Universidade de São Paulo Instituto de Física

Inclusão de MRI e informação Multigrid a Priori para inferência Bayesiana de fontes de M/EEG

Leonardo da Silva Barbosa

Orientador: Prof. Dr. Nestor Felipe Caticha Alfonso

Dissertação de mestrado apresentada ao Instituto de Física para a obtenção do título de Mestre em Ciências

Prof. Dr. Nestor Felipe Caticha Alfonso Prof. Dr. Said Rahnamaye Rabbani (IF-USP) Prof. Dr. Koichi Sameshima (FMUSP) São Paulo

2011

Barbosa, Leonardo da Silva

Inclusão de MRI e informação Multigrid a Priori para inferência Bayesiana de fontes de M/EEG. - São Paulo, 2011.

Dissertação (Mestrado) - Universidade de São Paulo. Instituto de Física - Depto. de Física Geral

Orientador: Prof. Dr. Nestor Felipe Caticha Alfonso

Área de Concentração: Física Geral

Unitermos: 1 Eletroencefalografia; 2. Magnetoencefalografia; 3. Física Médica

USP/IF/SBI-035/2011

Conteúdo

1	Des	crição	dos modelos estudados de EEG/MEG	12
	1.1	Origer	m do sinal	13
		1.1.1	Potencial de Ação e Potencial Pós-Sináptico	13
		1.1.2	Simetria na Organização das Fontes	15
	1.2	Proble	ema Direto : Determinação de V	16
		1.2.1	Aproximação para grandes distâncias	18
		1.2.2	Meio inomogêneo	19
		1.2.3	Modelo de 3 Esferas	19
		1.2.4	Métodos Numéricos	20
		1.2.5	Valores comuns para $\mathcal{K} \in \sigma_k \dots \dots$	22
	1.3	Proble	ema Inverso : Determinação de ${f J}$	22
		1.3.1	Linearização do problema	23
		1.3.2	Restrição da localização dos dipolos e definição de superfície s $\ .\ .\ .$.	23
		1.3.3	Modelagem Bayesiana	25
		1.3.4	Distribuições a Priori Hierárquicas	27
		1.3.5	Bayes Variacional	28
2	P_{TO}	blema	Inverso Abordagem Multi-Grid	32
-	0.1	M14:		0 _ 20
	2.1	Multi-	escala - Downsampling	32
	2.2	O pro	blema direto para cada uma das 5 superfícies	47
	2.3	Multi-	escala - Bayes Variacional	48
	2.4	Exem	plo de Simulação	50

3	Resultados		61
	3.1	Diferentes superfícies iniciais	62
	3.2	Diferentes intensidades de ruído	69
4	Con	clusão	77

Lista de Figuras

1.1	Esquematização da troca de informação entre neurônios: 1 - Potencial de Ação,	
	2 - Neurônio pré-sináptico, 3 - Sinapse, 4 - Neurônio pós-sináptico, 5 - Proteína	
	com permeabilidade seletiva sensível à mensageiros químicos [Wikimedia, b] [Wi-	
	kimedia, a]	14
1.2	esquema demonstrando as correntes dos PEPS e PIPS [Mann,] \hdots	15
1.3	fontes de corrente em um meio Ω com condutividade σ homogenia $\ \ldots \ \ldots \ \ldots$	18
1.4	k superfícies de 1 a $\mathcal{K},$ separando meios homogêneos, onde cada um possui uma	
	condutividade σ_k	20
1.5	O ângulo entre o vetor ${\bf r}$ que indica a posição onde o potencial esta sendo calculado	
	e o vetor \mathbf{r}^i que indica onde o dipolo i esta localizado é descrito por $\gamma.$ O ângulo	
	entre o dipolo \mathbf{q}^i e o vetor \mathbf{r}^i é descrito por α . O ângulo formado pelo plano	
	definido por ${\bf r^i}$ e ${\bf r}$ e o plano formado por ${\bf r^i}$ e ${\bf q}$ é descrito por β (fonte [Mosher	
	et al., 1999])	21
1.6	$3\ {\rm cortes}\ {\rm em}\ {\rm uma}\ {\rm imagem}\ {\rm tridimensional}\ {\rm de}\ {\rm MRI}$: coronal, sagital e horizontal,	
	$respectivamente. \ (Fonte: Free surfer \ http://surfer.nmr.mgh.harvard.edu/ \ (Athi-$	
	noula A. Martinos Center for Biomedical Imaging))	24
1.7	Visualização da superfície cortical com 566296 faces e 283152 vértices, gerada pelo	
	Freesurfer, a direita. Zoom de uma parte frontal do hemisfério esquerdo a esquerda.	25
1.8	Descrição do algoritmo Bayes Variacional	31
2.1	Hemisfério esquerdo, da esquerda para direita: dobrado (ou "real"), inflado e	
	esférico. As bordas dos triângulos estão transparentes e as cores dos mesmos	
	representam a curvatura na superfície original.	33
2.2	Icosaedro	34

2.3	Segunda ordem do algoritmo RE com 80 faces. A primeira coluna representa a	
	configuração inicial, dada pela divisão de cada aresta da ordem anterior no meio,	
	adicionando um novo vértice vermelho, enquanto os vértices que vieram da ordem	
	anterior, os azuis. Na coluna da direita estão representadas as configurações finais	
	do algoritmo RE. Os vértices azuis foram mantidos imóveis enquanto os vermelhos	
	sofreram pequenas mudanças de posição até que a variância parasse de diminuir.	
	Novamente os valores das legendas significam a área média na superfície original	
	representada por cada nova face, e entre parenteses a variância, em $mm^2.~\ldots$.	35
2.4	Terceira ordem.	36
2.5	O mesmo da figura 11, porém para a 4^a ordem com 1280 faces. É possível ver que	
	as faces laranjas somem, ou seja, a variância diminuiu. Vemos que isso corresponde	
	a pequenas distorções no grid na superfície esférica	37
2.6	5^a ordem com 5120 faces. Nesta escala a variância praticamente não diminuiu,	
	indicando que a divisão já ocorreu em uma condição inicial bem homogênea. $\ $	38
2.7	Representação da N^a faces que um determinado vértice \mathbf{v}^o_a qualquer faz parte. \mathbf{v}^o_b	
	e \mathbf{v}_{c}^{o} representam dois possíveis outros vértices que formam uma das faces, e as	
	setas indicam os versores representados por $\frac{\mathbf{v}_{o}^{c}+\mathbf{v}_{c}^{c}-2\mathbf{v}_{a}^{o}}{ \mathbf{v}_{o}^{c}+\mathbf{v}_{c}^{c}-2\mathbf{v}_{a}^{o} }$ na fórmula abaixo 2.4	40
2.8	Esquematização do algoritmo RE. A representação desejada escolhida foi 10240 $$	
	faces, o que equivale a repetir o passo 3 quatro vezes	42
2.9	Visualização das divisões das faces realizadas no passo 3 do algoritmo RE. Ordem	
	1 para Ordem 2	43
2.10	Visualização das divisões das faces realizadas no passo 3 do algoritmo RE. Ordem	
	2 para Ordem 3	44
2.11	Visualização das divisões das faces realizadas no passo 3 do algoritmo RE. Ordem	
	3 para Ordem 4	45
2.12	Visualização das divisões das faces realizadas no passo 3 do algoritmo RE. Ordem	
	4 para Ordem 5	46
2.13	Potencial elétrico gerado por dois dipolos, em V. Os eixos X e Y estão em mm. O	
-	campo visualizado em volta é uma interpolação dos valores conhecidos nos vértices	
	do polígono desenhado acima de tudo (estes vértices são os sensores).	48

