface des dendrites d'un autre neurone. Ces molécules ont pour effet d'ouvrir des canaux par où passeront les ions à travers la membrane créant une différence des charges entre les extrémités de ce neurone générant un courant électrique. C'est précisément ce courant électrique qui engendre le champ magnétique mesuré en MEG.

Les courants qui circulent dans les neurones sont appelés les courants primaires parce qu'ils sont les véritables courants initiaux. Cependant, ce ne sont pas les seuls courants à être présents. En effet, la conservation des charges électriques impose que des courants de conduction (secondaires) soient créés pour rétablir l'équilibre des charges qui se seront déplacées. La densité du courant ainsi créé est de l'ordre de  $100 \text{ nA/mm}^2$ . Il faudra donc l'activation simultanée de près de dix mille neurones pour engendrer un courant qui soit mesurable. Le champ magnétique est dû à l'ensemble de tous ces courants. Ce champ magnétique mesuré à l'extérieur du crâne sera de l'ordre de la centaine de femtoTeslas ( $10^{-13} \text{ T}$ ) qui est un milliard de fois plus petit que le champ magnétique terrestre ( $10^{-4} \text{ T}$ ). Il a fallu attendre le développement de senseurs supraconducteurs à la fin des années 60 pour que les premiers champs magnétiques cérébraux puissent être mesurés [9].

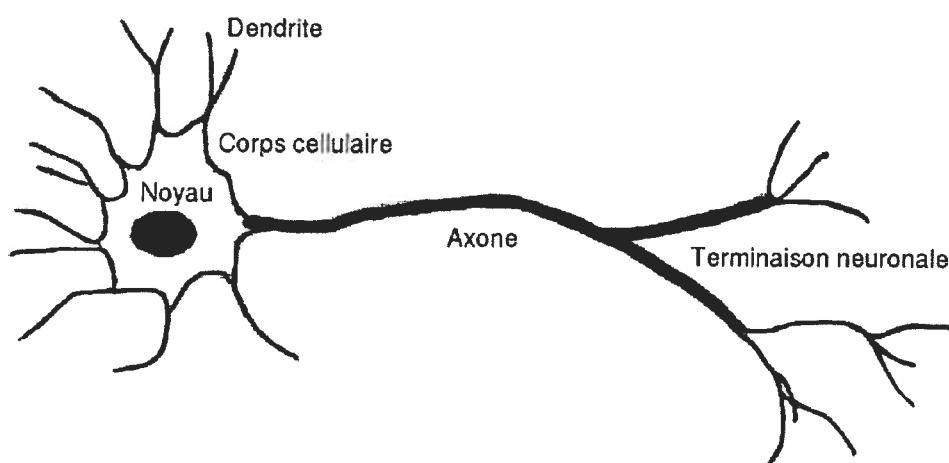


FIG. 1.1. Illustration des différentes parties anatomiques d'un neurone. (Source : [fr.wikipedia.org/wiki/Neurone](http://fr.wikipedia.org/wiki/Neurone))

Étant donné la faible intensité des champs magnétiques, il est nécessaire d'isoler le magnétoencéphalographe du champ magnétique terrestre. La façon la plus efficace est d'utiliser une chambre blindée. Ensuite, pour enlever les champs magnétiques restants, un gradiomètre est utilisé. Il s'agit de deux bobines superposées (ou côte-à-côte) qui mesurent la variation de champ magnétique entre les deux bobines. Ceci permet d'enlever les champs magnétiques qui passent au travers de la chambre blindée, puisque ceux-ci varient lentement dans l'espace. Pour de plus amples détails sur l'instrumentation, nous référons à l'incontournable article de revue de Hämäläinen *et al.* [17].

La MEG a pour but d'aider les scientifiques à étudier le fonctionnement du cerveau. Entre autres, dans le cadre des neurosciences cognitives, il est nécessaire d'identifier l'activité liée au traitement d'un stimulus pour accomplir une tâche. Par contre, cette activité est en général noyée dans de l'activité cérébrale spontanée. La technique des réponses évoquées consiste à répéter une tâche un grand nombre de fois et de moyenniser ces signaux pour isoler le plus possible l'activité électrique liée à la tâche cognitive. Dans le chapitre 6, on utilise des données moyennées sur 400 répétitions. Il n'est pas toujours possible de répéter une expérience des centaines de fois. Lorsqu'on étudie des pathologies, l'épilepsie par exemple, il n'est pas possible de répéter la mesure. On doit alors travailler avec

les signaux bruts. Ces signaux sont aussi bruités par la prise de mesure. Cependant, ce bruit de mesure est pratiquement éliminé lors du moyennage fait pour les réponses évoquées.

## 1.2. CALCUL DU CHAMP MAGNÉTIQUE INDUIT PAR UNE SOURCE DE COURANT

Connaître l'origine des signaux cérébraux en MEG nécessite d'abord de trouver le champ magnétique produit par une source de courant à l'intérieur de la tête. Il s'agit donc de résoudre les équations de Maxwell à l'extérieur de la tête sachant qu'il y a des courants électriques  $\mathbf{J}$  à l'intérieur [17, 42].

Pour simplifier le problème, quelques approximations sont possibles. Tout d'abord, la perméabilité magnétique des tissus cérébraux est la même que celle du vide ( $\mu = \mu_0$ ). Cette hypothèse est acceptable pour la plupart des matériaux. Les équations de Maxwell dans ce cas sont :

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon}, \quad (1.1a)$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}, \quad (1.1b)$$

$$\nabla \cdot \mathbf{B} = 0, \quad (1.1c)$$

$$\nabla \times \mathbf{B} = \mu_0 \left( \mathbf{J} + \epsilon \frac{\partial \mathbf{E}}{\partial t} \right). \quad (1.1d)$$

Les termes en dérivée par rapport au temps peuvent être négligés puisque le spectre des fréquences des signaux électrophysiologiques en MEG est typiquement de l'ordre de 100 Hz. C'est ce qui est appelé l'approximation quasi-statique. Pour accepter cette hypothèse, faisons les constatations suivantes. Le courant  $\mathbf{J}$  est la somme des courants de conduction et de polarisation :

$$\mathbf{J} = \sigma \mathbf{E} + \frac{\partial \mathbf{P}}{\partial t}, \quad (1.2)$$

où  $\mathbf{P} = (\epsilon - \epsilon_0)\mathbf{E}$  est la polarisation et  $\epsilon$  la permittivité des tissus du cerveau. Supposons que  $\sigma$  et  $\epsilon$  sont des constantes et considérons une onde à la fréquence  $f$  :

$$\mathbf{E} = \mathbf{E}_0(\mathbf{r})e^{2\pi i f t}.$$

Avec les équations (1.1d) et (1.2),

$$\nabla \times \mathbf{B} = \mu_0 \left[ \sigma \mathbf{E} + (\epsilon - \epsilon_0) \frac{\partial \mathbf{E}}{\partial t} + \epsilon \frac{\partial \mathbf{E}}{\partial t} \right]. \quad (1.3)$$

On pourra négliger le terme en dérivée par rapport au temps si  $|\epsilon \partial \mathbf{E} / \partial t| \ll |\sigma \mathbf{E}|$  ou bien  $2\pi f \epsilon \ll \sigma$ . En mettant les valeurs numériques valides [17] pour les tissus cérébraux :  $\sigma = 0.3 \Omega^{-1} \text{m}^{-1}$ ,  $\epsilon = 10^5 \epsilon_0$  et  $f = 100 \text{ Hz}$ , on trouve  $2\pi f \epsilon / \sigma = 2 \times 10^{-3} \ll 1$ . Il est donc justifié de négliger  $\partial \mathbf{E} / \partial t$  dans les équations.

L'autre terme qui dépend de la dérivée par rapport au temps est  $\partial \mathbf{B} / \partial t$ . Il doit être négligeable devant  $\mathbf{E}$ . À l'aide des équations (1.1b) et (1.3), on peut écrire :

$$\begin{aligned} \nabla \times \nabla \times \mathbf{E} &= -\frac{\partial}{\partial t} \nabla \times \mathbf{B}, \\ &= -\mu_0 \frac{\partial}{\partial t} \left( \sigma \mathbf{E} + \epsilon \frac{\partial \mathbf{E}}{\partial t} \right), \\ &= -2\pi i f \mu_0 (\sigma + 2\pi i f \epsilon) \mathbf{E}. \end{aligned}$$

Les solutions à cette dernière équation ont une longueur d'onde caractéristique

$$\lambda_c = |2\pi f \mu_0 \sigma (1 + 2\pi i f \epsilon / \sigma)|^{-1/2}.$$

Avec les valeurs numériques appropriées,  $\lambda_c = 65 \text{ m}$  ce qui est évidemment beaucoup plus grand que les dimensions de la tête d'un humain normal. Donc, la contribution de  $\partial \mathbf{B} / \partial t$  à  $\mathbf{E}$  est très petite et nous pouvons aussi la négliger.

Ainsi, dans le régime quasi-statique, les équations à résoudre sont

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon}, \quad (1.4a)$$

$$\nabla \times \mathbf{E} = 0, \quad (1.4b)$$

$$\nabla \cdot \mathbf{B} = 0, \quad (1.4c)$$

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{J}. \quad (1.4d)$$

où le courant électrique  $\mathbf{J}$  est composé de deux sources :

$$\mathbf{J} = \mathbf{J}^p + \sigma \mathbf{E}, \quad (1.5)$$

avec

- $\mathbf{J}^p$  : courant engendré par le mouvement des charges électriques (courant primaire ou intracellulaire) et

–  $\sigma \mathbf{E}$  : courant de conduction (courant secondaire, extra-cellulaire ou volumique) et  $\sigma = \sigma(\mathbf{r})$ , la conductivité du milieu.

De l'équation (1.4b),  $\mathbf{E}$  provient d'un potentiel électrique  $V$  tel que  $\mathbf{E} = -\nabla V$ . L'équation (1.4d) est la loi de Biot-Savart, le champ magnétique  $\mathbf{B}$  est alors donné par :

$$\mathbf{B}(\mathbf{r}) = \frac{\mu_0}{4\pi} \int_{\Omega} [\mathbf{J}^i(\mathbf{r}') - \sigma(\mathbf{r}') \nabla' V(\mathbf{r}')] \times \frac{\mathbf{r} - \mathbf{r}'}{|\mathbf{r} - \mathbf{r}'|^3} dv' \quad (1.6)$$

où l'intégrale est faite sur tout le volume  $\Omega$  qui contient tous les courants (primaire et secondaire).

Pour résoudre ces équations, il faut tout d'abord définir la conductivité ( $\sigma(\mathbf{r})$ ) de la tête ici, représentée par le volume  $\Omega$ . La tête est faite essentiellement de quatre couches. En commençant par le centre, elle est constituée de matière cérébrale (le cerveau), ensuite de liquide céphalo-rachidien, puis il y a le crâne (l'os) et la peau. Chacune de ces couches a une conductivité différente. Il est généralement supposé que la conductivité des tissus cérébraux est constante et isotrope [15]. Cependant, cette hypothèse n'est pas justifiée pour la surface corticale. En effet, la structure en colonnes des neurones<sup>1</sup> la caractérisant rend la conduction dix fois plus grande perpendiculairement à la surface par rapport à parallèlement à celle-ci. Des modèles ont donc été développés pour s'affranchir de l'hypothèse d'isotropie des tissus. Dans [35], on offre une revue des différentes méthodes qui ont été proposées. Nous allons nous limiter au cas du conducteur homogène par morceaux, une hypothèse largement utilisée dans le domaine.

Supposons que le conducteur  $\Omega$  est constitué de  $N$  parties  $\Omega_j$  qui ont chacune une conductivité  $\sigma_j$  constante. La loi de Biot-Savart s'écrit, à partir de (1.6),

$$\mathbf{B}(\mathbf{r}) = \mathbf{B}_0(\mathbf{r}) - \frac{\mu_0}{4\pi} \sum_{j=1}^N \sigma_j \int_{\Omega_j} \nabla' V(\mathbf{r}') \times \frac{\mathbf{r} - \mathbf{r}'}{|\mathbf{r} - \mathbf{r}'|^3} dv' \quad (1.7)$$

où

$$\mathbf{B}_0(\mathbf{r}) = \frac{\mu_0}{4\pi} \int_{\Omega} \mathbf{J}^i(\mathbf{r}') \times \frac{\mathbf{r} - \mathbf{r}'}{|\mathbf{r} - \mathbf{r}'|^3} dv'.$$

Par le théorème de Stokes et l'identité vectorielle

$$\nabla' \times [V(\mathbf{r}') \times \nabla' (|\mathbf{r} - \mathbf{r}'|^{-1})] = \nabla' V \times \frac{\mathbf{r} - \mathbf{r}'}{|\mathbf{r} - \mathbf{r}'|^3},$$

<sup>1</sup>voir la section 2.2

les intégrales dans l'équation (1.7) deviennent

$$\begin{aligned} \int_{\Omega_j} \nabla' V(\mathbf{r}') \times \frac{\mathbf{r} - \mathbf{r}'}{|\mathbf{r} - \mathbf{r}'|^3} dv' &= \int_{\Omega_j} \nabla' \times [V(\mathbf{r}') \times \nabla' (|\mathbf{r} - \mathbf{r}'|^{-1})] dv', \\ &= \int_{S_j} V(\mathbf{r}') \mathbf{n}(\mathbf{r}') \times \frac{\mathbf{r} - \mathbf{r}'}{|\mathbf{r} - \mathbf{r}'|^3} dS'_j. \end{aligned}$$

La surface  $S_j$  est la frontière de  $\Omega_j$ ,  $dS'_j$  est l'élément de surface sur  $S_j$  et  $\mathbf{n}(\mathbf{r}')$  est la normale extérieure à la surface  $S_j$  au point  $\mathbf{r}'$ . Puisque les normales à l'interface de deux milieux de conduction différentes sont de signes opposés, l'équation (1.7) devient

$$\mathbf{B}(\mathbf{r}) = \mathbf{B}_0(\mathbf{r}) - \frac{\mu_0}{4\pi} \sum_{j=1}^N (\sigma'_j - \sigma''_j) \int_{S_j} V(\mathbf{r}') \mathbf{n}(\mathbf{r}') \times \frac{\mathbf{r} - \mathbf{r}'}{|\mathbf{r} - \mathbf{r}'|^3} dS'_j \quad (1.8)$$

où  $\sigma'_j$  et  $\sigma''_j$  sont les conductivités des régions intérieure et extérieure de  $S_j$ . Comme on peut le voir dans l'équation précédente, le champ magnétique dépend du potentiel  $V(\mathbf{r})$  sur chacune des surfaces. Or, pour connaître le potentiel, cela nécessite la résolution d'une équation intégrale qui n'a de solutions analytiques que pour certaines géométries [35] (sphérique, par exemple). Sinon, on doit la résoudre numériquement pour trouver  $\mathbf{B}$ . L'approximation généralement acceptée est de modéliser la tête comme étant composée de plusieurs sphères concentriques. Plusieurs travaux ont été développés pour s'affranchir de cette hypothèse [4, 15, 35]. Il a été montré qu'en MEG, cette approximation était valide du point de vue de la localisation des sources [4]. Cela est dû au fait que la MEG n'est pas très sensible à la différence de conductivité entre les tissus.

Donc, considérons le cas d'un conducteur constitué de plusieurs sphères concentriques. La composante radiale de  $\mathbf{B}$  se simplifie alors grandement puisque le deuxième terme de (1.8) s'annule complètement. En détails,

$$B_r(\mathbf{r}) = \mathbf{B}(\mathbf{r}) \cdot \mathbf{e}_r,$$

$$B_r(\mathbf{r}) = \mathbf{B}_0(\mathbf{r}) \cdot \mathbf{e}_r - \frac{\mu_0}{4\pi} \sum_{j=1}^N (\sigma'_j - \sigma''_j) \int_{S_j} V(\mathbf{r}') \left( \mathbf{n}(\mathbf{r}') \times \frac{\mathbf{r} - \mathbf{r}'}{|\mathbf{r} - \mathbf{r}'|^3} \right) \cdot \mathbf{e}_r dS'_j.$$

Les termes de la somme s'annulent parce que  $\mathbf{n}(\mathbf{r}') = \frac{\mathbf{r}'}{r'}$  et  $\mathbf{e}_r = \frac{\mathbf{r}}{r}$  et donc

$$B_r(\mathbf{r}) = \mathbf{B}_0(\mathbf{r}) \cdot \mathbf{e}_r = \frac{\mu_0}{4\pi} \int_{\Omega} \left( \mathbf{J}^i(\mathbf{r}') \times \frac{\mathbf{r} - \mathbf{r}'}{|\mathbf{r} - \mathbf{r}'|^3} \right) \cdot \mathbf{e}_r dv'. \quad (1.9)$$

Ainsi, la composante radiale de  $\mathbf{B}(\mathbf{r})$  est indépendante du profil  $\sigma(\mathbf{r})$  de la conductivité. Ce résultat est valide en tout point de l'espace. Or, les autres composantes de  $\mathbf{B}(\mathbf{r})$  peuvent être récupérées lorsque l'on mesure le champ magnétique hors du milieu conducteur, ce qui est le cas en MEG.

Si  $\mathbf{r}$  est à l'extérieur de  $\Omega$ , il n'y a aucun courant alors l'équation (1.4d) donne  $\nabla \times \mathbf{B} = 0$  et  $\mathbf{B}$  peut s'écrire en terme d'un potentiel magnétique :  $\mathbf{B}(\mathbf{r}) = -\nabla U(\mathbf{r})$ . Puisque  $U(\mathbf{r})$  doit s'annuler à l'infini, il peut s'écrire de la façon suivante :

$$\begin{aligned}
 U(\mathbf{r}) &= - \int_0^\infty \nabla U(\mathbf{r} + t\mathbf{e}_r) \cdot \mathbf{e}_r dt, \\
 &= \int_0^\infty \mathbf{B}_0(\mathbf{r} + t\mathbf{e}_r) \cdot \mathbf{e}_r dt. \quad \text{Avec (1.9) on a} \\
 &= \frac{\mu_0}{4\pi} \int_\Omega dv' \left( \mathbf{J}^p(\mathbf{r}') \times \int_0^\infty \frac{\mathbf{r} + t\mathbf{e}_r - \mathbf{r}'}{|\mathbf{r} + t\mathbf{e}_r - \mathbf{r}'|^3} dt \right) \cdot \mathbf{e}_r, \\
 &= -\frac{\mu_0}{4\pi} \int_\Omega dv' (\mathbf{J}^p(\mathbf{r}') \times \mathbf{r}') \cdot \mathbf{e}_r \int_0^\infty \frac{dt}{|\mathbf{r} + t\mathbf{e}_r - \mathbf{r}'|^3}, \\
 U(\mathbf{r}) &= -\frac{\mu_0}{4\pi} \int_\Omega dv' \frac{(\mathbf{J}^p(\mathbf{r}') \times \mathbf{r}') \cdot \mathbf{r}}{ra(a + \mathbf{a} \cdot \mathbf{e}_r)} \tag{1.10}
 \end{aligned}$$

où  $\mathbf{a} = \mathbf{a}(\mathbf{r}, \mathbf{r}') = \mathbf{r} - \mathbf{r}'$  et  $r$  et  $a$  la norme des vecteurs  $\mathbf{r}$  et  $\mathbf{a}$  respectivement. Ayant une expression pour  $U(\mathbf{r})$ , on peut en prendre le gradient pour trouver  $\mathbf{B}$ .

Jusqu'à maintenant, nous avons fait plusieurs hypothèses sur les conducteurs. Cependant, il est difficile d'aller plus loin sans faire de même pour les sources de courant  $\mathbf{J}^p$ .

### 1.3. LE DIPÔLE DE COURANT

On peut faire certaines hypothèses sur les sources de courant en observant les processus impliqués. Comme expliqué au début de ce chapitre, les neurones sont à l'origine des courants électriques cérébraux. Pour la MEG, les neurones situés sur la surface corticale constituent les sources de courant qui engendrent des signaux mesurables à l'extérieur du crâne. En effet, le champ magnétique décroît très rapidement avec la distance. De plus, bien que tous les courants contribuent aux champs électromagnétiques, des structures spatiales spécifiques doivent être présentes pour que la superposition des courants donne naissance à des signaux observables.



La surface corticale est composée de quelques dix milliards de neurones. Comment dégager de l'ordre et de l'information structurale devant cette complexité de neurones ? Il s'avère qu'en observant une coupe transversale de la surface corticale, on voit que les neurones sont organisés perpendiculairement à la surface. En plus de cette organisation spatiale, les neurones ont une activation synchrone puisqu'ils génèrent des champs magnétiques mesurables [35]. Cette cohérence spatio-temporelle permet de modéliser les sources de courant par des dipôles de courant. En effet, les observations de courants neuronaux suggèrent que la densité de courant typique de l'activité cérébrale est de  $100 \text{ nA/mm}^2$  [4]. Si la couche corticale est épaisse de 4 mm et que l'on considère que le dipôle modélise une surface de  $5 \text{ mm} \times 5 \text{ mm}$ , on arrive à des intensités de dipôles de courant de l'ordre de  $10 \text{ nA}\cdot\text{m}$  ce qui concorde avec les ordres de grandeur estimés à l'aide des champs magnétiques [17].

Cette modélisation de l'activité cérébrale par des dipôles de courant n'est justifiée que pour les sources corticales. Pour l'EEG, les sources profondes se manifestent dans le potentiel. Cependant, celles-ci ne permettent pas de mettre en évidence une cohérence spatiale et sont plutôt modélisées par des sources multipolaires. Sauf pour quelques travaux sur les distributions de courants multipolaires [36], le dipôle de courant est la matière première de toutes les méthodes de reconstruction. La modélisation par des dipôles de courants est la plus utilisée même si pour des sources profondes, cette modélisation n'est pas parfaite. Des études ont démontré que cette approximation est tout de même valable [12].

Un dipôle de courant est donc une ligne de courant infinitésimale à la position  $\mathbf{r}_0$  dont la densité de courant est exprimée comme

$$\mathbf{J}^p(\mathbf{r}) = \mathbf{Q} \delta(\mathbf{r} - \mathbf{r}_0) \quad (1.11)$$

où  $\mathbf{Q}$  contient l'intensité (unités d'Ampère-mètre) et l'orientation du dipôle et  $\mathbf{r}_0$  est la position du dipôle. Ce dipôle  $\mathbf{Q}$  peut être vu comme une ligne de courant qui part d'un point  $\mathbf{r}_1$  et arrête subitement à un autre point  $\mathbf{r}_2$  :  $\mathbf{Q} = \int \mathbf{J}^p dv'$ . La fonction  $\delta$  donne l'étendue de la source, c'est-à-dire, un point. Ainsi, l'expression (1.10) pour le potentiel magnétique  $U$  se réduit à

$$U(\mathbf{r}) = -\frac{\mu_0}{4\pi} \frac{(\mathbf{Q} \times \mathbf{r}_0) \cdot \mathbf{r}}{F} \quad (1.12)$$

avec  $\mathbf{a} = \mathbf{r} - \mathbf{r}_0$  et  $F = a(ra + r^2 - \mathbf{r}_0 \cdot \mathbf{r})$ . De cette expression, on déduit le champ magnétique

$$\mathbf{B}(\mathbf{r}) = -\nabla U(\mathbf{r}) = \frac{\mu_0}{4\pi F^2} \left( F \mathbf{Q} \times \mathbf{r}_0 - ((\mathbf{Q} \times \mathbf{r}_0) \cdot \mathbf{r}) \nabla F \right) \quad (1.13)$$

avec

$$\nabla F = \left( \frac{a^2}{r} + \frac{\mathbf{a} \cdot \mathbf{r}}{a} + 2a + 2r \right) \mathbf{r} - \left( a + 2r + \frac{\mathbf{a} \cdot \mathbf{r}}{a} \right) \mathbf{r}_0,$$

ce qui donne l'expression du champ magnétique en tout point à l'extérieur du volume conducteur. La figure 1.2 contient une illustration des différents vecteurs contenus dans l'expression pour le champ magnétique généré par un dipôle.

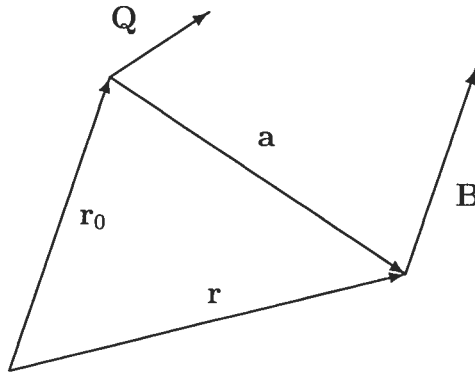


FIG. 1.2. Schéma montrant les vecteurs de l'expression pour le champ magnétique  $\mathbf{B}$  généré par un dipôle de courant  $\mathbf{Q}$  situé à  $\mathbf{r}_0$  à un point d'observation  $\mathbf{r}$  (voir équation (1.13)).

L'équation (1.13) donne lieu à plusieurs remarques importantes. Tout d'abord, le champ magnétique  $\mathbf{B}$  est indépendant de la conductivité des tissus cérébraux dans le cas d'une géométrie sphérique. Nous n'avons alors pas besoin de connaître la valeur des conductivités pour chaque conducteur de  $\Omega_j$ . Cette propriété rend la MEG très intéressante : n'étant pas sensible aux courants secondaires, elle nous permet de détecter directement les courants primaires, vraies signatures de l'activité cérébrale. Ensuite,  $\mathbf{B}$  dépend linéairement de l'intensité  $\mathbf{Q}$  du dipôle, ce qui n'est pas une surprise puisque, par la loi de Biot-Savart,  $\mathbf{B}$  dépend linéairement de l'intensité du courant. Par contre,  $\mathbf{B}$  dépend non-linéairement de la position  $\mathbf{r}_0$ . Donc, une petite variation de  $\mathbf{r}_0$  peut entraîner de grandes variations dans la valeur du champ magnétique, rendant cette expression sensible aux erreurs de position des dipôles. Ensuite, si  $\mathbf{Q} = Q\mathbf{e}_r = Q\frac{\mathbf{r}_0}{r_0}$  (un dipôle radial), le champ magnétique est nul. La détermination des courants à partir de la valeur de  $\mathbf{B}$  à

l'extérieur (même en un nombre infini de points) est donc impossible. En particulier, la composante radiale de  $\mathbf{Q}$  ne génère pas de champ magnétique. Ceci n'est qu'un cas particulier d'un résultat plus général d'Helmholtz qui, dès 1853 [48], démontrait que différentes configurations de courants peuvent engendrer le même champ magnétique.

C'est le sujet de cette thèse : comment retrouver la configuration de courants qui a donné le champ magnétique mesuré par les capteurs à l'extérieur de la tête. Tel est l'énoncé du problème inverse de la MEG. Ce problème est difficile puisqu'il n'admet pas de solution unique par les arguments mentionnés plus tôt. Il faut alors des méthodes qui permettent de discriminer certaines solutions pour pouvoir n'en choisir qu'une seule. Cette discrimination se fait en intégrant de l'information pertinente sur l'activation cérébrale. Différentes approches ont été développées pour ce faire et elles sont discutées dans le prochain chapitre. Ensuite, dans les quatre derniers chapitres, le principe du maximum d'entropie en moyenne est présenté et appliqué au problème inverse en MEG ce qui constitue le travail original dans cette thèse.

## Chapitre 2

---

### PROBLÈME INVERSE EN MAGNÉTOENCÉPHALOGRAPHIE

Le problème inverse, que l'on cherche à résoudre ici, consiste à trouver les sources de courants qui ont généré le champ magnétique observé à l'extérieur de la tête. Ce problème n'admet pas de solution unique en vertu des lois de la physique. Il est dit mal posé. En plus, la MEG ne dispose que d'un nombre limité, parfois très faible, de mesures spatiales. Les appareils magnétoencéphalographiques fabriqués de nos jours ne disposent que d'une centaine de capteurs distribués autour du scalp. Les autres méthodes d'imagerie, comme la TEP ou l'IRMf, disposent de cent à mille fois plus de données. Tout de même, nous verrons qu'il est possible d'extraire beaucoup avec cette information spatiale limitée. Dans cet exposé, nous nous limiterons aux méthodes de reconstruction à temps fixé puisque la faiblesse principale de la MEG est son indétermination spatiale. La reconstruction temporelle sera abordée dans le chapitre 6.

Les problèmes inverses mal posés ne se retrouvent pas seulement en MEG. En fait, le travail du physicien est de résoudre ce genre de problème. Aussitôt que des données sont prises sur un système, il utilise les lois de la physique et des techniques mathématiques et statistiques pour connaître les propriétés de ce système et faire des prédictions sur son comportement.

Il existe présentement deux approches pour résoudre le problème inverse en MEG. La première, le modèle dipolaire, suppose que l'activation peut être expliquée par un ensemble relativement petit de dipôles dont il faut estimer, à partir du champ magnétique mesuré, les paramètres les décrivant chacun. L'autre approche, le modèle distribué, consiste à distribuer à l'intérieur du volume conducteur un grand nombre de dipôles dont la position et l'orientation sont prédéterminés ne laissant que l'intensité des dipôles à estimer. L'approximation dipolaire est ici utilisée pour modéliser toutes les sources même si elle est moins justifiée pour

des sources profondes. Ces deux approches sont brièvement discutées dans les prochaines sections.

Dans cette thèse, nous utiliserons le modèle distribué. Ceci permettra de construire un modèle de source qui comprend de l'information pertinente, information nécessaire pour la résolution du problème inverse.

## 2.1. MODÈLES DIPOLAIRES

Ce sont actuellement les méthodes les plus employées pour les usages cliniques de la MEG. Elles offrent l'avantage d'être rapides et l'interprétation de leurs résultats est directe. Elles considèrent que l'activité électrique cérébrale est concentrée en un petit nombre d'aires dont la dimension est petite par rapport à leur distance aux capteurs. L'approximation des sources de courant avec peu de dipôles distincts est ainsi justifiée. Bien que ces hypothèses schématisent le portrait véritable de l'activité cérébrale, il a été démontré que les résultats étaient en général compatibles avec les connaissances sur l'anatomie et la physiologie du cerveau [15]. La résolution de ces méthodes appliquées à la MEG est en général de quelques millimètres.

Dans l'équation (1.11), il y a 6 paramètres qui définissent un dipôle : la position ( $\mathbf{r}_0$ ), l'orientation ( $\hat{\mathbf{Q}}$ ) et l'intensité ( $Q$ ). S'il y a  $N$  dipôles dans le modèle, cela donne  $6N$  paramètres à estimer dont  $3N$  (pour la position) qui ne dépendent pas linéairement des données.

La difficulté avec ces méthodes est que l'on doit savoir *a priori* le nombre de sources dans le modèle. On pourrait avoir tendance à mettre un grand nombre de sources mais le problème devient ambigu lorsque le nombre de paramètres à estimer est plus grand que le nombre de mesures. De plus, l'estimation des paramètres non-linéaires est très sensible au bruit présent dans les données.

Pour contourner ce problème, des méthodes comme MUSIC (Multiple Signal Classification) pratiquent un balayage spatial [37] qui consiste à passer tout le volume du cortex (voxel par voxel) pour trouver la probabilité de présence d'une source. Cette probabilité est exprimée en trouvant la contribution à l'espace signal de chaque source supposée. Le principal avantage de cette méthode est que chaque source est trouvée individuellement plutôt que toutes les sources à la fois [2].

Ces méthodes ont tous le même désavantage d'avoir besoin du nombre de sources *a priori* et l'utilisateur doit s'assurer que le nombre de paramètres à estimer reste plus petit que le nombre de mesures.

## 2.2. MODÈLES DISTRIBUÉS

Ces méthodes contrastent avec les méthodes dipolaires en considérant un grand nombre de dipôles disposés uniformément dans une partie ou la totalité du volume cérébral. De plus, comme les méthodes distribuées ne font aucune hypothèse sur le nombre de sources, cette technique est appropriée lorsqu'il y a des sources multiples. Cependant, le nombre de variables dans le problème devient plus grand que le nombre de mesures. Lors des premiers essais au milieu des années 1980, l'espace des sources était constitué d'un plan parallèle aux capteurs [18] et on minimisait l'énergie de la solution. Cette méthode ne permet pas de rendre compte des sources focales et donne toujours une solution très lissée. On a alors dû se tourner vers des méthodes de régularisation. De telles méthodes ont été développées en traitement de signal et d'images pour, par exemple, débiter des signaux d'où le nom de régularisation. Clarke *et al* furent les premiers à introduire ce concept en MEG [8]. Ils modélisent chacune des sources par une fonction continue. Cependant, ils s'efforçaient de ne pas introduire d'information anatomique ou fonctionnelle.

Dale et Sereno émirent ensuite l'idée que l'espace source pouvait être contraint en apportant de l'information anatomique provenant de l'IRM. Dans [11], ils argumentent que, comme les neurones du cortex sont organisés en colonnes, les dipôles résultant sont perpendiculaires au cortex. Donc, si la forme de la surface corticale est connue, la position et l'orientation des sources peuvent être contraintes en divisant la surface en parcelles qui sont suffisamment petites pour justifier l'utilisation d'un dipôle. Le problème est donc réduit à estimer l'intensité de chaque dipôle. L'image IRM est utilisée pour connaître la forme du cortex. Cette étape peut sembler laborieuse puisqu'on doit obtenir une image volumique IRM pour chaque sujet. Or, cette image IRM est acquise de façon routinière et est maintenant presque toujours incluse avec les données MEG.

Avec ces hypothèses, essayons de revoir les équations développées dans le chapitre 1 pour pouvoir en dégager une relation simple. Regardons ce que devient

l'équation (1.13). Plaçons  $N$  dipôles sur la surface corticale. Le champ  $\mathbf{B}$  est donc

$$\mathbf{B}(\mathbf{r}) = \sum_{i=1}^N \frac{\mu_0}{4\pi F_i^2} (F_i \mathbf{Q}^i \times \mathbf{r}_0^i - ((\mathbf{Q}^i \times \mathbf{r}_0^i) \cdot \mathbf{r}) \nabla F_i)$$

où  $\mathbf{r}_0^i$  est la position du dipôle  $i$  et  $F_i$  est l'expression pour  $F$  définie plus tôt mais avec  $\mathbf{r}_0 = \mathbf{r}_0^i$ . Or, la position  $\mathbf{r}_0^i$  et l'orientation  $\hat{\mathbf{Q}}^i$  pouvant être fixées à l'aide d'une image IRM, il ne reste plus que l'intensité  $Q^i$  de chaque dipôle à reconstruire. Comme le champ magnétique  $\mathbf{B}$  dépend linéairement de  $Q^i$ , on peut alors reformuler le problème comme

$$\mathbf{b} = G \mathbf{q}. \quad (2.1)$$

Le vecteur  $\mathbf{b}$  contient les valeurs des champs magnétiques mesurés sur les  $d$  capteurs de l'appareil magnétoencéphalographique,  $\mathbf{q} = [Q^1, Q^2, \dots, Q^N]^t$  est le vecteur des intensités et  $G$  est la matrice  $d \times N$  définie à l'aide de l'équation (1.13) dont les éléments sont

$$G_{ji} = \frac{\mu_0}{4\pi F_{ji}^2} \left( F_{ji} \hat{\mathbf{Q}}^i \times \mathbf{r}_0^i - ((\hat{\mathbf{Q}}^i \times \mathbf{r}_0^i) \cdot \mathbf{r}_j) \nabla F_{ji} \right).$$

Le facteur  $F_{ji} = a_{ji}(r_j a_{ji} + r_j^2 - \mathbf{r}_0^i \cdot \mathbf{r}_j)$ ,  $a_{ji} = r_j - r_0^i$  avec  $\mathbf{r}_0^i$ ,  $\hat{\mathbf{Q}}^i$  est la position et l'orientation du dipôle  $i$  et  $\mathbf{r}_j$  est la position du capteur  $j$ . On remarque que la dimension de  $\mathbf{b}$  est la même que le nombre de capteurs. Or, le champ magnétique est un vecteur et, en principe, on devrait pouvoir mesurer les trois composantes de  $\mathbf{B}$  en chaque point de mesure. Il s'avère que les appareils magnétoencéphalographiques ne mesurent qu'une composante du champ magnétique, en général la composante radiale, en utilisant des gradiomètres pour enlever les champs magnétiques parasites persistant à l'intérieur de la chambre blindée.

Le problème est donc ramené à celui de l'inversion de la matrice  $G$ . Si celle-ci était inversible, la solution serait triviale. Cependant, ce problème linéaire fait partie des problèmes inverses mal posés pour les raisons suivantes. Comme mentionné dans la section 1.3, le problème électromagnétique est lui-même sous-déterminé puisque plusieurs configurations de sources peuvent donner le même champ magnétique à l'extérieur du volume conducteur [48]. De plus, le nombre de dipôles équivalents nécessaires pour couvrir adéquatement la surface corticale est normalement très élevé, de l'ordre de plusieurs milliers, soit au moins un ordre de grandeur plus grand que le nombre de capteurs dans les magnétoencéphalographes. Ainsi, le problème linéaire (2.1) est-il sous-déterminé. De plus, les mesures  $\mathbf{b}$  sont

en général bruitées. Il est donc impossible de trouver une solution parfaite si les données ne le sont pas. Enfin, la solution n'est pas unique du point de vue des lois de la physique. En effet, comme nous en avons discuté dans le chapitre précédent, il y a des configurations de sources qui ne donnent aucun champ magnétique à l'extérieur du crâne. Cet aspect est réflété dans le modèle de sources que nous avons choisi puisqu'un dipôle radial  $\mathbf{Q} \propto \mathbf{r}_0$  n'émet aucun champ magnétique. Ainsi, l'espace nul de la matrice  $G$  n'est, en général, pas trivial.

Le problème de reconstruction de sources tel que présenté ici offre de grandes similitudes avec la reconstruction et la restauration d'images. Différentes méthodes ont été développées pour résoudre ces problèmes [13]. Dans les prochaines sous-sections, nous allons décrire succinctement les principales méthodes utilisées jusqu'à maintenant.

### 2.2.1. Méthode de norme minimale

La méthode dite de norme minimale pose un principe que toutes les méthodes de reconstruction doivent respecter : l'attache aux données. En effet, c'est la seule information fonctionnelle disponible. Comme les données sont en général bruitées et que le modèle de sources distribuées qu'on impose ne correspond pas nécessairement à la réalité du point de vue physiologique, il se peut que  $\mathbf{b}$  ne soit pas dans l'image de  $G$ . Par contre, on peut toujours chercher à minimiser l'écart aux données, c'est-à-dire, trouver un  $\hat{\mathbf{q}}$  pour lequel la différence entre le champ magnétique donné par cette source ( $G\hat{\mathbf{q}}$ ) et les données  $\mathbf{b}$  est minimal :

$$\hat{\mathbf{q}} = \underset{\mathbf{q}}{\operatorname{argmin}} \|\mathbf{b} - G\mathbf{q}\|^2. \quad (2.2)$$

La solution générale à ce problème est  $\hat{\mathbf{q}} = G^+\mathbf{b}$  où

$$G^+ = \lim_{\lambda \rightarrow 0} (G^t G + \lambda \mathbb{I})^{-1} G^t \quad (2.3)$$

est l'inverse de Moore-Penrose. L'avantage de cette solution est qu'elle est analytique et les méthodes numériques de manipulation de matrices sont bien connues et très rapides pour les ordres de grandeur (de 100 à 1000) qui nous préoccupent. Cependant, elle est très sensible au bruit contenu dans  $\mathbf{b}$ . En effet, la matrice  $G^t G$  contient typiquement des valeurs propres très près de zéro<sup>1</sup>, ce qui amplifie

<sup>1</sup>Par exemple, pour l'espace des sources utilisé dans le chapitre 5, le rapport entre la plus grande et la plus petite valeur propre est de l'ordre de  $10^8$ .



énormément les effets du bruit lors de l'inversion. De plus, la solution n'est en général pas en accord avec les connaissances physiologiques que nous avons sur les processus neuronaux. Des discontinuités inacceptables sont typiquement présentes à cause de la sensibilité au bruit. Il a donc été proposé de régulariser la solution en incluant un terme de Tikhonov [46] :

$$\hat{\mathbf{q}} = \underset{\mathbf{q}}{\operatorname{argmin}} \|\mathbf{b} - G\mathbf{q}\|^2 + \lambda\|\mathbf{q}\|^2. \quad (2.4)$$

Cette dernière expression garantit que de grandes intensités sont d'autant plus pénalisées que le paramètre de régularisation  $\lambda$  est grand. En effet, la solution de 2.4,

$$\hat{\mathbf{q}} = (G^t G + \lambda\mathbb{I})^{-1} G^t \mathbf{b}$$

fait intervenir le paramètre  $\lambda$  qui atténue l'effet des petites valeurs propres de  $G^t G$ . Cette régularisation permet de trouver la solution avec une énergie minimale. Cependant, ces méthodes ne garantissent aucune continuité dans la solution. Alors, on peut imposer des contraintes, non pas directement sur  $\mathbf{q}$ , mais sur le gradient, ce qui garantit que la solution sera lisse [3] :

$$\hat{\mathbf{q}} = \underset{\mathbf{q}}{\operatorname{argmin}} \|\mathbf{b} - G\mathbf{q}\|^2 + \lambda\|\nabla\mathbf{q}\|^2.$$

Dans ce cas la solution analytique est

$$\hat{\mathbf{q}} = (G^t G + \lambda\nabla^t \nabla)^{-1} G^t \mathbf{b}$$

où  $\nabla^t \nabla$  est une approximation d'un laplacien dans l'espace des sources contenant  $\mathbf{q}$ .

Toutes les méthodes dérivées de la norme minimale souffrent du manque de justifications physiologiques. Il faut modifier quelque peu le critère de minimisation pour ce faire.

### 2.2.2. LORETA

Les précédentes méthodes de reconstruction dépendent exclusivement de l'information du problème direct contenue dans  $G$ . Or, il s'avère que les sources profondes, importantes en EEG, ne contribuent que très peu aux données alors qu'elles peuvent être importantes à identifier. En effet, des sources en surface avec une intensité faible sera privilégiée à une source profonde et d'intensité forte. C'est pour mieux reconstruire ces sources profondes que Pascual-Marqui *et. al.* [39] ont

introduit la méthode LORETA (LOW Resolution Electromagnetic Tomography). LORETA compense la faible contribution des sources profondes en normalisant la matrice  $G$ . De plus, LORETA fait l'hypothèse que les dipôles sont simultanément activés et synchrones. Cette hypothèse s'appuie sur des expériences qui ont montré, par des enregistrements sur des cellules du cerveau, que les neurones sont fortement synchronisés lorsqu'ils sont rapprochés. On peut alors utiliser cette information spatiale pour régulariser le problème inverse. Cette information est introduite à l'aide d'une matrice de corrélation qui correspond au laplacien. Dans sa version originale, LORETA calcule la densité de courant dans tout le volume du cortex. La méthode consiste à pondérer la matrice  $G$  par la norme de chacune de ses colonnes :

$$\hat{\mathbf{q}} = \underset{\mathbf{b}=G\mathbf{q}}{\operatorname{argmin}} \mathbf{q}^t \mathbf{A} \mathbf{q}$$

où  $\mathbf{A} = \mathbf{W}\mathbf{C}^t\mathbf{C}\mathbf{W}$  avec  $\mathbf{W}$  une matrice diagonale  $N \times N$  qui contient la norme de chaque colonne de  $G$  et  $\mathbf{C}$  est un laplacien spatial qui contient de l'information sur le voisinage de chaque dipôle. La solution  $\hat{\mathbf{q}}$  correspond à la configuration de courants la plus lisse capable d'expliquer les données. C'est pourquoi les solutions sont à basse résolution, d'où le nom de la méthode. Une solution analytique existe pour ce problème, ce qui est très intéressant puisqu'elle nous donne une façon simple et rapide de reconstruire l'activation tout en incorporant de l'information spatiale dans le laplacien  $\mathbf{C}$ . Ce laplacien joue le même rôle qu'une matrice de corrélation. On peut définir de fortes corrélations comme de faibles. Dans [39], les auteurs indiquent comment le laplacien est défini.

Comme nous l'avons dit, LORETA renonce à une solution focale et donne une solution lisse dans l'espace. Or, il s'avère que les densités de courants ne sont pas toujours lisses, elles peuvent contenir des discontinuités. Pour avoir des solutions réalistes, nous devons modifier quelque peu notre approche. Pour ce faire, les méthodes bayésiennes sont parmi les plus populaires en MEG. Ce formalisme se distingue des méthodes du type norme minimale parce qu'elles peuvent prendre en compte des contraintes plus générales comme nous le verrons dans le prochain paragraphe. Cette thèse explore une approche probabiliste assez proche des méthodes bayésiennes, s'appuyant sur le principe du maximum d'entropie.

### 2.2.3. Méthodes bayésiennes

Au cours de l'exposé des méthodes précédentes, nous raisonnions directement dans l'espace des sources. Les méthodes bayésiennes envisagent le problème d'un point de vue probabiliste et permettent une façon générale, formelle et quantitative d'incorporer de l'information additionnelle indépendamment des données en MEG. Il s'agit de trouver un estimateur qui maximise la probabilité *a posteriori* (MAP) de  $\mathbf{q}$  étant donné les données  $\mathbf{b}$  [2] :

$$\hat{\mathbf{q}} = \underset{\mathbf{q}}{\operatorname{argmax}} P(\mathbf{q}|\mathbf{b}). \quad (2.5)$$

Ce formalisme bayésien prend en compte l'attache aux données en spécifiant la loi de probabilité  $P(\mathbf{q}|\mathbf{b})$  plutôt qu'en imposant des contraintes directement sur la solution  $\hat{\mathbf{q}}$ . Il existe quelques variantes de ce formalisme mais elles sont peu utilisées en MEG [34]. Dans l'exposé qui suit, nous nous limiterons au Maximum *a posteriori*. Par la loi de Bayes (d'où le nom de la méthode), on peut écrire

$$P(\mathbf{q}|\mathbf{b}) = \frac{P(\mathbf{b}|\mathbf{q})P(\mathbf{q})}{P(\mathbf{b})}. \quad (2.6)$$

Dans cette dernière équation,  $P(\mathbf{b}|\mathbf{q})$  est la probabilité de  $\mathbf{b}$  étant donné les sources  $\mathbf{q}$ . Donc  $P(\mathbf{b}|\mathbf{q})$  contient le modèle direct qui exprime  $\mathbf{b}$  en fonction de  $\mathbf{q}$ . La probabilité  $P(\mathbf{q})$  contient la distribution de probabilité *a priori* et c'est dans cette probabilité que nous pourrions mettre toute l'information anatomique et fonctionnelle dont on dispose sur la solution. Pour le dénominateur dans (2.6),  $P(\mathbf{b})$  est indépendant de  $\mathbf{q}$  et on peut l'ignorer dans le processus de maximisation (2.5).

Quelles formes peuvent prendre ces distributions de probabilité ? Pour  $P(\mathbf{b}|\mathbf{q})$ , on doit mettre l'information sur le modèle direct qui exprime  $\mathbf{b}$  en fonction de  $\mathbf{q}$ . Or, on connaît la relation entre le  $\mathbf{b}$  et  $\mathbf{q}$  :  $\mathbf{b} = G\mathbf{q}$ . Ainsi, la probabilité  $P(\mathbf{b}|\mathbf{q})$  doit représenter l'écart à ce modèle. Comme le bruit est l'écart entre  $\mathbf{b}$  et le champ magnétique  $G\mathbf{q}$  donné par les courants  $\mathbf{q}$  et que le bruit est généralement modélisé par une gaussienne, on choisit

$$P(\mathbf{b}|\mathbf{q}) \propto \exp \left[ -\frac{1}{2\sigma^2} \|\mathbf{b} - G\mathbf{q}\|^2 \right]$$

où  $\sigma^2$  est la variance du bruit dans les données.

Pour  $P(\mathbf{q})$ , l'idée est de définir une "énergie"  $U(\mathbf{q})$ , dit potentiel de Gibbs, pour une configuration de sources donnée  $\mathbf{q}$ . Ainsi, par analogie avec la mécanique statistique,  $P(\mathbf{q})$  est écrite sous la forme

$$P(\mathbf{q}) \propto \exp[-\lambda U(\mathbf{q})].$$

On observe que si  $U(\mathbf{q}) < \infty$  alors  $P(\mathbf{q}) > 0$  et tous les états sont alors accessibles. Comme dans le cas physique, la configuration de probabilité maximale sera celle avec le potentiel minimal qui correspond à l'état d'équilibre du système. Le problème de maximisation de la probabilité (2.5) peut donc se réécrire sous la forme :

$$\hat{\mathbf{q}} = \underset{\mathbf{q}}{\operatorname{argmin}} \frac{1}{2\sigma^2} \|\mathbf{b} - G\mathbf{q}\|^2 + \lambda U(\mathbf{q}). \quad (2.7)$$

L'état d'équilibre initial défini par  $U(\mathbf{q})$  est ainsi modifié à cause de la contrainte d'attache aux données. On voit que le MAP généralise les méthodes de norme minimale.

L'hyperparamètre  $\lambda$  permet de pondérer les informations relatives des données par rapport au modèle *a priori* de sources qui est décrit dans le potentiel  $U(\mathbf{q})$ . Ainsi, lorsque  $\lambda \rightarrow 0$ , on tend à donner plus d'importance aux données et moins au modèle. Inversement, de grandes valeurs de  $\lambda$  seront appropriées lorsque nous ne pouvons faire confiance aux données, par exemple dans le cas où les données sont très bruitées. Ainsi, le modèle a plus d'importance et les données en ont moins.

Dans l'équation (2.7), on retrouve une expression similaire à (2.2) mais dans laquelle le terme de potentiel  $U(\mathbf{q})$  prend une toute nouvelle signification puisqu'il peut être déterminé à l'aide de propriétés probabilistes connues sur la solution. Plusieurs choix sont possibles. Nous allons discuter d'un cas particulier développé dans [3] et qui est intéressant dans le cadre de cette thèse puisque nous allons utiliser des idées analogues dans un autre contexte pour résoudre le problème inverse en MEG.

#### 2.2.4. S-MAP

La méthode S-MAP [3] (Spatial Maximum A Posteriori) vient de l'idée qu'avec une contrainte quadratique ( $U(\mathbf{q}) = \|\nabla\mathbf{q}\|^2$ ) les grands sauts sont pénalisés donnant une solution finale très lisse. Physiologiquement, cette solution n'est pas

appropriée puisque des gradients très forts (des discontinuités) peuvent être présents dans la solution. On peut utiliser une fonction convexe qui ne pénalise pas les grandes valeurs de gradient. Dans [16], les auteurs utilisent

$$U(\mathbf{q}) = \sum_{i=1}^N \sum_{\nu \in V(i)} \phi_{k,i,\nu}(\nabla_i^\nu \mathbf{q})$$

où la deuxième somme est faite sur tous les voisins  $\nu$  du dipôle  $i$  et

$$\phi_k(t) = \sqrt{1 + k^2 t^2} - 1.$$

Il est préférable de prendre une fonction convexe pour s'assurer que  $U(\mathbf{q})$  ne contient qu'un seul minimum local et qu'il corresponde au minimum global. Cependant, il est possible de prendre un potentiel qui n'a pas cette propriété. Par exemple, dans la version originale de S-MAP [3], le potentiel choisi n'est pas convexe. Dans ce cas, il est très difficile de trouver le véritable minimum global par des algorithmes d'optimisation numérique.

Pour bien comprendre de quelle façon la fonction  $\phi_k$  impose de moins grandes contraintes que le critère quadratique, regardons le comportement lorsque

$$t \rightarrow 0 \quad \text{alors} \quad \phi_k(t) \sim \frac{1}{2} k^2 t^2.$$

Cela se comporte comme une norme minimale lorsque les différences d'intensités entre dipôles voisins sont petites. Par contre, si

$$|t| \rightarrow \infty, \quad \text{alors} \quad \phi_k(t) \sim |kt|.$$

Ainsi, lorsque les valeurs de  $t$  sont grandes (grandes différences de valeurs entre des dipôles voisins), on ne pénalise plus comme  $t^2$  mais comme  $t$ . Dans la figure 2.1, on peut observer le comportement des potentiels. Au fur et à mesure que  $|t|$  augmente,  $\phi_k(t)$  se comporte linéairement. Alors que pour de petites valeurs de  $t$ ,  $\phi_k(t)$  est quadratique. Donc, des gradients forts sont acceptés rendant probable la présence de discontinuités dans la solution. Cette méthode intègre donc des informations locales sur la répartition des sources puisqu'on a des solutions lisses par morceaux.

De ces travaux ont découlé plusieurs méthodes de reconstruction. Dans [32], les auteurs démontrent qu'il est possible d'ajouter de l'information fonctionnelle provenant de l'IRMf à l'aide du formalisme S-MAP. Puis dans [16], les auteurs développent un algorithme itératif qui présente un problème inverse dont l'espace

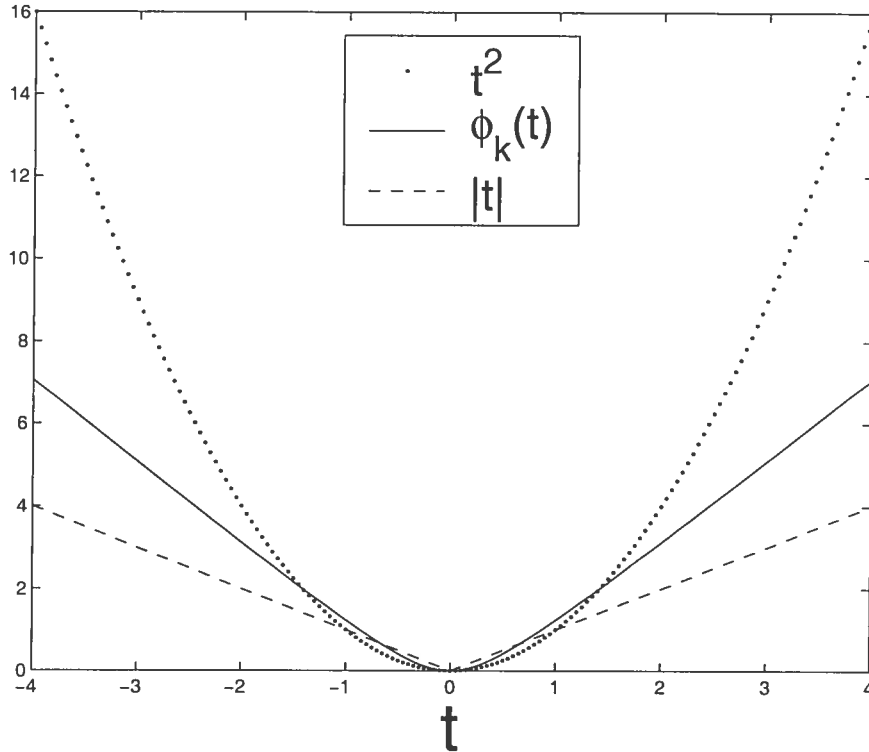


FIG. 2.1. Graphes des différents potentiels utilisé en reconstruction bayésienne.

solution est réduit à chaque itération. Il converge alors vers une solution focale en utilisant S-MAP comme méthode inverse. Dans les chapitres 4 et 5, nous développons des algorithmes analogues dans le cadre de cette thèse.

De plus, dans l'article original du S-MAP [3], les auteurs présentent la méthode ST-MAP qui permet d'utiliser les corrélations temporelles pour reconstruire l'activation sur tout un intervalle de temps en imposant une évolution temporelle lente en accord avec les données. Il a été montré que ces contraintes jouent un rôle important.

Le formalisme bayésien et son interprétation probabiliste est extrêmement riche. Par exemple, dans [43], on propose non pas de trouver la meilleure solution mais plutôt un ensemble de bonnes solutions qui, chacune, satisfait la contrainte d'attache aux données ainsi que l'information *a priori* que l'on a. Ainsi, les éléments qui sont communs à toutes ces bonnes solutions sont interprétés comme ayant une grande probabilité. Et dans [40], la loi  $P(\mathbf{q})$  est construite en explicitant

un processus binaire qui identifie l'activation du dipôle. L'activation est donc séparée de l'intensité. Nous aurons d'ailleurs un modèle de source semblable dans cette thèse dans le chapitre 4.

Les méthodes bayésiennes sont très puissantes. Le formalisme probabiliste permet d'intégrer de l'information sur la solution et de la comparer aux données MEG pour identifier une solution optimale. Cependant, la minimisation (2.7) peut être très laborieuse à effectuer. En effet, si le potentiel  $U(\mathbf{q})$  est non-convexe la solution peut ne pas être unique. De plus, en général, les modèles distribués doivent avoir un grand nombre de dipôles (au moins un millier) pour donner une représentation réaliste de la surface corticale. Alors, la minimisation se fait dans un espace de grande dimension qui, combiné à la non-convexité possible du potentiel, rend la résolution du problème de minimisation extrêmement difficile.

La méthode de la maximisation de l'entropie en moyenne (MEM) qui est appliquée dans cette thèse au problème inverse en MEG ne possède pas ces défauts. Dans sa formulation duale, la minimisation est faite dans l'espace des sources. De plus, le critère est toujours convexe. Ces deux éléments font du MEM une méthode efficace pour résoudre le problème inverse puisque, justement, la particularité de la MEG est d'avoir peu de données à sa disposition. Un exposé détaillé du MEM est fait dans le prochain chapitre. Mais avant, nous présentons une méthode qui permet de préconditionner le problème en réduisant l'espace des sources. Cette méthode est combinée au MEM dans les chapitres 5 et 6.

### 2.2.5. Présélection de sources

Une percée intéressante pour la résolution du problème inverse en MEG est venue de Mattout *et. al.* [33]. Étant donné que le caractère mal posé de ce problème provient d'une indétermination spatiale, séparer la localisation de l'estimation de l'amplitude peut faciliter la reconstruction des sources. Leur méthode, la Multivariate Source Preselection (MSP), a pour but d'isoler les sources les plus corrélées avec les données pour en sélectionner un sous-ensemble afin de réduire le support de l'activation.

Le principe consiste à considérer la matrice  $G$  comme un opérateur de l'espace source à l'espace des données et de définir l'ensemble

$$\mathcal{D} = \{G \mathbf{q} \text{ tel que } \mathbf{q} \in \mathbb{R}^N\} \quad (2.8)$$

comme l'image de cet opérateur. En principe, les données  $\mathbf{b}$  n'appartiennent pas à  $\mathcal{D}$  car elles sont bruitées. Cependant, pour que  $\mathbf{b} \notin \mathcal{D}$ , il faudrait que les colonnes de la matrice  $G$  engendrent un espace de dimension plus petite que le nombre de capteurs  $d$ . Or, si on regarde la provenance de la matrice  $G$ , on voit qu'elle est composée de  $N$  dipôles couvrant toute la surface du cortex et les orientations des dipôles couvrent presque toutes les directions. De plus, comme  $N$  est habituellement beaucoup plus grand que  $d$ , il est très peu probable que le problème soit dégénéré au point où un ensemble de moins de  $d$  vecteurs forme une base de  $\mathcal{D}$ . Puisque  $d$  vecteurs linéairement indépendants engendrent  $\mathbb{R}^d$  alors  $\mathcal{D} = \mathbb{R}^d$  et donc  $\mathbf{b} \in \mathbb{R}^d$ . L'ensemble  $D$  pourrait être plus petit que  $\mathbb{R}^d$  si les dipôles étaient corrélés entre eux. Cependant, dans cette thèse, nous utiliserons la version univariée de la MSP qui suppose tous les dipôles indépendants.

Donc, par hypothèse, il existe un ensemble de  $d$  vecteurs orthonormaux  $\mathbf{d}_i$  qui forment une base de  $\mathcal{D}$ . On peut exprimer  $\mathbf{b}$  en termes de cette base :

$$\mathbf{b} = D \Gamma$$

où  $D$  est la matrice dont les colonnes sont les  $\mathbf{d}_i$ . Puisque  $D$  est orthogonale ( $D^t D = \mathbb{I}$ ),

$$\Gamma = D^t \mathbf{b}.$$

On cherche maintenant à connaître la contribution de chaque colonne de  $G$  aux données. Ainsi, on fabrique le projecteur

$$P_\Gamma = \Gamma(\Gamma^t \Gamma)^{-1} \Gamma^t = D^t \mathbf{b}(\mathbf{b}^t \mathbf{b})^{-1} \mathbf{b}^t D.$$

Il projette un vecteur de l'espace engendré par les vecteurs de base de  $D$  vers l'espace des données. Ce projecteur est exprimé dans la base des vecteurs  $\mathbf{d}_i$ . Si on se replace dans la base dans laquelle sont exprimées la matrice  $G$  et les données  $\mathbf{b}$ , on aura

$$P_{\mathbf{b}} = D P_\Gamma D^t = \mathbf{b}(\mathbf{b}^t \mathbf{b})^{-1} \mathbf{b}^t.$$

Ce projecteur prend les colonnes de  $G$  pour les projeter dans l'espace des données. Notons que ce projecteur est aussi défini par la même expression lorsque les données sont prises à  $T$  différents temps représentant différentes réalisations du même phénomène physique.



Avec ce projecteur, on peut trouver la contribution normalisée de chaque colonne  $\mathbf{g}_i$  de  $G$  par

$$a_i = \frac{\|P_{\mathbf{b}} \mathbf{g}_i\|^2}{\|\mathbf{g}_i\|^2}.$$

Les coefficients  $a_i$  permettent de qualifier l'apport du dipôle  $i$  dans les données. Avec ces coefficients, on peut ordonner les dipôles selon leur contribution aux données. On peut alors réduire le support en n'en choisissant qu'un sous-ensemble à l'aide des  $a_i$ . Ces nombres sont compris entre 0 et 1 puisqu'ils sont normalisés. On peut alors les associer à une probabilité. Il suffit de trouver un seuil  $a$  avec une  $p$ -value :

$$P(a_i > a \mid \text{le dipôle } i \text{ est activé}) = p.$$

Dans la figure 2.2, on montre un histogramme des  $a_i$  avec le seuil choisi. On voit que la grande majorité des coefficients sont près de 0 ce qui indique que seule une faible quantité de dipôles sont probablement activés ce qui est effectivement le cas ici.

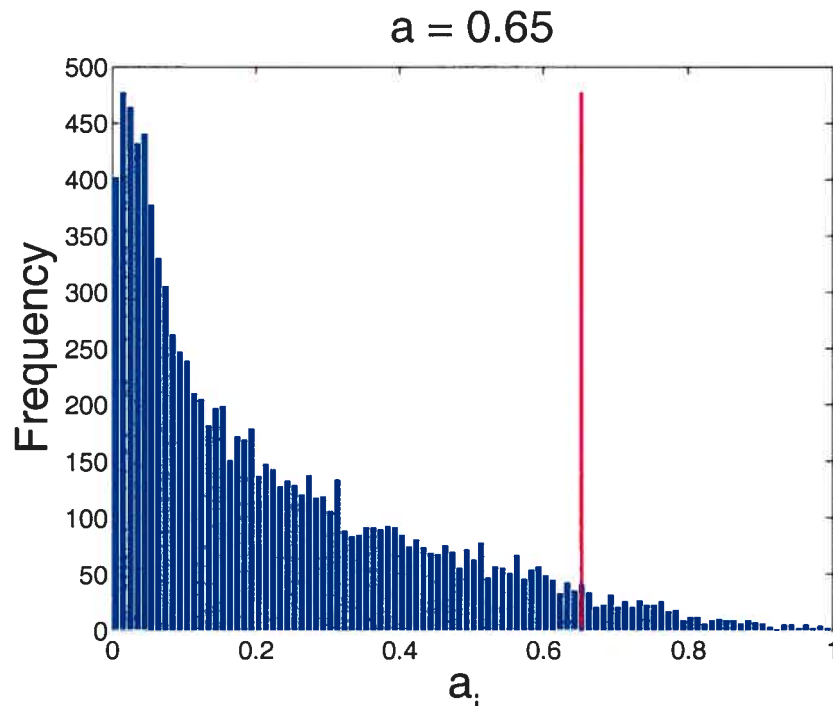


FIG. 2.2. Histogramme des  $a_i$  pour un exemple pris du chapitre 6. La ligne verticale rouge indique la valeur du seuil  $a = 0.65$  donné par une  $p$ -value de  $p = 0.05$ .

Ce seuil permet d'extraire un sous-ensemble des dipôles selon leur contribution aux données appelé support d'activation. Ce support réduit le nombre de sources donc la dimensionnalité du problème. Dans [31], l'auteur présente sa méthode en conjonction avec différentes méthodes de reconstruction. Nous verrons dans le chapitre 5 comment nous pouvons utiliser cette classification pour préconditionner le problème inverse que nous attaquerons avec la méthode du maximum d'entropie décrite dans le prochain chapitre. La MSP a aussi prouvé son efficacité sur des données réelles et elle sera utilisée à cet effet en conjonction avec le MEM dans le chapitre 6.

# Chapitre 3

---

## MAXIMUM D'ENTROPIE EN MOYENNE

La méthode du Maximum d'Entropie en Moyenne (MEM) repose sur une interprétation particulière de la mécanique statistique. Cette interprétation provient de la découverte par Boltzmann que le concept d'entropie est le véritable lien entre les lois microscopiques de la physique et les observations macroscopiques d'un système (pensons à l'irréversibilité). La mécanique statistique est habituellement présentée à partir des équations du mouvement et de quelques hypothèses additionnelles sur le comportement global du système. On en dégage ensuite le concept d'entropie par analogie avec les équations de la thermodynamique. Ce raisonnement est assez complexe et exige des techniques statistiques sophistiquées et des hypothèses non-intuitives comme l'ergodicité.

Or, dans [22], E.T. Jaynes affirme que l'entropie a une signification beaucoup plus profonde. Il développe une nouvelle interprétation de la mécanique statistique en se basant sur les travaux de Shannon [44] qui met l'accent sur la loi de probabilité qui décrit un système complexe. Il interprète l'entropie comme une mesure du manque d'information. Il considère cette entropie (et non les équations du mouvement) comme un concept premier. Ainsi, une distribution de probabilité d'un système à l'équilibre doit maximiser cette entropie pour maximiser le manque d'information tout en restant compatible avec les connaissances que l'on a sur le système. Avec ce principe, il est possible de dériver toutes les équations de la mécanique statistique des systèmes à l'équilibre. Le point le plus important de cette nouvelle interprétation n'est pas tant la simplification des concepts impliqués mais l'indépendance de ces équations par rapport aux équations microscopiques du mouvement. Elles jettent donc un regard beaucoup plus général sur la mécanique statistique. Il serait alors possible d'appliquer ce formalisme à d'autres problèmes comme cela est fait dans cette thèse pour la résolution du problème inverse en MEG.

### 3.1. FONDEMENTS DE LA MÉCANIQUE STATISTIQUE

La mécanique statistique<sup>1</sup> pose le problème de décrire les propriétés d'un système contenant un nombre macroscopique (ce nombre,  $N$ , est de l'ordre du nombre d'Avogadro  $\sim 6.022 \times 10^{23}$ ) de particules qui peuvent interagir entre elles. Le mouvement de ces particules est dicté par la mécanique hamiltonienne. En trois dimensions, un tel système a  $3N$  degrés de liberté qui spécifient un ensemble de  $6N$  coordonnées de positions et d'impulsion qui sont déterminées de façon unique par les équations de mouvement et les conditions initiales. Si on définit l'espace des phases  $\Gamma$ , de dimension  $6N$ , comme l'ensemble des positions et impulsions de chaque particule alors l'état du système est toujours donné par un unique point  $\mathbf{X}$  dans cet espace des phases.

Dans la réalité, il est impossible de prétendre pouvoir résoudre le système d'équations différentielles associé étant donné la très grande dimensionnalité du système. Ainsi, nous ne sommes jamais certains de l'état exact du système mais il est possible de calculer la probabilité de trouver le système dans un état particulier de l'espace des phases. Donc, l'état du système  $\mathbf{X}$  peut être vu comme une variable aléatoire et une distribution de probabilité  $\rho(\mathbf{X})$  est associée au système en accord avec les connaissances que l'on a sur le système. Le problème est donc réduit à trouver cette distribution de probabilité. Pour ce faire, on fait habituellement appel à la théorie ergodique.

#### 3.1.1. Théorie ergodique

La théorie ergodique s'efforce de donner un cadre axiomatique à la mécanique statistique. Le point central de cette théorie est le théorème ergodique qui relie la moyenne spatiale dans  $\Gamma$  et la moyenne temporelle d'une fonction de l'espace des phases.

Pour les systèmes étudiés en mécanique statistique, l'énergie est la seule quantité conservée. Ainsi le système sera contraint d'évoluer sur la surface  $S_E$  d'énergie constante  $E$ . Définissons la moyenne dans l'espace des phases d'une fonction

---

<sup>1</sup>On se restreindra, dans l'exposé qui suit, aux systèmes à l'équilibre décrits par la mécanique classique.

$f(\mathbf{X})$ ,  $f$  étant un observable du système, par

$$\langle f \rangle_S = \frac{1}{\sum(E)} \int_{S_E} f(\mathbf{X}) dS_E \quad (3.1)$$

où  $\sum(E)$  est l'aire de la surface  $S_E$  dans l'espace des phases  $\Gamma$ . De façon analogue, comme on connaît la trajectoire de  $\mathbf{X}(t)$  dans le temps, on peut définir la moyenne temporelle par

$$\langle f \rangle_T = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{t_0}^{t_0+T} f(\mathbf{X}(t)) dt. \quad (3.2)$$

Le système est dit ergodique si, au cours du temps, il passe arbitrairement près de chaque point de la surface  $S_E$ . Birkhoff a démontré [41, sec. 6.C.] en 1931 que pour un système ergodique toute fonction  $f$  dont les moyennes spatiale et temporelle existent alors elles sont égales ( $\langle f \rangle_S = \langle f \rangle_T$ ). Ce théorème, appelé théorème ergodique, est fondamental puisqu'il donne une correspondance entre la fraction de temps passée par le système dans une sous-région  $R_E$  de la surface  $S_E$  d'énergie constante  $E$  et l'aire de cette sous-région. En effet, si  $\tau_{R_E}$  est le temps passé dans la région  $R_E$  pour une durée totale de  $T$  alors, pour un système ergodique<sup>2</sup>,

$$\lim_{T \rightarrow \infty} \frac{\tau_{R_E}}{T} = \frac{\sum(R_E)}{\sum(E)} \quad (3.3)$$

où  $\sum(R_E)$  est l'aire de la sous-région  $R_E$  sur la surface  $S_E$ . On peut alors affirmer que le système passe un temps égal dans des régions de surface égale. Dans ce contexte, on peut définir la probabilité de trouver le système dans la région  $R_E$  à un temps arbitraire comme la fraction de la surface occupée par  $R_E$  par rapport à la surface totale :

$$P(R_E) = \frac{\sum(R_E)}{\sum(E)} = \int_{R_E} \frac{1}{\sum(E)} dS_E. \quad (3.4)$$

Ainsi, la densité de probabilité est donnée par

$$\rho(\mathbf{X}) = \frac{1}{\sum(E)}. \quad (3.5)$$

Cette dernière équation est l'équation fondamentale de la mécanique statistique et est appelée la distribution microcanonique de Gibbs. Une propriété remarquable de cette distribution est que tous les états du système sont équiprobables. En effet, la seule connaissance du système est l'énergie totale  $E$  alors le choix qui

<sup>2</sup>On n'a qu'à prendre  $f$  égale à la fonction indicatrice sur la région  $R_E$ .

ne privilégie aucun état et qui respecte cette connaissance *a priori* est la densité uniforme donne une probabilité égale à tous les états.

Une fois cette distribution  $\rho$  connue, il est possible de calculer les valeurs moyennes des quantités thermodynamiques. Cependant, il reste une quantité thermodynamique qui n'est pas définie de façon unique par ce formalisme : l'entropie. Il faut ajouter des hypothèses autres que l'ergodicité du système.

### 3.1.2. Entropie

Khinchin a écrit sur le concept d'entropie [25, p. 137] :

The notion of entropy is one of the most important physical notions from a theoretical as well as from a practical point of view. Very few other notions can compete with it in respect to the abundance of attempts to clarify its theoretical and philosophical meaning.

En effet, l'entropie se retrouve au centre de plusieurs contextes. Elle peut être définie par la thermodynamique (reliée à la chaleur), la mécanique statistique (reliée au nombre d'états quantiques accessibles au système). La théorie des grandes déviations [14] démontre que l'entropie est une quantité qui détermine le taux de convergence exponentielle de suites de lois de probabilité. Il existe différentes généralisations de l'entropie, les entropies de Rényi ou de Tsallis [23,24] par exemple, pour des utilisations dans des contextes différents. La théorie de l'information a donné un sens général à l'entropie. Pour cette thèse, l'entropie est utilisée comme une mesure de l'information contenue dans un système. Une telle quantité doit satisfaire certaines propriétés :

**Unicité:** L'entropie doit avoir un maximum unique.

**Invariance:** Le choix du système de coordonnées ne doit pas modifier ce maximum.

**Indépendance des systèmes:** Si on a de l'information indépendante sur deux systèmes indépendants, l'entropie totale des deux systèmes est alors la somme des entropies de chacun des sous-systèmes séparément.

**Indépendance des sous-systèmes:** Le maximum ne devrait pas changer si on considère un sous-ensemble d'un système avec une probabilité conditionnelle ou avec la probabilité du système complet réduite à ce sous-système.

Shore et Johnson [45] (voir aussi [5]) ont démontré qu'il n'existe qu'une seule fonction qui satisfait ces quatre propriétés et c'est l'entropie de Gibbs

$$\mathcal{S}(\rho) = -k_B \int \rho(\mathbf{X}) \log (C\rho(\mathbf{X})) d\mathbf{X} \quad (3.6)$$

où  $C$  est une constante qui permet de rendre l'expression à l'intérieur du logarithme sans unités<sup>3</sup> et  $k_B$  est la constante de Boltzmann égale à  $1.38 \times 10^{-23}$  J/K.

Par la deuxième loi de la thermodynamique, l'entropie croît toujours sauf dans le cas où le système a atteint un état d'entropie maximale. Ce système reste donc dans cet état puisque l'entropie ne peut plus augmenter. Si un système n'évolue plus, il est à l'équilibre. Ainsi, on peut donner l'état d'un système à l'équilibre en trouvant celui qui maximise l'entropie du système.

Par exemple, supposons que l'énergie moyenne du système est connue :

$$\langle E \rangle = \int E(\mathbf{X})\rho(\mathbf{X}) d\mathbf{X}. \quad (3.7)$$

En faisant le calcul de maximisation de l'entropie sous les contraintes que l'intégrale de  $\rho$  est l'unité et (3.7), on arrive à l'expression

$$\rho(\mathbf{X}) = \frac{1}{\mathcal{Z}} e^{\lambda E(\mathbf{X})}$$

où  $\mathcal{Z} = \int e^{\lambda E(\mathbf{X})} d\mathbf{X}$  est la fonction de partition du système et  $\lambda$  est un multiplicateur de Lagrange associé à la contrainte d'énergie moyenne. Ce paramètre peut être fixé à l'aide de la thermodynamique et est égale à  $\frac{1}{k_B T}$ . La dernière équation est exactement l'expression pour l'ensemble canonique de Gibbs.

On remarque que nous avons déduit la densité de probabilité sans jamais faire appel aux équations du mouvement microscopiques du système. En fait, si on s'attarde à la signification que Shannon donne à l'entropie, c'est-à-dire une mesure de désordre, on peut voir la distribution de probabilité qui maximise l'entropie comme celle qui maximise le désordre. En effet, si on n'a aucune information sur le système, la distribution qui maximise l'entropie (ou le désordre) est la distribution uniforme (3.5) retrouvant ainsi l'expression donnée par la théorie ergodique pour l'ensemble microcanonique. Ce formalisme rend donc inutile l'hypothèse d'ergodicité pour les systèmes en équilibre.

---

<sup>3</sup>Il est intéressant de noter que cette constante ne peut être définie par la mécanique classique. Il faut utiliser la mécanique quantique pour l'interpréter physiquement.

Avec le principe de maximisation de l'entropie, on se concentre seulement sur l'information disponible sur le système : les contraintes de la forme (3.7) sur la distribution de probabilité du système. Il n'est point nécessaire d'utiliser les équations du mouvement. Ainsi, la mécanique statistique peut être dérivée en utilisant l'inférence statistique.

### 3.2. PROBLÈMES INVERSES EN MEM

Comment peut-on appliquer ces principes de mécanique statistique au problème inverse en MEG ? Il s'agit de voir l'objet à reconstruire comme un état physique. Considérons un système physique qui contient un grand nombre de particules dont les états évoluent dans un espace d'états  $S_\mu$ . Cet espace  $S_\mu$  est constitué d'états plus ou moins probables selon une mesure de probabilité de référence  $\mu$  (analogue à  $\rho$  dans la section précédente) qui est déterminée par des contraintes que nous connaissons sur le système. Notons  $\mathbf{m}$ , la moyenne de  $\mu$ . L'état d'équilibre du système est alors décrit par le couple  $(\mathbf{m}, \mu)$ .

Si une nouvelle observation (ou information) est disponible sur le système, deux cas se présentent. L'observation peut être expliquée par la distribution  $\mu$  ou non. Dans le premier cas, le système est toujours correctement décrit par la loi de probabilité  $\mu$  et cette observation n'est pas une information vraiment pertinente puisqu'elle ne permet pas d'en savoir plus sur le système. Il n'y a donc aucun changement nécessaire pour caractériser les propriétés physiques du système. Dans le deuxième cas, le système n'est plus correctement décrit par  $\mu$ . On doit alors changer cette description avec une nouvelle distribution de probabilité  $\hat{P}$ . La mécanique statistique stipule que ce nouvel état d'équilibre est celui qui maximisera l'entropie du système tout en satisfaisant les nouvelles contraintes que l'ajout d'information a pu donner [14, sec. III.2].

Pour le problème inverse, on considère l'objet à reconstruire comme la moyenne d'un certain processus aléatoire  $\mathbf{Q}$  selon la mesure de référence  $\mu$  :

$$\langle \mathbf{Q} \rangle = E_\mu(\mathbf{Q}) = \int \mathbf{q} d\mu(\mathbf{q}).$$

Dans le cas qui nous intéresse, une observation sur le système est donnée par la contrainte d'attache aux données :

$$\mathbf{b} = G \mathbf{q}. \tag{3.8}$$



Si  $\langle \mathbf{Q} \rangle$  est une solution au problème inverse, alors l'équilibre du système est maintenu et son état ne change pas. Par contre, si  $\mathbf{b} \neq G \langle \mathbf{Q} \rangle$  alors on doit changer la description du système donnée par  $\mu$ . Le principe de la maximisation de l'entropie en moyenne stipule que la meilleure description possible est celle qui maximise l'entropie par rapport à la mesure de référence  $\mu$  tout en satisfaisant la contrainte (3.8). L'entropie appropriée dans ce cas, appelée entropie de Kullback, est une généralisation l'entropie de Gibbs (3.6) :

$$\mathcal{H}(P, \mu) = \begin{cases} - \int f(\mathbf{q}) \log f(\mathbf{q}) d\mu(\mathbf{q}) & \text{si } dP = f d\mu, \\ -\infty & \text{sinon.} \end{cases} \quad (3.9)$$

Cette fonctionnelle a la propriété que  $\mathcal{H}(P, \mu) \leq 0$  avec égalité si et seulement si  $P = \mu$ .

Donc le nouvel état d'équilibre  $\hat{P}$  sera donné en maximisant l'entropie de Kullback (équation (3.9)) par rapport à  $\mu$  sous la contrainte :

$$\mathbf{b} = G E_{\hat{P}}(\mathbf{Q}). \quad (3.10)$$

Si  $\mathcal{C} = \{\mathbf{q} \in \mathbb{R}^N | \mathbf{b} = G \mathbf{q}\}$  et  $\mathcal{P}$  l'ensemble des lois de probabilité dont la moyenne appartient à  $\mathcal{C}$ , alors

$$\hat{P} = \operatorname{argmax}_{P \in \mathcal{P}} \mathcal{H}(P, \mu) \quad (3.11)$$

et la moyenne de  $\mathbf{Q}$  sous  $\hat{P}$  est

$$\hat{\mathbf{q}} = E_{\hat{P}}(\mathbf{Q}). \quad (3.12)$$

Ainsi,  $\hat{P}$  satisfait la contrainte (3.10) par construction. Donc,  $\hat{P}$  englobe toute l'information de  $\mu$ , la loi de référence, tout en satisfaisant la contrainte sur la moyenne. De plus, on est assuré qu'il n'y a aucune autre information apportée puisque  $\hat{P}$  maximise l'entropie  $\mathcal{H}$ .

Dans la résolution de problèmes inverses mal-posés, souvent seule la solution  $\hat{\mathbf{q}}$  est extraite. Ainsi, l'étape (3.11) peut sembler superflue. Mais, comme nous le verrons dans le prochain chapitre, la mesure  $\hat{P}$  contient énormément d'information sur la solution que nous pouvons utiliser pour construire une nouvelle loi de référence.

Nous avons maintenant une méthode de régularisation de problèmes inverses mal posés. Elle a comme ingrédient une mesure de référence  $\mu$  et l'entropie par

rapport à cette loi. Nous allons tout d'abord expliquer comment on résoud le problème d'optimisation sous contraintes posé par l'équation (3.11). Ensuite, nous comparerons le MEM avec les autres méthodes de régularisation discutées dans le chapitre précédent. Nous verrons que le MEM comprend certaines de ces méthodes comme cas particulier. Puis, dans le prochain chapitre, nous construirons une loi de probabilité de référence  $\mu$  qui modélise l'activité cérébrale.

### 3.2.1. Expression duale du MEM

Dans cette section, la solution au problème (3.11) est explicitée. C'est un cas simple d'optimisation sous contraintes. Le lagrangien associé à (3.11) est

$$L(P, \boldsymbol{\lambda}, \lambda_0) = \mathcal{H}(P, \mu) + \boldsymbol{\lambda}^t (G E_P(\mathbf{Q}) - \mathbf{b}) + \lambda_0 \left( \int dP(\mathbf{q}) - 1 \right)$$

où  $\boldsymbol{\lambda}$  et  $\lambda_0$  sont les multiplicateurs de Lagrange associés aux contraintes

$$\mathbf{b} = G E_P(\mathbf{Q}) \quad \text{et} \quad \int dP(\mathbf{q}) = 1.$$

Puisque  $P$  doit être de la forme  $dP(\mathbf{q}) = f(\mathbf{q})d\mu(\mathbf{q})$  pour que  $\mathcal{H}(P, \mu) > -\infty$  alors  $L$  s'écrit comme

$$L(f, \boldsymbol{\lambda}, \lambda_0) = - \int f(\mathbf{q}) \log f(\mathbf{q}) d\mu(\mathbf{q}) + \boldsymbol{\lambda}^t \left( G \int \mathbf{q} f(\mathbf{q}) d\mu(\mathbf{q}) - \mathbf{b} \right) + \lambda_0 \left( \int f(\mathbf{q}) d\mu(\mathbf{q}) - 1 \right),$$

$$L(f, \boldsymbol{\lambda}, \lambda_0) = \int \left[ -f(\mathbf{q}) \log f(\mathbf{q}) + \boldsymbol{\lambda}^t G \mathbf{q} f(\mathbf{q}) + \lambda_0 f(\mathbf{q}) \right] d\mu(\mathbf{q}) - \boldsymbol{\lambda}^t \mathbf{b} - \lambda_0.$$

On doit trouver  $\hat{f}$ ,  $\hat{\boldsymbol{\lambda}}$  et  $\hat{\lambda}_0$  qui maximisent  $L$  et la distribution de probabilité *a posteriori* sera  $d\hat{P}(\mathbf{q}) = \hat{f} d\mu(\mathbf{q})$ . Fixons tout d'abord  $\boldsymbol{\lambda}$  et  $\lambda_0$ , puis maximisons  $L$  par rapport à  $f$ . Le lagrangien  $L$  sera maximal lorsque l'intégrand  $\mathcal{L}$  de  $L$  satisfera aux équations d'Euler-Lagrange<sup>4</sup> :

$$\frac{\partial \mathcal{L}}{\partial f} = 0 \quad \implies \quad -1 - \log \hat{f}(\mathbf{q}) + \boldsymbol{\lambda}^t G \mathbf{q} + \lambda_0 = 0$$

$$\hat{f}(\mathbf{q}) = \exp[\boldsymbol{\lambda}^t G \mathbf{q} + \lambda_0 - 1].$$

<sup>4</sup>Les équations d'Euler-Lagrange sont  $\nabla \cdot \left( \frac{\partial \mathcal{L}}{\partial \dot{f}} \right) - \frac{\partial \mathcal{L}}{\partial f} = 0$  où  $\dot{f} = \nabla f$ . Dans le cas qui nous intéresse,  $\frac{\partial \mathcal{L}}{\partial \dot{f}} = 0$ .