2.14	A esquerda dois plots do prior de α para dois valores de α_0 com $\gamma_0=10,$ a direita	
	dois plots do prior de J para dois α diferentes (não normalizadas) $\hfill\$	49
2.15	Visualização das correntes de um exemplo que geraram o sinal utilizado para a	
	inferência. 26 faces possuem corrente diferente de zero	52
2.16	Visualização das correntes estimadas pelo método VB. Podemos ver que as 3 faces	
	visíveis, que representam 90% do espectro de corrente, estão espalhadas longe das	
	posições onde realmente exite corrente	53
2.17	Visualização das correntes estimadas pelo método MG. Podemos ver que as 21	
	faces visíveis, que representam 68% do espectro de corrente, estão localizadas	
	próximas à região com atividade real.	54
2.18	Visualização da variância inicial e final usada no algoritmo VB	55
2.19	Visualização da variância inicial e final usada na primeira iteração do algoritmo	
	MG. Podemos ver que desde a escala mais baixa a localização da corrente já é	
	coerente com a posição real das fontes	56
2.20	Visualização da variância inicial e final usada na segunda iteração do algoritmo	
	MG. Nesta escala já ocorre uma divisão entre as duas fontes. \ldots \ldots \ldots \ldots	57
2.21	Visualização da variância inicial e final usada na terceira iteração do algoritmo	
	MG. Ocorre uma grande concentração em torno das fontes reais. \ldots	58
2.22	Visualização da variância inicial e final usada na quarta iteração do algoritmo MG.	
	A convergência na coluna da direita mostra a correção de localização que ocorreu	
	principalmente na fonte da parte superior, enquanto a da parte inferior quase não	
	sofre mudanças. O interessante é que a correção na parte superior foi da ordem	
	de uma ou duas faces da escala anterior	59
2.23	Visualização da variância inicial e final usada na última iteração do algoritmo MG.	
	Podemos ver claramente que a variância do Prior se distribui uniformemente em	
	torno da fonte real, mostrando como a informação trazida da escala anterior foi	
	modificada corretamente pelo método de ARD	60
3.1	Comparação dos ângulos entre o vetor corrente real e as estimativas dos dife-	
	rentes algoritmos : Bayes Variacional (VB), Norma Mínima (MNE), Multigrid	
	começando da superfície de ordem 1 (MG 1), ordem 2 (MG 2), ordem 3 (MG 3)	
	e ordem 4 (MG 4)	63

3.2	Histograma com as distâncias entre a localização do dipolo mais forte estimado e o dipolo mais forte real para o algoritmo comecando da superfície 1 (MG 1) e da	
	superfície 2 (MG 2).	64
3.3	Histograma com as distâncias entre a localização do dipolo mais forte estimado e o dipolo mais forte real para o algoritmo começando da superfície 3 (MG 3) e da superfície 3 (MG 3).	65
3.4	Histograma com as distâncias entre a localização do segundo dipolo mais forte esti- mado e o segundo dipolo mais forte real para o algoritmo começando da superfície 1 (MG 1) e da superfície 2 (MG 2)	67
3.5	Histograma com as distâncias entre a localização do segundo dipolo mais forte esti- mado e o segundo dipolo mais forte real para o algoritmo começando da superfície 3 (MG 3) e da superfície 4 (MG 4)	68
3.6	Parâmetro ρ_1 para os algoritmos VB, MG 1 e MG 3. Podemos ver como o algo- ritmo começando da superfície com menor resolução possível é mais resistente ao aumento do ruído.	71
3.7	Parâmetro ρ_2 para os algoritmos VB, MG 1 e MG 3. Podemos ver como o algo- ritmo começando da superfície com menor resolução possível é mais resistente ao aumento do ruído.	71
3.8	A distância entre o dipolo mais intenso estimado pelo algoritmo VB e o dipolo mais intenso real é a curva azul (quadrados). A distância entre o dipolo mais intenso estimado pelos algoritmos MG 1 e MG 3 e o dipolo mais intenso real são as curvas vermelhas (círculos) com linha sólida e tracejada, respectivamente. Podemos ver que a distância aumenta de maneira compatível entre os dois algoritmos	72
3.9	A distância entre o segundo dipolo mais intenso estimado pelo algoritmo VB e o segundo dipolo mais intenso real é a curva azul (quadrados). A distância entre o segundo dipolo mais intenso estimado pelos algoritmos MG 1 e MG 3 e o dipolo mais intenso real são as curvas vermelhas (círculos) com linha sólida e tracejada, respectivamente. Podemos ver que os algoritmos MG são constantemente melhores	
	que o algoritmo VB original, enquanto o MG 1 é mais robusto a ruído que o MG 3.	72

3.10) Erro entre os falsos positivos e o primeiro foco de corrente: média da Distância		
	Euclidiana em mm entre o máximo do primeiro foco de corrente e os falsos posi-		
	tivos. VB é o algoritmo Bayes Variacional original, MG 1 é o algoritmo Multigrid		
	começando da superfície $1 \ {\rm e} \ {\rm MG} \ 3 \ {\rm \acute{e}}$ o algoritmo Multigrid começando da superfície		
	3. Podemos ver claramente como o algoritmo MG é mais robusto ao ruído . $\ .$.	73	
3.11	Erro entre os falsos positivos e o segundo foco de corrente: média da Distância		
	Euclidiana em mm entre os falsos positivos e o máximo do segundo foco de cor-		
	rente. VB é o algoritmo Bayes Variacional original, MG 1 é o algoritmo Multigrid		
	começando da superfície $1 \ {\rm e} \ {\rm MG} \ 3 \ {\rm \acute{e}}$ o algoritmo Multigrid começando da superfície		
	3. Podemos ver claramente como o algoritmo MG é mais robusto ao ruído . $\ .$.	74	
3.12	Pontos no espaço ROC (Sensibilidade por 1 - Especificidade) para diferentes níveis		
	de ruído. VB é o algoritmo Bayes Variacional original e MG 1 é o algoritmo		
	Multigrid começando da superfície 1	75	
3.13	Pontos no espaço ROC (Sensibilidade por 1 - Especificidade) para diferentes níveis		
	de ruído. VB é o algoritmo Bayes Variacional original e MG 3 é o algoritmo		
	Multigrid começando da superfície 3	76	
-		<u>م</u>	
1	Controles nas figuras do Matlab.	85	

Resumo

A Neuroimagem Funcional evoluiu muito nos últimos anos com o aparecimento de técnicas como Positron Emission Tomography ou PET (Tomógrafo por Emissão de Pósitrons) e Functional Magnetic Ressonance Image ou fMRI (Imagem de Ressonância Magnética Funcional) [Belliveau et al., 1991]. Elas permitem a observação de atividade no cérebro com uma resolução de alguns milímetros, e devido a natureza do sinal medido, com uma resolução temporal da ordem de 5 segundos [Kim et al., 1997]. Magnetoencefalografia e Eletroencefalografia (M/EEG), por outro lado, possuem uma resolução temporal da ordem de milissegundos, já que o sinal é produzido pela movimentação do íons através das membranas celulares [Nunez and Srinivasan, 2006]. Porém a sua resolução espacial é muito baixa já que tipicamente são problemas mal postos, com muito mais variáveis do que dados. Um equipamento de M/EEG de alta resolução possui da ordem de O(200) canais, que permitem medidas do campo magnético (para o MEG) ou do potencial elétrico (para o EEG) em O(200) posições em torno da cabeça. Para uma escala com resolução de ordem l existem $(\frac{L}{I})^3$ variáveis, onde $L \approx 15cm$.

Neste trabalho procuramos estudar métodos para aumentar a resolução espacial das técnicas de EEG, pois o mapeamento funcional do cérebro humano esta intimamente relacionado à localização da atividade no espaço bem como no tempo [Friston, 2009] (muitas relativo ao momento de um estímulo externo). Todo o trabalho de localização de fontes para EEG pode ser facilmente estendido para MEG.

Métodos Bayesianos são o cenário natural para lidar com problemas mal postos [Wipf and Nagarajan, 2009]. Existem, essencialmente, duas direções nas quais os algoritmos Bayesianos podem ser melhoradas, através da construção de uma melhor verossimilhança ou uma distribuição *a Priori*. Embora reconheçamos que avanços importantes podem ser feitos no direção anterior, aqui nos concentramos na segunda.