La contrainte de normalisation  $\int f(\mathbf{q}) d\mu(\mathbf{q}) = 1$  fixe  $\lambda_0$  :

$$1 - \lambda_0 = \log \int \exp(\boldsymbol{\lambda}^t G \mathbf{q}) d\mu(\mathbf{q}).$$

Cette fonction est très importante et nous la noterons

$$\mathcal{F}_\mu^*(\boldsymbol{\xi}) = \log \int \exp(\boldsymbol{\xi}^t \mathbf{q}) d\mu(\mathbf{q}). \quad (3.13)$$

C'est le logarithme de la transformée de Laplace de la loi  $\mu$ . Ainsi,

$$f(\mathbf{q}) = \exp [\boldsymbol{\lambda}^t G \mathbf{q} - \mathcal{F}_\mu^*(G^t \boldsymbol{\lambda})].$$

Le lagrangien devient

$$D(\boldsymbol{\lambda}) = L(\exp [\boldsymbol{\lambda}^t G \mathbf{q} - \mathcal{F}_\mu^*(G^t \boldsymbol{\lambda})], \boldsymbol{\lambda}, 1 - \mathcal{F}_\mu^*(G^t \boldsymbol{\lambda})),$$

$$D(\boldsymbol{\lambda}) = \mathcal{F}_\mu^*(G^t \boldsymbol{\lambda}) - \boldsymbol{\lambda}^t \mathbf{b}.$$

La loi *a posteriori* s'écrira :

$$d\hat{P}(\mathbf{q}) = \hat{f}(\mathbf{q})d\mu(\mathbf{q}) = \exp [\hat{\boldsymbol{\lambda}}^t G \mathbf{q} - \mathcal{F}_\mu^*(G^t \hat{\boldsymbol{\lambda}})] d\mu(\mathbf{q})$$

où  $\hat{\boldsymbol{\lambda}}$  est la valeur qui minimise le critère  $D(\boldsymbol{\lambda})$ . La fonction  $\mathcal{F}_\mu^*(\mathbf{s})$  est strictement convexe<sup>5</sup> ce qui assure que la fonction  $D(\boldsymbol{\lambda})$  est aussi strictement convexe et qu'il existe une unique valeur  $\hat{\boldsymbol{\lambda}} \in \mathbb{R}^n$  telle que

$$D(\boldsymbol{\lambda}) \geq D(\hat{\boldsymbol{\lambda}}), \quad \forall \boldsymbol{\lambda} \in \mathbb{R}^n,$$

qui minimise  $D$ . Et donc, formellement :

$$\hat{\boldsymbol{\lambda}} = \underset{\boldsymbol{\lambda}}{\operatorname{argmin}} \mathcal{F}_\mu^*(G^t \boldsymbol{\lambda}) - \boldsymbol{\lambda}^t \mathbf{b} \quad (3.14)$$

qui est appelé critère dual du MEM. Ayant la loi *a posteriori*  $\hat{P}$ , la moyenne de cette loi est l'objet reconstruit  $\hat{\mathbf{q}}$ . On peut l'exprimer en fonction du gradient de

<sup>5</sup>Une fonction  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  est strictement convexe si  $\forall x, y \in \mathbb{R}^n$  et  $\lambda \in [0, 1]$  alors  $f(\lambda x + (1 - \lambda)y) < \lambda f(x) + (1 - \lambda)f(y)$ . Graphiquement, cela signifie que, sur le segment reliant deux points, la valeur d'une fonction convexe est toujours plus petite que l'interpolation linéaire de la fonction entre ces deux points.

$\mathcal{F}_\mu^*$ . En partant de l'équation (3.12), on a

$$\begin{aligned}
\hat{\mathbf{q}} &= \int \mathbf{q} d\hat{P}(\mathbf{q}) = \int \mathbf{q} \hat{f}(\mathbf{q}) d\mu(\mathbf{q}), \\
&= \int \mathbf{q} \exp \left[ \hat{\boldsymbol{\lambda}}^t G \mathbf{q} - \mathcal{F}_\mu^*(G^t \hat{\boldsymbol{\lambda}}) \right] d\mu(\mathbf{q}), \\
&= \exp \left[ -\mathcal{F}_\mu^*(G^t \hat{\boldsymbol{\lambda}}) \right] \int \mathbf{q} \exp \left[ \hat{\boldsymbol{\lambda}}^t G \mathbf{q} \right] d\mu(\mathbf{q}), \\
&= \exp \left[ -\mathcal{F}_\mu^*(G^t \hat{\boldsymbol{\lambda}}) \right] \nabla_{\boldsymbol{\xi}} \left[ \int \exp(\boldsymbol{\xi}^t \mathbf{q}) d\mu(\mathbf{q}) \right] \Big|_{\boldsymbol{\xi}=G^t \hat{\boldsymbol{\lambda}}}, \\
&= \exp \left[ -\mathcal{F}_\mu^*(G^t \hat{\boldsymbol{\lambda}}) \right] \nabla_{\boldsymbol{\xi}} \exp \left[ \mathcal{F}_\mu^*(\boldsymbol{\xi}) \right] \Big|_{\boldsymbol{\xi}=G^t \hat{\boldsymbol{\lambda}}}, \\
\hat{\mathbf{q}} &= \nabla_{\boldsymbol{\xi}} \mathcal{F}_\mu^*(\boldsymbol{\xi}) \Big|_{\boldsymbol{\xi}=G^t \hat{\boldsymbol{\lambda}}}. \tag{3.15}
\end{aligned}$$

Donc, l'objet reconstruit est le gradient du logarithme de la transformée de Laplace de  $\mu$  évalué au point  $G^t \hat{\boldsymbol{\lambda}}$ . Dans cette formulation duale se trouve une propriété remarquable du MEM. Alors que dans les méthodes présentées dans le chapitre 2, il fallait minimiser dans l'espace des sources, ici on minimise dans l'espace des données. En effet, la fonction  $D(\boldsymbol{\lambda})$  tire ses valeurs dans un espace de même dimension que celui des données. Comme le problème inverse est généralement dégénéré alors la dimension de l'espace des données est beaucoup plus petit que celui des sources. Il y a un rapport d'un ordre de grandeur entre ces deux quantités. La minimisation numérique est alors extrêmement simplifiée pour le MEM.

### 3.2.2. Critère primal du MEM

Il est possible de trouver une expression semblable à (3.11) pour obtenir directement l'objet reconstruit  $\hat{\mathbf{q}}$  au lieu de la loi *a posteriori*  $\hat{P}$ . Définissons la transformée de Cramér  $\mathcal{F}_\mu$  de la loi  $\mu$  :

$$\mathcal{F}_\mu(\mathbf{q}) = \max_{\boldsymbol{\xi}} \{ \boldsymbol{\xi}^t \mathbf{q} - \mathcal{F}_\mu^*(\boldsymbol{\xi}) \}. \tag{3.16}$$

Dans [29], il est montré que la solution  $\hat{\mathbf{q}}$  est le minimum de cette fonction dans l'ensemble  $\mathcal{C}$  :

$$\hat{\mathbf{q}} = \underset{\mathbf{q} \in \mathcal{C}}{\operatorname{argmin}} \mathcal{F}_\mu(\mathbf{q}) \tag{3.17}$$

qui est une expression de la même forme que l'équation (3.11). On appelle ce critère la formulation primale (par opposition à la formulation duale, équation (3.14))

du MEM. Le passage de la minimisation dans l'espace des lois de probabilité (3.11) à une minimisation dans l'espace des sources (3.17) se nomme un principe de contraction [14].

On peut faire une analogie avec la mécanique statistique et la thermodynamique. En regardant la définition (équation (3.13)) de  $\mathcal{F}_\mu^*$ , nous voyons que c'est l'analogie de l'énergie libre d'Helmholtz (le logarithme de la fonction de partition). Or,  $\mathcal{F}_\mu$  est précisément défini (équation (3.16)) comme la transformée de Legendre de l'énergie libre d'Helmholtz qui est l'énergie libre de Gibbs. Ainsi, on peut voir  $\hat{\mathbf{q}}$  comme l'état qui minimise l'énergie et c'est donc un état d'équilibre. Cette minimisation est faite dans l'espace des sources  $\mathcal{C}$  et non dans l'espace des distributions de probabilité  $\mathcal{P}$ . Ainsi on perd l'avantage computationnel de la formulation duale.

Cette formulation nous permettra de voir que le MEM généralise tout un ensemble de méthodes. Dans la section 3.3, nous verrons qu'en choisissant des lois de référence  $\mu$  appropriées, nous pouvons retrouver les critères présentés dans les sections précédentes. Ceci jette un regard nouveau sur ces méthodes.

### 3.2.3. Bruit en MEM

L'équation de départ (3.8) ne prend pas en compte le bruit dans les mesures. Si des erreurs se glissent dans les mesures on écrit habituellement

$$\mathbf{b} = G \mathbf{q} + \mathbf{n}$$

où  $\mathbf{n}$  est un vecteur de  $d$  composantes qui est habituellement modélisé par une loi normale de moyenne nulle non-corrélée aux données. On peut se ramener à la forme de l'équation (3.8) en écrivant

$$\mathbf{b} = [G \ \mathbb{I}] \begin{bmatrix} \mathbf{q} \\ \mathbf{n} \end{bmatrix}.$$

En posant  $A = [G \ \mathbb{I}]$  et  $\mathbf{z}^t = [\mathbf{q}^t \ \mathbf{n}^t]$ , on retrouve une équation de la forme (3.8) :

$$\mathbf{b} = A \mathbf{z}.$$

On peut alors appliquer le MEM à la variable aléatoire  $\mathbf{Z}^t = [\mathbf{Q}^t \ \mathbf{N}^t]$  et remplacer  $\mu(\mathbf{q})$  par  $\bar{\mu}(\mathbf{q}, \mathbf{n}) = \mu(\mathbf{q})\nu(\mathbf{n})$  où la mesure de probabilité  $\nu$  est celle du bruit. La plupart du temps, une loi normale est choisie mais on pourrait très bien prendre

quelque chose de moins général si de l'information supplémentaire sur la loi du bruit est disponible. Il se pourrait aussi que le bruit soit corrélé aux sources et donc, dans ce cas, on ne pourrait factoriser la loi de référence  $\bar{\mu}$ .

### 3.3. LIENS AVEC LES AUTRES MÉTHODES DE RECONSTRUCTION

Comment se compare le MEM avec les autres méthodes de reconstruction ? Il s'avère que le MEM généralise plusieurs de ces méthodes. Tout est dans le choix de la mesure de référence  $\bar{\mu}$  et l'expression (3.17). En effet, si le bruit n'est pas corrélé aux intensités  $\mathbf{q}$ , alors on factorise la loi  $\bar{\mu}$  en deux parties :

$$\bar{\mu}(\mathbf{q}, \mathbf{n}) = \mu(\mathbf{q})\nu(\mathbf{n}).$$

Pour avoir l'estimé  $(\hat{\mathbf{q}}, \hat{\mathbf{n}})$  donné par le MEM pour cette loi de référence, on doit trouver une expression pour la transformée de Cramér de  $\bar{\mu}$ . Il s'avère que, pour une loi factorisée, la transformée de  $\bar{\mu}$  est simplement la somme des transformées de  $\mu$  et  $\nu$  :

$$\mathcal{F}_{\bar{\mu}}(\mathbf{q}, \mathbf{n}) = \mathcal{F}_{\mu}(\mathbf{q}) + \mathcal{F}_{\nu}(\mathbf{n}).$$

De plus, comme le bruit est le plus souvent modélisé comme une loi normale  $\mathbf{N} \sim N(0, \Sigma)$ , alors on peut trouver une expression explicite pour

$$\mathcal{F}_{\nu}(\mathbf{n}) = \frac{1}{2} \mathbf{n}^t \Sigma^{-1} \mathbf{n}.$$

Donc, le critère (3.17) devient

$$\hat{\mathbf{z}} = (\hat{\mathbf{q}}, \hat{\mathbf{n}}) = \underset{(\mathbf{q}, \mathbf{n}) \in \mathcal{C}}{\operatorname{argmin}} \frac{1}{2} \mathbf{n}^t \Sigma^{-1} \mathbf{n} + \mathcal{F}_{\mu}(\mathbf{q})$$

où

$$\mathcal{C} = \{(\hat{\mathbf{q}}, \hat{\mathbf{n}}) \in \mathbb{R}^{N+d} \mid \mathbf{b} = G \mathbf{q} + \mathbf{n}\}.$$

Comme  $\mathbf{n}$  est simplement l'écart aux données ( $\mathbf{n} = \mathbf{b} - G \mathbf{q}$ ), on peut écrire

$$\hat{\mathbf{q}} = \underset{\mathbf{q}}{\operatorname{argmin}} \frac{1}{2} (\mathbf{b} - G \mathbf{q})^t \Sigma^{-1} (\mathbf{b} - G \mathbf{q}) + \mathcal{F}_{\mu}(\mathbf{q}).$$

Dans ce dernier calcul, le lecteur pourrait imaginer que les méthodes bayésiennes ne sont qu'un simple cas particulier du MEM puisque la dernière expression est semblable à (2.7). Or, il existe quelques différences conceptuelles importantes entre ces deux [34]. Tout d'abord, l'objet à reconstruire  $\hat{\mathbf{q}}$  est la moyenne d'une loi de probabilité pour le MEM alors que, pour le formalisme bayésien, il est

un échantillon d'un vecteur aléatoire pour lequel une loi de probabilité *a priori* est définie. Cette différence très importante se retrouve dans les bases conceptuelles de la mécanique statistique. Pour cette dernière, nous décrivons les propriétés macroscopiques d'un système par la valeur moyenne sur tout l'espace des phases. Or rien ne justifie *a priori* ce choix. On aurait pu tout aussi bien prendre la valeur qui a la probabilité maximale. Mais, le grand nombre de particules fait en sorte que les distributions de probabilité des quantités thermodynamiques du système possèdent un mince pic très prononcé autour de la valeur moyenne à cause de la loi des grands nombre rendant les deux versions équivalentes [22]. Cependant, rien ne justifie un tel choix pour des problèmes qui n'ont pas cette propriété. C'est le cas de certains problèmes discrets pour lesquels la moyenne n'est pas un élément du domaine de départ (en analyse d'images binaires, par exemple). Dans ces cas, le MEM ne peut être utilisée alors que l'approche bayésienne reste applicable.

Puisque l'origine du MEM provient de la mécanique statistique, on peut donner une interprétation physique à la transformée de Cramér  $\mathcal{F}_\mu$ . En effet, par comparaison avec les expressions de la thermodynamique,  $\mathcal{F}_\mu^*$  peut être vu comme le logarithme de la fonction de partition qui est l'énergie libre d'Helmholtz. Sa transformée de Legendre  $\mathcal{F}_\mu$  est l'énergie libre de Gibbs du système et comme il n'y a aucune cinétique, cette énergie provient seulement du potentiel. Cette interprétation est corroborée par la similarité entre le MEM et le formalisme bayésien. Dans l'équation (2.7), le potentiel de Gibbs  $U(\mathbf{q})$  est minimal au point d'équilibre du système ce qui correspond à l'énergie libre de Gibbs  $\mathcal{F}_\mu$  qui doit être minimale à l'équilibre.

L'avantage du MEM se manifeste dans la formulation duale qui permet de trouver l'objet à estimer  $\hat{\mathbf{q}}$  en optimisant la fonction  $D(\lambda)$  qui prend ses valeurs dans un espace de dimension  $d$  plutôt que de minimiser une fonction dans l'espace des  $\mathbf{q}$  qui est de dimension  $N$ ;  $N$  est en général beaucoup plus grand que  $d$  pour les problèmes inverses mal posés (ce qui est évidemment le cas en MEG). Cette distinction est aussi importante au niveau conceptuel. En ayant  $d$  équations qui définissent les contraintes alors qu'il y a  $N$  variables à fixer, il ne reste que  $N - d$  véritables degrés de liberté dans le problème. La minimisation se faisant dans l'espace dual (ou des contraintes), cela permet de travailler dans un espace restreint et ensuite de laisser l'entropie propager cette information dans les degrés de liberté restants de façon à n'introduire aucun biais, c'est-à-dire la distribution

avec l'entropie maximale (et donc de désordre maximal). Dans le formalisme bayésien, il n'est pas possible de transposer cette argumentation puisqu'il n'y a pas de formulation duale.

De plus, la fonction  $D(\lambda)$  est convexe pour pratiquement toute mesure de probabilité de référence  $\mu$  (voir [29] pour des critères plus précis). Cette propriété nous assure qu'il existe un seul minimum local qui correspond au minimum global de  $D(\lambda)$ , cette propriété ne dépendant pas de  $\mu$ . Ceci rend les calculs numériques plus fiables puisqu'on assure ainsi de trouver la solution optimale sans risque de tomber dans un minimum local. Dans le formalisme bayésien, ce n'est pas le cas en général puisque la propriété de convexité pour  $U(\mathbf{q})$  n'est pas toujours satisfaite et impose des contraintes fortes sur le choix de  $U(\lambda)$ , contraintes qui n'existent pas lors du choix de la loi de référence  $\mu$  en MEM.

Toutefois, le MEM ne résout pas tous les problèmes. Par exemple, il est difficile de manipuler les hyperparamètres de la loi de référence et de les estimer à l'aide des données, nous traitons dans cette thèse ce problème par des méthodes différentes. Pour les méthodes bayésiennes, des outils ont déjà été développés [34]. Une autre différence importante est que le MEM ne peut prendre en compte que des contraintes convexes alors que pour le formalisme bayésien, ce problème ne se pose pas. Par contre, le MEM, à l'aide de la loi de référence  $\mu$ , peut respecter facilement des contraintes comme la positivité en restreignant le domaine de définition de  $\mu$ .

Ainsi le MEM et les méthodes bayésiennes permettent d'attaquer des problèmes de façon différente. L'interprétation des *a priori* et des résultats est différente pour chacune des méthodes. Les méthodes bayésiennes ont déjà été appliquées à différents problèmes ce qui n'est pas encore le cas pour le MEM. Par exemple, le problème inverse en MEG a été traité avec l'approche bayésienne dans plusieurs travaux comme nous l'avons souligné au chapitre précédent. Pour le MEM toutefois, il n'y a que peu de travaux appliqués à la MEG. Ils seront présentés dans le prochain chapitre qui applique le MEM à la MEG.



# Chapitre 4

---

## MAXIMUM D'ENTROPIE ET MAGNÉTOENCÉPHALOGRAPHIE

### Premier article

Au chapitre 1, il a été établi que le problème inverse en MEG pouvait être formulé linéairement dans le contexte des méthodes distribuées. Au chapitre 3, nous avons présenté le principe de la maximisation de l'entropie comme méthode de résolution de problèmes inverses linéaires. Il est maintenant temps d'appliquer cette dernière à la MEG.

Les travaux présentés dans ce chapitre ont fait l'objet d'un poster présenté à Brighton en Angleterre lors de la conférence *Human Brain Mapping* en juin 2001 [26]. Puis, l'article inclus dans cette thèse a été publié dans la revue *IEEE Transactions on Biomedical Engineering*, volume 51, numéro 3, en mars 2004, pages 427–442 [1].

Cet article démontre l'utilité du MEM appliqué au problème linéaire inverse en MEG. Comme nous l'avons montré dans le chapitre précédent, l'ingrédient principal du MEM est une loi de probabilité de référence qui permet d'introduire des informations anatomiques et fonctionnelles sur la solution dans la résolution du problème inverse.

Le choix de cette loi est donc crucial. Pour la définir, quelques hypothèses sont faites sur l'activation des dipôles :

- (1) l'activation d'un dipôle ne peut être ambiguë : il est pleinement activé ou non,
- (2) les dipôles sont regroupés en région et les dipôles dans une même région sont activés ou non tous à la fois.

Un modèle semblable a déjà été utilisé dans un contexte bayésien. Dans [40], les auteurs introduisent un processus binaire qui pondère l'intensité qui, lorsqu'il est

estimé avec les données, permet d'identifier les dipôles activés sans se soucier de l'intensité de l'activation ce qui n'est pas sans rappeler la MSP présentée dans la section 2.2.5.

Ces hypothèses sont introduites dans la loi de référence à l'aide de variables aléatoires discrètes de Markov cachées (c'est-à-dire non-observées) qui définissent l'état d'activation de chaque région. Un modèle bâti à partir de variables cachées a déjà été développé dans le contexte de l'IRMf [19]. Cette modélisation permet de séparer l'état d'activation de l'intensité du dipôle. Un modèle de mixture est présenté et permet de donner des corrélations entre les dipôles non pas directement sur les intensités mais entre les états d'activation. La spécification de la loi de référence est grandement simplifiée par cette modélisation discrète des corrélations.

Dans le modèle présenté ici, la deuxième hypothèse va encore plus loin puisqu'on impose les corrélations entre les régions et non directement sur les dipôles. Nous introduisons ces corrélations à l'aide de modèles graphiques sur les variables cachées. Un modèle graphique est une façon très visuelle de représenter une loi de probabilité [49]. Pour inférer correctement cette information à la solution, il faut pouvoir donner une estimation de ces corrélations. Elles ont été l'objet de travaux en IRMf [30] où les auteurs estiment les corrélations entre pixels et estiment le graphe représentant le mieux ces corrélations. Cependant, la correspondance entre les signaux générateurs de l'IRMf (signal neurochimique) et les courants électriques neuronaux n'est pas encore parfaitement comprise. Alors, l'intégration des modèles graphiques issus de l'IRMf à des modèles en MEG est très délicate. Ceci ne sera pas fait dans cette thèse mais des travaux sur ce sujet sont en cours dans notre groupe de recherche. Dans ce travail, les régions sont simplement considérées comme statistiquement indépendantes. Cependant, des corrélations subsistent tout de même entre les dipôles d'une même région puisque la variable d'activation décrit l'activation de tous les dipôles appartenant à la région.

Le MEM ainsi défini permet d'estimer non seulement les intensités de chaque dipôle mais aussi la probabilité d'activation de chaque région via la loi *a posteriori* donnée par le MEM. Cette dernière permet de définir une nouvelle loi de référence et ainsi de réestimer les intensités des dipôles avec cette nouvelle loi. Suivant ce

raisonnement, un algorithme itératif a été développé et permet d'améliorer les performances de reconstruction du MEM.

D'autres travaux ont été publiés sur l'application du principe du maximum d'entropie au problème inverse en MEG. Dans [20], les valeurs des intensités sont ramenées entre 0 et 1 ce qui permet d'interpréter l'objet à reconstruire comme une distribution de probabilité. L'entropie utilisée est donc calculée directement avec les intensités :

$$S = - \sum_i q_i [\log(q_i/A) - 1].$$

Or cette entropie ne permet pas de reconstruire des dipôles dont l'intensité  $q_i$  est plus petite que 0. Pour contourner ce problème, les auteurs définissent une nouvelle entropie qui n'est plus vraiment reliée à l'entropie de Shannon et donc leur principe de maximisation perd le sens que lui donne la théorie de l'information. De plus, leur formalisme permet difficilement d'introduire des corrélations entre les dipôles comme cela est fait dans cette thèse avec les régions d'activation.

Un travail plus raffiné a été fait par Clarke *et al.* dans [7, 8, 21]. Leur formulation diffère de celle présentée dans cette thèse en plusieurs points. Tout d'abord, ils présentent le problème inverse pour retrouver la densité de courant  $\mathbf{J}^p$  et non seulement une approximation par dipôles. De plus, ils n'incluent aucune information physiologique sur la solution. Enfin, les données sont utilisées par l'ajout d'une contrainte d'écart global de la forme

$$\chi^2 = \|\mathbf{b} - G\mathbf{q}\|^2.$$

La contrainte n'est plus sur chacune des mesures  $\mathbf{b}$ . Ces travaux démontrent toutefois la puissance et la versatilité du MEM puisqu'il peut être utilisé dans divers contextes sous différentes hypothèses. Dans ce travail, nous avons introduit des corrélations spatiales pour régulariser le problème inverse.

# Biomagnetic Source Detection by Maximum Entropy and Graphical Models

publié sous *IEEE Transactions on Biomedical Engineering*, volume 51, numéro 3, mars 2004, pages 427–442.

ERVIG LAPALME, *Centre de Recherches Mathématiques et Département de Physique de l'Université de Montréal*

CÉCILE AMBLARD, *Laboratoire LabSAD, Grenoble, France*

JEAN-MARC LINA, *Centre de Recherches Mathématiques et Département de Génie Électrique de l'École de Technologie Supérieure, Montréal*

## Abstract

This article presents a new approach for detecting active sources in the cortex from magnetic field measurements on the scalp in magnetoencephalography (MEG). The solution of this ill-posed inverse problem is addressed within the framework of Maximum Entropy on the Mean (MEM) principle introduced by Clarke and Janday. The main ingredient of this regularization technique is a reference probability measure on the random variables of interest. These variables are the intensity of current sources distributed on the cortical surface for which this measure encompasses all available prior information that could help to regularize the inverse problem. This measure introduces hidden Markov random variables associated with the activation state of predefined cortical regions. MEM approach is applied within this particular probabilistic framework and simulations show that the present methodology leads to a practical detection of cerebral activity from MEG data.

**Keywords:** Inverse Problem, Hidden Markov Models, Maximum Entropy on the Mean, Magnetoencephalography.

## 4.1. INTRODUCTION

Most recent theoretical works in neuroscience about brain function emphasize either localization or distributed properties of the brain activity. Despite the high level of accuracy in spatial localization of activation sources, the usual fMRI (functional magnetic resonance imaging) suffers from its poor time resolution and the lack of a precise relation between the detected blood flow and true neuronal activity. Still, fMRI delivers a huge quantity of anatomical and physiological information about the brain. With recent progress in the measurement of very low intensity magnetic fields, magnetoencephalography (MEG) has become a valuable functional brain imaging technique. Here, superconductive coils measure the magnetic field outside the head generated by electric activity (modelled with local macro-currents). In contrast to fMRI, this technique offers very good time resolution but poor spatial resolution. The aim of the present work is to develop a methodology for solving the so-called “MEG inverse problem”, *i.e.* a suitable framework for finding the activation sources that explain the MEG data and makes use of reliable *prior* information coming from other cerebral imaging modalities as, for instance, the fMRI.

The need for external information about the solution is unavoidable. Among various sources making this inverse problem ill-posed are an infinite number of electric current distributions which can lead to the same MEG measurements. In order to reduce the space of source distributions and to be able to select the best candidate, we must constrain the solution by using regularization criteria. Thus, a solution of the “MEG inverse problem” should address the three following components: the origin of the data, the constraints and the criterion that defines the “best compromise” between data explanation and *prior* knowledge expressed through the constraints.

Following previous investigations based on a probabilistic description of current intensities distributed on the cortical surface [3,4], we consider a model based on Markovian hypotheses [1]. Our main contribution is to express this hypothesis in terms of hidden random variables that describe the local activation state. This set of discrete variables characterizes the direction (“outgoing state” or “ingoing state”) of the currents or the absence of activity (“silent” state) inside a region. The states are not directly observed but play the role of explanatory variables for the current intensities located inside the regions. We claim that this probabilistic

model encompasses most of the prior information coming from other modalities, either anatomical information (the spatial support of the probabilistic law on the amplitudes) or functional information (the structure of the joint probabilistic law on the hidden state variables).

Maximum Entropy will be the regularization scheme of the present work. Mainly due to Jaynes's seminal works in the 50's [11,12], this general framework infers a probabilistic law that maximizes the missing information. Zero entropy being reached by a predefined probabilistic law (in general the uniform law), the Maximum Entropy solution consists in modifying this reference law in such a way that only available information is taken into account. As a particular case, we show that the Maximum Entropy on the Mean (MEM) formalism is suitable to the MEG problem. We present this formalism as an integration tool from which we can infer both the currents' amplitude and the hidden activation states subjected to spatial correlations through a graphical model. Being explicitly written in the reference law, this graphical model expresses prior knowledge that can be extracted from fMRI [15] or other functional analysis modalities.

The structure of the paper is as follows. MEG direct and inverse problems are briefly presented in Section 2. In Section 3, we review the principle of the Maximum Entropy on the Mean; emphasis is on the relevance of this inference technique for solving an ill-posed problem with *prior* knowledge about the solution. We also discuss the relationship between the present formalism and the probabilistic approach advocated by Clarke et al. [4].

In Section 4, we propose a probabilistic framework for the *prior* knowledge available in cerebral imaging. As said previously, the main novelty in this model is the use of a hidden state variable that describes local activation which *explains* the observation *via* the associated dipole intensities. We progressively refine the statistical dependencies between those explanatory variables. The simplified version is a Finite Mixture model that is evaluated in Section 5. An iterative scheme based on the hidden activation state variables is also introduced and is shown to solve the detection problem with accuracy and minimal *prior* knowledge about the smoothness of the solution. Finally, we prove the efficiency of the method by evaluating the reconstruction and detection performance on simulated data.

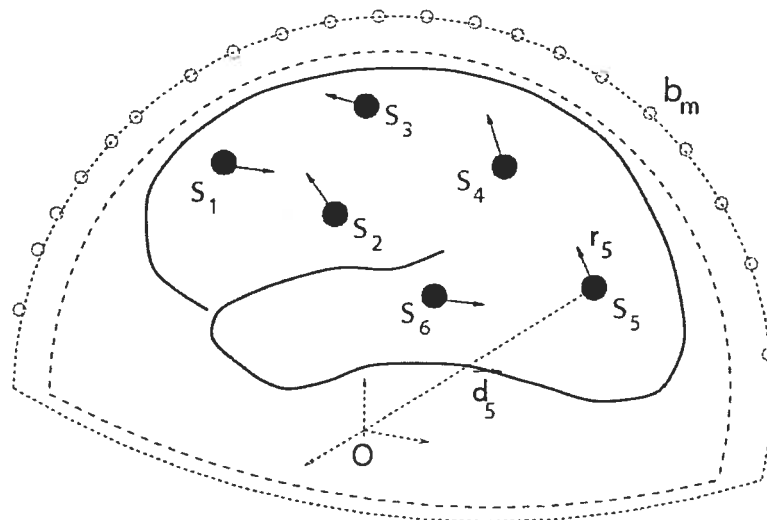


FIGURE 4.1. Schematic representation of the source brain model. Dipoles are located at sites “ $i$ ”. Position ( $d_i$ ) and orientation of the dipoles are fixed; the dipole intensities (or amplitudes)  $r_i$  are random quantities to be estimated from the magnetic field  $\mathbf{b}$ . The present approach consists in introducing *latent* (or *hidden*) state variables  $S_i$  that describe locally the activation state of the cortical strip.

## 4.2. THE FORWARD AND INVERSE PROBLEMS

In MEG, data amount to a finite number of measurements of the magnetic field induced by the electric neuronal activity and collected outside the head. Data are linked to neuronal activity by Maxwell’s equations. The quasistatic approximation of Maxwell’s equations is sufficient since the neuromagnetism frequencies are below 100 Hz so that terms containing a time derivative are negligible. In this case, the sensitivity of a detector to current distribution can be summarized in a lead field function which depends (non-linearly) on the position and orientation of the current sources and detectors [19]. However, the magnetic field is a linear function of the current sources’ amplitudes. In a spherical approximation of the conductor, radial sources do not induce magnetic fields on the surface of the conductor. The presence of non-vanishing currents in the kernel of the linear operator that maps the current sources to the external magnetic field is the principal reason why this problem is said ill-posed when inversion is considered.

The estimation of the current distribution from a discrete finite set of data is impossible. As illustrated in Fig.4.1, the problem can be simplified by assuming that there is a finite number of sources, say  $d$ . Typically, when looking for the activity of a set of neurons (a macrocolumn [9]), sources are represented by a dipolar current. Each dipole is characterized by six parameters: position (3), orientation (2) and intensity (1). We reduce the number of parameters to be estimated by assuming that the dipoles' positions and the orientations are known while leaving their intensity free. Within such a distributed sources model, the problem can be written as a linear system:

$$\mathbf{b} = A \mathbf{r}, \quad (4.1)$$

where the  $n$ -dimensional vector  $\mathbf{b}$  is the magnetic field,  $n$  being the number of sensors. The  $d$ -dimensional vector  $\mathbf{r}$  represents the source intensities and operator  $A$  the lead field matrix (we use bold face notation for vectors or multivariate quantities and ordinary face for scalar quantities and matrices). Each column of  $A$  represents the contribution of a particular dipole to the magnetic field on the sensors. The larger the number of considered dipoles is, the more accurate the reconstruction will be. However, for technical reasons, the number of measurements, say  $m$ , is currently limited to a hundred whereas the robustness of the distributed sources model relies on a very large number of dipoles typically a few thousands. This is the second reason for the "ill-posedness" of this inverse problem.

A third source of difficulty in solving this problem is the presence of randomness in the data. Assuming an additive measurement noise modeled by an *i.i.d* random contribution with zero mean, the following process equation

$$\mathbf{b} = A \mathbf{r} + \epsilon, \quad (4.2)$$

must be considered rather than (4.1).

Let us denote by  $\mathbf{b} = \mathbf{m}$  some MEG observations. Most of the statistical techniques for MEG sources reconstruction amounts to minimizing a parametrized cost function, *i.e.*

$$\tilde{\mathbf{r}} = \arg \min_{\mathbf{r}} U(\mathbf{r}, \mathbf{b} = \mathbf{m}). \quad (4.3)$$



In general,  $U$  is composed of two terms: a goodness-of-fit expression where data attachment is imposed and a weighted regularization term that enforces smoothness and thus constraints the set of possible solutions,

$$U(\mathbf{r}, \mathbf{b} = \mathbf{m}) = \|\mathbf{m} - A\mathbf{r}\|_{\Sigma}^2 + \lambda \|\mathbf{r}\|_M^2 \quad (4.4)$$

The  $n \times n$  matrix  $\Sigma$  and  $d \times d$  matrix  $M$  are metrics for the data and the model parameters space respectively ( $\|\mathbf{r}\|_M^2 = \mathbf{r}^t M \mathbf{r}$ , for instance).  $\lambda$  is the usual regularization parameter that determines the relative strength of the smoothing constraint expressed in  $M$ . This framework leads, in general, to least-squares solutions of the form  $\tilde{\mathbf{r}} = \mathbf{G} \mathbf{m}$  where the  $n \times d$  matrix  $\mathbf{G}$  is the generalized inverse of the  $d \times n$  operator  $A$ :

$$\mathbf{G} = \Sigma^{-1} A^t [A \Sigma^{-1} A^t + \lambda M^{-1}]^{-1} \quad (4.5)$$

Most of the inverse solutions developed for the MEG problem are of this form. Among them, LORETA [17] and S-MAP [1] make use of slightly different regularization criteria but both are based on the weighted Laplacian of the currents.

Due to the absence in the model of prior information on local cortical activity and correlation between regional simultaneous activity, these techniques do not necessarily lead to a unique realistic solution. In the next section, we revisit the probabilistic approach of the Maximum Entropy technique already advocated by Clarke [3, 4] with special emphasis on the *prior* information we may introduce in order to regularize the MEG inverse problem.

### 4.3. PRIMAL AND DUAL FORMULATIONS OF THE MEM

The magnetic field  $\mathbf{b}$  is measured with sensors at various locations outside the head, during  $N$  repetitions of a given task or stimuli. After a convenient registration of the  $N$  time sequences for each sensor, signals are averaged through a simple normalized summation. The resulting time sequences for each sensor constitute the data. More precisely, the present work deals with data at a specific time. Denoted by  $\mathbf{m}$ , these data are thus sample statistics of the mean value of the instantaneous magnetic field on the sensors and, for  $N$  large enough, data is assumed to be equal to the first moment, *i.e.*  $\mathbf{m} = E[\mathbf{b}]$  or

$$\mathbf{m} = A E[\mathbf{R}] \quad (4.6)$$

if  $\mathbf{R}$  denotes a multidimensional continuous random variable that describes the intensities of  $d$  dipoles distributed on the cortical surface. Finally, a residual zero-mean Gaussian noise may be added to insure a more realistic description of the data, *i.e.*  $\mathbf{m} = E[\mathbf{b}] + \epsilon$  or

$$\mathbf{m} = A E[\mathbf{R}] + \epsilon \quad (4.7)$$

This experimental framework of the MEG supports the forthcoming probabilistic method to infer dipoles intensity from the MEG data.

#### 4.3.1. The Maximum of Entropy on the Mean

$\mathbf{R}$  being the  $d$ -dimensional continuous random variable that describes the dipole intensities, this variable has the probability law  $dp(\mathbf{r}) = \rho(\mathbf{r}) d\mathbf{r}$  where  $\mathbf{r} \in \mathbb{R}^d$ . Let us denote by  $d\mu(\mathbf{r}) = \rho_0(\mathbf{r}) d\mathbf{r}$  the law that represents *all* (or, at least, most of) the knowledge about the current distribution *in absence* of the MEG data.  $\rho_0$  describes the *priors* and, without any further information, this density is the best description of the currents. Given this reference probability distribution  $d\mu(\mathbf{r})$ , we consider the Kullback's  $\mu$ -entropy defined by [13]

$$S_\mu(dp) = - \int \log \frac{dp}{d\mu} dp = - \int f(\mathbf{r}) \log f(\mathbf{r}) d\mu(\mathbf{r}) \quad (4.8)$$

where  $f$  is the  $\mu$ -density of  $dp$ , *i.e.*  $dp(\mathbf{r}) = f(\mathbf{r})d\mu(\mathbf{r})$  and  $S_\mu(dp) = -\infty$  otherwise. Being a pseudo-distance between the reference distribution  $d\mu$  and any  $\mu$ -density  $dp$ , this entropy is always negative,

$$-\infty \leq S_\mu(dp) \leq S_\mu(d\mu) = 0,$$

and measures the amount of new information brought by the data, with respect to the prior. It will be used as a criterion for the choice of the probability distribution.

In order to take into account the MEG data, of the form (4.6) for instance, we define by  $\mathcal{C}_m$  the set of probability density measures on  $\mathbf{R}$  that “explain” the observations in average:

$$dp \in \mathcal{C}_m : \quad \mathbf{m} - A E_p[\mathbf{R}] = 0. \quad (4.9)$$

This is connected with what we said in the beginning of Section 3. The underlying process of dipole activation is modeled by the multivariate random variable  $\mathbf{R}$ . During a MEG experiment, several realizations of this process are made

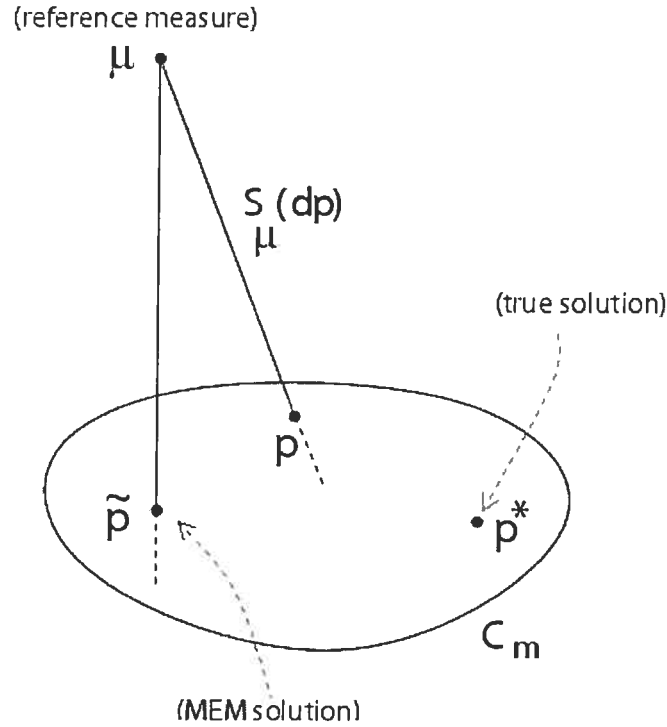


FIGURE 4.2. MEM: the solution  $\tilde{p}$  is the “closest” density to the reference law ( $\mu$ ) that explains the data (the “true solution” is also in  $C_m$ ).

and the measurements are collected. Then, we reconstruct the activity with the mean of the trials. This is what we are doing here in defining the set  $C_m$  in the equation (4.9).

As illustrated in Fig.4.2, the “Maximum Entropy on the Mean” solution [8] consists in selecting  $d\tilde{p}$  in  $C_m$  that maximizes the  $\mu$ -entropy (4.8) and thus maximizes the missing information:

$$d\tilde{p} = \arg \max_{dp \in C_m} \mathcal{S}_\mu(dp). \quad (4.10)$$

This solution, if it exists, is unique and describes only what we “know”: the prior knowledge encompassed in the reference measure  $\mu$ , and the measurements that define the space  $C_m$  as in (4.9). The estimated sources are then simply given by

$$\tilde{\mathbf{r}} = E_{\tilde{p}}[\mathbf{R}]. \quad (4.11)$$

It is worth stressing the role of  $\mu$  since it characterizes, as much as the constraints that define  $C_m$ , the optimal solution in (4.11). The optimal criterion that solve

the trade-off between goodness-of-fit and prior knowledge is quite similar to the familiar Bayesian point of view [16].