Neste trabalho nós introduzimos um método multiescala para construir uma melhor

distribuição *a Priori*. Uma idéia similar foi estudada dentro do contexto mais simples de fMRI [Amaral et al., 2004]. Muitos novos problemas aparecem ao lidar com o caráter vetorial do EEG. O mais importante, é a construção de um conjunto de superfícies renormalizadas que aproximam a região cortical onde a fonte de atividade esta localizada e o problema relacionado de definir as variáveis relevantes para representar o cérebro em uma escala com menor resolução.

A validação do novo algoritmo é sempre um problema essencial. Nós apresentamos resultados que sugerem, em dados simulados, que nosso método pode ser uma alternativa válida para os atuais algoritmos, julgando ambos pela taxa de erros na localização de fontes bem como pelo tempo que eles levam para convergir.

Abstract

Functional Neuroimaging has evolved in the last few decades with the introduction of techniques such as *Positron Emission Tomography* or PET and *Functional Magnetic Ressonance Image* or fMRI [Belliveau et al., 1991]. These allow observing brain activity with a resolution of a few millimeters and, due to the nature of the signal, a time resolution of the order of 5 seconds [Kim et al., 1997]. M/EEG, on the other hand, have a millisecond time resolution, since the signal is produced by the transport of ions through cell membranes [Nunez and Srinivasan, 2006]. However their space resolution is much lower since these are typically ill posed problems with many more unknowns than data points. A high resolution M/EEG has of the order of O(200)data channels, which allow measuring the magnetic or electric field at O(200) positions around the head. For a resolution scale of order l there are $O(\frac{L}{l})^3$ variables, where $L \approx 15cm$.

In this work we aim at studying methods to increase the spatial resolution of EEG techniques, since functional mapping of the human brain is intimately related to the localization of the activity in space as well as in time [Friston, 2009] (often relative to the time of external stimuli). Any advance in the inverse problem of source localization for EEG can rather easily be extended to deal with MEG.

Bayesian methods are the natural setting to deal with ill posed problems [Wipf and Nagarajan, 2009]. There are essentially two directions in which Bayesian algorithms can be improved, by building a better likelihood or a prior distribution. While we recognize that important advances can be done in the former direction we here concentrate in the latter.

In this work we introduce a multiscale method to build an improved prior distribution. A similar idea has been studied within an easier context of fMRI [Amaral et al., 2004]. Several new problems appear in dealing with the vectorial character of EEG. The most important, is the construction of a set of renormalized lattices that approximate the cortex region where the source activity is located and the related problem of defining the relevant variables in coarser scale representation of the cortex.

Validation of a new algorithm is always an essential problem. We present results which suggest on simulated data, that our method might be a valid alternative to current algorithms, judged both by the rate of errors in source localization as well as by the time it takes to converge.

Capítulo 1

Descrição dos modelos estudados de EEG/MEG

Faz muitos anos que o estudo de pequenas variações de potencial elétrico na superfície da cabeça é realizado em seres humanos, sendo de particular interesse devido a sua natureza não invasiva. Este tipo de estudo é chamado de Eletroencefalografia ou EEG, e a primeira pessoa a realizar experimentos desta natureza em humanos foi Hans Berger [Berger, 1929]. Em seu experimento ele utilizou folhas metálicas como eletrodos, uma na testa e uma na parte de trás da cabeça. Hoje em dia os aparelhos de EEG modernos possuem 256 eletrodos dispostos com espaços aproximadamente iguais entre si.

O estudo de sinais magnéticos associados com as correntes intra-celulares no cérebro é chamado de Magnetoencefalografia ou MEG [Cohen, 1968].O campo magnético na parte externa da cabeça é extremamente fraco, e é capturado por dispositivos supercondutores do tipo *Superconducting Quantum Interference Device* (SQUID) (existem estudos utilizando detectores não criogênicos chamados de *spin exchange relaxation-free* ou SERF).

O problema de M/EEG é dividido em duas partes: o problema de calcular o potencial elétrico ou o campo magnético devido a uma distribuição conhecida de correntes, chamado de Problema Direto (conhecido na literatura como *Forward Problem*), e o problema de calcular a distribuição de correntes para um determinado potencial elétrico ou campo magnético medido, chamado de Problema Inverso.

1.1 Origem do sinal

Daremos a seguir uma descrição muito simplificada da origem do sinal medido pelo M/EEG. Não pretendemos com isso explicar o funcionamento do cérebro, mas apenas dar uma idéia de qual fenômeno esta mais relacionado com os campos medidos pelo M/EEG, ou seja, explicitar que a principal fonte de campos medidos são as correntes pós-sinápticas e não o potencial de ação.

1.1.1 Potencial de Ação e Potencial Pós-Sináptico

Os neurônios são eletricamente ativos e são responsáveis pelo sinal medido pelo M/EEG. O cérebro possui aproximadamente 10^{10} neurônios, e estas células transmitem informação e se comunicam umas com as outras através de uma combinação de sinais químicos e elétricos. A figura 1.1 esquematiza esta comunicação. O neurônio da esquerda transmite informação para o neurônio da direita em forma de mensageiros químicos. A região onde ocorre a transmissão e recepção de mensageiros químicos entre os neurônios é chamada de sinapse. As moléculas liberadas na sinapse pelo neurônio pré-sináptico se acoplam às proteínas especializadas presentes na membrana do neurônio pós-sináptico, que por sua vez mudam a conformação e permitem um fluxo de íons que entram (ou saem) pela membrana. Por outro lado, estes íons saem (ou entram) constantemente por outras proteínas especializadas, que utilizam ATP para manter um gradiente químico através da membrana da célula. Esta diferença de concentração gera uma diferença de potencial elétrico constante entre a membrana do neurônio (da ordem de -75 mV). Sempre que as proteínas, que dependem de mensageiros químicos, permitem a passagem de íons, ocorre uma despolarização (ou hiperpolarização) local da célula, alterando localmente esta diferença de potencial. Esta alteração local de potencial elétrico é denominado Potencial Excitatório Pós Sináptico, ou PEPS no caso de uma despolarização, ou Potencial Inibitório Pós Sináptico, ou PIPS, no caso de uma hiperpolarização. Os PEPS e PIPS estão esquematizados na figura 1.2.

Quando uma quantidade suficientemente grande de PEPS ocorrem ao mesmo tempo (somatória espacial) ou em um intervalo suficientemente curto de tempo (somatória temporal), outras proteínas, que modificam a sua conformação por causa da mudança de potencial elétrico e estão localizadas no corpo celular próximo ao axônio (no cone de implantação), permitem a passagem de íons aumentando ainda mais a despolarização. Isso faz com que as proteínas do mesmo tipo que estão próximas também permitam a passagem de íons. Como estas proteínas



Figura 1.1: Esquematização da troca de informação entre neurônios: 1 - Potencial de Ação, 2 - Neurônio pré-sináptico, 3 - Sinapse, 4 - Neurônio pós-sináptico, 5 - Proteína com permeabilidade seletiva sensível à mensageiros químicos [Wikimedia, b] [Wikimedia, a].

dependente de potencial estão espalhadas por todo o axônio, este efeito em cascata se propaga pelo mesmo, e é denominado Potencial de Ação. Quando este potencial atinge os terminais no final do axônio, mensageiros químicos são liberados nas sinapses nestes terminais, passando a informação para o próximo neurônio. Por fim, quando a despolarização atinge um novo limiar, outras proteínas permitem a passagem de outros tipos de íons, polarizando novamente a célula e retornando esta ao seu estado de "repouso", com sua polarização original característica que é mantida pelas proteínas que utilizam ATP para bombear íons de diferentes tipos para dentro e para fora da célula [Bear et al., 2006].