The  $\mu$ -entropy is a strictly convex function of  $p$  to be maximized under constraint. From a Lagrangian point of view, such an optimization is equivalent to the maximization of an unconstrained strictly concave function [7, 18]. Indeed, from the Lagrangian associated with (4.10),

$$L(p, \boldsymbol{\lambda}, \lambda_0) = \int f(\mathbf{r}) \log f(\mathbf{r}) d\mu(\mathbf{r}) + \boldsymbol{\lambda}^t (\mathbf{m} - AE_p[\mathbf{R}]) + \lambda_0 \left( 1 - \int dp(\mathbf{r}) \right),$$

where  $\boldsymbol{\lambda}^t$  denotes as usual the transpose of the  $n$ -dimensional Lagrange multiplier vector  $\boldsymbol{\lambda}$ ; the optimal solution  $(\tilde{f}, \tilde{\boldsymbol{\lambda}}, \tilde{\lambda}_0)$  is then given by

$$L(\tilde{f}, \tilde{\boldsymbol{\lambda}}, \tilde{\lambda}_0) = \min_f \max_{\boldsymbol{\lambda}, \lambda_0} L(f, \boldsymbol{\lambda}, \lambda_0). \quad (4.12)$$

This solution defines an *a posteriori* density law in the exponential family of the reference law

$$d\tilde{p}(\mathbf{r}) = \exp \left( \tilde{\boldsymbol{\lambda}}^t A \mathbf{r} - F_\mu^*(A^t \tilde{\boldsymbol{\lambda}}) \right) d\mu(\mathbf{r}), \quad (4.13)$$

where  $F_\mu^*$  is the log of the partition function associated with the distribution  $d\mu$ ,

$$F_\mu^*(\boldsymbol{\xi}) = \log \int \exp(\boldsymbol{\xi}^t \mathbf{r}) d\mu(\mathbf{r}),$$

and  $\tilde{\boldsymbol{\lambda}}$  is the value that maximizes the Lagrange dual criteria:

$$\tilde{\boldsymbol{\lambda}} = \arg \max_{\boldsymbol{\lambda}} D(\boldsymbol{\lambda}) \quad (4.14)$$

with

$$D(\boldsymbol{\lambda}) = \boldsymbol{\lambda}^t \mathbf{m} - F_\mu^*(A^t \boldsymbol{\lambda}). \quad (4.15)$$

This dual formulation (4.14) of the primal problem (4.10) is computationally more attractive since,  $F_\mu^*$  being strictly convex if  $\mu$  is maximal [14], the dual criteria is strictly concave and leads to a unique maximum  $\tilde{\boldsymbol{\lambda}}$ . The “primal” solution  $\tilde{\mathbf{r}}$  given by (4.11) is linked to the “dual” solution through the so-called primal-dual relation:

$$\tilde{\mathbf{r}} = \nabla F_\mu^*(\boldsymbol{\xi})|_{\boldsymbol{\xi}=A^t \tilde{\boldsymbol{\lambda}}}. \quad (4.16)$$

Therefore, one can find a solution of an inverse problem by MEM using a classical optimization algorithm for which complexity does not depend on the prior law  $d\mu$ .

### 4.3.2. Noisy observations

As previously mentioned, observations are corrupted with noise and we may have to consider (4.7) instead of (4.6). MEM takes this random component into account (see for instance [14]) by considering noise as an extra source in the problem and solving the previous optimization with

$$dp \in \mathcal{C}_m : \quad \mathbf{m} - [A | I] \begin{bmatrix} E_p[\mathbf{R}] \\ \epsilon \end{bmatrix} = 0 \quad (4.17)$$

instead of (4.9).  $I$  is the  $n$  by  $n$  identity matrix. Assuming the noise is independent of the source intensities, the solution takes the form

$$dp(\mathbf{r}, \epsilon) = \exp \left( \tilde{\lambda}^t A \mathbf{r} - F_\mu^*(A^t \tilde{\lambda}) - F_{\mu_n}^*(\lambda) \right) d\mu(\mathbf{r}) d\mu_n(\epsilon) \quad (4.18)$$

where  $d\mu_n(\epsilon)$  is a known mesasure that describes the noise.  $\tilde{\lambda}$  is still defined by (4.14) but the dual criteria get an extra term

$$D(\lambda) = \lambda^t \mathbf{m} - (F_\mu^*(A^t \lambda) + F_{\mu_n}^*(\lambda)). \quad (4.19)$$

Finally, the MEM estimate of the random variable of interest is still given by (4.16).

### 4.3.3. MEM and Maximum Entropy formalisms.

The present formalism has the same foundations as the principle of Maximum Entropy discussed in [3, 4, 10] where the authors also consider the problem of finding a posterior density on the dipolar currents that explains the data. While both approaches are regularization techniques based on statistical entropy, they differ mainly in the definition of goodness-of-fit criteria or, in the language of the present work, they adopt the following definition of the set  $\mathcal{C}_m$ :

$$dp \in \mathcal{C}_m : \quad E [(\mathbf{m} - A\mathbf{R})^t (\mathbf{m} - A\mathbf{R})] = e \quad (4.20)$$

where  $e$  is some prescribed  $\chi^2$  error on the observations. Following the same reasoning as before, the Maximum Entropy solution is found to be of the form

$$d\tilde{p}(\mathbf{r}) = \exp \left( \tilde{\lambda} (\mathbf{b} - A\mathbf{r})^t (\mathbf{b} - A\mathbf{r}) - F^*(\tilde{\lambda}) \right) d\mu(\mathbf{r}) \quad (4.21)$$

where

$$F^*(\lambda) = \log \int \exp(\lambda(\mathbf{b} - A\mathbf{r})^t(\mathbf{b} - A\mathbf{r})) d\mu(\mathbf{r}),$$

and optimality is reached by  $\tilde{\lambda}$ , the solution of

$$\left. \frac{\partial F^*}{\partial \lambda} \right|_{\tilde{\lambda}} + e = 0.$$

Therefore, the main difference between the present MEM approach and [3] is in the definition of the set of probability  $\mathcal{C}_m$  in which the optimal solution is looked for. The two works also differ on the role of the reference measure: in [4],  $d\mu(\mathbf{r})$  is uniform or tends to be so by some limit process. This differs from the present formalism that emphasizes the statistical priors encompassed in the reference measure  $d\mu(\mathbf{r})$ .

#### 4.4. GRAPHICAL STATISTICAL MODELS AND MEM

As suggested by work in neuroscience [9], activation sources are organized in sets of macrocolumns perpendicular to the cortex surface. To take into account this *prior* information, we assume that the dipoles are grouped together in sets  $P_k$ ,  $k = 1, 2, \dots, K$ . Such a set represents a specific *cortical region* or *cortical parcel* characterized by a specific response under a given stimulus. In order to represent this activation state, we define a hidden random variable  $S_k$  for each set. This latent variable is not observable: it defines the common state of all  $d_k$  dipoles belonging to the region  $P_k$ .

##### 4.4.1. General Case

Let us assume  $s$  possible states,  $S_k = s_k \in \{0, 1, \dots, s-1\}$ , where  $s_k = 0$  represents a region with *silent* dipoles and  $s_k \neq 0$  describes various activation states of the region  $P_k$ . We note by  $\pi(S)$  the joint probability law of the  $K$ -dimensional state vector  $S$ . Thus, the joint reference probability measure on the dipole intensities can be written as:

$$d\mu(\mathbf{r}) = \sum_{\{S\}} d\mu(\mathbf{r}, S) = \sum_{\{S\}} d\mu(\mathbf{r}|S) \pi(S), \quad (4.22)$$

where the summation over  $\{S\}$  runs over all configurations of the multivariate random state  $S = \{S_1, S_2, \dots, S_K\}$ . If we further assume that the dipole amplitude depends only on the state  $S_k$  of the region  $P_k$  to which it belongs, then the joint conditional probability law factorizes as:

$$d\mu(\mathbf{r}|S) = \prod_{k=1}^K d\mu(\mathbf{r}_k|S_k), \quad (4.23)$$

where  $\mathbf{r}_k$  is the  $n_k$ -dimensional vector of the amplitude of the sources belonging to  $P_k$ . This model allows us to express a statistical coupling between sources through hidden states, *i.e.* through the joint probability  $\pi(S)$ . Let us denote by

$$F_{S_k}^*(\boldsymbol{\xi}_k) = \log \int \exp(\boldsymbol{\xi}_k^t \mathbf{r}_k) d\mu(\mathbf{r}_k|S_k), \quad (4.24)$$

the log-partition function of the conditional law for source amplitudes, associated with the region  $P_k$ . Since the log-Laplace transform is additively separable whenever  $\mu$  is multiplicatively separable, the log-partition function of the joint measure is

$$F_\mu^*(\boldsymbol{\xi}) = \log \left[ \sum_{\{S\}} \pi(S) \exp \sum_{k=1}^K F_{S_k}^*(\boldsymbol{\xi}_k) \right]. \quad (4.25)$$

This model leads naturally to the notion of scale. At the finest scale, *i.e.* the dipole scale, statistics are described by the conditional joint probability density  $\mu(\mathbf{r}_k|S_k)$  while at the larger scale of the regions, statistics are modeled through the latent variable  $S_k$  and joint probability  $\pi(S)$ . Now, let us investigate two specific cases for which we define the regions  $P_k$  and the statistical dependencies between them.

In view of the forthcoming analysis with real MEG data, we introduce the notion of *cerebral connectivity* in the model. First, at the scale of some particular region  $P_k$ , we consider a non-trivial joint probability density for the multivariate random amplitudes  $\mathbf{r}_k$  of the dipoles that “live” in  $P_k$ , given the state  $S_k = s_k \in \{0, 1, \dots, s-1\}$  of  $P_k$ . These probability densities will take into account the local statistical dependencies (binding) and the prior knowledge about average current intensities for different activation states. Second, at the larger scale of the hidden variable  $S_k$ , an undirected graph (or graph of dependencies) as illustrated in Fig.4.3, will encompass *prior* information about functional connectivity between

the cortical regions. This graph is a part of the priors (coming from the fMRI for instance, [15]) but could contain some unknown hyperparameters.

Basically, a graphical model is a *structure* (nodes and edges) that specifies a *set of conditional independence relations* between random variables: if the variable  $S_i$  does not depend directly on the variable  $S_j$ , then no edge connects the corresponding nodes in the graph. Therefore, independence assumptions are represented by graph connectivity. Given a graph and denoting by  $\gamma(k)$  the set of “parents” for the node  $k$  (a directed edge connects a parent to its child), the conditional law  $\pi(S_k|S_1, S_2, S_3, \dots, S_{k-1}, S_{k+1}, \dots, S_K)$  reduces to  $\pi(S_k|S_{n \in \gamma(k)})$  and we write

$$\pi(S) = \prod_{k=1}^K \pi(S_k|S_{n \in \gamma(k)}). \quad (4.26)$$

Note that  $\gamma(k)$  may be empty in which case  $\pi(S_k|S_{n \in \gamma(k)})$  is simply  $\pi(S_k)$ . With the model (4.26), the prior law is

$$d\mu(\mathbf{r}) = \sum_{\{S\}} \prod_{k=1}^K \pi(S_k|S_{n \in \gamma(k)}) d\mu(\mathbf{r}_k|S_k). \quad (4.27)$$

We obtain the following log-partition function of the reference law in the dual formulation of MEM

$$F_\mu^*(\boldsymbol{\xi}) = \log \sum_{\{S\}} \prod_{k=1}^K \pi(S_k|S_{n \in \gamma(k)}) \exp F_{S_k}^*(\boldsymbol{\xi}_k). \quad (4.28)$$

This is a very general framework in which we clearly distinguish the scale of statistical dependencies between regions (described with the graph) from the finer scale of the dipoles within the regions (described by the emission law and its log-partition function transform).

Using (4.13), the posterior law  $\tilde{p}$  can be written as

$$d\tilde{p}(\mathbf{r}) = \sum_{\{S\}} \tilde{\pi}(S) \prod_{k=1}^K d\tilde{\mu}(\mathbf{r}_k|S_k) \quad (4.29)$$

where the resulting joint state probability and state conditional current densities are given by

$$d\tilde{\mu}(\mathbf{r}_k|S_k) = \exp \left( \tilde{\boldsymbol{\lambda}}_k^t A \mathbf{r}_k - F_{S_k}^*(A^t \tilde{\boldsymbol{\lambda}}_k) \right) d\mu(\mathbf{r}_k|S_k), \quad (4.30)$$



and

$$\tilde{\pi}(S) = \frac{\prod_{k=1}^K \pi(S_k | S_{n \in \gamma(k)}) \exp F_{S_k}^*(A^t \tilde{\lambda}_k)}{\sum_{\{S\}} \prod_{k=1}^K \pi(S_k | S_{n \in \gamma(k)}) \exp F_{S_k}^*(A^t \tilde{\lambda}_k)}. \quad (4.31)$$

Since  $d\tilde{\mu}(\mathbf{r}_k | S_k)$  is the posterior conditional law estimated by MEM, the probability  $\tilde{\pi}(S)$  can be interpreted as a posterior joint probability on hidden states. Indeed, using (4.16), we can write the amplitude estimate in the following form

$$\tilde{r}_{i,l} = \sum_{\{S\}} \tilde{\pi}(S) \nabla_i F_{S_l}^*(\boldsymbol{\xi})|_{\boldsymbol{\xi}=A^t \tilde{\lambda}_l}, \quad i \in P_l, \quad l = 1, \dots, K. \quad (4.32)$$

where  $\nabla_i F_{S_l}^*(\boldsymbol{\xi})|_{\boldsymbol{\xi}=A^t \tilde{\lambda}_l}$  is the estimate of  $r_i$ ,  $i \in P_l$  for a given state  $S_l$  of  $P_l$ . From (4.31), we can compute a ‘‘posterior graphical model’’ associated with the conditional probabilities  $\tilde{\pi}(S_k | S_{n \in \gamma(k)})$  computed from  $\tilde{\pi}(S)$ :

$$\tilde{\pi}(S_k | S_{n \in \gamma(k)}) = \frac{\sum_{\{S\} \setminus k, \gamma(k)} \prod_{l=1}^K \pi(S_l | S_{n \in \gamma(l)}) \exp F_{S_l}^*(A^t \tilde{\lambda}_l)}{\sum_{\{S\} \setminus \gamma(k)} \prod_{l=1}^K \pi(S_l | S_{n \in \gamma(l)}) \exp F_{S_l}^*(A^t \tilde{\lambda}_l)}. \quad (4.33)$$

In this expression,  $\{S\} \setminus m$  denotes the configuration states of all nodes except node  $m$  and  $\gamma(k)$  represents the neighborhood of  $k$  as shown in the graph (see Fig.4.3). This inference on the graphical model can be used to update it: the new graph is always a sub-graph of the initial one. This avenue goes beyond the scope of the present work and will be amplified later with the use of graphical statistical information on connectivity, extracted from other modalities than MEG (fMRI for instance). In order to evaluate the potential of our method and to compare it with other approaches, we consider a simple model in which the graph has no edges and thus the regions are statistically uncorrelated in space.

#### 4.4.2. A degenerate case: the Finite Mixture Model

We consider  $d_k$  uncorrelated dipoles in the region  $P_k$  with possible hidden internal statistical dependencies but no hidden correlations between the regions (see Fig.4.4). In this case, the states joint probability factors as

$$\pi(S) = \prod_{k=1}^K \pi_k(S_k) \quad (4.34)$$

and the reference law (4.27) can be written as follows:

$$d\mu(\mathbf{r}) = \sum_{\{S\}} \prod_{k=1}^K \pi(S_k) d\mu(\mathbf{r}_k | S_k) = \prod_{k=1}^K \sum_{S_k=0}^{s-1} \pi(S_k) d\mu(\mathbf{r}_k | S_k). \quad (4.35)$$

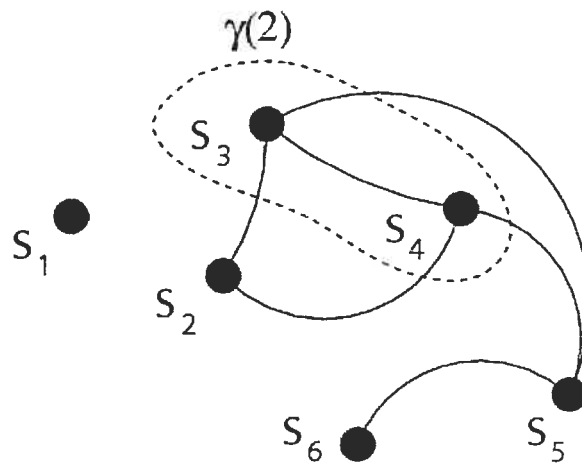


FIGURE 4.3. Graphical model: nodes, links and neighborhood. The set  $\gamma(2)$  is the statistical neighborhood of node 2.

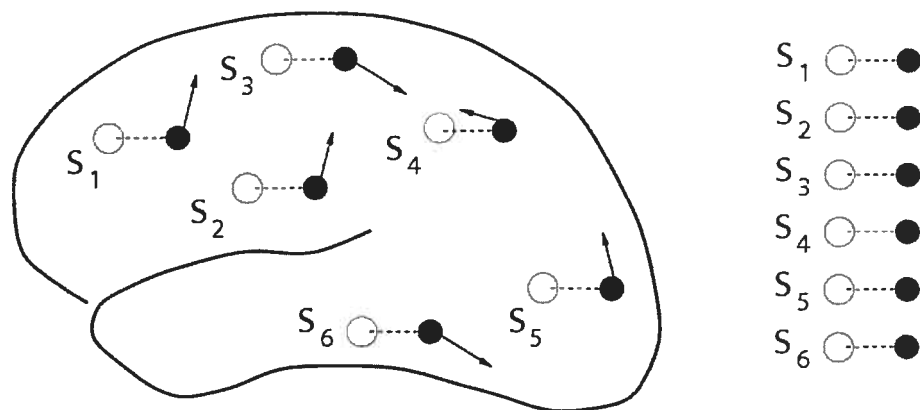


FIGURE 4.4. Mixture model for the activation sources and associated graphical model. A white node represents a region and its state variable. It “emits” the intensities of the  $n_k$  dipoles (black nodes) that belong to the corresponding region  $P_k$ .

Let us be more specific about hidden states. We set  $s = 3$  since, in general, dipoles can be found in one of the following states (Notice that we do not consider region with ambiguous directional currents):  $S_k = +1$  corresponds to *outgoing*

currents in  $P_k$ ,  $S_k = -1$  corresponds to *ingoing* currents in  $P_k$  and  $S_k = 0$  corresponding to the *absence* of activity in  $P_k$ . Let us denote by  $1 - \alpha_k$  and  $\alpha_k \beta_k$  the probabilities that  $S_k = 0$  and  $S_k = +1$  respectively. The general reference law is of the form

$$d\mu(\mathbf{r}) = \prod_{k=1}^K [(1 - \alpha_k) \delta(\mathbf{r}_k) + \alpha_k (1 - \beta_k) \mu_{-,k}(\mathbf{r}_k) + \alpha_k \beta_k \mu_{+,k}(\mathbf{r}_k)] d(\mathbf{r}) \quad (4.36)$$

where

$$\alpha_k = \pi(S_k = -1) + \pi(S_k = +1), \quad (4.37)$$

$$\beta_k = \frac{\pi(S_k = +1)}{\pi(S_k = -1) + \pi(S_k = +1)}. \quad (4.38)$$

We specify the emission laws by choosing the conditional probability density

$$\mu(\mathbf{r}_k | S_k = 0) = \delta(\mathbf{r}_k), \quad (4.39)$$

$\mu(\mathbf{r}_k | S_k = \pm 1) = \mu_{\pm,k}(\mathbf{r}_k)$  being arbitrary.

This mixture model greatly simplifies the estimated joint probability law on the hidden state variables (4.33) since it factors into

$$\tilde{\pi}(S) = \frac{\prod_{k=1}^K \pi(S_k) \exp F_{S_k}^*(A^t \tilde{\lambda}_k)}{\sum_{\{S\}} \prod_{k=1}^K \pi(S_k) \exp F_{S_k}^*(A^t \tilde{\lambda}_k)} = \prod_{k=1}^K \tilde{\pi}(S_k) \quad (4.40)$$

and leads to new marginals for the state variables:

$$\tilde{\pi}(S_k) = \frac{\pi(S_k) \exp F_{S_k}^*(A^t \tilde{\lambda}_k)}{\sum_{\{S_k\}} \pi(S_k) \exp F_{S_k}^*(A^t \tilde{\lambda}_k)}. \quad (4.41)$$

In terms of the mixture's hyperparameters, this result with (4.37-4.38) leads to the following new  $\alpha_k$ 's and  $\beta_k$ 's:

$$\begin{cases} \tilde{\alpha}_k = \frac{\alpha_k \beta_k + \alpha_k (1 - \beta_k) \exp[-\Delta_{+-}]}{\alpha_k \beta_k + (1 - \alpha_k) \exp[-\Delta_{+0}] + \alpha_k (1 - \beta_k) \exp[-\Delta_{+-}]}, \\ \tilde{\beta}_k = \frac{\beta_k}{\beta_k + (1 - \beta_k) \exp[-\Delta_{+-}]}, \end{cases}$$

where we define the "free energy" gaps

$$\Delta_{+-} = F_{S_k=+1}^*(A^t \tilde{\lambda}_k) - F_{S_k=-1}^*(A^t \tilde{\lambda}_k), \quad (4.42)$$

$$\Delta_{+0} = F_{S_k=+1}^*(A^t \tilde{\lambda}_k) - F_{S_k=0}^*(A^t \tilde{\lambda}_k).$$

Notice that  $F_{S_k=0}^*(\xi) = 0$  in the particular case of a Dirac measure for the null state as chosen in (4.39). The  $\tilde{\alpha}_k$ 's and  $\tilde{\beta}_k$ 's depend on observed data and can be interpreted as some *posterior* estimate of hyperparameters. (Later on, we will use a simplified version of this result in the classification of regions according to their “activity”). In this work, we will restrict ourselves to the case where the  $\beta_k$ 's are all set equal to 1 (assuming for instance that no inhibitory states are present). The more general case where the  $\beta_k$ 's are between 0 and 1 is under investigation and will be published in further works. Then, Equations (4.37-4.38) become

$$\begin{cases} \tilde{\alpha}_k = G_{\Delta_{+0}}(\alpha_k) \\ \tilde{\beta}_k = 1 \end{cases}, \quad (4.43)$$

with

$$G_{\Delta}(x) = \frac{x}{x + (1-x) \exp[-\Delta]}$$

The function  $G_{\Delta}(x)$  is plotted in Fig.4.5. We distinguish three cases that lead to nice interpretation of the log-partition function of the reference probability law. For instance, if  $\Delta_{+0} > 0$ , i.e.

$$F_{S_k=1}^*(A^t \tilde{\lambda}_k) > F_{S_k=0}^*(A^t \tilde{\lambda}_k)$$

then  $\tilde{\alpha}_k > \alpha_k$  and, consequently, the MEM solution increase activation in the region labelled by  $k$ . Conversely,  $F_{S_k=1}^*(A^t \tilde{\lambda}_k) < F_{S_k=0}^*(A^t \tilde{\lambda}_k)$  tends to reduce the activity in the region  $P_k$ . This result will be of main relevant to the next section.

## 4.5. SIMULATIONS AND ITERATIVE MEM

### 4.5.1. Lattice, phantoms and simulated data

The lattice displayed on Fig.4.6 consists of 1024 points gathered in patches (16 planes of 64 points each). This assembly of patches tries to model the geometry of some cortical surface as in [1]. A patch can be subdivided into regions of equal size with a well-defined activity state. The region size (or the number of dipoles per region) is a parameter of our model; it may vary from 64 (a whole patch defines a region characterized by a unique activation state) to 1 (each individual dipole has its own activation state).

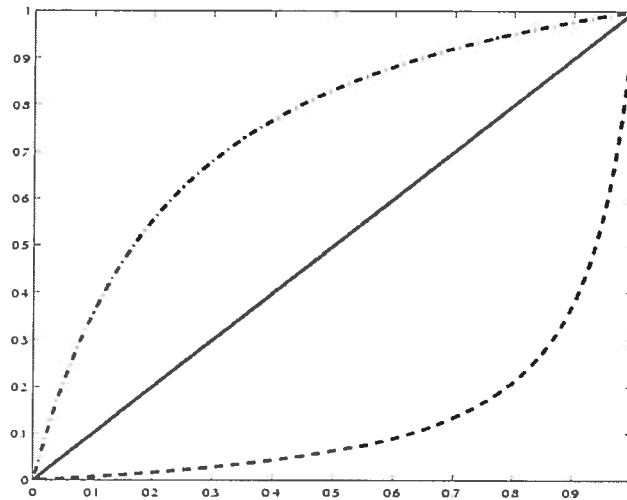


FIGURE 4.5.  $G_u(x) = \frac{x}{x+(1-x)e^{-u}}$  for  $u > 0$  (dashed-dot),  $u = 0$  (plain) and  $u < 0$  (dashed).

“Phantoms” are built from the previous lattice as follows: we split the 16 patches into regions of various size and select some active regions among the totality of regions that constitute the whole lattice. The intensities of each selected active dipole are randomly generated with mean equal to 1 (in regions with outgoing currents) or  $-1$  (in regions with ingoing currents); dipoles intensity in the “silent regions” is set equal to zero.

For a given activity configuration, data are generated from many realisations of the intensities, using the direct forward problem in MEG. Using the spherical model described in [19], we further compute the average magnetic field for each of the 130 measurement points uniformly distributed outside the head. This set of data constitutes our observation. Subsequent addition of a zero mean Gaussian noise yields data with an SNR ratio of about 30dB.

Three phantoms will be considered. *Phantom0* shown in Fig.4.6 consists of two active regions: a region of size 64 that corresponds to patch “3” is set in the so-called ‘outgoing currents’ state, and a smaller region of size 16 is set in the ‘ingoing currents’ state (in the patch “13”). The remaining dipoles have no activity at all. The two other dipoles distributions, *Phantom1* and *Phantom2*, are shown at the top of Fig.4.12 and Fig.4.13 respectively. They involve only either outgoing currents or “silent”dipoles. The difference between these two phantoms

is the presence of a small active region in patch “11” in *Phantom2*. Of course, the main challenge of these simulations is to reconstruct phantoms with various activation area of different size.

Various approaches will be investigated with the same lattice but with different region sizes in the reconstruction model. Four scales will be considered: 64 ( $K = 16$  regions), 16 ( $K = 64$  regions), 4 ( $K = 256$  regions) and 1 ( $K = 1024$  regions). In order to distinguish the reconstructions for these different scales, we will write final amplitudes as  $\tilde{\mathbf{r}}_{|K}$  and the corresponding hyperparameters as  $\tilde{\alpha}_{k|K}$  and  $\tilde{\beta}_{k|K}$ .

Let us remark that reconstruction involving regions of size 16 fits exactly the scale of activity defined in *Phantom0* and *Phantom1*. Consequently, this scale in the reconstruction model takes into account the knowledge of this prior information. We will relax this apparent prior assumption by using regions of various size: from 1 (each dipole has its own activity state) to 64 (the scale in the reconstruction model is much more larger than the activity scale we are looking for).

#### 4.5.2. Reference measure on the dipole amplitude

Various models and joint distributions for dipole intensities inside a region  $P_k$  are possible. With limited information about the dipolar emission law, we use the Normal law  $N(\rho, \Sigma)$  of mean  $\rho$  and covariance  $\Sigma$  for which the log-partition function is simply given by (see Fig.4.7)

$$F_{\mu_g}^*(\boldsymbol{\xi}) = \rho^t \boldsymbol{\xi} + \frac{1}{2} \boldsymbol{\xi}^t \Sigma \boldsymbol{\xi}.$$

In the following simulations, the conditional emission reference laws in (4.36) are of the form

$$\mu_{\pm,k}(\mathbf{r}_k) : \mathbf{r}_k | (S = \pm) \sim N(\rho_{\pm}, \Sigma_{\pm}) \quad (4.44)$$

where  $\rho_{\pm}$  is the  $d_k$ -dimensional vector  $\pm \rho(1, 1, 1, \dots, 1)^t$  and  $\Sigma_{\pm} = \sigma \mathbf{I}_{d_k \times d_k}$ . The MEM solution is of the form

$$d\tilde{p}(\mathbf{r}) = \prod_{k=1}^K \left[ (1 - \tilde{\alpha}_k) \delta(\mathbf{r}_k) + \tilde{\alpha}_k (1 - \tilde{\beta}_k) \tilde{p}_{-,k}(\mathbf{r}_k) + \tilde{\alpha}_k \tilde{\beta}_k \tilde{p}_{+,k}(\mathbf{r}_k) \right] d(\mathbf{r}) \quad (4.45)$$

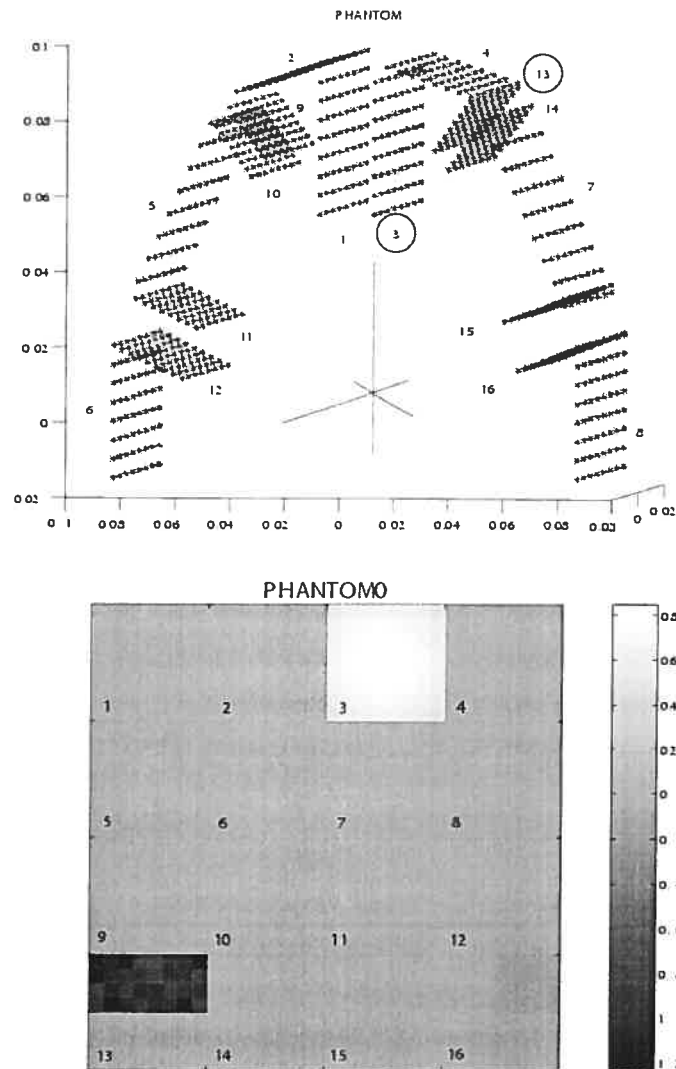


FIGURE 4.6. “Cortical mesh” used in simulations and an example of phantom. **Upper:** the “cortical surface” is modeled as an assembly of 16 patches with 64 nodes each. **Lower:** *Phantom0*. Intensities  $r$  of the 1024 dipoles are coded in grayscale. The patch 3 is active (with outgoing currents) and patch 13 is partially active (upper half dipoles are ingoing currents; lower half dipoles are “silent” with vanishing currents). Notice that this phantom mixes different scales of activation.

with the conditional emission laws equal to

$$\begin{cases} \tilde{p}_{+,k}(\mathbf{r}_k) : \mathbf{r}_k | (S = +1) \sim N(\boldsymbol{\rho}_+ - \Sigma_+(A^t \tilde{\boldsymbol{\lambda}})_k, \Sigma_+) , \\ \tilde{p}_{-,k}(\mathbf{r}_k) : \mathbf{r}_k | (S = -1) \sim N(\boldsymbol{\rho}_- - \Sigma_-(A^t \tilde{\boldsymbol{\lambda}})_k, \Sigma_-) . \end{cases}$$

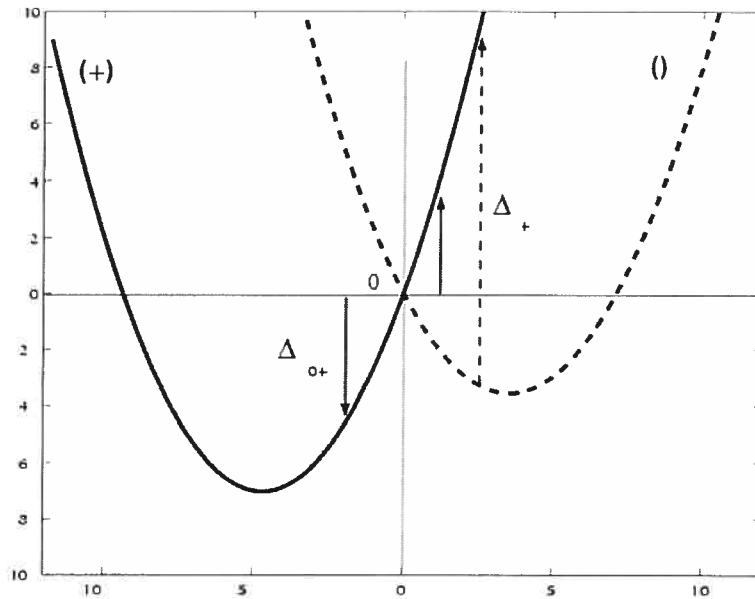


FIGURE 4.7.  $F_{\mu_g}^*(\xi)$  for Normal laws  $N(m_+ > 0, \sigma_+)$  and  $N(m_- < 0, \sigma_-)$  (line and dashed line respectively).

### 4.5.3. MEM estimation

Assuming that no prior information about states is available, we consider a reconstruction lattice with  $K$  regions of definite size and activity state and we set the  $\alpha_k$ 's and  $\beta_k$ 's equal to  $1/2$  in the reference measure (4.36). This non-informative setting of hyperparameters completely defines the prior law in the MEM optimization method. As described before, the MEM algorithm consists in maximizing a strictly concave function  $D(\lambda)$ . In Fig.4.8, we show MEM estimations of the current amplitude for various values of  $K$ , from large scale ( $K = 16$ ) to the smallest scale ( $K = 1024$ ). First, we observe that most active dipoles of the phantom have been well detected. Second, despite the favorable size of regions used in this reconstruction, we obtain dipoles ambiguously activated in "silent patches". Elimination of those spurious false-positive events which is of primary importance in practice can be performed in various ways. For example, we can take advantage of some prior knowledge or assumptions about the hyperparameters  $\alpha_k$  and thus introduce bias in the reference law. However, in absence



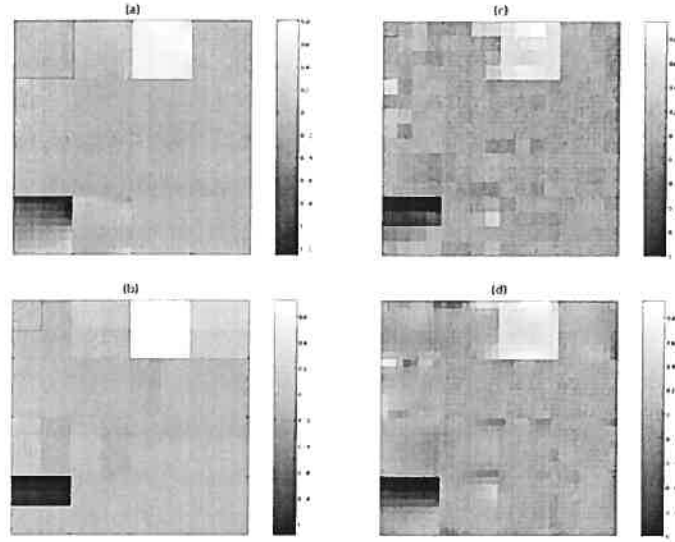


FIGURE 4.8. MEM reconstruction of *Phantom0* with various activation scales:  $\tilde{\mathbf{r}}_{K=16}$  (a),  $\tilde{\mathbf{r}}_{K=64}$  (b),  $\tilde{\mathbf{r}}_{K=256}$  (c) and  $\tilde{\mathbf{r}}_{K=1024}$  (d). The scale of the region is indicated with a box in the upper left corner of each display. Data are noise free.

of any hypothesis about the optimal scale at which cortical activity may occur, we can consider a weighted sum of the various reconstructions,

$$\tilde{\mathbf{r}}^* = \sum_{K=K_{min}}^{K_{max}} \rho_K \tilde{\mathbf{r}}_{|K}, \quad \text{with} \quad \sum_{K=K_{min}}^{K_{max}} \rho_K = 1 \quad (4.46)$$

Such a result is illustrated on Fig.4.9. It relaxes any prior assumption on scale and smoothes intensities over the patches (higher scale in the model) but it still preserves the internal structure (as seen in patch number 13).

- Remark: As illustrated on Fig.4.12 and Fig.4.13, LORETA and S-MAP algorithms are considered in order to compare the various reconstruction techniques. Simulations are performed with two phantoms that exhibit different scales of distributed activation. Beside the notion of scale, the presence of noise in the data (in the magnetic field outside the head) will also be discussed. Let us comment on how scale may enter in the LORETA and S-MAP techniques. In LORETA, a scale is introduced in the discrete expression of the laplacian that expresses the regularisation term. This is the  $M$  matrix in (4.5). S-MAP is very

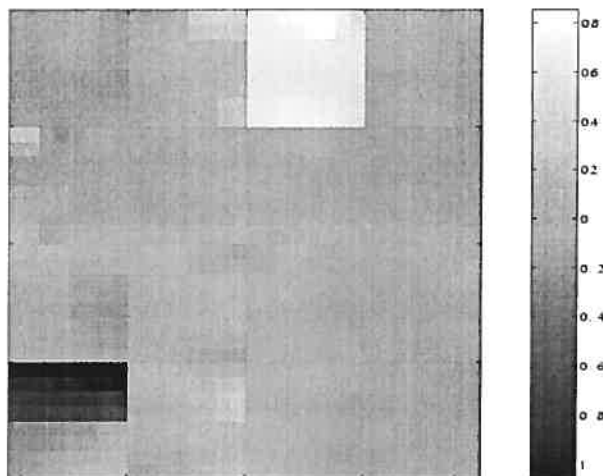


FIGURE 4.9. MEM reconstruction  $\tilde{\mathbf{r}}^*$  without privileged scale.  $\rho_K = 1/4$  and  $K = 16, 64, 256$  and  $1024$  in (4.46).

similar to LORETA but introduces a specific parameter that defines the strength of the correlation between any dipole and its neighbors within a region.

#### 4.5.4. Learning the states: Iterative MEM algorithms

Without loss of generality, let us assume for simplicity that there are no “in-going currents” states, *i.e.*  $\beta_k = 1$ . Following the results obtained in Section 4.2, an iterative scheme is set up. It alternates the usual MEM reconstruction with an estimation of hyperparameters based directly on (4.43):  $\alpha_k \longrightarrow \tilde{\alpha}_k = G_{\Delta_{0+}}(\alpha_k)$ . Notice that the  $\beta_k$ 's remain equal to the initial value 1 as expected. This iterative algorithm (denoted by *MEMiter* and described below) is illustrated in Fig.4.10. A new space of probability densities,  $\mathcal{C}_\mu$ , contains a family of reference laws in which members are iteratively selected for successive MEM optimization. This scheme nicely regularizes the initial MEM reconstruction, as we shall see in the next section. For example, graphs of Fig.4.11 display the values of the  $\alpha_k$ 's along the successive iterations. Since the algorithm may run forever if  $F_{S_k=1}^*(A^t \tilde{\lambda}_k) \sim 0$ , a stopping criterion for *MEMiter* could be based on the values of the  $\alpha_k$ 's such that  $F_{S_k=1}^*(A^t \tilde{\lambda}_k) \notin [-\epsilon, \epsilon]$  for some  $\epsilon > 0$ . In the following description of the *MEMiter* algorithm, thresholds  $\epsilon_{min}$  and  $\epsilon_{max}$  define a decision criterion that “freezes” the region activation either in the *silent* or *active* state:

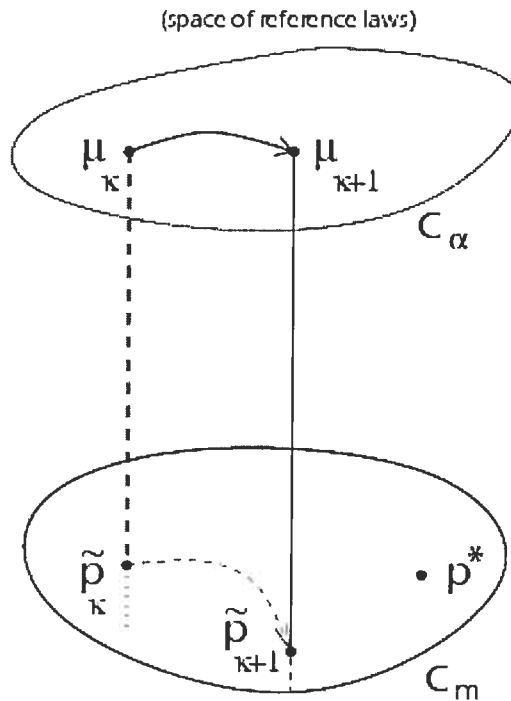


FIGURE 4.10. MEMIter: the reference law “lives” in a space of prior models.

- (1)  $j = 0$ : initialization of the hyperparameters :  
 $\alpha_k^{(j=0)} = 1/2, k = 1, \dots, K,$
- (2) MEM reconstruction and  $\alpha_k^{(1)} = \tilde{\alpha}_k^{(0)},$
- (3) Let  $\sigma > 0$ . Repeat until  $k$  exists such that  
 $|\alpha_k^{(j+1)} - \alpha_k^{(j)}| > \sigma :$ 
  - (a) Update of the reference law with:
  - (b) if  $\tilde{\alpha}_k^{(j)} < \epsilon_{min}$  then  $\alpha_k^{(j+1)} = 0,$
  - (c) else if  $\tilde{\alpha}_k^{(j)} > \epsilon_{max}$  then  $\alpha_k^{(j+1)} = 1,$
  - (d) else  $\alpha_k^{(j+1)} = \tilde{\alpha}_k^{(j)}.$
  - (e) MEM reconstruction and hyperparameter estimate.
  - (f)  $j \rightarrow j + 1.$

#### 4.5.5. Results and Evaluation

The performance of the reconstruction algorithms is evaluated through the computation of three quantities to be maximized:

- (1) The Mean Square Error (measured in  $dB$ ): we measure the ability to reconstruct the dipoles amplitudes with

$$MSE = -10 \log(\|\tilde{\mathbf{r}} - \mathbf{r}\|^2 / \|\mathbf{r}\|^2)$$

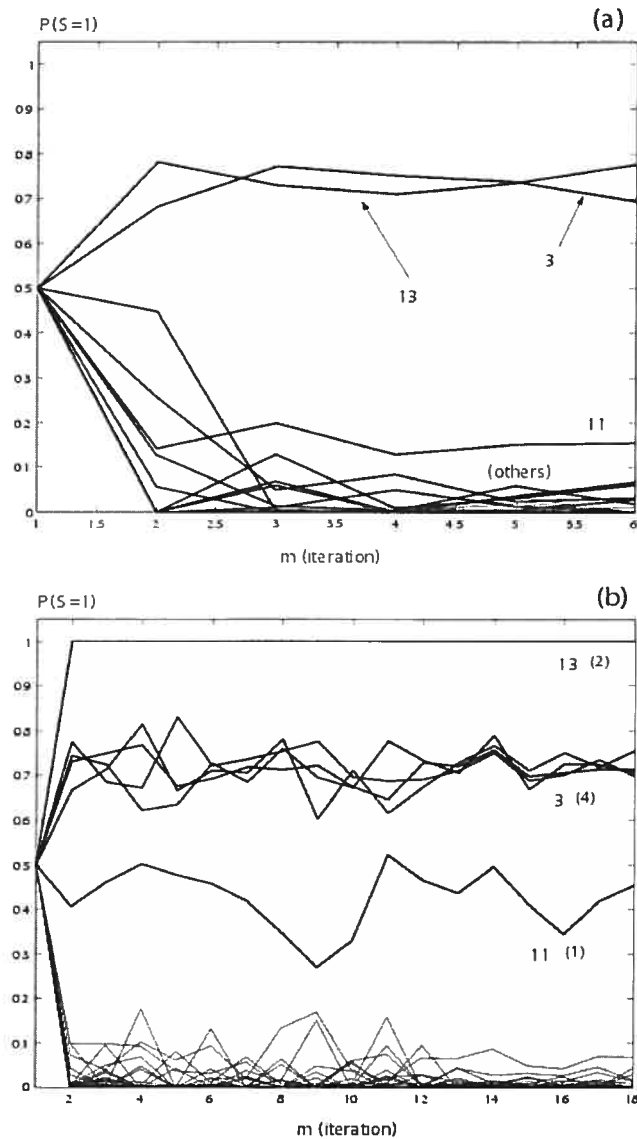


FIGURE 4.11. Progression of the iterative MEM algorithm applied to *Phantom2* reconstruction (see Fig.4.13): we plot  $\alpha_k^{(j)}$  versus  $j$ . Plot (a) and plot (b) are associated with two different scales in the reconstruction model: (a) corresponds to  $K = 16$  (see MEMiter(a) in Fig.4.13) and (b) corresponds to  $K = 64$  (see MEMiter(b) in Fig.4.13)). Numbers indicate the index  $k$  of patches as defined in Fig.4.6. In (b), a number in parenthesis indicates how many active regions are found the patch. In (a) or (b), detected active regions correspond to the original phantom.