Este fenômeno é extremamente transiente e dura da ordem de 1 ms, podendo ser modelado por um quadrupolo magnético (ver [Wikswo, 1983] e [Tanzer, 2006]) cujo potencial elétrico decai com $\frac{1}{r^3}$. Por outro lado os PEPS e PIPS podem durar dezenas de milissegundos e são modelados por um dipolo magnético que decai com $\frac{1}{r^2}$. Por esse motivo, acredita-se que o sinal do M/EEG é devido principalmente a soma dos PEPS e PIPS que ocorrem nas sinapses. Se o Potencial de Ação influencia ou não no campo medido na superfície da cabeça esta atualmente em debate [Nummenmaa, 2008].



Figura 1.2: esquema demonstrando as correntes dos PEPS e PIPS [Mann,]

1.1.2 Simetria na Organização das Fontes

A soma dos PEPS e PIPS de apenas um neurônio não é suficiente para gerar um campo ou potencial na superfície da cabeça suficientemente grande, de forma que possa ser distinguido de ruído. Para isso é necessário que dezenas ou centenas de milhares de neurônios estejam alinhados e possuam uma simetria tal que a soma de seus PEPS e PIPS produza um potencial macroscopicamente relevante. Esta soma de fato ocorre principalmente na quinta camada do córtex [Okada, 1981], onde neurônios piramidais se distribuem de uma maneira razoavelmente simétrica, paralelos uns aos outros (onde dezenas de milhares de neurônios correspondem a algo da ordem de 1 mm^2 da superfície cortical [Rockel et al., 1980]). Outras estruturas mais profundas também possuem simetrias conhecidas, porém devido a rápida queda do potencial elétrico e do campo magnético com a distância, estas estruturas contribuem muito pouco para o sinal medido na superfície da cabeça [Nunez and Srinivasan, 2006].

Esta distribuição de dipolos denominada *corrente de dipolos* é o modelo de fontes normalmente adotado nos problemas de M/EEG, e será também utilizado nesse trabalho.

1.2 Problema Direto : Determinação de V

Uma vez definida as características do sinal gerado pelo M/EEG, precisamos representar matematicamente os campos gerados por essas correntes, o que é feito a partir das equações de Maxwell macroscópicas

$$\boldsymbol{\nabla} \times \mathbf{H} = \frac{\partial \mathbf{D}}{\partial t} + \mathbf{J},\tag{1.1}$$

$$\boldsymbol{\nabla} \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t},\tag{1.2}$$

$$\boldsymbol{\nabla} \cdot \mathbf{H} = 0, \tag{1.3}$$

$$\boldsymbol{\nabla} \cdot \mathbf{E} = \rho. \tag{1.4}$$

A relação entre os campos é descrita por

$$\mathbf{H} = \mu \mathbf{B},\tag{1.5}$$

$$\mathbf{D} = \epsilon \mathbf{E},\tag{1.6}$$

$$\mathbf{J} = \sigma \mathbf{E}.\tag{1.7}$$

onde μ é a permitividade do meio, ϵ a permeabilidade e σ a condutividade.

Devido a baixa velocidade dos portadores de carga dentro da cabeça e pela natureza de μ , ϵ e σ dos tecidos (puramente resistivo, ou com efeitos capacitivos desprezíveis [Nunez and Srinivasan, 2006]), é possível verificar que os campos do M/EEG podem ser descritos por uma aproximação *Quasi-estática* das equações de Maxwell [Plonsey, 1969].

Desta forma os termos que variam com o tempo nas equações 1 e 2 são iguais a zero

$$\boldsymbol{\nabla} \times \mathbf{H} = \mathbf{J},\tag{1.8}$$

$$\boldsymbol{\nabla} \times \mathbf{E} = 0. \tag{1.9}$$

Devido a equação ~1.9podemos escrever o campo vetorial ${\bf E}$ em termos de um campo escalarV

$$\mathbf{E} = -\boldsymbol{\nabla} V. \tag{1.10}$$

Por fim, uma última suposição será feita para descrever o modelo direto: as correntes dentro da cabeça podem ser separadas em uma corrente macroscópica chamada de *corrente* primária $\mathbf{J}_{\mathbf{p}}^{\mathbf{i}}$ e uma corrente induzida ou corrente de volume $\mathbf{J}_{\mathbf{e}}$ [Geselowitz, 1967]

$$\mathbf{J}_{\mathbf{e}} = \sigma \mathbf{E} = -\sigma \nabla V,$$

$$\mathbf{J} = \mathbf{J}_{p}^{i} - \sigma \nabla V, \qquad (1.11)$$

onde todas as correntes estão dentro de uma superfície fechada Ω . Isto pode ser feito se modelarmos as nossas fontes como *fontes mesoscópicas*, compostas por grupos de milhares de neurônios, de forma que as correntes descritas na figura 1.2 de um neurônio no grupo *i* se somem apenas com as correntes de neurônios do mesmo grupo. Esta modelagem é feita devido à evidencia fisiológica de que os neurônios da terceira camada do córtex se agrupam em colunas razoavelmente distantes umas das outras. Ainda, todas as correntes estão dentro de uma região definida pela superfície da cabeça.

Desta forma podemos modelar cada grupo como um *dipolo magnético*, sendo que não existem fontes de corrente entre eles, mas apenas *correntes de volume*, como mostra a figura 1.3 [Nunez and Srinivasan, 2006].

Utilizando a equação 1.11, a equação de continuidade 1.12

$$\boldsymbol{\nabla} \cdot \mathbf{J} + \frac{\partial \rho}{\partial t} = 0 \tag{1.12}$$

e levando em conta que a cabeça é macroscopicamente neutra $(\frac{\partial \rho}{\partial t} = 0)$, podemos escrever a equação que relaciona o potencial elétrico medido na superfície da cabeça com as *correntes* primárias \mathbf{J}_{p}^{i}

$$\boldsymbol{\nabla} \cdot \mathbf{J}_p^i = \boldsymbol{\nabla} \cdot \boldsymbol{\sigma} \boldsymbol{\nabla} V. \tag{1.13}$$

Considerando o meio homogêneo e isotrópico, a equação 1.13 se torna do tipo de Poisson

$$\boldsymbol{\nabla} \cdot \mathbf{J}_{p}^{i} = \sigma \boldsymbol{\nabla}^{2} V \tag{1.14}$$



Figura 1.3: fontes de corrente em um meio Ω com condutividade σ homogenia

e usando a condição de contorno que não existem cargas fora da superfície Ω , podemos calcular o potencial com referência no infinito em função da *corrente primária*

$$V_0^i(\mathbf{r}) = \frac{1}{4\pi\sigma} \int_{\Omega} \mathbf{J}_p^i(\mathbf{r}') \cdot \frac{\mathbf{r} - \mathbf{r}'}{|\mathbf{r} - \mathbf{r}'|^3} d\Omega.$$
(1.15)

1.2.1 Aproximação para grandes distâncias

Para distâncias grandes em relação ao tamanho das fontes, podemos aproximar \mathbf{J}_p^i como

$$\mathbf{J}_{p}^{i} = \mathbf{q}^{i}\delta(\mathbf{r} - \mathbf{r}^{i}), \tag{1.16}$$

onde \mathbf{q}^i é a densidade de momento de dipolo magnético, ou momento de dipolo magnético por metro associado á posição $i \in \delta(\mathbf{r} - \mathbf{r}^i)$ é a função Delta de Dirac calculada na posição \mathbf{r}^i . Esta posição \mathbf{r}^i é a posição da fonte i. Descobrir \mathbf{q}^i é a principal preocupação deste trabalho.

Desta forma $V_0^i,$ que é o potencial na posição rgerado por $\mathbf{q}^i,$ se torna

$$V_0^i(\mathbf{r}) = \frac{1}{4\pi\sigma} \frac{\mathbf{q}^i \cdot \mathbf{d}^i}{(d^i)^3},\tag{1.17}$$

onde $\mathbf{d}^i = \mathbf{r} - \mathbf{r}^i$ e $d^i = |\mathbf{r} - \mathbf{r}^i|$.