- (2) The Data Fit Error (measured in  $dB$ ): we measure the goodness-of-fit performance with

$$DFE = -10 \log(\|\mathbf{m} - A\tilde{\mathbf{r}}\|^2 / \|\mathbf{m}\|^2)$$

- (3) A ROC area: the area under the ROC (Receiver Operating Characteristics [6]) curve is an estimation of the reconstruction method's ability to *classify* active (or non-active) dipoles. The closer to 1 this value is, the higher the performance of method in terms of activation detection.

Comparison between MEM techniques and the two other regularization approaches LORETA and S-MAP, will focus on the notion of activation scale in the prior model and the presence of noise in data.

- **Simulations with specific scale** in the reconstruction model:

Fig.4.12 and Fig.4.13 show the sources amplitude reconstruction computed with MEM, MEMiter, LORETA and S-MAP. In each case, a scale has been introduced either explicitly (LORETA and S-MAP) in spatial correlation or implicitly (MEM and MEMiter) through hidden variables. Various scales, *i.e.* regions with different size, have been used and denoted by (a), (b), (c) and (d) when it was feasible. Let us point out that for  $K = 1024$ , *i.e.* the case for which a region is a unique dipole, LORETA and S-MAP are nothing but a Least Mean Square reconstruction and are found to fail in the present problem. The spatial “coherence” induced by the hidden state variable does not mean that dipoles’ amplitudes in the same region share the same probability law. Actually, they only share the hyperparameters that define the mixture law, but still remain spatially uncorrelated. Spatial correlation *inside* a region would amount to defining a non diagonal  $\Sigma_{\pm}$  in (4.44) and would introduce more spatial regularization.

There are three main comments we could say about the results in the Table 4.1. First, iterative MEM improves the performance achieved with the initial MEM reconstruction, at least for the small values of  $K$ . For  $K = 1024$ , where all dipoles have their own state variable and there is no spatial regularization involved in the MEM scheme; the reconstruction exhibits many false positive fluctuations. Conversely, when non trivial spatial coherency expressed through the hidden state variables is used in the probabilistic model, the iterative MEM drastically improves the reconstruction.

Second, we note that LORETA gives solutions that fit data with great precision. However, as we can see on Fig.4.12 and Fig.4.13, those solutions are far from the true configurations. This point illustrates the ill-posedness of the “MEG inversion” problem: many (and even an infinite number of) sources configurations may explain the data. Thus, only small  $DFE$  values are meaningful and

TABLE 4.1. Performance of MEM, MEMiter, LORETA and S-MAP. No Noise.

K		MEM	MEMiter
16 (a)	MSE	19.89 (14.51)	29.57 (18.26)
	DFE	77.65 (83.18)	119.20 (73.18)
	ROC area	<b>0.953</b> (0.948)	<b>0.953</b> (0.934)
64 (b)	MSE	37.10 (18.54)	<b>49.01 (33.05)</b>
	DFE	115.62 (106.39)	139.42 (117.40)
	ROC area	<b>0.953</b> (0.950)	<b>0.953 (0.951)</b>
256 (c)	MSE	0.42 (-0.09)	12.08 (9.04)
	DFE	193.53 (206.96)	184.34 (178.12)
	ROC area	0.924 (0.917)	0.877 (0.863)
1024 (d)	MSE	-5.83 (-2.32)	-9.30 (-6.34)
	DFE	168.26 (233.09)	174.62 (276.0)
	ROC area	0.915 (0.907)	0.903 (0.862)

K		LORETA	S-MAP
16 (a)	MSE	-26.70 (-24.96)	21.39 (18.05)
	DFE	296.99 (297.56)	99.58 (98.52)
	ROC area	0.940 (0.939)	<b>0.953</b> (0.950)
64 (b)	MSE	-11.36 (-11.84)	24.11 (17.76)
	DFE	278.50 (278.07)	159.15 (153.29)
	ROC area	0.940 (0.939)	<b>0.953 (0.951)</b>
256 (c)	MSE	-15.11(-15.15)	17.93 (18.16)
	DFE	274.33 (273.56)	<b>301.27 (302.32)</b>
	ROC area	0.932 (0.931)	<b>0.953 (0.951)</b>

**Note:** The reconstruction algorithms are applied to *Phantom 1* and *Phantom 2* (values in parenthesis). (a), (b), (c) and (d) refer to the labels used in Fig.4.12 and Fig.4.13. Values in bold faces are highest scores.

we may disregard some reconstruction schemes. As a matter of fact, most of the regularization schemes used in this work give high *DFE* scores and thus high goodness-of-fit.

Finally, on the basis of *MSE* and ROC analysis results, MEMiter gives the best results as long as we know some optimal scale to work with. Furthermore, the *MSE* significantly varies with scale in our model, whereas LORETA and S-MAP

TABLE 4.2. Performance MEM, MEMiter, LORETA and S-MAP (scale average). No Noise.

$\tilde{\mathbf{r}}^*$		MEM	MEMiter
(e)	MSE	17.73 (13.47)	<b>31.51 (25.37)</b>
	DFE	98.72 (104.97)	137.39 (94.64)
	ROC area	<b>0.953 (0.951)</b>	<b>0.953 (0.951)</b>

$\tilde{\mathbf{r}}^*$		LORETA	S-MAP
(e)	MSE	-18.21 (-17.59)	27.08 (23.65)
	DFE	<b>288.03 (287.14)</b>	121.49 (120.42)
	ROC area	0.938 (0.937)	<b>0.953 (0.951)</b>

**Note:** Performance of the reconstruction algorithms given by (4.46) with MEM, MEMiter, LORETA and S-MAP, applied to *Phantom 1* and *Phantom 2* (values in parenthesis). Values in bold faces are highest scores. For corresponding amplitude reconstruction, see cases (e) on Fig.4.12 and Fig.4.13.

are less sensitive to  $K$ . This difference is related to the role played by the scale parameter in the various algorithms. In our model, the scale is directly associated with the notion of local cortical activity whereas it enters more technically in LORETA and S-MAP at the level of the spatial regularization.

- **Simulations without a specific scale** in the reconstruction model:

Let us consider case (e) in Fig.4.12 and Fig.4.13 and related performance scores as shown in Table 4.2. In each case, this reconstruction is a weighted average of the scale-dependent reconstruction algorithm and thus any assumption about specific scale gets relaxed in the model. MEMiter clearly delivers the best performance. The role of entropy appears in these results since it gives the best tradeoff between the information provided by data and the true *prior* information.

- **Simulations with noisy data:** (see Table 4.3)

Reconstruction of dipoles amplitude from noisy data is shown on Fig.4.14 and Fig.4.15. The prior model for reconstruction introduces the noise through a reference measure of the form (4.18). The noise measure has one parameter that can be set using a standard estimation on noise variance. In the present simulations, the noise level is moderately strong in order to obtain still interesting results with LORETA. Performance of the MEM techniques obviously degrades with higher

level of noise but still simulations has shown quite acceptable results. Indeed, current investigation with noisy data are performed with realistic anatomical models and, more importantly, makes use of some preprocessing data analysis that restricts the space of priors and initialize more carefully the hyperparameters  $\alpha_k$ 's. This extension is out of the scope of the present work and presented elsewhere.

Most of the previous observations about performance are still valid and, again, favor the MEM approach. Let us stress at this point that our regularization scheme does not involve any local spatial correlation between amplitudes and thus differs from S-MAP on this particular constraint: S-MAP does more and better because it arbitrarily smoothes intensities. So, in some sense, by choosing uncorrelated dipoles in the regions, we deliberately disadvantage the MEM approach that, at the end, offers the same performance as the more constrained (but ad hoc) regularization schemes such as S-MAP.

#### 4.6. CONCLUSION

Detection of cerebral activation from magnetoencephalography measurements requires *a priori* information in order to solve unambiguously the underlying inverse problem. In this work, this necessary regularization is studied within the framework of maximum entropy on the mean and extends previous work based on entropy regularization [3, 4]. The method relies mostly on the definition of a *prior* model. We propose prior laws that represent the physical properties of electric activity through a normal probability law on the distributed dipolar current intensities. We further assume that currents are grouped within cortical regions that are statistically connected through a graphical model [5]. In this spirit, we introduce a hidden explanatory variable associated with each region in order to describe the local state of functional activity on the cortical surface of the brain. Using a distributed source model for the dipolar currents, we have developed an iterative and unsupervised procedure that progressively eliminates false-positive candidates from the solution by classifying dipoles according to their activity state. In some way, this work promotes a “hybrid regularization” scheme [2]: a *functional regularization* expressed through the joint probability on the hidden state variables (and subject to some hypothetical graphical model), and *anatomical smoothness regularization* expressed through the dipolar current emission laws. Therefore, this model is able to bring together various types of



TABLE 4.3. Performance MEM, MEMiter, LORETA and S-MAP. Noisy data.

K		MEM	MEMiter
16 (a)	MSE	3.66 (1.88)	26.10 ( <b>20.61</b> )
	DFE	77.55 (74.68)	90.66 (70.44)
	ROC area	0.951 (0.933)	<b>0.953</b> (0.931)
64 (b)	MSE	-4.55 (-4.86)	<b>43.07</b> (1.99)
	DFE	78.45 (78.58)	73.80 (58.43)
	ROC area	0.935 (0.931)	<b>0.953</b> (0.910)
256 (c)	MSE	-8.45 (-7.99)	8.65 (11.23)
	DFE	65.73 (66.44)	51.08 (64.26)
	ROC area	0.939 (0.934)	0.872 (0.898)
1024 (d)	MSE	-8.64 (-9.07)	-6.75 (-8.01)
	DFE	78.50 (75.87)	69.81 (73.28)
	ROC area	0.863 (0.884)	0.815 (0.853)
$\tilde{r}^*$ (e)	MSE	-2.02 (-2.31)	<b>26.83</b> ( <b>17.81</b> )
	DFE	75.68 (75.30)	69.08 (72.32)
	ROC area	<b>0.953</b> (0.940)	<b>0.953</b> ( <b>0.950</b> )

K		LORETA	S-MAP
16 (a)	MSE	-40.26 (-31.07)	21.39 (4.49)
	DFE	229.73 (237.53)	99.58 (62.71)
	ROC area	0.842 (0.876)	<b>0.953</b> (0.928)
64 (b)	MSE	-39.19 (-29.32)	24.07 (17.71)
	DFE	231.15 (241.71)	159.15 (153.29)
	ROC area	0.807 (0.878)	<b>0.953</b> ( <b>0.951</b> )
256 (c)	MSE	-43.73 (-35.95)	8.10 (3.51)
	DFE	231.17 (241.88)	<b>233.63</b> ( <b>231.19</b> )
	ROC area	0.777 (0.804)	0.951 (0.949)
$\tilde{r}^*$ (e)	MSE	-26.97 (-30.21)	23.40 (17.54)
	DFE	<b>231.26</b> ( <b>241.85</b> )	121.48 (84.68)
	ROC area	0.817 (0.869)	<b>0.953</b> ( <b>0.950</b> )

**Note:** Performance of the reconstruction algorithms (MEM, MEMiter, LORETA and S-MAP) applied to *Phantom 1* and *Phantom 2* (values in parenthesis). These reconstructions were made using a single realization of noise. (a), (b), (c), (d) and (e) refer to the labels used in Fig.4.14 and 4.15. Values in bold faces are highest scores.

*prior* knowledge: the physics of electro-neural processes, the anatomo-physiology of the cortex and information from other functional imaging modalities that may define, for instance, the graphical model [15]. This complementary scheme also addresses different prior information that may show up on different scales of the cortex. The generalization of our approach toward hierarchical graphical models is on going and, finally, assessment of our technique with real and dynamical MEG data is currently under investigation and will be reported in the near future.

#### ACKNOWLEDGEMENTS

The authors wish to thanks B. Goulard, H. Benali and M. Goldstein for their help during the progression of this work. Numerical simulations have been made possible thanks to the help of L. Garnero who kindly provided us with advice, data and codes on the forward problem. Also, comments from anonymous referees have been very helpful.

This work is supported in part by the Natural Sciences and Engineering Research Council of Canada (NSERC) and Atlantic Nuclear In Quebec (ANIQ) within the Network for Computing and Mathematical Modeling (NCM<sub>2</sub>).

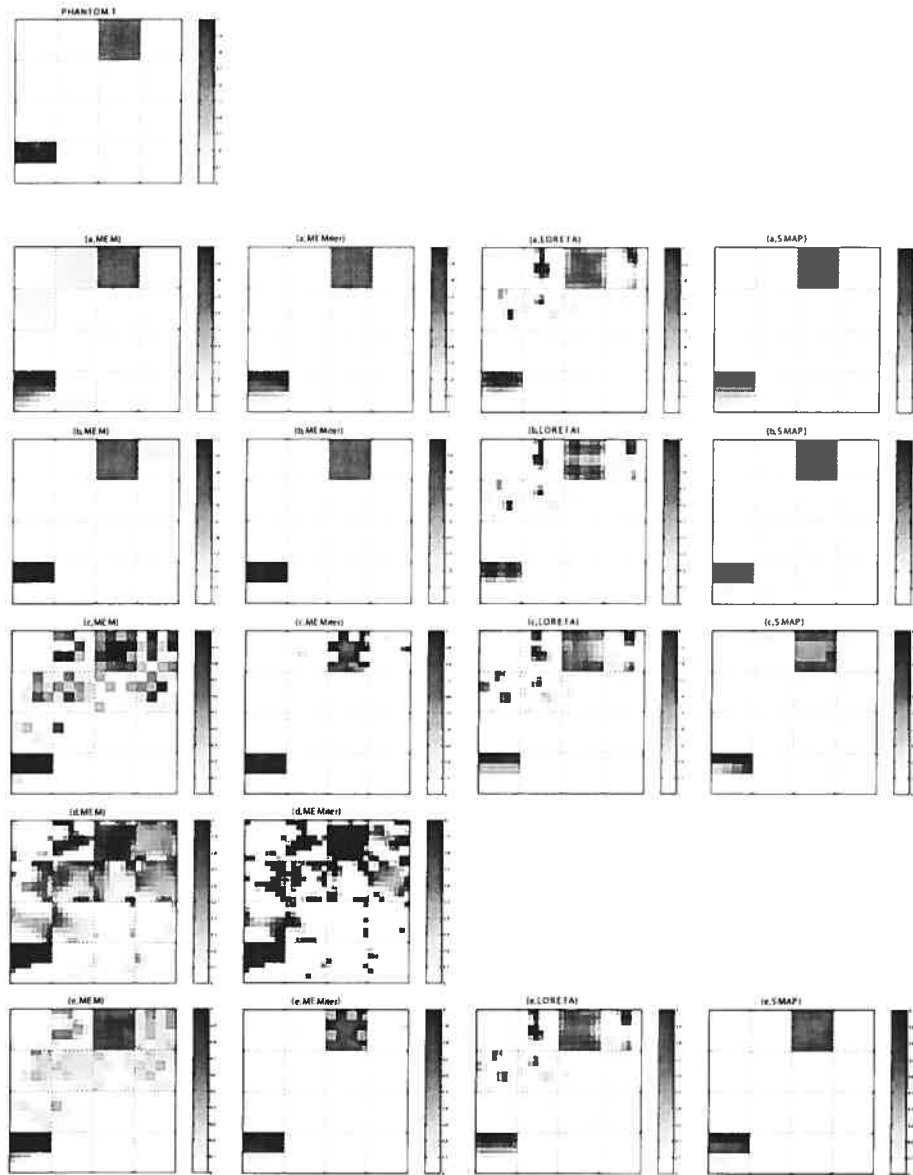


FIGURE 4.12. Phantom 1 (top, left) with 4 reconstructions from **non-noisy** observations: MEM, MEMiter, LORETA and S-MAP. (a):  $K = 16$  (64 dipoles per region); (b):  $K = 64$  (16 dipoles per region); (c):  $K = 256$  (4 dipoles per region); (d):  $K = 1024$  (1 dipole per region). (e) is the reconstruction of the form (4.46) with  $K = 16, 64$  and  $256$ .

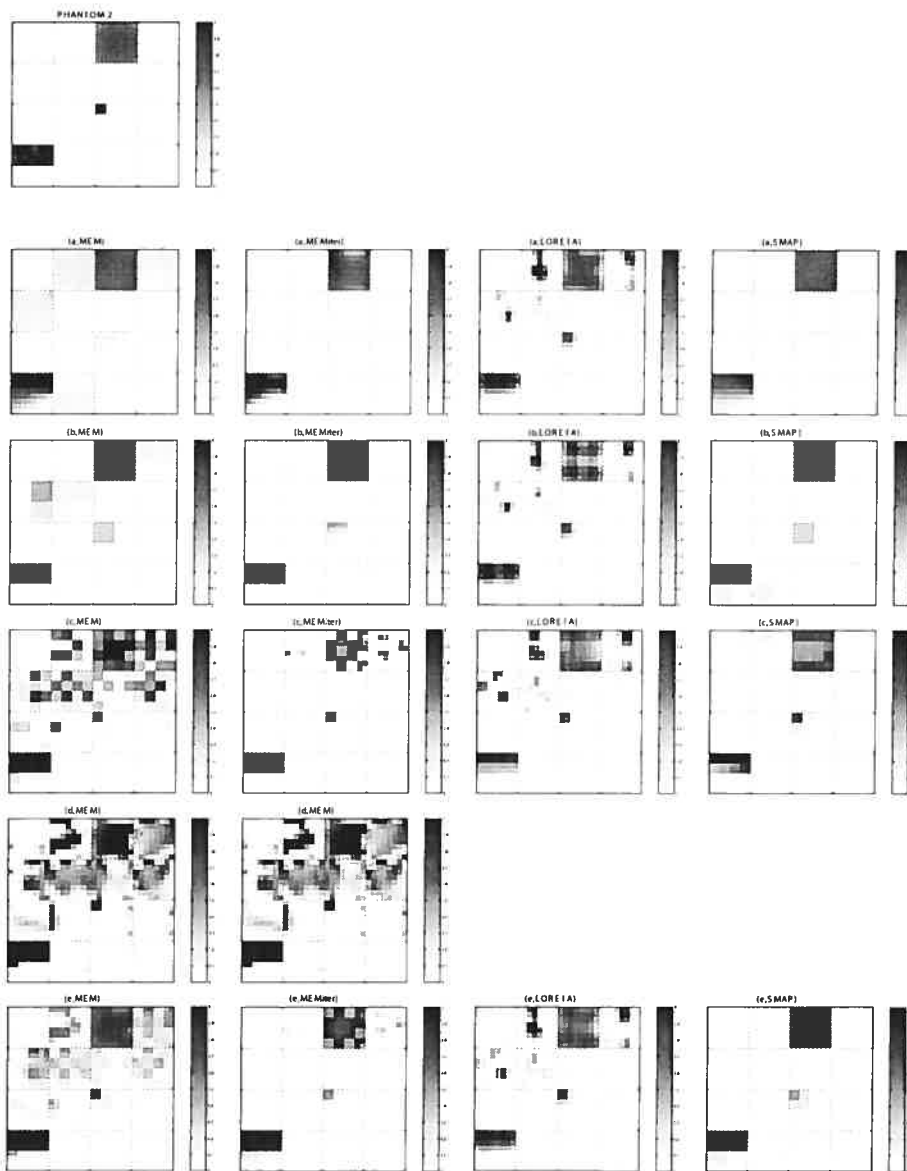


FIGURE 4.13. Phantom 2 (top, left) with 4 reconstructions from **non-noisy** observations: MEM, MEMiter, LORETA and S-MAP. (a):  $K = 16$  (64 dipoles per region); (b):  $K = 64$  (16 dipoles per region); (c):  $K = 256$  (4 dipoles per region); (d):  $K = 1024$  (1 dipole per region). (e) is the reconstruction of the form (4.46) with  $K = 16, 64$  and  $256$ .

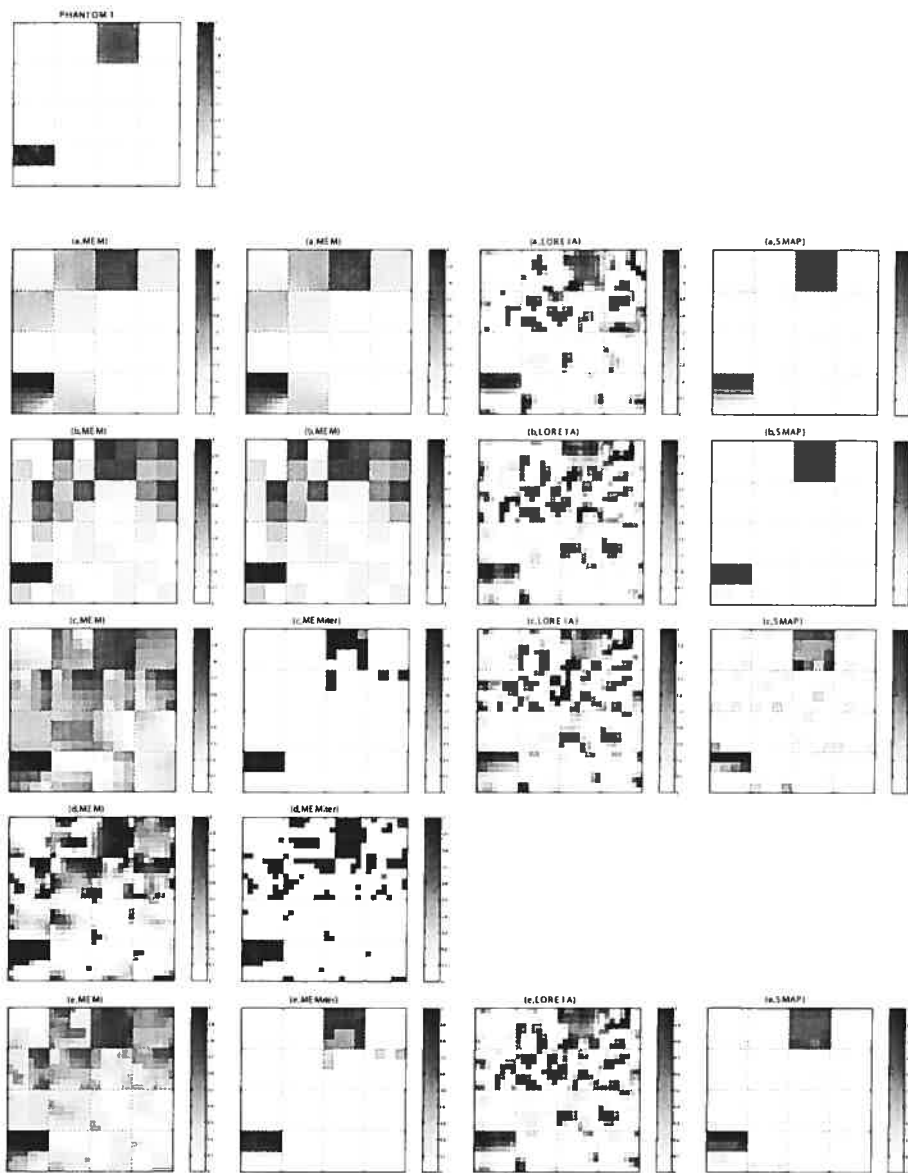


FIGURE 4.14. Phantom 1 (top, left) with 4 reconstructions from noisy observations (30 dB): MEM, MEMiter, LORETA and S-MAP. (a):  $K = 16$  (64 dipoles per region); (b):  $K = 64$  (16 dipoles per region); (c):  $K = 256$  (4 dipoles per region); (d):  $K = 1024$  (1 dipole per region). (e) is the reconstruction of the form (4.46) with  $K = 16, 64$  and  $256$ .

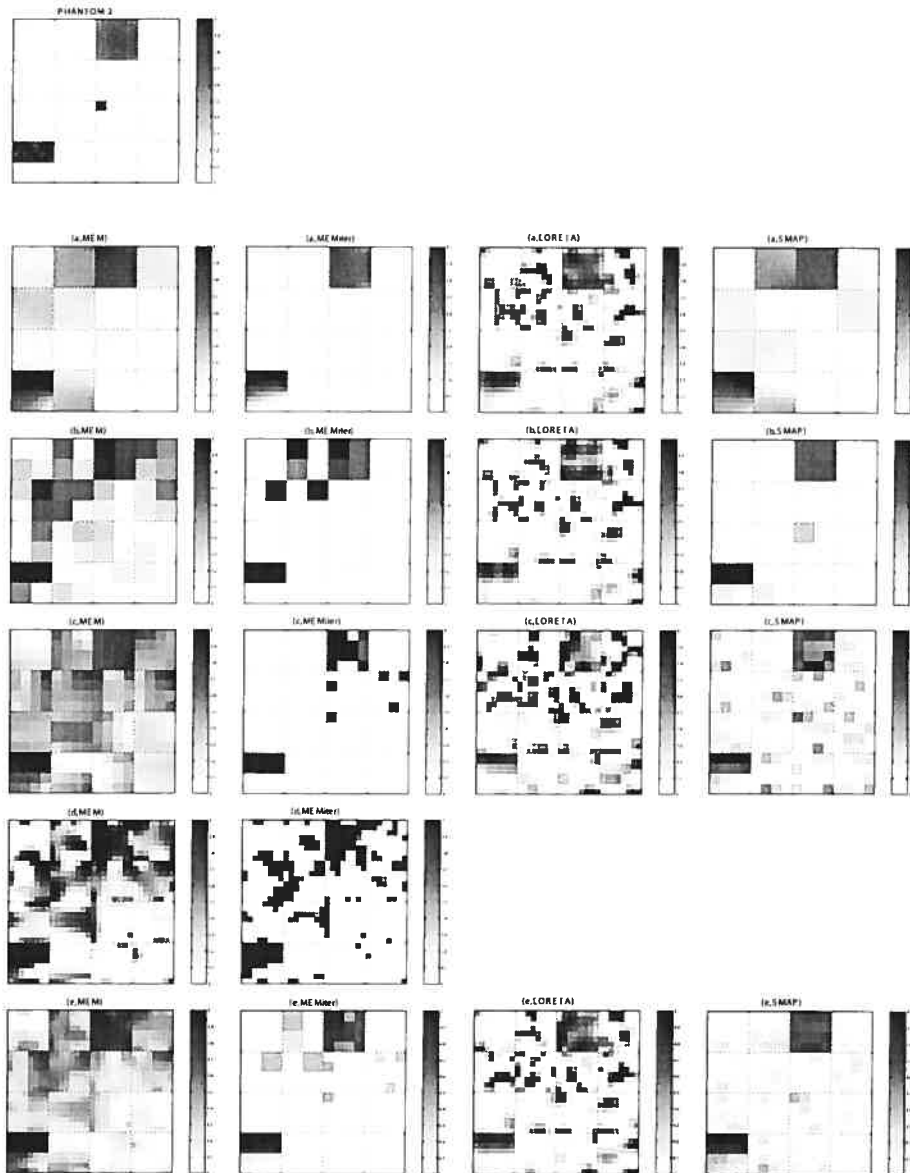


FIGURE 4.15. Phantom 2 (top, left) with 4 reconstructions from **noisy** observations (30 dB): MEM, MEMiter, LORETA and S-MAP. (a):  $K = 16$  (64 dipoles per region); (b):  $K = 64$  (16 dipoles per region); (c):  $K = 256$  (4 dipoles per region); (d):  $K = 1024$  (1 dipole per region). (e) is the reconstruction of the form (4.46) with  $K = 16, 64$  and  $256$ .

## BIBLIOGRAPHY

- [1] S. Baillet and L. Garnero. A bayesian approach to introducing anatomo-functional priors in the MEG-EEG inverse problems. *IEEE Transactions on Biomedical Engineering*, 44(3):374–385, 1997.
- [2] E. Bienenstock. On the dimensionality of the cortical graphs. *J. Physiol*, 90:251–256, 1996.
- [3] C.J.S Clarke. Probabilistic methods in a biomagnetic inverse problem. *Inverse Problems*, 5(6):999–1012, Décembre 1989.
- [4] C.J.S Clarke and B.S. Janday. The solution of the biomagnetic inverse problem by maximum statistical entropy. *Inverse Problems*, 5(4):483–500, Août 1989.
- [5] R. Dahlhaus, M. Eichler, and J. Sandkühler. Identification of synaptic connections in neural ensembles by graphical models. Technical report, Univ. Heidelberg, 1995.
- [6] E.R. DeLong, D.M. DeLong, and D.L. Clarke-Pearson. Comparing the areas under two or more correlated receiver operating characteristic (ROC) curves : a nonparametric approach,. *Biometrics*, 44:837–845, 1988.
- [7] R.S Ellis. *Entropy, large deviations and statistical mechanics*, volume 271 of *Grundlehren der mathematischen Wissenschaften*, chapter 1–2, page 364. Springer-Verlag, New-York, 1985.
- [8] F. Gamboa and E. Gassiat. Bayesian methods and maximum entropy method for ill-posed problems. *The Annals of Statistics*, 25(1):328–350, 1997.
- [9] L. Garnero, S. Baillet, J.P. Lachaux, and B. Renault. Data operating in PET/MEG/MRI experiment. *Human Brain Mapping*, 1:84, 1995. Suppl 1.
- [10] A.A. Ioannides, J.P.R. Bolton, and C.J.S Clarke. Continuous probabilistic solutions to the biomagnetic inverse problem. *Inverse Problems*, 6(4):523–542, Août 1990.
- [11] E.T. Jaynes. Information theory and statistical mechanics. *Physical Review*, 106(4):620–630, Mai 1957.
- [12] E.T. Jaynes. Where do we stand on maximum entropy? In D. Levine and M. Tribus, editors, *the Maximum Entropy Formalism*. MIT Press, 1979.
- [13] S. Kullback and R.A. Leibler. On information and sufficiency. *Ann. Math. Statist.*, 22:79–86, 1951.
- [14] G. Le Besnerais, J.-F. Bercher, and G. Demoment. A new look at entropy for solving linear inverse problems. *IEEE Transactions on Information Theory*, 45(5):1565–1578, Juillet 1999.
- [15] G. Marrelec, H. Benali, and M. Péligrini-Issac. Bayesian estimation of local and global independence relations for gaussian graphical models. Technical report, INSERM-U494, 2001.
- [16] A. Mohammad-Djafari. A comparison of two approaches: maximum entropy on the mean and bayesian estimates for inverse problems. *arXiv:physics 0111122*, 2001.
- [17] R. D. Pascual-Marqui, C. M Michel, and D. Lehmann. Low resolution electromagnetic tomography: a new method for localizing electrical activity in the brain. *Int. Journal of Psychophysiology*, 18:49–65, 1994.
- [18] R.R. Rockafeller. Conjugate duality and optimization. *SIAM, Conference Board of the Mathematical Science Regional Conference series in Applied Mathematics*, 16, 1974.

- [19] J. Sarvas. Basic mathematical and electromagnetic concepts of the biomagnetic inverse problem. *Phys. Med. Biol.*, 32(1):11–22, 1987.



# Chapitre 5

---

## ESTIMATION DE LA LOI DE RÉFÉRENCE Second article

L'article contenu dans ce chapitre a été soumis à la revue NeuroImage le 8 juin 2004 [28].

Dans le chapitre 4, nous avons démontré l'efficacité du MEM appliqué à la MEG. Sa flexibilité nous a permis d'introduire de l'information anatomique et fonctionnelle sur la solution à l'aide d'une loi de probabilité de référence. Cette loi contient plusieurs hyperparamètres :

- $\alpha_k$  : probabilité d'activation de la région  $k$ ,
- $\mu_k$  : intensité moyenne des dipôles de la région  $k$  lorsque celle-ci est activée,
- $R_k$  : les dipôles appartenant à chaque région.

Dans le dernier chapitre, ces paramètres étaient fixés au départ. Entre autres, les régions sont définies très rigidelement et il s'avère que c'est l'élément de régularisation le plus important. Cependant, lorsqu'on s'attaque à des problèmes réalistes, il est difficile de définir, *a priori*, de telles régions d'activation.

Dans cette deuxième publication, nous nous sommes inspiré des méthodes bayésiennes qui utilisent les données MEG pour définir la loi de référence. Ici, les données seront utilisées pour définir les régions à l'aide de la MSP, décrite dans la section 2.2.5, qui identifie les dipôles contribuant le plus aux données MEG. Ces dipôles sélectionnés servent de germes pour une méthode de type croissance de régions. Ces régions seront ensuite utilisées pour regrouper les dipôles.

À partir de ces régions, il est possible d'estimer les  $\alpha_k$  et les  $\mu_k$  donnant une expression complète de la loi de référence pour pouvoir ensuite appliquer le MEM qui estime l'activation qui a généré les signaux MEG. Dans le même esprit qu'au chapitre 4, des algorithmes itératifs sont présentés afin d'améliorer la solution donnée par le MEM.

Cette méthode est comparée au MEM sans estimation des régions ainsi qu'à LORETA à l'aide de simulations utilisant un cortex provenant d'une image IRM réelle. Nous montrons que nos techniques sont très performantes malgré la difficulté du problème inverse. Puisque ces travaux assouplissent les hypothèses sur la solution, ils permettront une application du MEM aux données MEG réelles ce qui est fait dans le chapitre 6 en considérant les corrélations temporelles.

# Data-Driven Cortex Parcelling and Entropic Inference in MEG

Article soumis à *NeuroImage* le 8 juin 2004.

ERVIG LAPALME, *Centre de Recherches Mathématiques et Département de Physique de l'Université de Montréal*

JEAN-MARC LINA, *Centre de Recherches Mathématiques et Département de Génie Électrique de l'École de Technologie Supérieure, Montréal*

JÉRÉMIE MATTOU, *Wellcome Department of Imaging Neuroscience, London, UK*

## Abstract

In [1], the authors introduced the maximum entropy on the mean (MEM) as a new methodological framework for solving the Magnetoencephalography (MEG) inverse problem. The main ingredient of the MEM is a reference probability law which enables one to include all kind of prior information on the active source distribution to be estimated. However, defining the reference law highly relies upon a prior parcelling of the cortical surface that is required as an input. In this paper, we present a data-driven approach for parcelling out the cortex into regions that are functionally coherent. This parcelling consists of a data-driven clustering (DDC) of the dipoles of a distributed representation of the cortical surface and furthermore provides an initial estimate of the reference law parameters. It is based on the recently developed Multivariate Source Prelocalization (MSP) principle. On MEG simulated data, the DDC proves to further improve the MEM inverse approach, as evaluated considering two different iterative algorithms and using classical criteria as well as ROC curve analysis. The MEM solution is also compared to a classical LORETA-like inverse approach. The data-driven clustering allows to make the most of the MEM formalism whose main trumps lie in the flexible probabilistic way of introducing priors and in the notion of coherent regions of activation. The latter reduces the dimensionality of the problem, thus narrowing down the gap between the two types of inverse methods, the popular dipolar approaches and the so-called distributed ones.

**Keywords:** Maximum Entropy on the Mean, Cortex Parcelling, Clustering, Multivariate Source Prelocalization, Inverse Problem, Magnetoencephalography.

## 5.1. INTRODUCTION

### 5.1.1. The MEG inverse problem

Non-invasive characterization of human brain activity remains one of today's most challenging issues in terms of image and signal processing. In particular, inferring the cortical source locations and intensities of a signal acquired outside the head, such as given by Magnetoencephalography (MEG) and Electroencephalography (EEG), is an ill-posed inverse problem that still requires further investigations and methodological developments. Solving this problem is of great importance since MEG and EEG are the only neuroimaging techniques whose data directly reflect the electromagnetic nature of the neuronal activity. Moreover, compared to functional Magnetic Resonance Imaging (fMRI) and Positron Emission Tomography (PET), MEG and EEG provide an instantaneous measure of the whole brain activity.

A large number of articles have been yet published regarding the physics and the models for explaining the MEG and EEG signal generation. We refer the reader to [16] for a thorough review of this so-called forward problem.

In the present paper, given a forward solution, we address the MEG inverse problem which consists in localizing the brain current sources of the measured magnetic field outside the head. Although an ongoing controversy about EEG and MEG respective localization power [4], MEG is known to be much less sensitive to the anisotropy of the skull electric conductivity, hence providing a more focal field topography and more tractable forward solutions [3]. Nevertheless, the proposed methodology can be equally well applied to MEG and EEG data.

During the past two decades, two families of methods have been proposed for solving the MEG/EEG inverse problem. The first approaches consist in looking for a very few putative active regions (typically less than five), each of them being modeled by an Equivalent Current Dipole (ECD) [21]. These methods are still very popular because of being easy to use and providing solutions that are easy to compare and interpret. However, they highly rely upon the user prior experience for specifying the number of active regions, whereas the more recent distributed methods do not require such a hard prior [7]. The latter type of approach indeed consists of displaying a large amount of fixed dipoles all over the cortical surface. Such a model is usually based upon each subject segmented structural MRI, hence

making MEG/EEG much closer to a true imaging technique. A dense enough cortical description enables one to derive an estimate of the spatial extension of each activated area. However, when using distributed methods, the inverse problem becomes highly degenerated because of the huge number of unknown parameters (the dipole moments). External quantitative priors are needed to derive a unique and realistic inverse solution. Most of the existing distributed solutions differ according to the type of constraints they incorporate. Various constraints such as anatomical, physiological, mathematical and functional prior knowledge have been yet considered [2, 6, 9, 10, 12, 19]. One of the most popular distributed approach is the Low Resolution Tomography Algorithm (LORETA) also referred to as the smoothest inverse solution [18] (see section 5.1.2).

In this paper we introduce a new methodological framework which somehow conciliates one type of inverse approach with the other. Through the use of a flexible and realistic distributed representation of the solution space, this new approach combines the advantage of a realistic and flexible model with the notion of region of activation represented by a common reduced set of parameters, as achieved by ECD models.

### 5.1.2. Notations and classical inverse solutions

Considering a distributed source model made of  $n$  dipoles of fixed location and orientation (perpendicular to the cortical surface [7]), the generative model for a single time sample of MEG data is given by the linear system

$$\mathbf{M} = G \mathbf{Q} + \epsilon, \quad (5.1)$$

where  $\mathbf{M}$  indicates the observation random variable vector on the  $d$  sensors,  $\mathbf{Q}$  the multivariate random variable made of the  $n$  dipole amplitudes and  $\epsilon$  an additive measurement noise which is usually described by a zero-mean Gaussian independent and identically distributed (i.i.d.) process. The  $d \times n$   $G$  matrix is known as the lead-field operator which embodies the physical link between each dipole intensity ( $\mathbf{Q}$ ) and the observed magnetic fields ( $\mathbf{M}$ ). It is estimated by solving the Maxwell's equations of electromagnetism using a particular model of the physical and geometrical properties of the head tissues [16].

Localizing the brain sources of MEG activity using a distributed model then consists of estimating  $\mathbf{Q}$  by inverting equation (5.1). Because of the large number

of needed dipoles compared to the number of sensors (a few thousand versus a few hundred), this system is highly under-determined. It is moreover instable due to the noise that corrupts the data. Therefore regularization is needed to derive a unique and stable solution. A regularized solution is the only source parameter distribution that satisfies a given trade-off between the goodness of the data fit and an external prior or regularization term.

The goodness of the data fit can be expressed as a strict data attachment constraint so that the ensemble  $\mathcal{C}_{\mathbf{m}}$  of the putative solutions  $\mathbf{q}$  is defined by

$$\mathcal{C}_{\mathbf{m}} = \{\mathbf{q} \text{ s.t. } G \mathbf{q} = \mathbf{m}\},$$

where  $\mathbf{m}$  is the given realization of the observation variable.

A more general and relaxed constraint may be considered through the allowance of a maximum variance  $\delta$  so that

$$\mathcal{C}_{\mathbf{m},\delta} = \{\mathbf{q} \text{ s.t. } \|G \mathbf{q} - \mathbf{m}\|^2 < \delta\}.$$

Then, in principle, if a solution does exist, one can use the Penrose pseudo-inverse operator  $G^+$  of  $G$  in order to compute the solution  $\mathbf{q}^* = G^+ \mathbf{m}$  of  $\mathcal{C}_{\mathbf{m},\delta}$  when  $\delta \rightarrow 0$ .  $G^+$  verifies  $G G^+ G = G$  and is given by

$$G^+ = \lim_{\lambda \rightarrow 0} G^t (G G^t + \lambda \mathbb{I})^{-1}, \quad (5.2)$$

where  $\mathbb{I}$  indicates the  $d \times d$  identity matrix.

In practice, such an exact solution is highly unstable and not even computable because of the ill-conditioned matrix to be inverted. In addition to the goodness of data fit constraint, a regularization criterion  $U(\mathbf{q})$  has to be considered so that  $\mathbf{q}^* = \underset{\mathbf{q} \in \mathcal{C}_{\mathbf{m},\delta}}{\operatorname{argmin}} U(\mathbf{q})$ . Any regularized inverse approach is then characterized by the type of prior knowledge or particular hypotheses expressed through the choice of  $U(\mathbf{q})$ .