1.2.2 Meio inomogêneo

Se consideramos o meio isotrópico, porém inomogêneo, onde a condutividade do meio muda em \mathcal{K} superfícies bem definidas e com variações suficientemente suaves, sendo que todas as fontes \mathbf{J}_p^i se localizam dentro da superfície mais interna k = 0 (figura 1.4), podemos calcular o potencial no infinito [Geselowitz, 1967]

$$\sigma_{\mathcal{K}}V^{i}(\mathbf{r}) = \sigma_{0}V_{0}^{i}(\mathbf{r}) + \frac{1}{4\pi}\sum_{k=0}^{\mathcal{K}}\int_{S_{k}}(\sigma_{k} - \sigma_{k}^{+})V(\mathbf{r}')\frac{\mathbf{d}^{i}}{(d^{i})^{3}}\cdot\mathbf{n}dS_{k}^{'},$$
(1.18)

onde o $V^i(\mathbf{r})$ representa o potencial do lado de fora da última superfície, mais exterior, representando assim o potencial medido no eletrodo no lado externo da cabeça, σ_k representa a condutividade da parte interna à superfície $k \in \sigma_k^+$ representa a condutividade na parte externa à superfície k. Podemos interpretar o primeiro termo da direita na equação 1.18 como sendo o potencial em \mathbf{r} para um meio homogêneo, gerado por uma fonte em \mathbf{r}^i que esta dentro da superfície mais interna S_0 . E a somatória no segundo termo como a correção devido as fontes secundárias induzidas nas superfícies de separação dos meios.

1.2.3 Modelo de 3 Esferas

Se considerarmos \mathcal{K} esferas, com condutividades σ_k com $k \in [1.\mathcal{K}]$, com $\sigma_{\mathcal{K}}^+ = 0$, temos as condições de contorno necessárias para resolver a equação 1.18 analiticamente. A solução desta equação em coordenadas esféricas é dada pelos Polinômios de Legendre [Zhang, 1995] [Mosher et al., 1999]

$$V(\mathbf{r}, \mathbf{r}^{i}, \mathbf{q}^{i}) = \frac{q^{i}}{4\pi\sigma_{\mathcal{K}}r^{2}} \sum_{n=1}^{\infty} \frac{2n+1}{n} (\frac{r^{i}}{r})^{n-1} f_{n}[n\cos\alpha P_{n}(\cos\gamma) + \cos\beta\sin\alpha P_{n}^{1}(\cos\gamma)], \quad (1.19)$$

onde $q^i = |\mathbf{q}^i|$, $r^i = |\mathbf{r}^i|$, $r = |\mathbf{r}|$, α é o ângulo ente $\mathbf{r}^i \in \mathbf{q}^i$, γ é o ângulo entre $\mathbf{r}^i \in \mathbf{r}$, β é o ângulo entre o plano definido por $\mathbf{r}^i \in \mathbf{q}^i$ e o plano definido por $\mathbf{r}^i \in \mathbf{r}$ (figura 1.5), $P_n \in P_n^1$ são os Polinômios de Legendre e os Polinômios Associados de Legendre, respectivamente,

$$f_n = \frac{n}{n m_{22} + (1+n) m_{21}},\tag{1.20}$$

e m_{21} e m_{22} são dados por

$$\begin{bmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{bmatrix} = \frac{1}{(2n+1)^{(\mathcal{K}-1)}} \prod_{k=1}^{\mathcal{K}-1} \begin{bmatrix} n + \frac{(n+1)\sigma_k}{\sigma_{k+1}} & (n+1)(\frac{\sigma k}{\sigma_{k+1}} - 1)(\frac{r^i}{r^k})^{2n+1} \\ n(\frac{\sigma k}{\sigma_{k+1}} - 1)(\frac{r^k}{r^i})^{2n+1} & (n+1) + \frac{n\sigma_k}{\sigma_{k+1}} \end{bmatrix}.$$
 (1.21)



Figura 1.4: k superfícies de 1
a $\mathcal{K},$ separando meios homogêneos, onde cada um pos
sui uma condutividade σ_k

1.2.4 Métodos Numéricos

Se as superfícies tiverem formas mais realistas, descrevendo o formato da cabeça e do crânio (obtidas a partir de imagens de MRI), é possível resolver o problema numericamente usando *Border Element Method* - BEM e *Finite Element Method* - FEM [Cuffin, 1995] [Oostendorp and van Oosterom, 1989] [Phillips, 2001]

Para isso é preciso discretizar as superfícies S_k e aproximar a integral em dS_k por uma somatória de M integrais sobre os elementos que as compõe

$$\sigma_{\mathcal{K}} V^{i}(\mathbf{r}) = \sigma_{0} V_{0}^{i}(\mathbf{r}) + \frac{1}{4\pi} \sum_{k=0}^{\mathcal{K}} \sum_{m=1}^{M} \int_{\Delta_{m}} (\sigma_{k} - \sigma_{k}^{+}) V(\mathbf{r}') \frac{\mathbf{d}^{i}}{(d^{i})^{3}} \cdot \mathbf{n} dS_{km}^{'}, \qquad (1.22)$$

como foi feito na referência [Phillips, 2001] utilizando triângulos como elementos que discreti-



Figura 1.5: O ângulo entre o vetor \mathbf{r} que indica a posição onde o potencial esta sendo calculado e o vetor \mathbf{r}^i que indica onde o dipolo *i* esta localizado é descrito por γ . O ângulo entre o dipolo \mathbf{q}^i e o vetor \mathbf{r}^i é descrito por α . O ângulo formado pelo plano definido por \mathbf{r}^i e \mathbf{r} e o plano formado por \mathbf{r}^i e \mathbf{q} é descrito por β (fonte [Mosher et al., 1999])

zam as superfícies S_k . Porém outras superfícies isoparamétricas podem ser usadas no lugar do triângulos, utilizando funções quadráticas ou cubicas para interpolar os elementos que compõe a discretização. Ainda é possível introduzir anisotropia no meio, considerando σ_k como tensores [Tanzer, 2006].

Não iremos entrar em detalhes sobre a implementação dos métodos numéricos para calcular o *forward model*. Todos estes métodos estão implementados e disponíveis em um software open-source chamado Fieldtrip : http://fieldtrip.fcdonders.nl/ (Donders Institute for Brain, Cognition and Behaviour, The Netherlands). Além do algoritmo de Phillips, estão implementados os algoritmos de [Oostendorp and van Oosterom, 1989], [Gramfort et al., 2010] e vários outros.

É interessante notar que no trabalho de [Tanzer, 2006] uma nova forma de modelar as fontes foi proposta, utilizando *Whitney Elements* [Bossavit, 1988]. Desta maneira no lugar de uma densidade superficial de dipolos, se consegue uma densidade volumétrica de corrente. Esta distribuição, ao contrário da normalmente utilizada no problema de M/EEG, possui a propriedade de convergir para a solução analítica correta quando os elementos de corrente se tornam infinitesimais. Porém o custo desta modelagem é aumentar 3 vezes o número de parâmetros desconhecidos para cada elemento estimado.

1.2.5 Valores comuns para $\mathcal{K} \in \sigma_k$

Normalmente a cabeça é modelada como tendo 3 superfícies: Ar/Escalpo, Escalpo/Crânio e Crânio/Cérebro. Alguns modelos mais preciso levam em consideração uma quarta superfície Crânio/LCR e LCR/Cérebro, onde LCR é Líquido Cefalorraquidiano. A seguir uma tabela com os valores típicos das condutividades citadas acima (fonte [Nunez and Srinivasan, 2006])

Material	Resistividade $\left(\frac{1}{\sigma}\right)$ (Ω cm)
Água do mar	20
LCR	64
Sangue	150
Córtex (5KHz)	230
Cortex (5Hz)	350
Substância Branca (média)	650
Osso (100Hz)	8.000-16.000
Membrana Ativa (neurônio da lula)	2×10^7
Membrana Passiva (neurônio da lula)	10^{9}
Ar	$10^{10} - 10^{11}$

1.3 Problema Inverso : Determinação de J

Determinar a corrente a partir do potencial V, como dito anteriormente, é um problema mal posto. Existem duas abordagens mais comuns para o problema. A primeira consiste em escolher um numero fixo de dipolos que devem ter suas posições, orientações e intensidades encontradas. Esta abordagem normalmente é chamada de Corrente de Dipolos Equivalente (*Equivalent Current Dipoles* ou ECD). Um dos problemas com este método é que o número de dipolos é desconhecido. Alguns métodos restringem a busca de cada dipolo em um compartimento exclusivo do cérebro, bem como tentam estimar qual o número de dipolos a serem procurados [Mosher and Leahy, 1998]. Porém além disso, este é um problema com muitos mínimos, logo é um problema de otimização não-linear em um espaço de 6 dimensões. Quando o número de dipolos é maior que algo da ordem O(5), este método começa a ficar intratável. Ainda, é difícil justificar fisiologicamente uma escolha arbitrária de localização de um dipolo.

A segunda abordagem consiste em escolher N posições fixas para os dipolos. Estas posições são extraídas de imagens de sMRI. Isto lineariza o problema, já que a parte não-linear depende apenas da posição dos dipolos. Ainda existem evidências fisiológicas que tornam possível extrair também a orientação dos dipolos a partir das imagens de sMRI. Isso reduz o problema não linear em um espaço de 6 dimensões para um problema linear em um espaço de 1 dimensão. Esta abordagem permite estimar um número de fontes da ordem de O(5000), é conhecida com Espaço de Fontes (*Source Space*) e será utilizada neste trabalho de agora em diante.

1.3.1 Linearização do problema

Como demostrado em [Mosher et al., 1999] e [Mosher et al., 1992] para métodos analíticos e em [Oostendorp and van Oosterom, 1989] e [Tanzer, 2006] para dois dos métodos numéricos, é possível separar a solução do potencial em uma parte linear que depende de \mathbf{q}^i e uma parte não linear que depende de \mathbf{r}^i . Uma vez fixadas as posições \mathbf{r}^i de todos os dipolos, bem como as suas orientações, todo o problema depende apenas da intensidade $|\mathbf{q}^i|$. Apenas fixar a posição já é suficiente. Se a orientação não estiver fixada, teremos um valor para cada direção do dipolo. Desta forma, podemos escrever o problema como um sistema linear

$$\mathbf{V} = \mathbf{G}\mathbf{J} + \boldsymbol{\xi},\tag{1.23}$$

onde V é uma matriz [M x T], sendo M o número de sensores e T o tempo (ou o número de medidas tomadas, normalmente da ordem de 500 por segundo), J é uma matriz [N x T], onde N é o número de dipolos, G é uma matriz [N x M] e ξ é o ruído na medida, assumido de agora em diante como sendo gaussiano. Ou seja, o ruído em um canal é tomado como gaussiano e não depende de qualquer outra coisa. Caso isso não seja mais verdade, todo o método de localização de fontes introduzido neste trabalho continua válido, sendo necessário apenas construir uma nova verossimilhança para o modelo Bayesiano.

 \mathbf{G} é chamado na literatura de *Lead Field*, e representa as características geométricas e fisiológicas da cabeça. Ou seja, dado um dipolo J_i ele mede qual é a *distância elétrica* até o eletrodo que esta medindo o potencial V_i . Cada linha de \mathbf{G} representa a distancia de todos os dipolos até um sensor *i* na mesma linha, e cada coluna representa a distância de uma dipolo *i* à todos os sensores.

1.3.2 Restrição da localização dos dipolos e definição de superfícies

Para fixar a posição e direção dos dipolos, utilizamos uma imagem de MRI como a que esta abaixo

As imagens de sMRI são feitas através da ressonância dos prótons das moléculas de água presentes no corpo. De acordo com a densidade de moléculas de água em cada voxel é possível reconstruir uma imagem tridimensional da cabeça. Foram desenvolvidos algorítimos para



Figura 1.6: 3 cortes em uma imagem tridimensional de MRI : coronal, sagital e horizontal, respectivamente. (Fonte : Freesurfer http://surfer.nmr.mgh.harvard.edu/ (Athinoula A. Martinos Center for Biomedical Imaging))

identificar os limites dos tecidos, gerando assim superfícies. Estas superfícies são representadas por um conjunto de **vértices** $\mathbf{v}_k^{\mathcal{O}}$ e **faces** $\mathbf{f}_n^{\mathcal{O}}$, onde \mathcal{O} representa a superfície original gerada a partir da imagem de sMRI. Para a superfície utilizada neste trabalho, $k \in [1..283152]$ e $n \in$ [1..566296].

Nós utilizamos um software chamado Freesurfer (http://surfer.nmr.mgh.harvard.edu/ (Athinoula A. Martinos Center for Biomedical Imaging)) implementado por [Dale and Sereno, 1993] [Fischl et al., 1999a]. Além de criar as superfícies este software infla as mesmas, e gera novas superfícies como é possível ver na Figure 9. Isto é feito minimizando a deformação ([Schwartz et al., 1989]), criando um espaço comum para comparação de superfícies corticais de indivíduos diferentes [Fischl et al., 1999b].

A superfície que aparece na figura 1.7 é a mais importante. Ela representa a superfície do córtex em sua conformação original e é usada, independente do modelo direto escolhido, para definir a posição e orientação dos dipolos. Como discutido no início, normalmente a orientação do dipolo é escolhida sempre como sendo perpendicular a cada triangulo que compõe a superfície (cada face), e a posição do dipolo é escolhida como sendo a mediana do triangulo.

Porém a superfície gerada a partir da imagem de sMRI normalmente possui uma granularidade muito maior do que é possível inferir a partir de apenas 200 ou 300 medidas de campo/potencial [Nummenmaa, 2008] [Sato and Yoshioka, 2008], pois estes algoritmos estão preocupados em representar da maneira mais fidedigna possível a imagem original. A solução normalmente utilizada para o problema é a dizimação da superfície, ou seja, escolher apenas uma



Figura 1.7: Visualização da superfície cortical com 566296 faces e 283152 vértices, gerada pelo Freesurfer, a direita. Zoom de uma parte frontal do hemisfério esquerdo a esquerda.

fração das faces como fontes, de maneira que as faces escolhidas fiquem o mais homogeneamente espalhadas. Isso é um problema porque a orientação da face não representa a região que esta em volta, o que pode gerar um viés incorreto na estimativa das fontes. Uma forma de atacar esse problema foi proposta por [Lin et al., 2006], fazendo uma média das normais das faces que estão mais próximas de cada ponto escolhido na dizimação. As faces mais próximas foram encontradas usando o algoritmo de Dijkstra.

Além das médias das normais, nós fizemos uma média das medianas de cada face da superfície original para cada mudança de escala, como descrito no próximo capítulo.