As a simple example, the classical Minimum Norm criterion defines  $U(\mathbf{q}) = \|\mathbf{q}\|^2$  so that  $\mathbf{q}^* = G_{MN}^+ \mathbf{m}$  and

$$G_{MN}^+ = G^t (G G^t + \lambda \mathbb{I})^{-1}, \quad (5.3)$$

$\lambda$  being set to a finite and strictly positive value.

LORETA, another well-known regularization approach, enforces a quadratic smoothness constraint between neighbouring sources in space. The corresponding

criterion is given by

$$U(\mathbf{q}) = \mathbf{q}^t W \mathbf{q} \quad (5.4)$$

where  $W$  indicates the spatial laplacian. The optimal solution  $\mathbf{q}^* = \mathbf{G}_{Lo}^+ \mathbf{m}$  is easily computed since

$$\mathbf{G}_{Lo}^+ = G^t (G W^{-1} G^t + \lambda \mathbb{I})^{-1}. \quad (5.5)$$

The regularization parameter or hyperparameter  $\lambda$  tunes the relative weight of the two constraints: the goodness of the data fit ( $\mathbf{q}^* \in \mathcal{C}_{m,\delta}$ ) and the smoothness prior ( $U(\mathbf{q}^*) \simeq 0$ ).

### 5.1.3. The entropic inference

In [1], the authors revisited the probabilistic approach based on the Maximum Entropy on the Mean (MEM) principle, early introduced by Clarke and Janday in the context of the biomagnetic inverse problem [5]. Apart from providing a promising alternative framework to the well-known bayesian probabilistic methodology, this recent approach also proposes an improvement to the traditional distributed source model. By introducing the notion of region of activation whose parameteres are entered in the model as hidden variables, the dimension of the problem is reduced and so is the number of parameters to be estimated. However, the crucial point then lies in the definition of such regions. As a first step and because this previous paper was mainly dedicated to the presentation of the new methodological framework, regions of activation were arbitrarily defined, both in terms of location and spatial extension. In the present paper, we address this critical issue and propose to further improve the MEM-based inverse reconstruction by following an empirical Bayes-like approach and use the data itself to drive a cortex parcelling into anatomically and functionally coherent regions. We therefore apply the Multivariate Source Prelocalization (MSP), which also allows to set some activation parameters associated with each cortical parcel.

This paper is organized as follows. Section 5.2.1 is divided into two parts. First part is devoted to the general MEM principle and its application to the MEG inverse problem. In the second part, the data-driven clustering (DDC) method is introduced. The proposed methodology has been evaluated on MEG simulated data according to the procedure described in section 5.3.1. The results

are presented in section 5.4 and both the new inverse approach and its performances are discussed in the last section.

## 5.2. METHOD

### 5.2.1. Maximum entropy on the mean

#### 5.2.1.1. *The general principle*

Let us consider the following compact form of data generative model (5.1),

$$[G \ \mathbb{I}] \begin{bmatrix} \mathbf{Q} \\ \boldsymbol{\epsilon} \end{bmatrix} = \mathbf{M}. \quad (5.6)$$

The random multivariables  $[\mathbf{Q}, \boldsymbol{\epsilon}]^t$  can be described by their joint probability density  $dP(\mathbf{q}, \mathbf{e})$ . Now, under the reasonable assumption that the noise and the signal due to the cortical sources are statistically independent, this density can be factorized as the product of its marginals  $dP(\mathbf{q}) d\nu(\mathbf{e})$ . Averaging equation (5.6) over  $dP$  leads to

$$G E_P(\mathbf{Q}) = \mathbf{m} \quad (5.7)$$

where  $E_P(\mathbf{Q})$  indicates the expectation of  $\mathbf{Q}$  and  $\mathbf{m}$  the expected observation.

In practice,  $\mathbf{m}$  is approximated by the evoked magnetic response which corresponds to the averaged observation over a large number of repetition of the same task or stimulation. Then, as shown in [1], the MEM formalism is suitable for solving the MEG inverse problem stated as follows: what is the optimal probability density  $dP$  on the sources so that (5.7) is satisfied?

Although it is not explicitly modeled in equation (5.7), it is worth noticing that the MEM approach implicitly accounts for the noise through the joint probability density  $dP$ . Noise is taken into account in the regularization process as a second order statistical contribution.

Using MEM formalism, the ensemble  $\mathcal{C}_{\mathbf{m}}$  of all the probability distributions that satisfy (5.7) is given by

$$\mathcal{C}_{\mathbf{m}} = \{dP(\mathbf{q}, \mathbf{e}) = f(\mathbf{q}, \mathbf{e}) dP_{\text{ref}}(\mathbf{q}) d\nu(\mathbf{e}) \text{ s.t. } G E_P(\mathbf{Q}) = \mathbf{m}\}$$

where the source probability density  $dP$  is related to a reference or prior probability distribution  $dP_{\text{ref}}(\mathbf{q})$  through the function  $f(\mathbf{q}, \mathbf{e})$ . As in any regularization



approach,  $dP_{\text{ref}}(\mathbf{q})$  embodies all the prior knowledge on the source distribution. It corresponds to the prior probability distribution in a Bayesian formalism [2]. However, contrary to a classical regularization scheme, the MEM approach first requires from the source distribution to match the prior probability law best, under the secondary constraint of fitting the observed data. This reference probability distribution is thus all the more essential. It defines the initial model to start with in order to fit the data.

As in an empirical Bayes framework, a significant contribution of the present work consists of using the data to infer some properties of the reference probability law.

Given  $dP_{\text{ref}}(\mathbf{q})$ , the MEM solution  $d\mathcal{P}^*$  of the inverse problem is obtained by solving the following optimization problem

$$d\mathcal{P}^* = \underset{d\mathcal{P} \in \mathcal{C}_m}{\text{argmin}} \mathcal{K}(d\mathcal{P}, dP_{\text{ref}} d\nu), \quad (5.8)$$

where  $\mathcal{K}(d\mathcal{P}, dP_{\text{ref}} d\nu)$  indicates the Kullback's information divergence and plays the role of a pseudo-distance (or cost function) to be minimized between the reference law  $dP_{\text{ref}} d\nu$  and any  $d\mathcal{P}$  of  $\mathcal{C}_m$ . It is defined by

$$\begin{aligned} \mathcal{K}(d\mathcal{P}, dP_{\text{ref}} d\nu) &= \int \log \frac{d\mathcal{P}}{dP_{\text{ref}} d\nu} d\mathcal{P} \\ &= \int f(\mathbf{q}, \mathbf{e}) \log f(\mathbf{q}, \mathbf{e}) dP_{\text{ref}}(\mathbf{q}) d\nu(\mathbf{e}), \end{aligned} \quad (5.9)$$

Solving (5.8) then amounts to maximize the Shannon entropy  $\mathcal{S} = -\mathcal{K}(d\mathcal{P}, dP_{\text{ref}} d\nu)$  under some constraint [11]. The optimal solution being of the form  $d\mathcal{P}^* d\nu$ , the source estimate is finally given by

$$\mathbf{q}^* = E_{P^*}(\mathbf{Q}). \quad (5.10)$$

This solution explains the data *in average* and maximizes the information carried by the underlying probability density. One should furthermore emphasize that  $dP_{\text{ref}}$  corresponds to the zero-entropy solution. As mentioned previously, it may be considered as a prior density but, in general, it expresses some initial assumption that may be modified through the MEM iterative scheme (see equation (5.11) below). It is therefore natural to somehow exploit the data in order to define  $P_{\text{ref}}$ .

### 5.2.1.2. The dual formulation

The “Primal formulation” of the MEM presented in the previous section can be made computationally more attractive by considering its “Dual formulation”. Indeed, it can be shown [11] that the solution of equation (5.8) is the unique probability distribution in the exponential family of the reference law defined by

$$d\mathcal{P}^*(\mathbf{q}, \mathbf{e}) = \exp(\boldsymbol{\lambda}^{*t} G \mathbf{q} - F^*(G^t \boldsymbol{\lambda}^*, \boldsymbol{\lambda}^*)) dP_{\text{ref}}(\mathbf{q}) d\nu(\mathbf{e}), \quad (5.11)$$

where  $\boldsymbol{\lambda}^{*t}$  indicates the transpose of the  $d$ -dimensional Lagrange multiplier vector  $\boldsymbol{\lambda}^*$  and  $F^*$  the “log-partition function” also known as the “free energy” associated with the joint reference distribution

$$F^*(\boldsymbol{\xi}, \boldsymbol{\zeta}) = \log \int \exp(\boldsymbol{\xi}^t \mathbf{q} + \boldsymbol{\zeta}^t \mathbf{e}) dP_{\text{ref}}(\mathbf{q}) d\nu(\mathbf{e}) = F_Q^*(\boldsymbol{\xi}) + \frac{\sigma^2}{2} \boldsymbol{\zeta}^t \boldsymbol{\zeta}, \quad (5.12)$$

In (5.12), the last equality has been derived assuming that the noise is Gaussian and *i.i.d.* ( $\mathbf{e} \sim \mathcal{N}(0, \sigma^2)$ ).  $F_Q^*$  is the *free energy* associated with the sources

$$F_Q^*(\boldsymbol{\xi}) = \log \int \exp(\boldsymbol{\xi}^t \mathbf{q}) dP_{\text{ref}}(\mathbf{q}). \quad (5.13)$$

Note that it only depends upon the reference distribution.

As a standard result in MEM, it can be shown that (5.11) solves the equation (5.8) if and only if  $\boldsymbol{\lambda}^*$  maximizes the dual strictly concave criterion  $D(\boldsymbol{\lambda}) = \boldsymbol{\lambda}^t \mathbf{m} - F^*(G^t \boldsymbol{\lambda}, \boldsymbol{\lambda})$ , *i.e.*

$$\boldsymbol{\lambda}^* = \underset{\boldsymbol{\lambda}}{\operatorname{argmax}} \left( \boldsymbol{\lambda}^t \mathbf{m} - F_Q^*(G^t \boldsymbol{\lambda}) - \frac{\sigma^2}{2} \boldsymbol{\lambda}^t \boldsymbol{\lambda} \right). \quad (5.14)$$

This is the essence of the dual formulation from which the “primal” solution  $\mathbf{q}^*$  given by (5.10) can now be expressed as

$$\mathbf{q}^* = \nabla F_Q^*(\boldsymbol{\xi})|_{\boldsymbol{\xi}=G^t \boldsymbol{\lambda}^*}. \quad (5.15)$$

Consequently, within the MEM framework, the solution of an inverse problem can be achieved using a classical optimization algorithm.

Indeed, the dual criterion  $D(\boldsymbol{\lambda})$  is always strictly concave and leads to a unique maximum  $\boldsymbol{\lambda}^*$  [11]. Furthermore, the optimization problem, as stated by equation (5.14) is being solved in the  $d$ -dimensional sensor space and does not depend on the number of dipoles in the model. This makes the MEM approach particularly suitable for solving the MEG inverse problem in a distributed model framework.

### 5.2.1.3. The reference probability law

As in [1], we here define the notion of region of activation by considering that the cortical surface is divided into  $K$  parcels. Each parcel is made of one or several neighbouring dipoles of the distributed model that are expected to have a common functional behaviour. The activation state of a given parcel  $k$  is then defined by a binary hidden random variable  $S_k$ .  $S_k = 0$  (the zero-state) tells the absence of an active dipole in parcel  $k$  whereas  $S_k = 1$  (the active-state) indicates that at least one dipole is active in the parcel. According to this hidden-state model, any reference probability distribution on the sources can be expressed as a mixture so that

$$dP_{\text{ref}}(\mathbf{q}) = \sum_{\{S\}} dP_{\text{ref}}(\mathbf{q}, S) = \sum_{\{S\}} \pi_{\text{ref}}(S) dP_{\text{ref}}(\mathbf{q}|S), \quad (5.16)$$

where  $\{S\} = \{S_1, S_2, \dots, S_K\}$  indicates all the possible states of the hidden variable and  $\pi_{\text{ref}}(S)$  is a joint probability distribution that incorporates the correlations between all the various parcel states.

To fully characterize the reference probability law, we need to explicit both  $dP_{\text{ref}}(\mathbf{q}|S)$  and  $\pi_{\text{ref}}(S)$ , for each state of the hidden variable. We therefore assume that

- the activation state of region  $k$  only depends on the local state variable  $S_k$  (so-called Markov hypothesis) which leads to

$$dP_{\text{ref}}(\mathbf{q}|S) = \prod_{k=1}^K dP_{\text{ref}}(\mathbf{q}_k|S_k),$$

where  $\mathbf{q}_k$  represents the intensities of the dipoles in parcel  $k$ ;

- the activation of the different regions are statistically independent, such that

$$\pi_{\text{ref}}(S) = \prod_{k=1}^K \pi_{\text{ref}}(S_k).$$

Denoting  $\alpha_k = \pi_{\text{ref}}(S_k = 1)$ , these assumptions produce the following rewriting of mixture model (5.16)

$$dP_{\text{ref}}(\mathbf{q}) = \prod_{k=1}^K [(1 - \alpha_k) dP_{\text{ref}}(\mathbf{q}_k|S_k = 0) + \alpha_k dP_{\text{ref}}(\mathbf{q}_k|S_k = 1)]. \quad (5.17)$$

The conditional laws in (5.17) now remained to be specified. We therefore consider the two following probability distributions for respectively defining the *zero* and the *active* states:

- the absence of activation ( $S_k = 0$ ) is described by “silent dipoles” and the corresponding local conditional law is modeled by a Delta function ( $\delta(x)$  is null everywhere except at 0 where  $\delta(0) = \infty$ ) such that

$$dP_{\text{ref}}(\mathbf{q}_k | S_k = 0) = \delta(\mathbf{q}_k) d\mathbf{q}_k.$$

- the active-state ( $S_k = 1$ ) is described by a local spatial normal distribution of mean  $\mu_k$  and covariance  $\Sigma_k$

$$dP_{\text{ref}}(\mathbf{q}_k | S_k = 1) \sim N(\mu_k, \Sigma_k).$$

According to those definitions, we end up with the following and fully parametrized model of the reference probability distribution

$$dP_{\text{ref}}(\mathbf{q}) = \prod_{k=1}^K \left\{ (1 - \alpha_k) \delta(\mathbf{q}_k) + \frac{\alpha_k}{(2\pi)^{\frac{n_k}{2}} |\text{Det}(\Sigma_k)|^{\frac{1}{2}}} \exp \left[ -\frac{1}{2} (\mathbf{q}_k - \mu_k)^t \Sigma_k^{-1} (\mathbf{q}_k - \mu_k) \right] \right\} d\mathbf{q}_k, \quad (5.18)$$

where  $n_k$  is the number of dipoles in parcel  $k$ .

This leads to the following and easily computable form of the free energy (see equation (5.13))

$$F_Q^*(G^t \boldsymbol{\lambda}) = \sum_{k=1}^K \log [(1 - \alpha_k) + \alpha_k \exp[F_k^*(G_k^t \boldsymbol{\lambda})]], \quad (5.19)$$

where  $G_k$  is the  $d \times n_k$  lead-field matrix restricted to parcel  $k$  and  $F_k^*$  is the prior free energy of parcel  $k$ , given by

$$\begin{aligned} F_k^*(G_k^t \boldsymbol{\lambda}) &= \log \int \exp(\boldsymbol{\lambda}^t G_k \mathbf{q}) dP_{\text{ref}}(\mathbf{q}_k | S_k = 1) \\ &= \frac{1}{2} \boldsymbol{\lambda}^t G_k \Sigma_k G_k^t \boldsymbol{\lambda} + \mu_k G_k^t \boldsymbol{\lambda}. \end{aligned} \quad (5.20)$$

#### 5.2.1.4. The algorithmic procedures

The reference probability distribution being of the form (5.18) and fully defined by the initial parameters  $\{\alpha_k, \mu_k, \Sigma_k\}_{k=1, \dots, K}$  (see sections 5.2.1.3 and 5.2.2.3), we hereafter describe the two different algorithmic procedures that may be used to solve the MEG inverse problem using the MEM formalism. Our Maximum Entropy based approach amounts to solve equation (5.14) using equations (5.19) and (5.20). Then, given the optimal value  $\boldsymbol{\lambda}^*$  and using equation (5.15), we derive the following source intensity estimate

$$\mathbf{q}_k^* = \mu_k^* + \alpha_k^* \Sigma_k G_k^t \boldsymbol{\lambda}^* \quad (5.21)$$

where

$$\alpha_k^* = \frac{\alpha_k}{\alpha_k + (1 - \alpha_k) \exp[-F_k^*(G_k^t \boldsymbol{\lambda}^*)]}, \quad (5.22)$$

$$\mu_k^* = \alpha_k^* \mu_k. \quad (5.23)$$

Note that equation (5.21) clearly illustrates the 2-levels multiresolution scheme of the model. Indeed, first term  $\mu_k^*$  describes the parcel intensity, which is common to each dipole of the parcel, whereas the second term, which depends on  $\boldsymbol{\lambda}^*$ , enables the solution to differentiate the dipole activities within a parcel.

In [1], the authors proposed an iterative computation scheme for solving equation (5.21). It consists of updating the parameters of the reference probability law according to (5.22). Denoted by MEMiter<sub>1</sub>, this algorithm depends on initial parameter values (see section 5.2.2.3) which are then modified as follows

$$\text{MEMiter}_1 : \begin{cases} \alpha_k^{(l)} \\ \mu_k^{(l)} \end{cases} \rightarrow \begin{cases} \alpha_k^{(l+1)} = \frac{\alpha_k^{(l)}}{\alpha_k^{(l)} + (1 - \alpha_k^{(l)}) \exp[-F_k^*(G_k^t \boldsymbol{\lambda}^{*(l)})]}, \\ \mu_k^{(l+1)} = \mu_k^{(l)}, \end{cases} \quad (5.24)$$

where  $(l)$  indicates the iteration index and  $\boldsymbol{\lambda}^*$  corresponds to the current solution of (5.14) given the parameters  $\{\alpha_k^{(l)}, \mu_k^{(l)}, \Sigma_k = \sigma_k^2 \mathbb{I}\}_{k=1, \dots, K}$ . This algorithm updates the hidden state probability while keeping constant the conditional probability that drives the source intensities.

As an alternative solution, we here propose a second algorithm MEMiter<sub>2</sub> which proceeds the other way around. Contrary to MEMiter<sub>1</sub>, this new approach iteratively updates the reference law using the mean of the previously estimated

conditional reference probability, while maintaining a constant balance mixture of zero and active states

$$\text{MEMiter}_2 : \begin{cases} \alpha_k^{(l)} = \frac{1}{2} \\ \mu_k^{(l)} \end{cases} \rightarrow \begin{cases} \alpha_k^{(l+1)} = \frac{1}{2}, \\ \mu_k^{(l+1)} = 2 \frac{\mu_k^{(l)} + \sigma_k^2 \tilde{\mathbf{g}}_k^t \boldsymbol{\lambda}_k^{*(l)}}{1 + \exp[-F_k^*(G_k^t \boldsymbol{\lambda}_k^{*(l)})]}, \end{cases} \quad (5.25)$$

where  $\tilde{\mathbf{g}}_k$  indicates the averaged forward-field of parcel  $k$

$$\tilde{\mathbf{g}}_k = \frac{1}{n_k} \sum_{i \in \text{Parcel } k} \mathbf{g}_i, \quad (5.26)$$

$\mathbf{g}_i$  indicating dipole  $i$  forward-field (the  $i^{\text{th}}$  column of  $G$ ).

In (5.25), the mean  $\mu_k^{(l+1)}$  of the reference probability associated with parcel  $k$  is set equal to averaged intensity estimated from the previous iteration, *i.e.*  $\alpha_k^{(l+1)} \mu_k^{(l+1)} = (\sum q_k^{*(l)})/n_k$  where  $q_k^*$  is given by (5.15). We thus assume that, within a parcel, the dipole amplitudes are i.i.d (see section 5.2.2.3). Compare to MEMiter<sub>1</sub>, MEMiter<sub>2</sub> exploits the successive entropic estimates in order to iteratively optimize the reference probability distribution.

In both cases the iterative process ends when convergence is achieved.

### 5.2.2. Data-driven cortex parcelling

Solving the MEG inverse problem using the MEM approach as described in section 5.2.1 requires both the definition of  $K$  cortical regions or parcels of activation associated with the hidden state variable  $S$  and the initialization of the parameters  $\{\alpha_k^{(0)}, \mu_k^{(0)}\}_{k=1, \dots, K}$  that characterize the local prior or reference probability law. As this initial step is crucial for the convergence of the algorithm towards a stable and realistic solution, a natural way and an essential contribution of the present work consists in developing a data-driven cortex parcelling and initialization procedure. In the spirit of an empirical bayesian approach, the data are used in order to derive both the parcels and some prior information to be incorporated in the reference probability distribution. We therefore exploit the recently proposed Multivariate Source Prelocalization (MSP) to estimate probability-like coefficient of activation associated with each dipole of the distributed cortical model [14, 15].

### 5.2.2.1. *The multivariate source prelocalization*

Let us consider the following question: is it possible to sort the dipoles of the distributed model according to their contribution to the data, assuming that they are indeed actives? In other word, is it possible to estimate their relative influence for explaining the data, irrespective of their own, true and absolute quantitative contribution?

The recently developed Multivariate Source Prelocalization (MSP) enables one to address this issue. By comparing the normalized dipole lead-fields or scalp contributions to the effective measured field topographies, MSP does not aim at solving the inverse problem but focusses on the sole localization ambiguity associated with it. This quantitative and multivariate comparison leads to an estimate of each dipole contribution to the data, which can directly be used to define and weight the activation region of the MEM based model. We here implemented the simplest version of the MSP approach described hereafter.

Considering a full cortical distributed model with  $n \gg d$ , one can reasonably assume that the source space has enough degrees of freedom to fully explain the measured data including noise. Any observation  $\mathbf{m}$  thus belongs to the space spanned by the forward-field matrix  $G$ . Then, by projecting each dipole forward-field  $\mathbf{g}_i$  onto the *ray-space* spanned by  $\mathbf{m}$ , one allows to compute the normalized contribution  $a_i$  of each distributed source  $i$  to the data

$$\|P_{\mathbf{m}} \mathbf{g}_i\|^2 = a_i \|\mathbf{g}_i\|^2, \quad (5.27)$$

where

$$P_{\mathbf{m}} = \mathbf{m}(\mathbf{m}^t \mathbf{m})^{-1} \mathbf{m}^t \quad (5.28)$$

projects the source onto the data space. The dipoles can then be sorted and selected according to the estimated activation coefficients  $a_i$ . Given a low  $p$ -value (say  $p = 0.1$ ), the corresponding threshold  $a_c$  can be derived such that

$$P(a_i < a_c \mid \text{dipole } i \text{ is active}) = p.$$

Given  $a_c$ , any dipole  $i$  that do not satisfy  $a_i > a_c$  is removed from the model, unless it belongs to the close neighbourhood of a selected dipole. A neighbourhood is being defined by a parcel according to the data-driven process described in the next paragraph.

### 5.2.2.2. *The data-driven clustering*

Cortical parcels are needed for the MEM based reconstruction and the optimal use of the hidden state variables. A natural way of building such parcels is by clustering the dipoles of the distributed model according to some anatomical as well as functional criteria. This guarantees a clear correspondance between the two resolution levels of the model, the parcels and the dipoles. Furthermore, it should lead to an optimal anatomo-functional description of the solution space. We therefore use the aforementioned activation coefficients (resp. the cortical distance between sources) in order to enforce functional (resp. anatomical) coherence within parcel. Since the activation coefficients are estimated from the data by MSP, this parcelling approach of the cortical surface is referred to as a data-driven clustering (DDC).

Let us denote by  $C$  the group of remaining dipoles after having performed the MSP and associated selection procedure. As a starting point for the proposed iterative DDC, each dipole of  $C$  defines an initial and single cluster. Each dipole  $i$  is successively considered, in the decreasing order of its activation coefficient  $a_i$ , and defines a seed of a new cluster into which the neighbouring dipoles are integrated. Any surrounding dipole is being considered as a neighbour as soon as its distance to the dipole seed is inferior to a given limit  $D$ . Neighbouring dipoles that do not belong to the initial selected subset  $C$  might be incorporated in the cluster and thus in the model. However, if a surrounding dipole already belongs to a cluster that contains more than one dipole, it is not considered for incorporation into another cluster. This prevents for overlapping parcels. DDC is complete when all the dipoles of  $C$  have been considered as cluster seeds. Figure 5.1 illustrates this DDC process.

Nevertheless, some clusters might only contain one or two dipoles. Those few clusters are aggregated with the closest neighbouring cluster. Such a procedure can be seen as a simple smoothing or spatial regularization procedure. Note finally that the limit distance  $D$  is an hyperparameter of this approach whose influence is illustrated in section 5.3. This parameter  $D$  is the resolution scale.



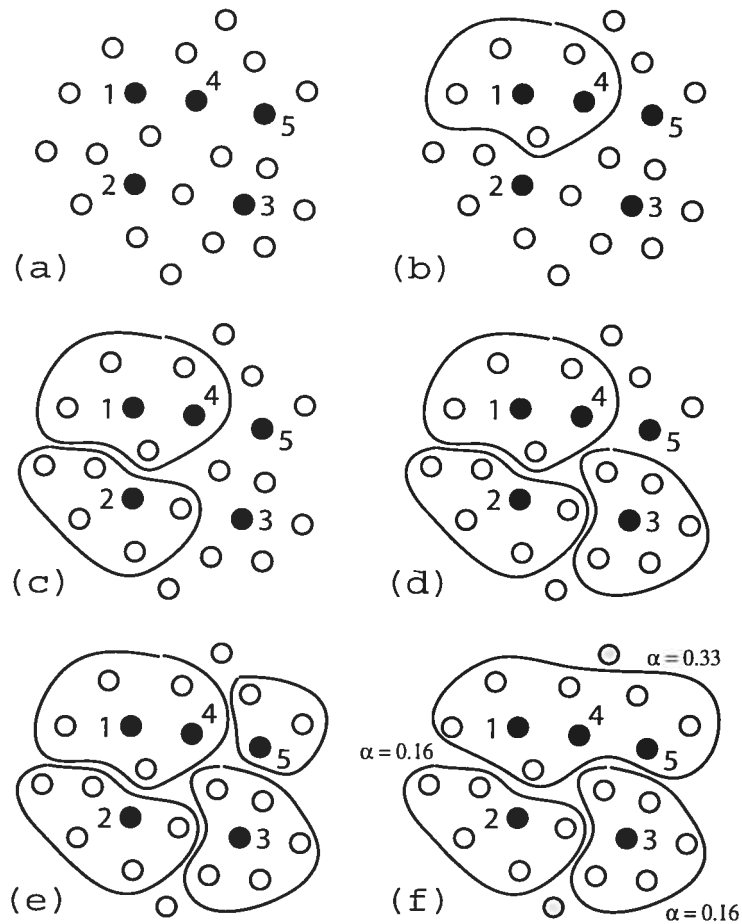


FIGURE 5.1. Illustration of the clustering process. The top left image contains the initial set of dipoles. The filled circles are in  $S$  and are ordered with their  $a_i$ . The top-right image shows the first constructed region. The following images illustrate the construction of the three other regions. Image (f) shows the aggregation process and the final situation with 4 regions together with the initial  $\alpha_k$ 's as discussed in section 5.2.2.3. There remains one dipole which does not belong to any region.

### 5.2.2.3. The initialization of the parcel parameters

Each parcel  $k$  is characterized by three parameters: the mixture parameter or probability  $\alpha_k$  of being active and the mean  $\mu_k$  and the variance  $\Sigma_k$  of the gaussian law that defines the reference probability distribution within the parcel.

We here consider two types of initialization for parameter  $\alpha_k$ :

- *a non-informative initialization* where  $\alpha_k$  is set equal to 1/2 (see the description of the MEMiter<sub>2</sub> algorithm in section 5.2.1.4). This initialization will be always considered when not performing the DDC beforehand.
- *an informative initialization* which uses the MSP derived information and sets  $\alpha_k$  equal to the the ratio of the number of selected dipoles by the number of dipoles in parcel  $k$ . This initialization will be used when having previously performed the DDC.

In order to initialize the reference probability distribution within parcel  $k$ , we compute the matrix  $\tilde{G}$  which columns  $\tilde{\mathbf{g}}_k$  are the averaged forward-field within the parcel  $k$  and then derive the corresponding pseudo-inverse solution such that  $\boldsymbol{\mu} = \tilde{G}^+ \mathbf{m}$ , where  $\tilde{G}^+$  denotes the pseudo-inverse of  $\tilde{G}$  as given by equation (5.3) and  $\boldsymbol{\mu}$  is the vector constructed with the  $\mu_k$ . This procedure amounts to introduce, as an initial local prior, the estimate of the amplitude of a virtual ECD associated with parcel  $k$  and whose forward-field is given by the average forward-field  $\tilde{\mathbf{g}}_k$  over the parcel.

Finally, we set  $\Sigma_k = \sigma_k^2 \mathbb{I}$ , with  $\sigma_k^2 = 10^{-2} \mu_k^2$ , so that the dipoles of parcel  $k$  are i.i.d.

### 5.3. APPLICATION

#### 5.3.1. The MEG data simulation

Since the MEG sources are widely believed to be restricted to the pyramidal neuron cells of the cortical strip [17], a common approach within the distributed model framework consists in constraining the dipoles to be distributed onto the cortical surface extracted from a structural MRI [7]. After the segmentation of the MRI volume, dipoles are typically located at each node of a triangular mesh of the white/grey matter interface [13]. Furthermore, since the apical dendrites of these cortical neurons are organized perpendicularly to the surface, the corresponding dipoles are also constrained to have this particular orientation.

In order to simulate MEG data, a 3D high resolution (voxel size: 0.9375 mm  $\times$  0.9375 mm  $\times$  1.5 mm) MRI volume from a healthy volunteer was segmented. The boundary between white and grey matter was approximated with small triangles whose vertices provided 7,000 dipole positions uniformly spread all over the cortex. The spatial resolution was high enough to well describe the cortical

topology, since the mean space between two neighbouring dipoles was less than 5 mm.

We calculated the forward operator  $G$  corresponding to this dipole mesh. Since head tissues are non-magnetic, estimating matrix  $G$  by solving the electromagnetic equations does not require a precise description of the geometrical and physical properties of the head. We therefore simply designed a single shell spherical head model, which allowed us to calculate an analytical solution of the equations of magnetic field propagation [20].

MEG data were simulated over 150 sensors uniformly spread all around the head, by artificially activating two simultaneous extended sources with respectively a positive and negative intensity.

For the source configuration chosen,  $R = 10$  distributions of intensities ( $\mathbf{q}_r$ ) were simulated driven by a gaussian random process whose mean was set to  $\pm 1$ , depending on the polarity of the simulated source. The data were then produced by convolving these single time point intensities with the forward matrix. Each corresponding single-trial simulated observation  $\mathbf{m}_r$  was corrupted by noise so that  $\mathbf{m}_r = G\mathbf{q}_r + \rho\boldsymbol{\eta}$ , where  $\boldsymbol{\eta} \sim N(0, 1)$  and  $\rho$  indicates the noise intensity.  $\rho$  was set such that two signal to noise ratios (SNR) were considered:  $\text{SNR} = \infty$  (no noise) and  $\text{SNR} = 20$  dB. Finally, two quantities were computed from these simulated single-trial data and use in the reconstruction process:

- the evoked or averaged observation  $\mathbf{m} = \frac{1}{R} \sum_r \mathbf{m}_r$  to be entered in the reconstruction process (see equation (5.14)),
- the noise variance  $\sigma^2 = \frac{1}{R-1} \sum_r (\mathbf{m}_r - \mathbf{m})^2$  (see equation (5.14)).

Applied on the averaged observation  $\mathbf{m}$ , the two proposed iterative algorithms (see section 5.2.1.4) were compared to a LORETA-like solution and (see section 5.1) and to a non-iterative MEM using an arbitrary non data-driven clustering (see figure 5.2(b)) and evaluated using the evaluation criteria described in the following paragraph.

LORETA was applied with the data-driven clustering. The parcels were present in the laplacian operator  $W$  in equation (5.5) by enforcing the dipoles within a parcel to be correlated.

Finally, whatever the approach, the solution space was restricted to parcels  $k$  such that  $\alpha_k \neq 0$ .

### 5.3.2. The evaluation criteria

We here use three different quantitative criteria to evaluate the inverse source estimates. The two first quantities are classical criteria and aims at respectively evaluating the Mean Square Error (MSE) between the true source amplitudes and the estimated ones and the Data Fit Error (DFE) which measures the residuals between the true simulated data and the data part explained by the reconstructed sources. These two criteria are defined as follows and expressed in decibel (dB):

$$\text{MSE} = -10 \log \left( \frac{\|\mathbf{q}^* - \mathbf{q}_f\|^2}{\|\mathbf{q}_f\|^2} \right), \quad (5.29)$$

$$\text{DFE} = -10 \log \left( \frac{\|\mathbf{m} - G \mathbf{q}^*\|^2}{\|\mathbf{m}\|^2} \right), \quad (5.30)$$

where  $\mathbf{q}^*$  is the estimated amplitude distribution and  $\mathbf{q}_f$  is the true one used to generate the MEG data.

We finally introduce the use of the Receiver Operating Characteristics (ROC) analysis to evaluate the source detection ability of an inverse approach. ROC analysis have been extensively used for nuclear medicine applications and consist of displaying the true positive rate (sensitivity) as a function of the false positive rate (1 - specificity). Each point of the ROC curve corresponds to a particular threshold value on the estimated amplitude distribution. To complete the ROC curve, the minimum threshold (no threshold) as well as the maximum threshold (no remaining active source) have to be considered. Then, the area under the curve (AUC) indicates the probability that the considered inverse method would correctly distinguish an active dipole from a non-active one [8].

In practice, since the number of active dipoles is much smaller than the number of non-active ones and in order to compute an unbiased estimate of the sensitivity and the specificity, we randomly select as much non-active dipoles as active ones in order to perform a fair ROC analysis.

Note that this criterion complements the two others since it evaluates the detection power of the inverse method but not its estimation performance. Indeed, only the energy but the polarity of the source is taken into account.

## 5.4. RESULTS

Figure 5.2 illustrates the data-driven clustering. One should qualitatively emphasize that the obtained parcels are homogeneous in size. Moreover, quite a lot of dipoles have been rejected so that two main cortical areas remain in the model. Note that each of these two areas encompasses one of the two true simulated sources. At that stage, the DDC enabled us to restrict the dimension of the problem by focussing on large brain cortical regions that are most likely to be activated. These large areas are furthermore split into functionally coherent parcels as shown in the MSP map. The DDC appears not to be much affected by noise.

On figure 5.3, the initial  $\alpha_k$  and  $\mu_k$  values computed from the DDC and MSP procedure (see section 5.2.2.3) are displayed over the cortical surface. The true active regions do correspond to high estimated values of  $\alpha_k$ , indicating a high probability of being activated. However, these  $\alpha$ -maps do exhibit other high probable active regions that do not contribute to the data. These false positive sources will have to be removed by the iterative MEM procedure. Regarding the initial estimate of the mean amplitude  $\mu_k$  of each parcel  $k$ , the first criterion relates to the polarity of the intensity. Indeed, this polarity remains the same during the MEM iterative process and cannot be corrected. Figure 5.3 shows that the polarity of the sources has been correctly estimated, except for the highest value of the clustering hyperparameter ( $D = 15$  mm). This might reflect a limitation of the approximated averaged forward model. Indeed, the larger the parcel, the less realistic our approximation of an averaged forward-field for explaining the whole contribution of the parcel.  $D = 15$  mm may thus indicate an upper limit for the parcel size one should consider.

Figure 5.4 shows the non-iterative MEM estimations using different values of  $D$ . Whatever this value, the true active sources always belong to the final set of reconstructed regions. However, all solutions do also find false positive sources that differ according to the value of the clustering hyperparameter. This result argues in favour of considering several hyperparameter values and then taking the averaged solution. This should both reduce the dependency on a single value of the hyperparameter and increase the specificity of the source reconstruction. Such an averaging procedure is illustrated on figure 5.5 where eleven source reconstructions were performed for each hyperparameter value between  $D = 5$  mm

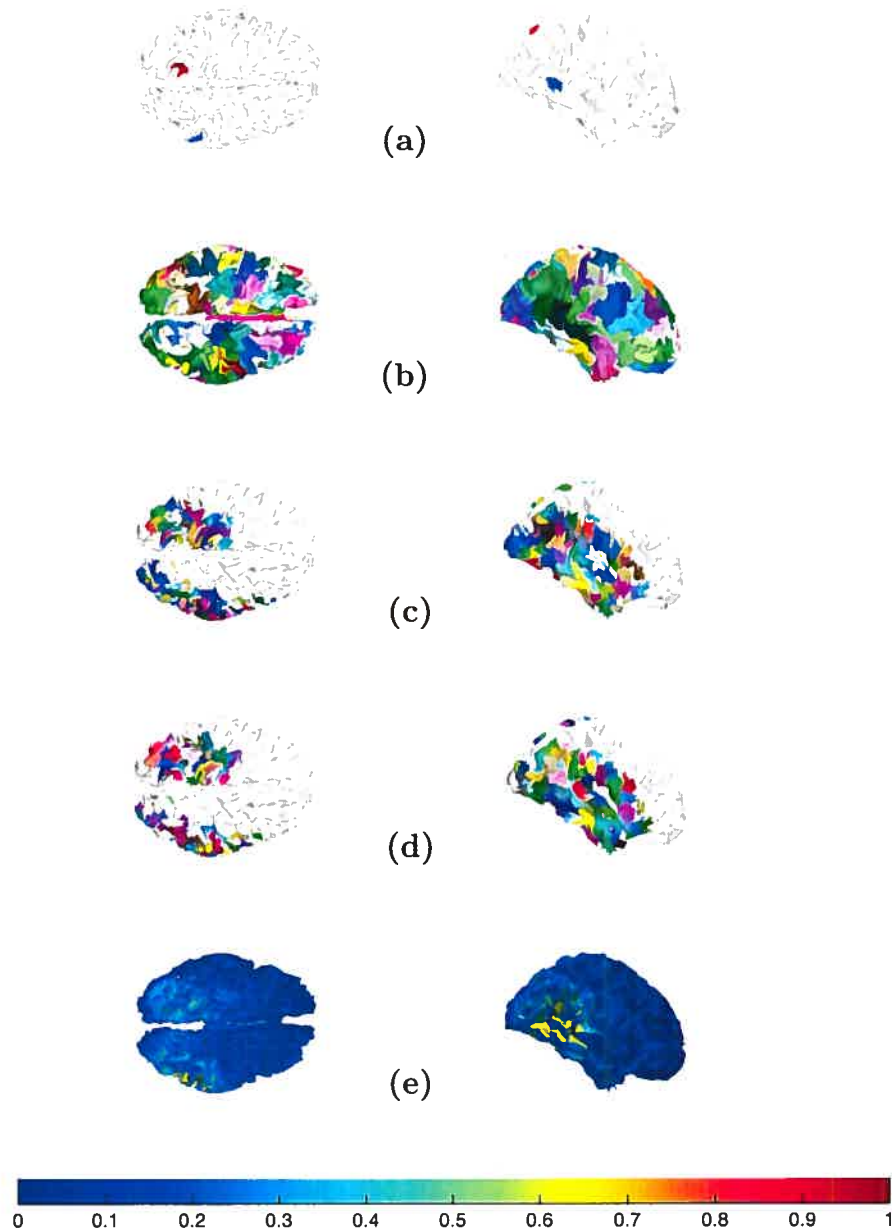


FIGURE 5.2. (a): Activation of the phantom on the cortex surface. The red dipoles are positive and the blue ones are negative. (b): The set of regions we use in the simulations for the reconstruction without the data-driven clustering. The regions are the connected sets of the same color. (c) and (d): Examples of data-driven clustering with the two different sets of data used in these simulations for a characteristic scale of  $D = 10$  mm. The big white region contains the dipoles that don't belong to any region. The active regions are the connected sets of the same color. The data has been generated with the activity of the phantom. (c): There were no added noise. The clustering in the bottom images (d) has been created with 20 dB noisy data. (e): The MSP map ( $a_j$  in equation (5.27)). The color legend corresponds to the color for the MSP map image. Two cortices on the same row are just two views of the same thing.

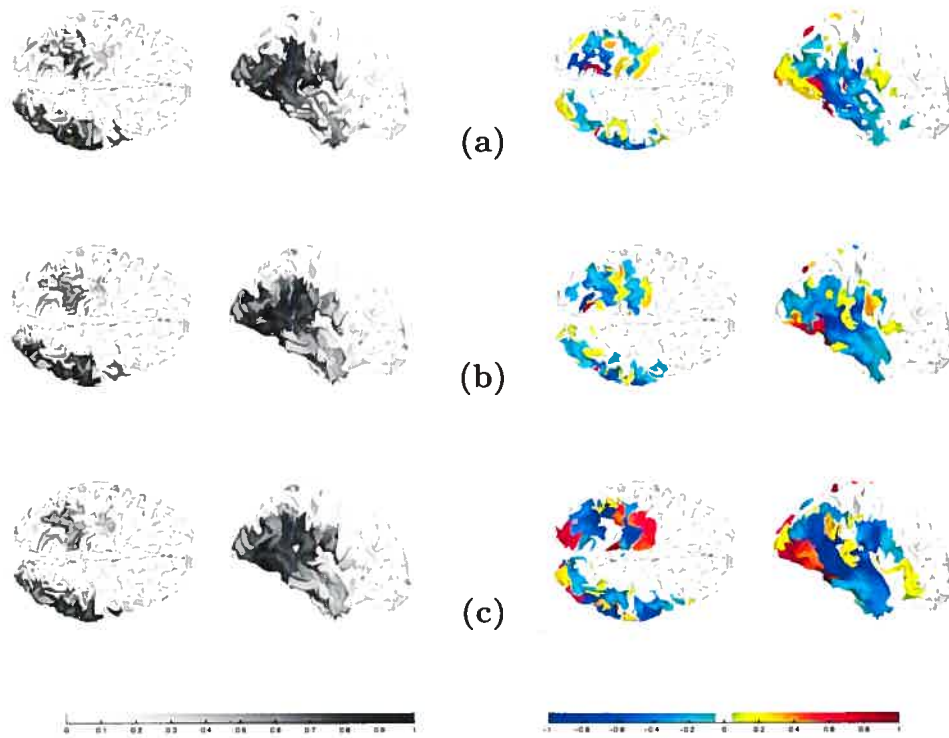


FIGURE 5.3. Image of the initial  $\alpha_k$  for all the regions with shades of gray. The darker is the region, the greater is  $\alpha_k$ . The first row (a) shows the  $\alpha_k$  the smallest reconstruction scale (5 mm). The second row (b) with a greater reconstruction scale (10 mm). And the third row (c), the largest reconstruction scale (15 mm). For the left images, the white means  $\alpha_k = 0$  and black  $\alpha_k = 1$ . For the right images, the intensities displayed are between -1 and +1. The cold colors (royal to light blue) are for intensities lower than 0 and the hot colors (yellow to red) are for intensities greater than 0.

and  $D = 15$  mm. Most of the false positive sources are indeed removed by this filtering. The best qualitative reconstruction is given by the MEMiter<sub>2</sub> algorithm when coupled with the DDC procedure. It minimizes the number of false positive sources and is the only solution that reconstruct the negative simulated source.

On figure 5.5, one can furthermore compare the iterative and non-iterative MEM solutions, either when initialized with DDC or with a non data-driven clustering [1]. In the latter case, the extension of the active sources is hardly estimated and much more false positive sources were reconstructed. Note that these solutions based upon an arbitrary clustering did not benefit from a dipole selection. Finally, the LORETA-like approach applied on the noisy data and

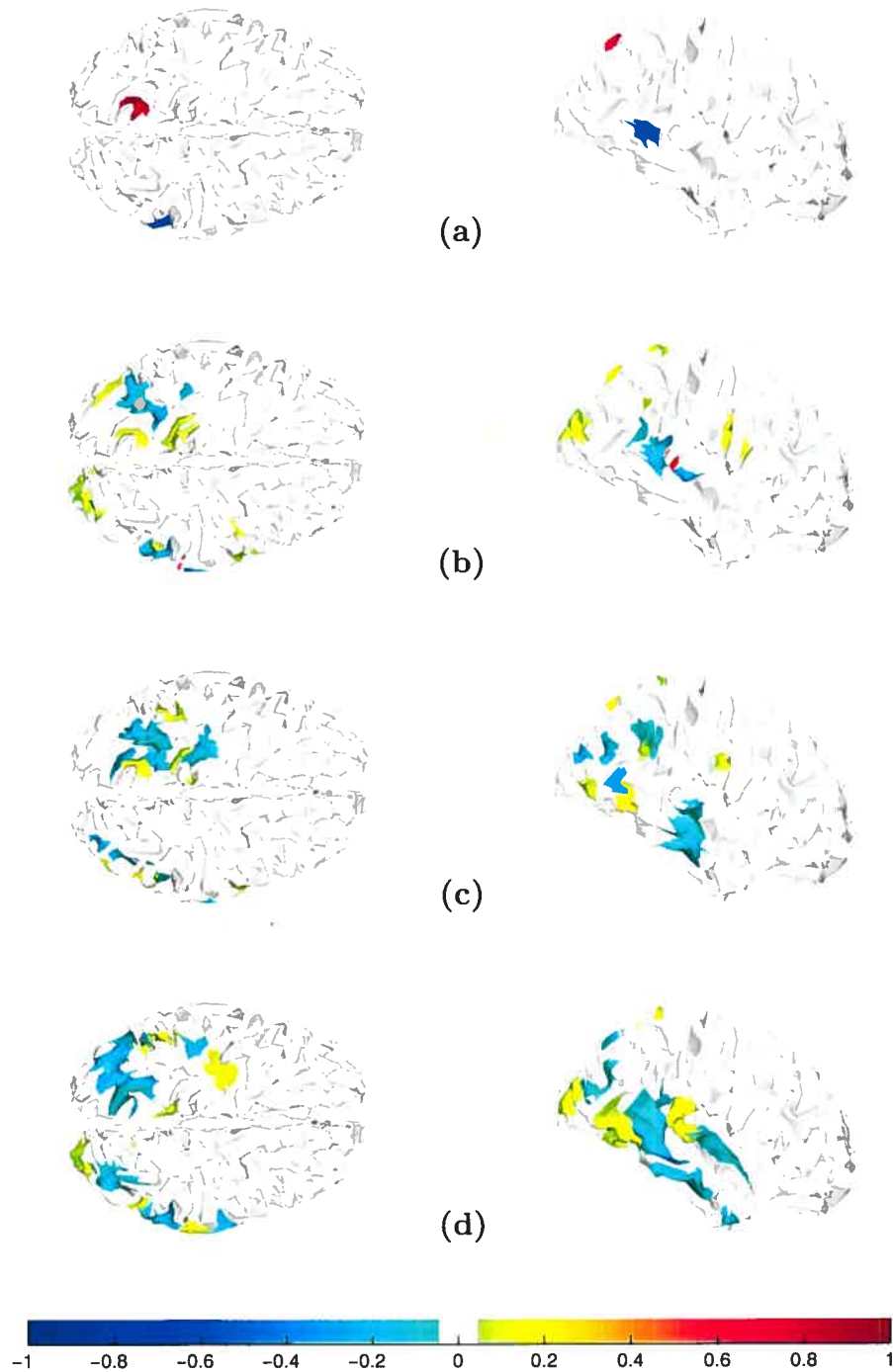


FIGURE 5.4. Results of the reconstruction over different activation scales using noisy data with non-iterative MEM. Top row (a) displays the phantom. Second row (b) shows the results at a reconstruction scale of 5 mm (the smallest scale used). Third (c), 10 mm. Four (d), 15 mm (the largest scale used). The colour legend is the same as in the previous figure.



together with the DDC does not produce a focal solution and moreover mislocate the maxima of the activity. This solution seems to be more sensitive to noise than the MEM ones. Note that the MEM approaches already implicitly include a sort of smoothing (LORETA-like) prior through the use of parcels but then use a more general and flexible criterion such that more focal solutions can be obtained.

The same conclusions can be drawn from the quantitative evaluation of the same inverse solutions. Table 5.1 gives the evaluation criterion values for each of the algorithm, when considering either non-noisy or noisy data. MEMiter<sub>2</sub> appears to give the best results, although it leads to a small MSE value with non-noisy data. This small value indicates a large error in the estimation of the source amplitude distribution. This is due to the fact that, when dealing with non-noisy data, the MEM approach hardly tolerates any departure from the data fit (as supported by the DFE criterion value). Then, if the model does not fit the data perfectly, MEM will tend to compensate this departure by overestimating the source amplitudes. On noisy data, such a departure is tolerated and this behaviour is no more observed.

## 5.5. DISCUSSION

In this paper, we further investigated and assessed the usefulness of the Maximum Entropy on the Mean (MEM) principle for regularizing and solving the MEG inverse problem. The proposed framework is a general probabilistic approach which proved suitable for the introduction of needed multiple prior knowledge on the solution. These constraints can be specified in a flexible way since their nature and form may be of all kind, providing that they can be expressed in terms of probability distribution. The influence of these prior is furthermore locally and softly controlled, thanks to the definition of regions of activation and associated hidden variables. Note also that, because of the convenient dual formulation of the regularization criterion, the iterative computational process only deals with low dimension variables, in sensor space.

However, the most important contribution of this paper lies in the proposed data-driven clustering (DDC). Based on the Multivariate Source Prelocalization, the DDC allows to define anatomically and functionally coherent regions and provides initial estimates of the prior probability distribution parameters. All

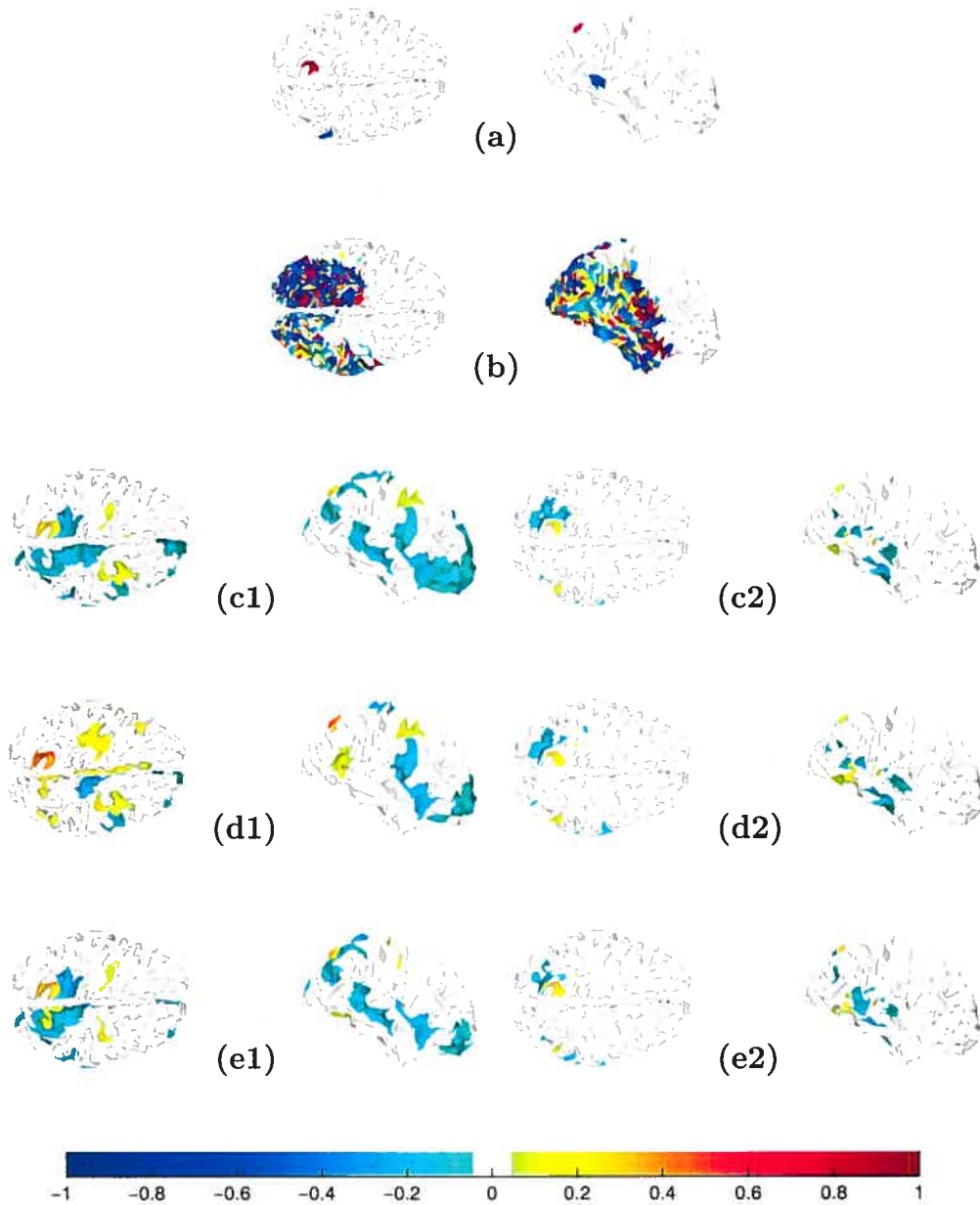


FIGURE 5.5. Results of the reconstruction for the methods discussed in this paper with 20 dB noisy data. (a) is the phantom as in figure 5.2 for easy comparisons. (b) is the results with LORETA and DDC. (c1) is non-iterative MEM without DDC. (c2) is non-iterative MEM with DDC. (d1) is MEMiter<sub>1</sub> without DDC. (d2) is MEMiter<sub>1</sub> with DDC. (e1) is MEMiter<sub>2</sub> without DDC. (e2) is MEMiter<sub>2</sub> with DDC. The colour legend is the same as in the previous figure.

No noise in data:

Method	DDC	MSE (dB)	DFE (dB)	ROC
MEM	no	<b>2.0</b>	13.6	<b>0.74</b>
	yes	1.7	28.9	<b>0.74</b>
MEM <sub>iter<sub>1</sub></sub>	no	-29.0	8.2	0.65
	yes	-1.6	39.0	0.73
MEM <sub>iter<sub>2</sub></sub>	no	-0.8	40.2	0.73
	yes	-10.5	<b>44.7</b>	<b>0.74</b>
loreta	yes	-0.5	0.5	0.65

20 dB noise in data:

Method	DDC	MSE (dB)	DFE (dB)	ROC
MEM	no	0.2	16.6	<b>0.74</b>
	yes	0.4	17.3	0.73
MEM <sub>iter<sub>1</sub></sub>	no	-1.7	10.4	0.61
	yes	0.8	18.7	0.69
MEM <sub>iter<sub>2</sub></sub>	no	-0.1	17.7	<b>0.74</b>
	yes	<b>1.7</b>	<b>20.0</b>	<b>0.74</b>
loreta	yes	-28.8	-23.7	0.655

TABLE 5.1. Results of the reconstruction for non-iterative (non-iter), iterative MEM (iter<sub>1</sub>, iter<sub>2</sub>) and LORETA for non-noisy and noisy (20 dB) MEG data. The acronym DDC stands for data-driven clustering.

the particularity and the efficiency of the MEM based inverse approach relies upon the definition of such regions. Using the data and a distributed source model of the cortex is a natural way of providing such a model and DDC appears to be particularly useful for improving the MEM reconstruction.

But more generally, the DDC coupled with the MEM approach fits into an overall tendency in the recent literature to improve the now traditional distributed model, both in terms of tractability and efficiency [14,22]. Indeed, parcelling out the cortex enables one to link neighbouring dipoles according to their functional properties, hence reducing the number of unknown parameters to be estimated. This type of approach tends to fill the gap between ECD and distributed

inverse methods. Each hidden variable estimated by the MEM characterizes a whole cortical region as if modeled by a single ECD. The prior amplitude of each parcel source was here initially calculated using an average forward field which may be seen as an ECD contribution to the data. On the other hand, our approach enables the user to deal with a lot more region than usual ECD methods allows to and still provides an estimate of the source anatomical extension.

Such hybrid methods that attempt to conciliate the advantages of complementary models, deserve to be further developed since they might provide performant alternative solutions that are easy to perform and interpret. The DDC approach could be coupled with any distributed inverse method. However, the proposed MEM framework is particularly suitable for dealing with cortical parcels and proved here to be further improved by such a coupling. These performance now need to be extensively evaluated, both in space and time, which will be addressed in a subsequent paper. Finally, DDC interestingly illustrates another aspect of the usefulness of the Multivariate Source Prelocalization approach.

## 5.6. ACKNOWLEDGEMENTS

J-M.L. would like to thank NSERC program for financial support. E.L. would like to thank Philippe St-Jean for providing him the computer accessibility to do the simulations presented in this work.

## BIBLIOGRAPHY

- [1] C. Amblard, E. Lapalme, and J-M Lina. Biomagnetic source detection by maximum entropy and graphical models. *IEEE Transactions on Biomedical Engineering*, 55(3):427–442, March 2004.
- [2] S. Baillet and L. Garnero. A bayesian approach to introducing anatomo-functional priors in the EEG/MEG inverse problem. *IEEE Trans. Biomed. Eng.*, 44:374–385, 1997.
- [3] S. Baillet, J. C. Mosher, and R. M. Leahy. Electromagnetic brain mapping. *IEEE Sign. Proc. Mag.*, 18:14–30, 2001.
- [4] G. L. Barkley. Controversies in neurophysiology. MEG is superior to EEG in localization of interictal epileptiform activity: Pro. *Clin. Neurophysiol.*, 115:1001–1009, 2004.
- [5] C.J.S. Clarke and B.S. Janday. The solution of the biomagnetic inverse problem by maximum statistical entropy. *Inverse Problems*, 5(4):483–500, Août 1989.
- [6] A. M. Dale, A. K. Liu, B. R. Fischl, R. L. Buckner, J. W. Belliveau, J. D. Lewine, and E. Halgren. Dynamic statistical parametric mapping: combining fMRI and MEG for high-resolution imaging of cortical activity. *Neuron*, 26:55–67, 2000.
- [7] A. M. Dale and M. Sereno. Improved localization of cortical activity by combining EEG and MEG with MRI surface reconstruction: a linear approach. *J. Cognit. Neurosci.*, 5:162–176, 1993.
- [8] E.R. DeLong, D.M. DeLong, and D.L. Clarke-Pearson. Comparing the areas under two or more correlated receiver operating characteristic (ROC) curves: a nonparametric approach. *Biometrics*, 44:837–845, 1988.
- [9] I. F. Gorodnitsky, J. S. George, and B. D. Rao. Neuromagnetic source imaging with FOCUSS: a recursive weighted minimum norm algorithm. *Electroencephalogr. Clin. Neurophysiol.*, 95:231–251, 1995.
- [10] M. S. Hämmäläinen and R. J. Ilmoniemi. Interpreting magnetic fields of the brain—minimum norm estimates. *Med. Biol. Eng. Comput.*, 32:35–42, 1994.
- [11] G. Le Besnerais, J-F. Bercher, and G. Demoment. A new look at entropy for solving linear inverse problems. *IEEE Transactions on Information Theory*, 45(5):1565–1578, July 1999.
- [12] A. K. Liu, J. W. Belliveau, and A. M. Dale. Spatiotemporal imaging of human activity using functional MRI constrained magnetoencephalography data: Monte Carlo simulations. *Proc. Natl. Acad. Sci. U.S.A.*, 95:8945–8950, 1998.
- [13] J. F. Mangin. From 3D magnetic resonance images to structural representations of the cortex topography using topology preserving deformations. *J. Math. Imag. Vision*, 5:297–318, 1995.
- [14] J. Mattout, M. Péligrini-Issac, A. Bellio, J. Daunizeau, and H. Benali. Localization Estimation Algorithm (LEA): a supervised prior-based approach for solving the EEG/MEG inverse problem. In C. J. Taylor and J. A. Noble, editors, *Information Processing in Medical Imaging (IPMI 2003)*, LNCS 2732, pages 536–547. Springer-Verlag, Berlin, 2003.

- [15] J. Mattout, M. Péligrini-Issac, A. Bellio, J. Daunizeau, and H. Benali. Better conditioning the MEG/EEG inverse methods: the multivariate source prelocalization approach. *Proceedings of the International Symposium on Biomedical Imaging*, 2004.
- [16] J. C. Mosher, R. M. Leahy, and P. S. Lewis. EEG and MEG: Forward solutions for inverse methods. *IEEE Trans. Biomed. Eng.*, 46:245–259, 1999.
- [17] P. L. Nunez and R. B. Silberstein. On the relationship of synaptic activity to macroscopic measurements: does co-registration of EEG with fMRI make sense? *Brain Topogr.*, 13:79–96, 2000.
- [18] R.D. Pascual-Marqui, C.M. Michel, and D. Lehmann. Low resolution electromagnetic tomography: a new method for localizing electrical activity in the brain. *Int. Journal of Psychophysiology*, 18:49–65, 1994.
- [19] C. Phillips, M. D. Rugg, and K. J. Friston. Anatomically informed basis functions for EEG source localization: Combining functional and anatomical constraints. *Neuroimage*, 16:678–695, 2002.
- [20] J. Sarvas. Basic mathematical and electromagnetic concepts of the biomagnetic inverse problem. *Phys. Med. Biol.*, 32:11–22, 1987.
- [21] M. Scherg and D. von Cramon. Evoked dipole source potentials of the human auditory cortex. *Electroenceph. Clin. Neurophysiol.*, 65:344–360, 1986.
- [22] N.J. Trujillo-Barreto, E. Aubert-Vázquez, and P.A. Valdés-Sosa. Bayesian model averaging in EEG/MEG imaging. *NeuroImage*, 21(4):1300–1319, April 2004.

## Chapitre 6

---

### RECONSTRUCTION SPATIO-TEMPORELLE Troisième article

Cet article inclus dans cette thèse a été soumis à la revue IEEE Transactions on Biomedical Engineering le 20 juin 2004 [27].

La précision temporelle est le principal intérêt de la MEG face aux autres méthodes d'imagerie plus populaires comme l'IRMf ou la PET qui ont, par contre, l'avantage de disposer d'une énorme quantité de données permettant une excellente résolution spatiale. Toutefois, leur précision temporelle n'étant pas aussi bonne, il n'est pas possible de sonder l'activité du cerveau aux échelles temporelles de l'activité cérébrale comme le permettent la MEG et l'EEG.

Dans nos précédents travaux, seule une régularisation spatiale était pratiquée. Même si nos méthodes extraient beaucoup d'information spatiale des données MEG, seule une méthode intégrant l'information temporelle permettra d'identifier le moment et l'endroit des étapes du traitement de l'information du cerveau.

Dans ce chapitre, nous traitons de la régularisation spatio-temporelle dont la difficulté principale est d'intégrer les corrélations temporelles offertes par les données MEG dans un modèle spatial. Avec notre formalisme, cela correspond à spécifier une loi de référence qui inclut ces corrélations.

Plutôt que d'intégrer directement les corrélations temporelles dans la loi de référence, nous allons spécifier les hyperparamètres de la loi pour inférer des corrélations entre les solutions à différents temps. Comme nous l'avons fait au chapitre précédent, nous présentons une méthode de parcellisation du cortex. Cette méthode utilise la MSP et les données MEG pour donner un ensemble de régions d'activation valides sur toute la fenêtre temporelle considérée. Avec ces régions, on estime les hyperparamètres  $\alpha_k$  et  $\mu_k$  de manière analogue à ce qui est

fait dans le chapitre 5 sauf qu'un lissage temporel est appliqué pour éviter que la solution soit discontinue dans le temps.

Après la présentation de la méthodologie, nous appliquons ces méthodes à des données MEG simulées et nous les comparons avec deux autres versions du MEM non "temporalisées" pour mettre en évidence le travail fait par la régularisation temporelle.

Ensuite, la méthode est appliquée à des données MEG prises lors d'une expérience somato-sensorielle qui consistait à donner de légères décharges électriques sur chacun des doigts d'un sujet et à mesurer les champs magnétiques produits par l'activation cérébrale. Nous verrons que nos résultats concordent avec les connaissances neurophysiologiques.

Le principal avantage du formalisme présenté dans ce chapitre est le peu d'information *a priori* introduite dans le modèle. La seule hypothèse sur les sources étant que les dipôles sont regroupés en régions, que leurs activations sont corrélées et que ces propriétés sont constantes dans le temps. On estime ensuite les régions et les autres hyperparamètres de la loi de référence. Le MEM devient alors une méthode puissante pour régulariser le problème inverse en MEG.



# Spatio-temporal biomagnetic source reconstruction with maximum entropy

Article soumis à *IEEE Transactions on Biomedical Engineering* le 20 juin 2004.

ERVIG LAPALME, *Centre de Recherches Mathématiques et Département de Physique de l'Université de Montréal*

JEAN-MARC LINA, *Centre de Recherches Mathématiques et Département de Génie Électrique de l'École de Technologie Supérieure, Montréal*

BERNARD GOULARD, *Centre de Recherches Mathématiques et Département de Physique de l'Université de Montréal*

## Abstract

In [1, 8], some work has been done to spatially regularize the magnetoencephalographic inverse problem using the maximum entropy on the mean method. A parcelling method was presented and a method of inference for a reference measure was developed. In this work, we extend it to a spatio-temporal regularization applied to time signals coming from magnetoencephalography (MEG) taking this way full of the great time resolution offered by MEG signals. It consists in integrating information coming from a large window of time samples giving better statistics to infer from. We demonstrate the effectiveness of the method by simulations and also with real MEG data coming from somatosensory experiments.

**Keywords:** Magnetoencephalography, Maximum Entropy on the Mean, Parcelling, Inverse Problems, Regions of Activation, Brain Imaging, Spatio-Temporal Regularization, Multivariate Source Preselection.

## 6.1. INTRODUCTION

Magnetoencephalography (MEG) consists in measuring the magnetic field created by electric currents flowing in neurons to detect the presence of cerebral activation. The great advantage of MEG over other functional brain imaging methods is its time accuracy. For example, functional magnetic resonance imaging have a time resolution one second at best since it relies on the hemodynamic response of neuronal activation. On its side, MEG measures the magnetic fields directly generated by the neural electric currents which are propagated instantaneously through the brain tissues yielding signals with strong time correlations. The time resolution is thus only limited by the sampling rate. However, they cannot give unequivocally the exact location and the number of neural areas giving rise to the observed electromagnetic field: computation of electric neural sources able to create electromagnetic fields variations observed on the scalp has an infinite number of solutions. This is the well known inverse problem. In order to take full advantage of the real time experiments made available by MEG, we must couple the inverse solution to other approaches which include time correlations between successive signals. This is the aim of the present paper.

Since the beginning of MEG, inverse methods for finding the source of neuronal activity has been developed. The first methods to appear were dipole fit methods aiming at identifying the location and orientation of one or several dipoles that generate the magnetic field measured by the MEG captors. The major drawback of these methods is that the number of equivalent dipoles is left to the user and the solution is very sensitive to noise variations in the MEG data due to the non-linearity dependance of the dipole position parameters. The second class of methods is based on the distributed modeling of the cortex. The surface of the cortex is covered by hypothetical dipoles [3]. This way, the position and orientation of each dipole are known. So, compared to the dipole fit methods, the problem is simplified because there only remains to estimate the electric current intensity of each dipole. But some complexity is added on with the number of dipoles used to cover the cortex surface. This number in the model is only limited by the intensity estimation method at work and the computer processing power.

With this distributed modeling, we can calculate the influence of each activated dipole on the magnetic fields measured outside the head. Using a spherical modelization of the brain tissue (a widely made assumption), it is possible to

show that the magnetic field depends linearly on dipole intensities and is independent of the conductivity profile of the brain tissues (see [12]). So, if  $\mathbf{q}$  is an  $n$ -dimensional vector representing the intensity of each dipole, the problem is then reduced to

$$\mathbf{m} = G \mathbf{q}, \quad (6.1)$$

a linear matrix problem. The vector  $\mathbf{m}$  is a  $d \times T$  matrix containing the magnetic field values measured at  $d$  captors located outside the skull taken at  $T$  different time steps. The  $n \times T$  vector  $\mathbf{q}$  is the dipole intensities at each time step. The matrix  $G$  includes the physics of magnetic fields (Biot-Savart law). This matrix represents the effects of the neuronal activation on the MEG captors.

Reducing the inverse problem to a linear one is an important simplification. But the problem is still ill-posed for two main reasons. First, to have a realistic representation of the cortex surface, the number of dipoles is usually much bigger than the number of captors ( $n \gg d$ ). Second, due to the spherical symmetry the radial vector dipole component stays silent on the captors. Hence, we cannot reproduce the exact activation pattern even if we have perfect data.

Several methods have been developed in the several past years [1, 2, 6, 8, 11] to solve this inverse problem. Most of them had addressed only the single time reconstruction (without any temporal regularization). So, these methods do not really take advantage of the temporal accuracy of MEG since they do not consider the time correlation between successive signals. The work presented in this paper aims at using these correlations to estimate the temporal activation. Some researchers had already addressed this problem [2], but we take a novel approach based on previous work in [1, 8]. That approach uses the concept of region of activation. These regions are constructed using the MEG data and give clusters of dipoles which anatomically fit the activation pattern. That work was applied only on single time MEG data. In this paper, we want to use the temporal regularity to constrain the inverse problem better.

Here is an outline of the structure of this paper. Section 6.2 recalls what MEM is. Then, section 6.3 shows how to integrate the temporal information to make a data-driven parcelling. Section 6.4 follows with an exposition of temporal regularization with MEM. Sections 6.5 and 6.6 demonstrate the effectiveness of the methods introduced in this paper: MEG data are taken from simulation first and then from a realistic framework.

## 6.2. MAXIMUM ENTROPY ON THE MEAN

The maximum entropy on the mean method (MEM) relies on a stochastic description of the MEG problem. Given a multivariate random variable  $\mathbf{Q}_t$  representing the  $n$  dipoles intensities at time  $t$  and described by the probability law:

$$d\hat{P}_t(\mathbf{q}) = \rho_t(\mathbf{q})d\mathbf{q}, \quad (6.2)$$

MEM aims at identifying a particular  $\hat{P}_t$  based on information we have on the solution. Information comes from two main sources:

- (1) anatomical and functional properties of the dipoles on the cortex and
- (2) the MEG data  $\mathbf{m}$  itself.

In the MEM, these two sources of information are taken into account through different ways. The first is modeled by a time-dependent reference probability law  $P_t^{\text{ref}}$  on  $\mathbf{Q}_t$ . It embodies all the prior information on the solution. This probabilistic framework gives us a maximum of flexibility in expressing the constraints as well as correlations or even deterministic constraints such as intervals of confidence for the activation intensity. As to the second source of information, we can hardly measure a random variable but we can estimate its mean. This is usually what is done during a MEG evoked field experiment. Several samples are taken and we estimate the activation with the mean which thus must satisfy

$$G E_{\hat{P}_t}(\mathbf{Q}_t) = \mathbf{m}_t \quad (6.3)$$

where  $\mathbf{m}_t$  represents the MEG data measured at time  $t$ . The challenge is to find such a  $\hat{P}_t$ : starting with a reference law  $P_t^{\text{ref}}$  and a data attachment constraint (6.3), we have to find a new law  $\tilde{P}_t$  which approximates  $\hat{P}_t$ . To compare two probability laws, we use the Kullback's information:

$$\mathcal{K}(P_t, P_t^{\text{ref}}) = \int \log \frac{dP_t}{dP_t^{\text{ref}}} dP_t. \quad (6.4)$$

For the law  $\tilde{P}_t$ , we want to introduce as little bias as possible<sup>1</sup>. This leads us to minimize the information in  $\tilde{P}_t$  coming from  $P_t^{\text{ref}}$ :

$$\tilde{P}_t = \underset{P_t \in \mathcal{A}_t}{\text{argmin}} \mathcal{K}(P_t, P_t^{\text{ref}}) \quad (6.5)$$

<sup>1</sup>It is possible to justify this choice. It has its origins in statistical mechanics and information theory. We refer the reader to [7,9,13] for an in depth presentation.

for some set  $\mathcal{A}_t$  which represents the acceptable posterior probability laws. This is where we express the data attachment constraint:  $\mathcal{A}_t$  is the set of probability laws with average satisfying (6.3), *i.e.*

$$\mathcal{A}_t = \{P_t \mid G E_{P_t}(\mathbf{Q}_t) = \mathbf{m}_t\}.$$

This set is time-dependent. Without the constraint  $P_t \in \mathcal{A}_t$  in (5.8), we would obtain  $\tilde{P}_t = P_t^{\text{ref}}$  since  $\mathcal{K}$  is convex,  $\mathcal{K}(P_t, P_t^{\text{ref}}) \geq 0$  and  $\mathcal{K}(P_t^{\text{ref}}, P_t^{\text{ref}}) = 0$  leaving us with the solution we started with. The MEM solution  $\tilde{P}_t$  integrates all the prior information contained in  $P_t^{\text{ref}}$  and the MEG data via the direct problem formulation (6.3). The estimated sources are then given by:

$$\tilde{\mathbf{q}}_t = E_{\tilde{P}_t}(\mathbf{Q}_t). \quad (6.6)$$

In [9], the authors defined the  $\mu$ -entropy  $H_\mu(P)$  as  $-\mathcal{K}(P, \mu)$ . So  $\tilde{P}_t$  is the  $P_t \in \mathcal{A}_t$  that gives the maximum entropy  $H_{\mu=P_t^{\text{ref}}}(P)$  value. That is why it is called a maximum entropy principle. If we see the entropy as a measure of the lack of information,  $\tilde{P}_t$  can be viewed as the probability measure which takes the minimum of information subject to the constraint (6.3). The "on the mean" qualifier means that we want to explain the data with  $\tilde{P}_t$  on the mean.

With the maximum entropy on the mean, information introduced by the prior law  $P_t^{\text{ref}}$  is crucial because it determines the solution all the way. How to define such a probability law? This is the subject of the next section.

### 6.2.1. Prior law

In previous work [1], was introduced the notion of region of activation to distribute dipoles in an anatomically coherent manner. Suppose that the region definition remains constant for the time sequence considered and that the activation of each region is described by a time-dependent hidden random variable. Let us denote by  $K$  the number of regions, by  $S_{k;t}$  the discrete hidden state random variable representing the activity in the region  $k$  at time  $t$  and by  $n_k$  the number of dipoles in the region  $k$ . If a region is activated, then the dipoles belonging to that region are activated and conversely. Let  $\alpha_{k;t} = \pi_{k;t}(S_{k;t})$ , the probability of activation of region  $k$  and  $\mu_{k;t}$  the mean activation of each dipole at time  $t$  in that

region then, following [1, 8], the chosen prior law is of the form

$$dP_t^{\text{ref}}(\mathbf{q}) = \prod_{k=1}^K \left\{ (1 - \alpha_{k;t}) \delta(\mathbf{q}_k) + \frac{\alpha_{k;t}}{\sqrt{(2\pi)^{n_k} |\text{Det}(\Sigma_{k;t})|}} \exp \left[ -\frac{1}{2} (\mathbf{q}_k - \mu_{k;t})^+ \Sigma_{k;t}^{-1} (\mathbf{q}_k - \mu_{k;t}) \right] \right\} d\mathbf{q}_k. \quad (6.7)$$

The symbol  $+$  represents the matrix transpose operator. This law simply expresses that each region is a mixture of laws representing no activation at all ( $\delta(\mathbf{r}_k)$ ) and activation ( $N(\mu_{k;t}, \Sigma_{k;t})$  with  $\mu_k \neq 0$ ). The prior law  $P_t^{\text{ref}}$  depends on several parameters: the dipoles contained in each region and the hyperparameters  $\alpha_{k;t}$ ,  $\mu_{k;t}$  and  $\Sigma_{k;t}$  for each region  $k$ . In [8], they are estimated with only single time data. The next section extends this method to time sequences of MEG data.

### 6.3. DATA-DRIVEN PARCELLING

Let us first address the problem of estimating the regions since it is a part of the prior model. The idea is to use a source selection method [10] to locate high probably activated dipoles and select some of them with a given threshold. Then, we will construct geometrical regions of activation with these selected dipoles. In [8], the method is working only with MEG data at a single time. Here we recall shortly the source selection method and we show how to use this information to construct regions.

#### 6.3.1. Source preselection

The idea behind source selection is to partially solve the inverse problem without estimating the intensity of each dipole but rather to identify dipoles that are the most correlated with the MEG data. This gives a support of activation. This support can then be used as a prior information for conventional dipole intensities estimates by reducing the number of degrees of freedom in the problem. Here is how it goes.

We want to identify dipoles that correlate most with the data in a window of  $T$  time steps. The data is contained in the  $d \times T$  matrix  $\mathbf{m}$ . The window width  $T$  must be chosen to allow us to assume that a stationary statistical process is

underlying the neuronal activity. In [10], the authors suggest a value of around  $T = 10$ , a value that we shall use throughout this paper. Windows are displayed in Figure 6.7(b). Let us assume that the data  $\mathbf{m}$  can be explained by the columns  $\mathbf{g}_j$  of  $G$  *i.e.* for each time step stands a linear combination of the columns of  $G$  giving  $\mathbf{m}$ . This is not true in general because noise may be present in the data  $\mathbf{m}$ . But, it is justified in the distributed modelling framework since the number of dipoles is quite larger than the number of captors ( $n \gg d$ ). So, we consider the projector  $P_{\mathbf{m}}$  on the space spanned by the columns of  $\mathbf{m}$ :

$$P_{\mathbf{m}} = \mathbf{m}(\mathbf{m}^+ \mathbf{m})^{-1} \mathbf{m}^+. \quad (6.8)$$

This projector can be used to quantify the contribution of each lead-field to the data by computing the normalized coefficients  $a_j$ :

$$a_j = \frac{\|P_{\mathbf{m}} \mathbf{g}_j\|^2}{\|\mathbf{g}_j\|^2}. \quad (6.9)$$

If we deal with temporal sequences with  $T$  larger than permitted then we cannot assume the stationarity of the process. However we can segment the data in several time windows for which the width is of reasonable size to justify the use of the source selection on each window. Formally, we divide the  $T$  time samples into  $L$  windows of equal (or roughly so) widths (see, for example, Figure 6.7(b)). For each time window  $l \in \{1, \dots, L\}$ , we make a source selection to find coefficients  $a_{j;l}$  for each dipole  $j$ . We are led to find a global  $a_j$  valid for all times. We choose the following definition

$$a_j = \max_{1 \leq l \leq L} a_{j;l}. \quad (6.10)$$

This way, we do not neglect a dipole that happens to be activated in only one sub-window. This is a conservative definition that tries to keep most of the possible candidates in the solution. Having now a single  $a_j$  value for each dipole, we can select a support of activation  $S$  by ordering the dipoles according to their  $a_j$ . This subset of dipoles is then set up by fixing a threshold  $a$  given a small  $p$ -value such that

$$P(a_j > a \mid \text{dipole } j \text{ is in the solution}) = p.$$

These selected dipoles ( $a_j > a$ ) form the support of activation  $S$ . In Figure 6.1, we show several histograms of  $a_j$ . The look is very different if whether one or two regions are activated. The left figure shows the  $a_{j;l}$  extract with the data from  $t = 1, \dots, 10$  of the phantom in Figure 6.3 and the middle figure is

for  $t = 51, \dots, 60$ . The histograms are shown together with the associated  $a$  threshold.

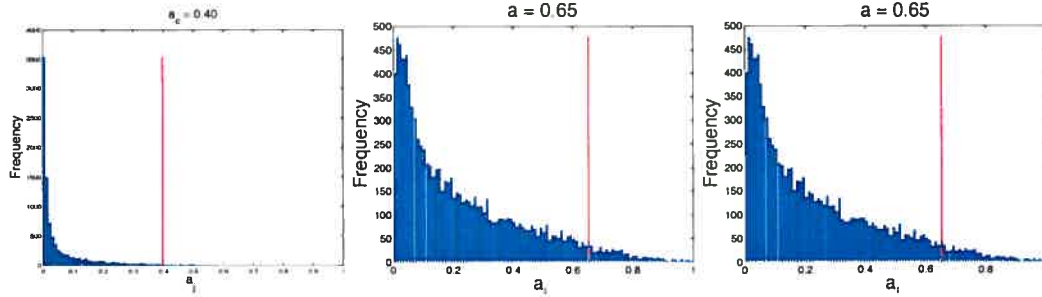


FIGURE 6.1. Histograms of the  $a_{j;l}$ . Left,  $a_{j;l}$  with the data from  $t = 1, \dots, 10$  ( $l = 1$ ) for the phantom in Figure 6.3. Center,  $a_{j;l}$  with  $t = 51, \dots, 60$  ( $l = 6$ ). Right,  $a_j = \max_{1 \leq l \leq L} a_{j;l}$  valid for all time steps,  $t = 1, \dots, 60$ . The threshold  $a$ , which is different for each histogram, is indicated by the vertical red line.

### 6.3.2. Parcelling

This support of activation can be used to construct a parcelling. We follow [8]: from the source support described earlier, take the first dipole (the one with the highest  $a_j$  in the classification). Around this dipole, we select all those on the cortex surface which are nearer than a fixed distance  $D$ . All these will form the first region  $R_1$ . We must check if the region is connected relatively to the cortex triangulation. If not, we keep only the connected subset which contains the preselected dipole. Afterwards, we take the next dipole in the classification which is not contained in the previous constructed region  $R_1$ . Then, we keep on doing the same by selecting all dipoles nearer than a fixed distance  $D$  from this second dipole but do not belong to the first region. These will form the second region  $R_2$ . We still keep doing this until there are no more selected dipoles in the support. At the end, we have a set of regions  $R_k$ ,  $k = 1, \dots, K$ . The dipoles not selected by the previous algorithm are simply ignored. We say that there is no possible activation over there.

We now have estimated the most important element of the prior law  $\tilde{P}_t$  in equation (6.7): the regions. These regions are assumed to be constant in time. This surely is justified since the time scale at which we study cerebral phenomena,



in this paper is very small (no more than 80 milliseconds). But the region parcelling algorithm depends on the parameter  $D$  which is the activation scale. We shall see later on how we will get rid of this parameter. For now on, there only remains the hyperparameters  $\alpha_{k;t}$ ,  $\mu_{k;t}$  and  $\Sigma_{k;t}$  for each region to be set. That is the subject of the next section.

#### 6.4. HYPERPARAMETERS ESTIMATION

Now that we have a set of parcels  $R_k$ ,  $k = 1, \dots, K$  valid for all times, we need to give initial values for  $\alpha_{k;t}$ ,  $\mu_{k;t}$  and  $\Sigma_{k;t}$  to completely set the prior probability law  $P_t^{\text{ref}}$  in equation (6.7). In [8], the authors have defined  $\alpha_{k;t}$  by the density of selected dipoles in each region at time  $t$ . Here, we keep doing the same for each time window  $l$ .

For the time window  $l$ , we will find a support of activation<sup>2</sup>  $S_l$  based on the source selection method with the same  $p$ -value as before. After, for each region  $R_k$ , we shall compute

$$\alpha_{k;l} = \frac{\#\{i \in R_k | i \in S_l\}}{n_k}$$

where  $\#$  means the cardinality of the set. So,  $\alpha_{k;l}$  is the density of dipoles in the region  $k$  that are present in the support  $S_l$  for each time window  $l$ . But, we want a  $\alpha_{k;t}$  for each time  $t$ . We choose the step function:

$$\alpha_{k;t} = \alpha_{k;l} \quad \text{if } t \in \text{time window } l$$

for each region  $k$ . Doing this way gives a non-smooth  $\alpha_{k;t}$  in time as expressed by the dashed line in Figure 6.2. That would introduce unacceptable discontinuities in the solution. So we will smooth  $\alpha_{k;t}$  with a cubic spline as seen in Figure 6.2. So we shall deal with a smooth prior leading us to a smooth solution.

There remains the hyperparameters  $\mu_{k;t}$  to estimate. First, we implicitly imposed that all dipoles in the same region have the same mean activation amplitude. To estimate  $\mu_{k;t}$ , we will first compute the  $d \times K$ -matrix  $\bar{G}$  in which each column  $\bar{\mathbf{g}}_k$  is the mean lead-field of each region:

$$\bar{\mathbf{g}}_k = \frac{1}{n_k} \sum_{i \in R_k} \mathbf{g}_i.$$

<sup>2</sup>The support  $S_l$  is not equal to  $S$  since the later is valid for all time steps. This is not the case for  $S_l$ .