1.3.3 Modelagem Bayesiana

Podemos calcular a distribuição posterior de probabilidade de que uma corrente seja gerada por um determinado potencial utilizando a teorema de Bayes

$$P(J|VI) = \frac{P(V|JI)P_0(J|I)}{P(V|I)},$$
(1.24)

onde I é toda informação disponível, P(J|VI) é a probabilidade dos dipolos assumirem valores **J** dado que o potencial medido é **V**, P(V|JI) é a verossimilhança, $P_0(J|I)$ é a distribuição a priori e P(V|I) é a constante de normalização dada por

$$P(V|I) = \int P(V|\theta, I) P_0(\theta|I) d\theta, \qquad (1.25)$$

onde θ representa todos os parâmetros. Uma vez que supomos que o ruído seja gaussiano, sem correlações temporais ou espaciais (o que pode ser reforçado se realizarmos o mesmo experimento varias vezes e analisarmos a média para todas as execuções do experimento), podemos modelar a verossimilhança como

$$P(V|J\beta) = \left(\frac{1}{2\pi}\right)^{\frac{MT}{2}} |\beta \boldsymbol{\Sigma}_{\mathbf{G}}|^{\frac{T}{2}} \exp\left[-\frac{1}{2} \sum_{t=1}^{T} (\mathbf{V}(t) - \mathbf{GJ}(t))' \left(\beta \boldsymbol{\Sigma}_{\mathbf{G}}\right) \left(\mathbf{V}(t) - \mathbf{GJ}(t)\right)\right], \quad (1.26)$$

onde o símbolo ' significa a transposta da matriz e Σ_G é a matriz de covariância fixa a menos de uma constante, β . Esta matriz é estimada a partir do sinal **V**, e neste trabalho sera o inverso da variância de **V** : $\frac{1}{\sigma_V^2} = \frac{1}{\langle \mathbf{V}^2 \rangle - \langle \mathbf{V} \rangle^2}$. Se definirmos uma distribuição a priori não informativa uniforme

$$P_0(V) = \frac{1}{\alpha},\tag{1.27}$$

podemos calcular o Máximo a Posteriori ou MAP de P(J|V)

$$\frac{\partial \ln P(J|V)}{\partial J} = 0, \tag{1.28}$$

e encontramos o método de mínimos quadrados

$$\hat{J}_{MQ} = \left(G'\left(\beta \Sigma_G\right) G \right)^{-1} G'\left(\beta \Sigma_G\right) V.$$
(1.29)

Porém devido a natureza mal posta do problema, esta solução é muito ruim. Para melhorar esta estimativa, podemos colocar novas informações a priori. Estas informações normalmente assumem a forma de densidades exponenciais [Baillet et al., 2001] e podem ser escritas na forma

$$P_0(\mathbf{J}(t)|\alpha) = \left(\frac{\alpha}{2\pi}\right)^{\frac{NT}{2}} \exp\{-f(\mathbf{J}(t))\}.$$
(1.30)

Diferentes $f(\mathbf{J}(t))$ foram propostos nos últimos anos ([Hämäläinen and Ilmoniemi, 1994], [Pascual-marqui, 1999], [Friston et al., 2008], [Wipf and Nagarajan, 2009]), porém vamos nos concentrar na família

$$f(\mathbf{J}(t)) = \frac{1}{2} \mathbf{J}(t)' \Sigma_{\alpha} \mathbf{J}(t), \qquad (1.31)$$

onde $\Sigma_{\alpha} = diag(\alpha)$ é a matriz de covariância. Este método define que as correntes com menor norma tem maior probabilidade ou *Estimativa de Norma Mínima* (ou *Minimum Norm Estimate* MNE) e foi proposto a primeira vez por Hämäläinen [Hämäläinen and Ilmoniemi, 1994]. Se resolvermos a equação 1.28 com esta nova distribuição a priori, temos

$$\hat{\mathbf{J}}_{\mathbf{MNE}} = \mathbf{L}(\alpha, \beta) \mathbf{V}$$

$$= \left(\mathbf{G}'\left(\beta\boldsymbol{\Sigma}_{\mathbf{G}}\right)\mathbf{G} + \boldsymbol{\Sigma}_{\alpha}\right)^{-1}\mathbf{G}'\left(\beta\boldsymbol{\Sigma}_{\mathbf{G}}\right)\mathbf{V}$$
$$= \boldsymbol{\Sigma}_{\alpha}^{-1}\mathbf{G}'\left(\mathbf{G}'\boldsymbol{\Sigma}_{\alpha}^{-1}\mathbf{G} + (\beta\boldsymbol{\Sigma}_{\mathbf{G}})^{-1}\right)^{-1}\mathbf{V}$$
(1.32)

1.3.4 Distribuições a Priori Hierárquicas

Seguindo o trabalho de [Sato et al., 2004], podemos assumir que cada dipolo possui um α associado : $\alpha \rightarrow \alpha = \{\alpha_i...\alpha_N\}$, de modo que α não seja mais uma constante e $\Sigma_{\alpha} = diag(\alpha)$. Assim podemos associar uma distribuição a priori para cada α e tentar estimar quais valores de α maximizam a posteriori. Ou seja, definir uma densidade de probabilidade diferente para cada uma das variâncias (ou seu inverso) α , e tentar estimar estes hyper-hyper-parâmetros a partir dos dados. Isto foi proposto primeiramente por Neal para determinar os pesos de redes neurais ([Neal, 1997]) e chamado de Priors Hierárquicos ou Automatic Relevance Determination

$$P_0(J,\alpha,\beta) = P_0(J|\alpha,\beta)P_0(\alpha)P_0(\beta),$$

onde

$$P_{0}(\mathbf{J}(t)|\alpha\beta) = \left(\frac{1}{2\pi}\right)^{\frac{N}{2}} |\beta\Sigma_{\alpha}|^{\frac{1}{2}} \exp\left[-\frac{\beta}{2}\mathbf{J}(t)'\Sigma_{\alpha}\mathbf{J}(t)\right],$$
$$P_{0}(\mathbf{J}|\alpha\beta) = \prod_{t=1}^{T} P(\mathbf{J}(t)|\alpha\beta), \qquad (1.33)$$

 \mathbf{e}

$$P_0(\alpha) = \prod_{n=1}^N \Gamma(\alpha_n | \bar{\alpha_{0n}}, \gamma_{0n\alpha}).$$
(1.34)

A função $\Gamma(\alpha_n | \alpha_{0n}, \gamma_{0n\alpha})$ é a Distribuição Gama definida por:

$$\Gamma(\alpha|\bar{\alpha},\gamma) \equiv \alpha^{-1} \left(\frac{\alpha\gamma}{\bar{\alpha}}\right)^{\gamma} \Gamma(\gamma)^{-1} e^{-\frac{\alpha\gamma}{\bar{\alpha}}},$$

sendo $\Gamma(\gamma)\equiv\int_0^\infty dt e^t t^{\gamma-1}.$

No trabalho original de Sato et al. também foi associado um prior para os valores fora da diagonal de Σ_{α} , definidos como λ , porém vamos analisar apenas o comportamento para

 Σ_{α} diagonal. Assumimos o prior de β uniforme : $P_0(\beta) = \frac{1}{\beta}$ Desta forma, podemos escrever a probabilidade conjunta de **J** e α ,

$$P(V, J, \alpha, \beta) = P(V|J\beta)P_0(J|\alpha, \beta)P_0(\alpha)P_0(\beta), \qquad (1.35)$$

e a posterior é

$$P(J,\alpha,\beta|V) = \frac{P(V,J,\alpha,\beta)}{P(V)}.$$
(1.36)

Para comparar diferentes modelos e escolher qual conjunto de α maximiza a posterior, precisaríamos calcular a integral em 1.25, que precisaria ser resolvida numericamente. Se considerarmos 1000 dipolos, e discretizarmos estes como tendo apenas 2 valores possíveis, teríamos 2^{1000} operações para calcular, onde um computador realizando 10^{13} operações de ponto flutuante por segundo levaria mais que a idade do universo para calcular o resultado! Por isso precisamos de um método aproximado.

1.3.5 Bayes Variacional

Primeiro definimos o funcional da energia livre de uma distribuição $Q(J, \alpha, \beta)$ utilizando a divergência de Kullback-Leibler (e.g. [Opper and Winther, 2001])

$$F(Q) = KL\left[Q(J,\alpha,\beta) || P(V,J,\alpha,\beta)\right] = \int dJ \, d\alpha \, d\beta \, Q(J,\alpha,\beta) \, \ln\left[\frac{Q(J,\alpha,\beta)}{P(V,J,\alpha,\beta)}\right].$$
(1.37)

A idéia é maximizar o limite inferior da energia livre. Utilizando a equação 1.36 podemos reescrever a equação 1.37 como

$$F(Q) = \ln[P(V)] - KL[Q(J,\alpha,\beta)||P(J,\beta,\alpha|V)], \qquad (1.38)$$

de modo que encontrar uma distribuição $Q(J, \alpha, \beta)$ que maximiza o limite inferior da energia livre é equivalente a minimizar esta divergência de Kullback-Leibler no segundo termo da equação 1.38, e *não* depende do valor de P(V). Este método é conhecido como *Bayes Variacional* ou VB.