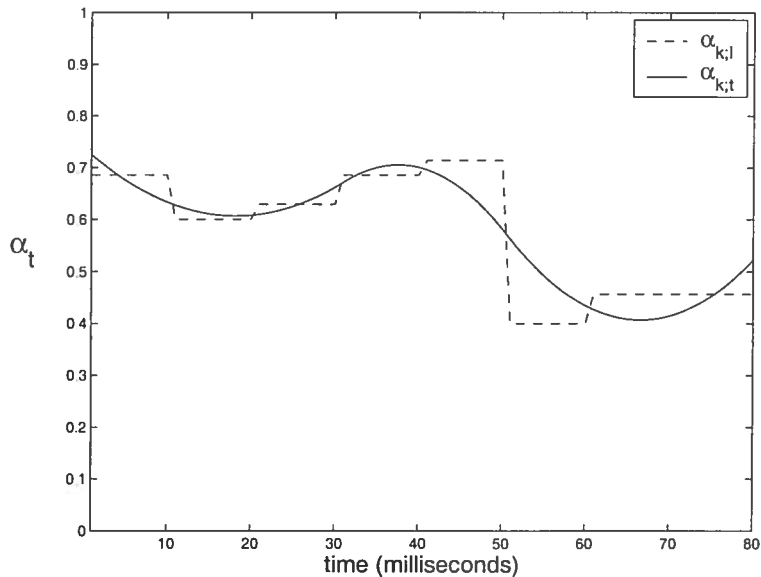


FIGURE 6.2. Example of spline smoothing of the  $\alpha_k$ 's. The dashed line is the discontinuous  $\alpha_{k;l}$ . We see the steps at  $t = 10, 20, 30, \dots$ . The solid line is the  $\alpha_{k;t}$  smoothed by a cubic spline.

We will construct the vector  $\boldsymbol{\mu}_t = (\mu_{k;t})_{k=1}^K$  and it must satisfy the equation:

$$\mathbf{m}_t = \bar{G} \boldsymbol{\mu}_t. \quad (6.11)$$

To solve this inverse problem, we shall take the simple minimum-norm estimate:

$$\bar{\boldsymbol{\mu}}_t = \bar{G}^+ [\bar{G}\bar{G}^+]^{-1} \mathbf{m}_t. \quad (6.12)$$

That gives the value for the hyperparameters  $\mu_{k;t}$  at each time  $t = 1, \dots, T$  and for each region  $k$ . The last hyperparameter  $\Sigma_{k;t}$  is arbitrarily set to  $10^{-2}\mu_{k;t}^2\mathbb{I}_{n_k}$  where  $\mathbb{I}_{n_k}$  is the  $n_k \times n_k$  identity matrix. This only says that dipoles in an activated regions are i.i.d. The factor  $10^{-2}$  is not very important. We empirically observed that the solution does not strongly depend on that value.

Now, all the requested information to define completely the prior law (6.7) is available. We can now estimate the activation so do we in the next two sections by testing this method on MEG data coming from both simulations and real experiment.

## 6.5. SIMULATIONS

### 6.5.1. Framework

To test the methods introduced in this paper, we have used a realistic cortex. This cortex contains 10 000 dipoles. The mean distance between two neighboring dipoles is 5 mm. Around this cortex, we placed 150 MEG captors. With this cortex, we created three different phantoms. For each phantom, we selected two distinct regions for a total of roughly 30 activated dipoles. The same time profile, with 60 time steps, was used for all phantoms. We illustrate the cortex and the temporal activation in Figure 6.3 for one of the phantoms. These phantoms generate MEG data  $\mathbf{m}_t$  on the 150 captors located around this cortex. We did not add on noise to the data because we consider that real MEG data are made of several trials which cancels out the measuring noise. In Figure 6.7(a), we can see real MEG data and observe that the noise is negligible in this framework.

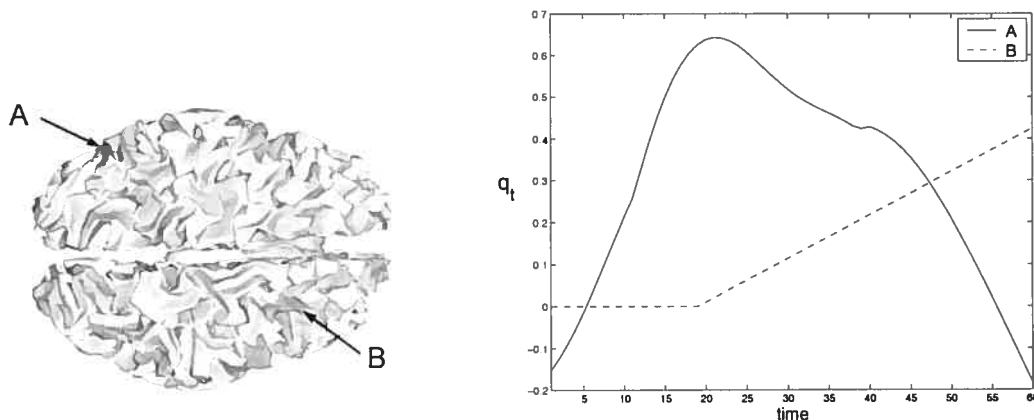


FIGURE 6.3. These figures present one of the phantoms used in the simulations. The left image shows the spatial distribution of the activation at  $t = 40$ . The right image contains the time evolution of the dipole intensity in the colored patches A and B.

We reconstruct the activity, the same way we did in [8]: 6 different characteristic scales of parcelling  $D$  have been used. They are ranging from  $D = 5$  mm to  $D = 10$  mm at each millimeter. We will average over all the reconstruction scales in order to get rid of them. We will then have a solution with no characteristic scale of activation. We will compare three reconstruction methods:

”*MEM+time+smooth*”: parcelling (section 6.3.2) with smooth  $\alpha_{k;t}$ ’s,

”*MEM+time-smooth*”: parcelling (section 6.3.2) with  $\alpha_{k;t} = \frac{1}{2}$ ,

”*MEM-time*”: parcelling as in [8] using the data for each time separately meaning no temporal regularization at all. The hyperparameters of the prior law are initialized accordingly.

The methods will be compared on the basis of three objective criteria. First, the root mean square error (RMSE) is the relative error between the solution  $\tilde{\mathbf{q}}_t$  and the phantom  $\mathbf{q}_t^i$  averaged over time:

$$\text{RMSE} = \frac{1}{T} \sum_{t=1}^T \frac{\|\tilde{\mathbf{q}}_t - \mathbf{q}_t^i\|}{\|\mathbf{q}_t^i\|}$$

where  $T$  is the number of time steps. Second, we calculate the data fit error (DFE) with roughly the same formula:

$$\text{DFE} = \frac{1}{T} \sum_{t=1}^T \frac{\|\mathbf{m}_t - G \tilde{\mathbf{q}}_t\|}{\|\mathbf{m}_t\|}.$$

The third criteria is the Receiver Operating Characteristics (ROC) analysis [4] to evaluate the source detection ability of each method. It consists in computing first the ROC curve which is a function of a threshold to classify the dipoles. The ROC score is the area under the curve. It indicates the probability of the reconstruction method to correctly distinguish between activated and non-activated dipoles. Since the number of active dipoles is outnumbered by the number of non-active dipoles, we randomly select an equal number of active and non-active dipoles to compute the ROC score. This way, the result is an unbiased estimate of the efficiency of the method. The ROC is calculated at each time step and then averaged.

### 6.5.2. Results

In this section, we discuss the results of the several reconstruction methods using the simulated MEG data presented in the section 6.5.1.

As to the parcelling. The Figure 6.4 shows the regions of a parcelling at fixed scale  $D = 9$  mm. We see that the segmented regions correctly envelop, on the cortex, the activated regions in the phantom displayed in Figure 6.3. The number of dipoles have been drastically reduced: in that particular example, only 2000 dipoles have been kept out the initial 10000 dipoles. This will greatly

help the reconstruction methods. In the right image of Figure 6.4, we see the regions that contain the phantom activated dipoles. This image helps to check if the parcels fits the spatial distribution of the activation. One of the activated regions is correctly identified. The other one is a bit fragmented but it is mainly concentrated in two regions. That's why we see several windows with the same time evolution. The two center images show the same activated region. Intensity is lower because the mean is taken over all dipoles in the region and many of them are zero in the lower right image.

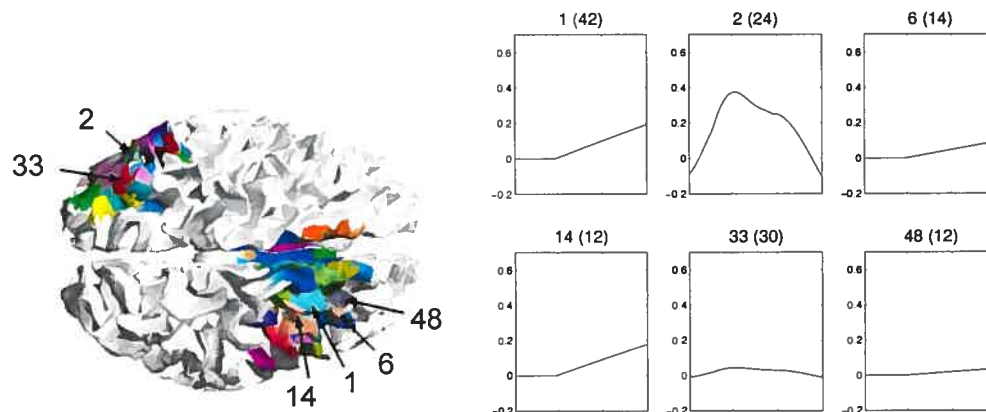


FIGURE 6.4. Parcelling with  $D = 9$  mm. On the left side, the cortex displays the reconstructed region. On the right, each image shows the phantom mean activation of each region which contains at least one activated dipole. The number on top of each box corresponds to the region in the parcelling while the number ( $n_k$ ) of dipoles in each region is in parentheses. Axes are the same as in Figure 6.3. This image can be used to observe the reconstruction performance of each method.

With this parcelling, we can then initialize the  $\alpha_{k;t}$  and  $\mu_{k;t}$  in order to define  $P_t^{\text{ref}}$  for a fixed  $D$  as described in the section 6.4. Since, it is difficult to estimate *a priori* the true activation scale, we take the average over several characteristic scales. We thus obtain the best possible reconstruction as we see in the following figures where the quality and accuracy of the reconstruction is appreciated with temporal regularization. In Figure 6.5, we see the reconstruction at a fixed time ( $t = 40$ ) where the two regions are active. The images show that the reconstruction with "MEM+time+smooth" is significantly better. Despite a small active region near one activated region, the location and size are nearly optimal. The

important thing is that all the active regions are activated in the reconstruction (no false-negative dipoles).

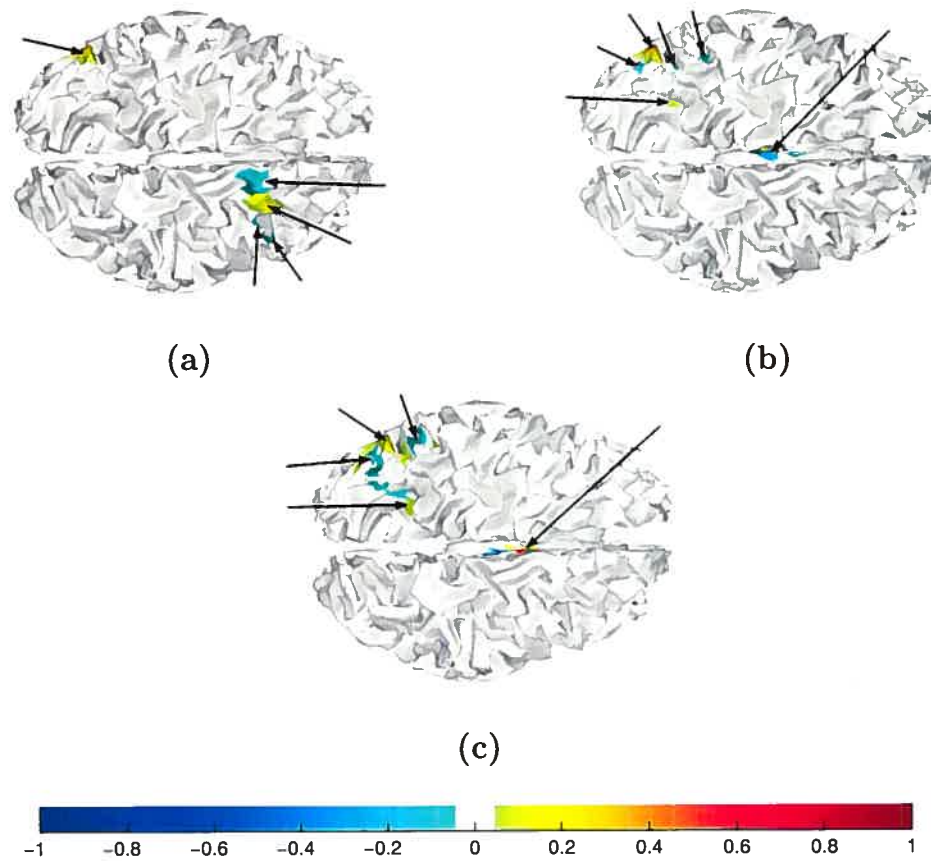


FIGURE 6.5. Cortex views of the reconstruction at  $t = 40$ . The phantom activation is in Figure 6.3. (a): "*MEM+time+smooth*". (b): "*MEM+time-smooth*". (c): "*MEM-time*".

Figure 6.6 displays the temporal evolution of the reconstructions. In the "*MEM+time+smooth*" case, the temporal course of the regions 1 and 2 is very well reproduced. For the first one, we see that the activation starts at the right time. "*MEM+time-smooth*" completely misses the first region whereas the activation in the first region is delayed in "*MEM-time*".

"*MEM+time+smooth*" worst results are the apparition of false-positive regions for the last time steps. They are correlated with the late activated region. There is no way to distinguish them from other activated regions. Nevertheless, temporal regularization gives the best results. Comparing with other methods,

not so many regions are activated and the solution is smooth in time. So, temporal regularization gives the expected improvements. In the table 6.1, we can see that the "*MEM+time+smooth*" is the best method. "*MEM+time+smooth*" efficiently exploits the temporal regularization given by the smooth  $\alpha_{k;t}$ .

Method	RMSE	DFE	ROC
" <i>MEM+time+smooth</i> "	$0.75 \pm 0.27$	$0.03 \pm 0.03$	$0.74 \pm 0.01$
" <i>MEM+time-smooth</i> "	$0.86 \pm 0.37$	$0.35 \pm 0.55$	$0.73 \pm 0.02$
" <i>MEM-time</i> "	$0.99 \pm 0.24$	$0.07 \pm 0.04$	$0.66 \pm 0.02$

TABLE 6.1. Objective comparison between the reconstruction methods. Each box displays the average over the three phantoms together with the standard deviation.

Before illustrating these methods on real MEG data, let us emphasize on the weak priors used in the regularization scheme. In the reference law, we modeled the activation as dipoles gathered together in regions. These regions were constructed with a data-driven parcelling at different scales. Then, hyperparameters of the reference law were estimated with the data using that parcelling. So the only strong prior is that dipoles do form regions and since we average the results over several activation scales, even this hypothesis is somewhat relaxed. It is remarkable that we obtained such good results with a method based on so little prior information, with a low number of data (150) together with a large number of dipoles in the model ( $10^4$ ).

## 6.6. REAL MEG DATA

In this last section, we present the results of the "*MEM+time+smooth*" applied to real MEG data in a somatosensory experiment recorded on a healthy human subject with a CTF MEG system with 150 channels at the Hôpital Pitié-Salpêtrière in Paris, France. The signals are the magnetic fields recorded after an electric stimulation of the right thumb. The MEG signals displayed in Figure 6.7(a) are the results of the mean over 400 repetitions of the same experiment. The signals have been sampled at 1 kHz. We have isolated the signal peak and took 80 samples around this peak. This data is shown in the Figure 6.7(b) where it is displayed together with the time windows. The reconstruction space is made

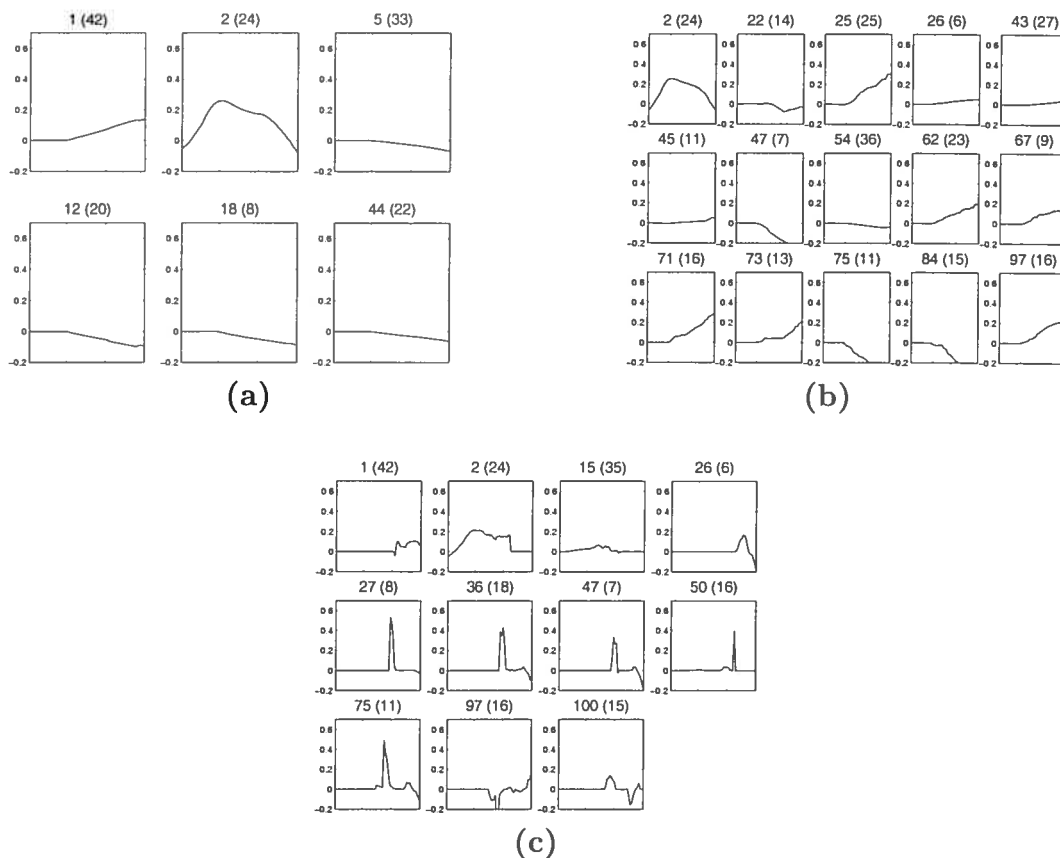


FIGURE 6.6. Temporal reconstructed activation.  
 (a): "MEM+time+smooth". (b): "MEM+time-smooth".  
 (c) "MEM-time". The axes are the same as in Figure 6.3.

on cortex surface that was extracted from a MRI image of the same subject used in the MEG experiment.

On Figure 6.7(c), we show the time evolution of dipoles circled on the cortex of Figure 6.7(d). They were selected because they are in regions with the highest amplitudes. The main activation is in the left hemisphere which is good. The activation peak is located where it is known as being the primary zone of activation for this kind of somatosensory experiment.

Let us stress that this result with real data makes use of very few hypotheses or ad hoc constraints. We just included correlations at different scales between dipoles to regularize the inverse problem. This is why there are other activated regions on the opposite face of the shown cortex but, since they are far away from the captors, they do not contribute that much to the data.



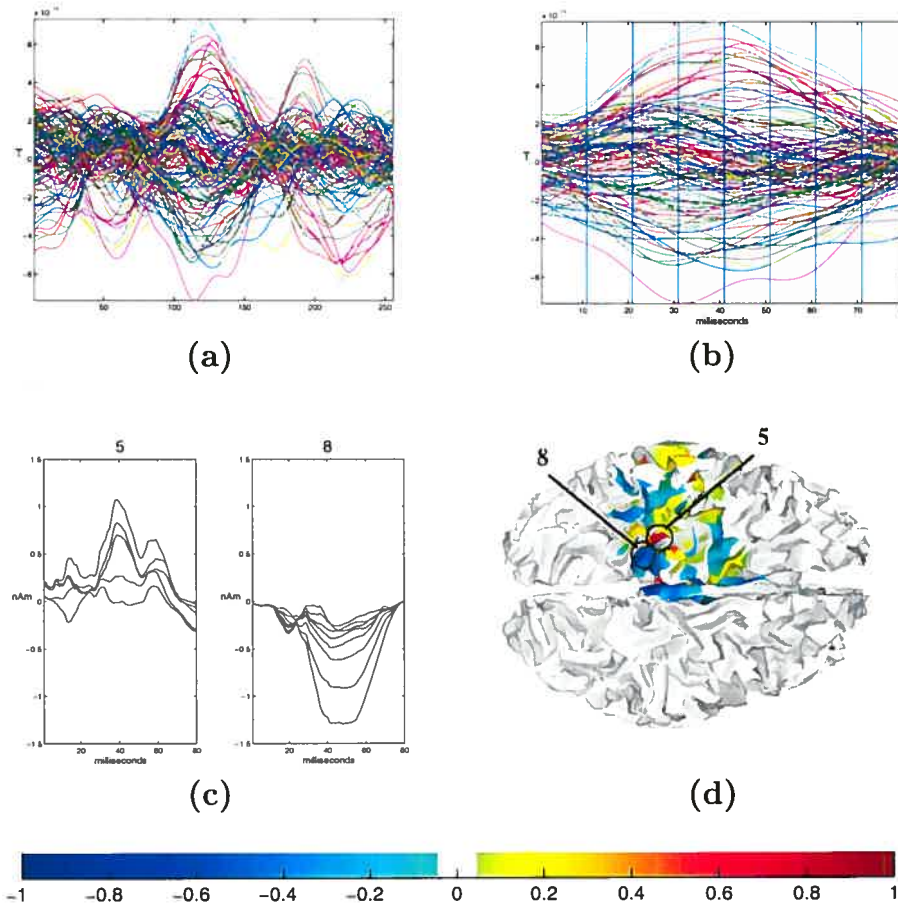


FIGURE 6.7. The images (a) and (b) show the real MEG data. In (a), it is the complete set for 256 milliseconds. In (b), we extracted 80 time samples around the signal peak to do the reconstruction and it is displayed with the time windows used in the parcelling process. The images (c) and (d) show the reconstruction made with "*MEM+time+smooth*". Time course of some dipoles is shown at left. The number on the top of each box in is the number of dipoles displayed. In (d), we show the spatial organization of the solution at  $t = 40$ . The dipoles displayed in the image (c) are the ones circled in image (d). The color bar at the very bottom is in units of  $10 \text{ nA} \cdot \text{m}$  and refers to the image (d).

## 6.7. CONCLUSION

The main advantage of MEG over other functional brain imaging modalities is its high time resolution. This resolution is only limited by the sampling rate

giving a time resolution for MEG. In previous works [1, 8], we addressed the inverse MEG problem at a single time step. We used the maximum entropy on the mean approach to spatially regularize the inverse problem. This approach relies on a reference probability measure  $P^{\text{ref}}$  which embodies all the information (anatomical and functional) we have on the solution. We presented such a reference measure in [1]. It relies on the strong constraint that dipoles are grouped together in activity regions. In [8], we presented a method to construct these regions using the data for a single time step. In the present work, we generalized that method to handle several time steps allowing us to use the time correlations offered by the MEG signals.

This parcelling method is shown to be very effective at identifying activated regions. Then, we demonstrated by simulations that MEM makes an excellent use of the information given by the parcelling to reconstruct dipole intensities. Moreover, the temporal course has been correctly reproduced by giving temporal priors in the reference measure  $P_t^{\text{ref}}$  using the hyperparameters  $\alpha_{k;t}$  and  $\mu_{k;t}$ .

We also have presented work based upon MEG data coming from a real experiment. The results are surely not as good as in the simulations but are far from bad. The prior information introduced via  $P_t^{\text{ref}}$  is very limited. So that demonstrates the power of the maximum entropy method to solve the MEG inverse problem where data is very limited compared to the source space. A great work on post-processing do able on the basis of the results. In this work, we did not talk about the *a posteriori* probability measure given by MEM. This measure could have been used to construct a new reference measure resulting in an iterative scheme as in [1, 8] in order to give more focal solutions. Such algorithms that we did not develop here could improve the already presented results. The formalism introduced in this paper is extremely flexible. It is suitable for a multiresolution approach similar to [5]. Furthermore, via the probability law  $\pi(S)$  of hidden state variables  $S_{k;t}$ , we can introduce functional correlations coming from other families of measurements such as fMRI.

## ACKNOWLEDGEMENTS

The authors thank Pr. Line Garnero for sending them the MEG data and cortex definition used in this paper and also Jean Daunizeau for useful discussions.

E.L. would also like to thank Philippe St-Jean for providing computer assistance.  
This work has been supported by NSERC fundings.

## BIBLIOGRAPHY

- [1] C. Amblard, E. Lapalme, and J-M Lina. Biomagnetic source detection by maximum entropy and graphical models. *IEEE Transactions on Biomedical Engineering*, 55(3):427–442, March 2004.
- [2] S. Baillet and L. Garnero. A bayesian approach to introducing anatomo-functional priors in the MEG-EEG inverse problems. *IEEE Transactions on Biomedical Engineering*, 44(3):374–385, 1997.
- [3] A.M. Dale and M.I. Sereno. Improved localization of cortical activity by combining EEG and MEG with MRI cortical surface reconstruction. *Journal of Cognitive Neuroscience*, 5(2):162–176, 1993.
- [4] E.R. DeLong, D.M. DeLong, and D.L. Clarke-Pearson. Comparing the areas under two or more correlated receiver operating characteristic (ROC) curves: a nonparametric approach. *Biometrics*, 44:837–845, 1988.
- [5] L. Gavit, S. Baillet, J.-F. Mangin, J. Pescatore, and L. Garnero. A multiresolution framework to MEG/EEG source imaging. *IEEE Transactions on Biomedical Engineering*, 48(10):1080–1087, Octobre 2001.
- [6] M. Hämäläinen, R. Hari, R.J. Ilmoniemi, J. Knuutila, and O.V. Lounasmaa. Magnetoencephalography – theory, instrumentation, and applications to noninvasive studies of the working human brain. *Reviews of Modern Physics*, 65(2):413–487, april 1993.
- [7] E.T. Jaynes. Information theory and statistical mechanics. *Physical Review*, 106(4):620–630, May 1957.
- [8] E. Lapalme, J-M Lina, and J. Mattout. Data-driven cortex parcelling and entropic inference in MEG. *submitted to NeuroImage*, June 2004.
- [9] G. Le Besnerais, J-F. Bercher, and G. Demoment. A new look at entropy for solving linear inverse problems. *IEEE Transactions on Information Theory*, 45(5):1565–1578, July 1999.
- [10] J. Mattout, M. Pélégriani-Issac, A. Bellio, J. Daunizeau, and H. Benali. Better conditioning the MEG/EEG inverse methods: the multivariate source prelocalization approach. *Proceedings of the International Symposium on Biomedical Imaging*, 2004.
- [11] R. D. Pascual-Marqui, C. M Michel, and D. Lehmann. Low resolution electromagnetic tomography: a new method for localizing electrical activity in the brain. *Int. Journal of Psychophysiology*, 18:49–65, 1994.
- [12] J. Sarvas. Basic mathematical and electromagnetic concepts of the biomagnetic inverse problem. *Phys. Med. Biol.*, 32(1):11–22, 1987.
- [13] J.E. Shore and R.W. Johnson. Axiomatic derivation of the principle of maximum entropy and the principle of minimum cross-entropy. *IEEE Transactions on Information Theory*, 26(1):26–37, Janvier 1980.

# CONCLUSION

---

Dans cette thèse, nous avons développé une nouvelle méthode pour résoudre le problème inverse en magnétoencéphalographie (MEG). Celle-ci n'étant pas très sensible à la conductivité des tissus cérébraux, on peut formuler le problème direct de façon simple. De plus, la résolution temporelle offerte par la MEG est excellente. Elle est donc une méthode toute indiquée pour étudier le cerveau en action. Cependant, le problème inverse associé à la MEG n'admet pas de solution unique. On doit alors faire appel à des méthodes de régularisation qui permettent d'en choisir une correspondant aux divers *a priori* physiologiques disponibles sur la solution. Donc, tout le problème réside dans la façon d'introduire cette information. Dans cette thèse, nous avons utilisé le principe de la maximisation de l'entropie en moyenne (MEM). Cette méthode est inspirée de travaux en mécanique statistique. L'ingrédient principal du MEM est une loi de référence qui englobe toute l'information que l'on peut avoir sur la solution au problème inverse.

Nous avons appliqué le MEM à la MEG en procédant en trois étapes. Nous présentons d'abord une méthode de régularisation spatiale avec un modèle de sources. On suppose que les dipôles sont regroupés en régions et que les dipôles d'une région sont tous activés ou désactivés en même temps. Tous nos travaux reposent sur ce modèle simple et intuitif.

Dans le chapitre 4, ces régions doivent être données en intrant ce qui n'est plus nécessaire pour les travaux présentés dans le chapitre 5 où une méthode de parcellisation est développée. Cette parcellisation utilise les données qui sont encore mises à profit pour estimer les hyperparamètres de la loi de référence. Ces deux chapitres complètent la régularisation spatiale faite avec le MEM pour résoudre le problème inverse en MEG.

Le principal avantage de la MEG sur les autres méthodes est sa résolution temporelle. Dans cette thèse, nous développons une méthode de régularisation spatio-temporelle qui utilise les techniques introduites dans les deux premiers articles. Elle met à profit les corrélations temporelles entre les signaux d'une

même séquence. En effet, la méthode de parcellisation présentée dans le chapitre 5 utilise les données de toute une séquence temporelle et les hyperparamètres sont estimés pour chaque pas de temps. Ainsi, la régularisation temporelle est faite sur les hyperparamètres de la loi de référence et non directement en introduisant des corrélations dans cette loi. On vérifie aussi cette méthode sur des données MEG réelles. Les résultats ne sont certes pas parfaits mais intéressants étant donné le peu de contraintes imposées à la solution.

Ces travaux ont permis de démontrer l'efficacité du MEM pour résoudre le problème inverse en MEG. Cependant, plusieurs points gagneraient à être explorés. Dans le modèle présenté dans le chapitre 4, la loi jointe des variables d'activation cachées permet l'introduction de corrélations entre les régions. Cette information pourrait provenir d'autres méthodes d'imagerie mais, dans ce travail, nous avons préféré nous restreindre au cas général où il n'y a aucune corrélation entre les régions. Cette fusion de données introduit une difficulté supplémentaire puisque si les corrélations sont mesurées par l'IRMf par exemple [30], ce n'est pas exactement le même phénomène qui est mesuré. L'IRMf mesure la signature chimique de l'activité cérébrale.

Les algorithmes de reconstruction des deux premiers articles sont un aspect qui mériterait d'être généralisé. En effet, ils travaillent tous à une résolution constante. Il serait possible de les modifier afin d'augmenter la résolution au fur et à mesure des itérations. Un travail semblable a été fait pour S-MAP dans [16]. Cette multirésolution atteindrait peut-être une résolution spatiale supérieure. De plus, des algorithmes semblables peuvent aussi être appliqués aux signaux temporels. L'expression de la loi *a posteriori* du MEM (équation (4.45) du chapitre 4) permet d'estimer l'évolution des hyperparamètres dans le temps pour éliminer des régions faussement positives par l'analyse du contenu temporel des signaux. Des algorithmes pourraient alors être développés en utilisant ces signaux temporels.

En dernier lieu, nous insistons sur le fait que la modélisation présentée dans cette thèse est due à l'excellente flexibilité du MEM à pouvoir prendre en compte différentes informations sur la solution. Cette flexibilité est évidente si on regarde la variété des problèmes considérés avec le MEM :

- cristallographie [38],
- traitement d'images en astronomie [29],
- en assimilation de données satellitaires en météorologie [47],

– traitement de la langue naturelle [6].

Le modèle introduit dans cette thèse mène à une méthode de reconstruction de sources qui pourra certainement être utilisée dans des études sur les processus cérébraux. Ce modèle mériterait d'être précisé par des spécialistes en neurophysiologie. Ceux-ci pourraient améliorer les *a priori* et ainsi obtenir une meilleure résolution spatiale et temporelle. De plus, la solution donnée par le MEM contient beaucoup plus d'informations que simplement les intensités des dipôles. En effet dans la loi *a posteriori* des probabilités d'activation et des corrélations peuvent être extraites des données. Cette information pourrait être très utile pour étudier les processus cognitifs du cerveau.

Bien sûr, la méthode développée dans cette thèse reste limitée par la MEG elle-même. Celle-ci ne peut mesurer que des phénomènes macroscopiques c'est-à-dire lorsque des milliers de neurones sont activés simultanément et de façon cohérente. Ces observations macroscopiques pourraient cacher des processus sous-jacents. Seul l'avenir nous dira comment les scientifiques pourront observer ces phénomènes. Nous espérons que les avancements présentés ici pourront contribuer à améliorer nos connaissances sur une des plus belles réalisations de la nature : le cerveau.

## BIBLIOGRAPHIE

---

- [1] C. Amblard, E. Lapalme, and J-M Lina. Biomagnetic source detection by maximum entropy and graphical models. *IEEE Transactions on Biomedical Engineering*, 55(3) :427–442, Mars 2004.
- [2] S. Baillet. Problème inverse en MEG et en EEG. In *Atelier de formation à la MEG et à EEG*, Paris, Juin 2001.
- [3] S. Baillet and L. Garnero. A bayesian approach to introducing anatomo-functional priors in the MEG-EEG inverse problems. *IEEE Transactions on Biomedical Engineering*, 44(3) :374–385, 1997.
- [4] S. Baillet, J.C. Mosher, and R.M. Leahy. Electromagnetic brain mapping. *IEEE Signal Processing Magazine*, 18(6) :14–30, Novembre 2001.
- [5] J.-F. Bercher and C. Heinrich. Construction de mesures de divergence. application à la résolution de problèmes inverses linéaires. In *15<sup>e</sup> Colloque GRETSI sur le Traitement du Signal et ses Applications*, Juan-les-Pins, Septembre 1995.
- [6] A. Berger, S.D. Pietra, and V. Della Pietra. A maximum entropy approach to natural language processing. *Computational Linguistics*, 22(1) :39–71, Mars 1996.
- [7] C.J.S Clarke. Probabilistic methods in a biomagnetic inverse problem. *Inverse Problems*, 5(6) :999–1012, Décembre 1989.
- [8] C.J.S Clarke and B.S. Janday. The solution of the biomagnetic inverse problem by maximum statistical entropy. *Inverse Problems*, 5(4) :483–500, Août 1989.
- [9] D. Cohen. Magnetoencephalography : evidence of magnetic fields produced by alpha rhythm currents. *Science*, 161 :784–786, 1968.
- [10] M.S. Cohen and S.Y. Bookheimer. Localization of brain function using magnetic resonance imaging. *Trends in Neurosciences*, 17(7) :268–277, 1994.
- [11] A.M. Dale and M.I. Sereno. Improved localization of cortical activity by combining EEG and MEG with MRI cortical surface reconstruction. *Journal of Cognitive Neuroscience*, 5(2) :162–176, 1993.
- [12] J.C. de Munck, B.W. van Dijk, and H. Spekreijse. Mathematical dipoles are adequate to describe realistic generators of human brain activity. *IEEE Transactions on Biomedical Engineering*, 35(11) :960–966, Novembre 1988.
- [13] G. Demoment. Image reconstruction and restoration : Overview of common estimation structures and problems. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 37(12) :2024–2036, Décembre 1989.
- [14] R.S Ellis. *Entropy, large deviations and statistical mechanics*, volume 271 of *Grundlehren der mathematischen Wissenschaften*, chapter 1–2, page 364. Springer-Verlag, New-York, 1985.



- [15] L. Garnero. Localisation des sources en MEG et en EEG. In *Atelier de formation à la MEG et à EEG*, Paris, Juin 2001.
- [16] L. Gavit, S. Baillet, J.-F. Mangin, J. Pescatore, and L. Garnero. A multiresolution framework to MEG/EEG source imaging. *IEEE Transactions on Biomedical Engineering*, 48(10) :1080–1087, Octobre 2001.
- [17] M. Hämäläinen, R. Hari, R.J. Ilmoniemi, J. Knuutila, and O.V. Luonasma. Magnetoencephalography – theory, instrumentation, and applications to noninvasive studies of the working human brain. *Reviews of Modern Physics*, 65(2) :413–487, Avril 1993.
- [18] M. Hämäläinen and R.J. Ilmoniemi. Interpreting magnetic fields of the brain : minimum norm estimates. *Med. Biol. Eng. Comp.*, 32 :35–42, 1994.
- [19] N.V. Hartvig and J.L. Jensen. Spatial mixture modelling of fMRI data. *Human Brain Mapping*, 11(4) :233–248, Décembre 2000.
- [20] M. Huang, R. Aaron, and C.A. Shiffman. Maximum entropy method for magnetoencephalography. *IEEE Transactions on Biomedical Engineering*, 44(1) :98–102, Janvier 1997.
- [21] A.A. Ioannides, J.P.R. Bolton, and C.J.S. Clarke. Continuous probabilistic solutions to the biomagnetic inverse problem. *Inverse Problems*, 6(4) :523–542, Août 1990.
- [22] E.T. Jaynes. Information theory and statistical mechanics. *Physical Review*, 106(4) :620–630, Mai 1957.
- [23] P. Jizba and T. Arimitsu. Generalized statistics : yet another generalization. arXiv :cond-mat/0312034, Décembre 2003.
- [24] P. Jizba and T. Arimitsu. On observability of Rényi’s entropy. arXiv :cond-mat/0307698, Décembre 2003.
- [25] A.I. Khinchin. *Mathematical foundations of statistical mechanics*. Dover, 1949.
- [26] E. Lapalme, C. Amblard, J. Mattout, H. Benali, B. Goulard, and J.M. Lina. Maximum entropy on the mean approach for solving the MEG inverse problem. *NeuroImage*, 13(6) :S62, Juin 2001.
- [27] E. Lapalme, J-M Lina, and B. Goulard. Spatio-temporal biomagnetic source reconstruction with maximum entropy. *soumis à IEEE Transactions on Biomedical Engineering*, Juin 2004.
- [28] E. Lapalme, J-M Lina, and J. Mattout. Data-driven cortex parcelling and entropic inference in MEG. *soumis à NeuroImage*, Juin 2004.
- [29] G. Le Besnerais, J.-F. Bercher, and G. Demoment. A new look at entropy for solving linear inverse problems. *IEEE Transactions on Information Theory*, 45(5) :1565–1578, Juillet 1999.
- [30] G. Marrelec, H. Benali, and M. Péligrini-Issac. Bayesian estimation of local and global independence relations for gaussian graphical models. Technical report, INSERM-U494, 2001.
- [31] J. Mattout. *Approches Statistiques multivariées pour la localisation de l’activation cérébrale en magnétoencéphalographie et en imagerie par résonance magnétique fonctionnelle. Vers une fusion d’informations multimodales*. PhD thesis, Université Paris 6, Juillet 2002.

- [32] J. Mattout, L. Garnero, M. Pélégriani-Issac, L. Gavit, and H. Benali. Functional MRI-derived priors for solving the EEG/MEG inverse problem. In *Biomag*, volume 12, pages 777–781, 2000.
- [33] J. Mattout, M. Pélégriani-Issac, A. Bellio, J. Daunizeau, and H. Benali. Better conditioning the MEG/EEG inverse methods : the multivariate source prelocalization approach. *Proceedings of the International Symposium on Biomedical Imaging*, 2004.
- [34] A. Mohammad-Djafari. A comparison of two approaches : maximum entropy on the mean and bayesian estimates for inverse problems. *arXiv :physics 0111122*, 2001.
- [35] J.C. Mosher, R.M. Leahy, and P.S. Lewis. EEG and MEG : Forward solutions for inverse problems. *IEEE Transactions on Biomedical Engineering*, 46(3) :245–259, Mars 1999.
- [36] J.C. Mosher, R.M. Leahy, D.W. Shattuck, and S. Baillet. MEG source imaging using multipolar expansions. In A. Kuba, M. Sámal, and A. Todd-Pokropek, editors, *Proceedings of the 16th Conference on Information Processing in Medical Imaging*, Springer Series in Computer Science, pages 15–28. Springer, Juin–Juillet 1999.
- [37] J.C. Mosher, P.S. Lewis, and R.M. Leahy. Multipole dipole modeling and localization from spatio-temporal dipole model. *IEEE Transactions on Biomedical Engineering*, 39(6) :541–557, Juin 1992.
- [38] J. Navaza. On the maximum entropy estimate of electron density function. *Crystallographica*, A-41 :232–244, 1985.
- [39] R.D. Pascual-Marqui, C.M. Michel, and D. Lehmann. Low resolution electromagnetic tomography : a new method for localizing electrical activity in the brain. *Int. Journal of Psychophysiology*, 18 :49–65, 1994.
- [40] J.W. Phillips, R.M. Leahy, and J.C. Mosher. MEG-based imaging of focal neuronal current sources. *IEEE Transaction on Medical Imaging*, 16(3) :338–348, Juin 1997.
- [41] L.E. Reichl. *A Modern Course in Statistical Physics*. Wiley-Interscience, 2nd edition, 1998.
- [42] J. Sarvas. Basic mathematical and electromagnetic concepts of the biomagnetic inverse problem. *Phys. Med. Biol.*, 32(1) :11–22, 1987.
- [43] D.M. Schmidt, J.S. George, and C.C. Wood. Bayesian inference applied to the electromagnetic inverse problem. *Human Brain Mapping*, 7(3) :195–212, 1999.
- [44] C.E. Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27 :379–423 and 623–656, Juillet et Octobre 1948.
- [45] J.E. Shore and R.W. Johnson. Axiomatic derivation of the principle of maximum entropy and the principle of minimum cross-entropy. *IEEE Transactions on Information Theory*, 26(1) :26–37, Janvier 1980.
- [46] A.N. Tikhonov and V.Y. Arsenin. *Solutions of Ill-Posed Problems*. John Wiley, New York, 1977.
- [47] B. Urban. Retrieval of atmospheric thermodynamical parameters using satellite measurements with a maximum entropy method. *Inverse Problems*, 12(5) :779–796, Octobre 1996.

- [48] H. von Helmholtz. Ueber einige gesetze der vertheilung elektrischer strome in korperlichenleitern, mit anwendung auf die thierisch-elektrischen versuche. *Ann. Phys Chem.*, 89 :211–233,53–377, 1853.
- [49] J. Whittaker. *Graphical models in applied multivariate statistics*. John Wiley & Sons, New York, Juillet 1990.