Uma forma de restringir a família de possíveis $Q(J, \alpha, \beta)$ é utilizar a idéia de aproximação de campo médio (e.g. [Parisi, 1988] e [Bishop, 2007]). Para isso, supomos que $Q(J, \alpha, \beta)$ é fatorável em $Q_{J,\beta}(J,\beta) Q_{\alpha}(\alpha)$, e usamos a equação 1.38 para encontrar uma distribuição $Q_{J,\beta}$ que maximiza F(Q) para Q_{α} fixa, e uma distribuição de Q_{α} que maximiza F(Q) para $Q_{J,\beta}$ fixa:

$$\left[\frac{\partial F(Q)}{\partial Q_{J,\beta}}\right]_{Q_{\alpha}fixa} = 0 \Rightarrow \hat{Q}_{J,\beta}(J,\beta) \propto \exp\left[\langle \ln P(V,J,\alpha,\beta) \rangle_{\alpha}\right], \tag{1.39}$$

$$\left[\frac{\partial F(Q)}{\partial Q_{\alpha}}\right]_{Q_{J,\beta}fixa} = 0 \Rightarrow \hat{Q}_{\alpha}(\alpha) \propto \exp\left[\langle \ln P(V, J, \alpha, \beta) \rangle_{J,\beta}\right], \tag{1.40}$$

onde $\langle X \rangle_Y$ significa a média de X com respeito a Y, mantendo a outra variável fixa.

Sem novas aproximações, é possível reescrever $Q_{J,\beta}(J,\beta)$ como

$$Q_{J,\beta}(J,\beta) = Q_J(J|\beta)Q_\beta(\beta),$$

$$Q_\beta(\beta) = \int dJ \, Q_{J,\beta}(J,\beta).$$
(1.41)

Usando a equação 1.39 e considerando o primeiro $Q_{\alpha}(\alpha)$ como sendo a equação 1.34 (ou seja, o prior de α inicial) podemos calcular $Q_J(J|\beta)$. O resultado é uma produtória de distribuições normais com diferentes médias e mesma matriz de covariância

$$Q_J(J|\beta) = \prod_{t=1}^T N(\hat{J}_{new}(t), (\beta \Sigma_{Jnew})^{-1}), \qquad (1.42)$$
$$\hat{J}_{new}(t) = \Sigma_{Jnew}^{-1} G' \Sigma_G V(t),$$
$$\Sigma_{Jnew} = G' \Sigma_G G + \Sigma_{\alpha}.$$

A inversão da matriz N x N, Σ_{Jnew} , pode ser evitada utilizando a seguinte identidade matricial

$$(G'\Sigma_G G + \Sigma_{\alpha})^{-1}G'\Sigma_G = \Sigma_{\alpha}^{-1}G'(G\Sigma_{\alpha}^{-1}G' + \Sigma_G^{-1})^{-1}$$
(1.43)

Agora podemos calcular a distribuição $Q_{\beta}(\beta)$ realizando a integral 1.41. O resultado é uma distribuição Gamma com parâmetros

$$\hat{\gamma}_{\beta_{new}} = \frac{MT}{2}$$

$$\hat{\beta}_{new} = \frac{\hat{\gamma}_{\hat{\beta}_{new}}}{\frac{1}{2}\sum_{t=1}^{'} \left((B(t) - G\hat{J}_{new}(t))' \Sigma_G(B(t) - G\hat{J}_{new}(t)) + \hat{J}_{new}(t)' \Sigma_\alpha \hat{J}_{new}(t) \right)}$$
(1.44)

A equação 1.42 e 1.44 estão representadas pelo passo 1 na figura 1.8. Ou seja, elas utilizam a matriz Σ_{α} para calcular o J e o β , denominados agora de \hat{J}_{new} e $\hat{\beta}_{new}$. Agora, estes valores serão utilizados para calcular os novos α que compõe a matriz Σ_{α} , como descrito a seguir.

Usando a equação 1.40 e o $Q_{J,\beta}(J,\beta)$ que acabamos de calcular, podemos calcular $Q_{\alpha}(\alpha)$. O resultado é um produto de N Gammas

$$Q_{\alpha}(\alpha) = \prod_{n=1}^{N} \Gamma(\alpha_n | \hat{\alpha}_{n_{new}}, \hat{\gamma}_{n\alpha_{new}})$$
(1.45)

sendo

$$\hat{\gamma}_{n\alpha_{new}} = \gamma_{0n\alpha} + \frac{T}{2}$$

$$\hat{\alpha}_{n_{new}} = \frac{\hat{\gamma}_{n\alpha_{new}}}{\frac{\alpha_{0n}}{\gamma_{0n\alpha}} + \frac{\hat{\beta}_{new}}{2} \sum_{t=1}^{T} \hat{J}_{new}(t)^2 + \frac{T}{2} (\Sigma_{Jnew})_{nn}}$$
(1.46)

onde α_{0n} e $\gamma_{0n\alpha}$ são os parâmetros iniciais da distribuição Gamma na equação 1.34, também utilizados na primeira vez que o $Q_{J,\beta}$ é calculado.

A equação 1.46 descreve a conta realizada no passo 2 da figura 1.8. Para testar a convergência, diferentemente dos trabalhos anteriores, não achamos necessário calcular a energia livre, e apenas testamos a convergência de Σ_{α} .

Resumindo, durante uma iteração, o \mathbf{J} e o β são calculados usando o $\hat{Q}_{J,\beta}(J,\beta)$. Em seguida, o $\hat{Q}_{\alpha}(\alpha)$ é usado para calcular um $\hat{\alpha}_{n_{new}}$. Caso os alfas de Σ_{α} tenham convergido, o algoritmo termina e o \mathbf{J} encontrado é aceito. Caso contrário, volta-se para o passo 1 e o \mathbf{J} é calculado novamente, porém agora usando o novo Σ_{α} composto pelos $\hat{\alpha}_{n_{new}}$ calculados no passo 2. O algoritmo é iterado até que os alfas tenham convergido. Devido a construção do problema, é garantido que o algoritmo irá convergir para um máximo local da energia livre.

Os valores utilizados para as variáveis $\alpha_{0n} \in \gamma_{0n\alpha}$ foram extensivamente estudados por [Nummenmaa, 2008]. Como notado por ele, o modelo é bastante sensível aos valores iniciais utilizados para estas variáveis. Ele tentou inclusive incluir $\alpha_{0n} \in \gamma_{0n\alpha}$ como distribuições (ou hyper-hyper-parâmetros), porém o modelo se mostrou instável, não convergindo mas indicando um valor cada vez maior para α_{0n} .

Por isso nós propomos uma nova maneira de acessar a massa de probabilidades descrita por $P(V, J, \alpha, \beta)$. No trabalho original de [Sato et al., 2004], em uma das simulações, foram usadas duas superfícies com granularidades diferentes, uma com 2.494 vértices, e uma com 24.917. O algoritmo foi rodado na primeira, com menos faces. O pico desta face foi usado para extrair superfícies da segunda superfície, com maior granularidade, porém o prior nesta superfície mais granular era uniforme. Nós propomos não apenas uma redução dimensional, mas quantas forem possíveis até termos a superfície mais simples possível capaz de representar o córtex.

Depois que esta superfície estiver definida, podemos rodar o algoritmo VB e usar os α_n de quando o algoritmo convergiu como α_{0n} da próxima superfície, e assim sucessivamente até a superfície mais granular.

Porém imaginamos que as superfícies devam ter um vínculo, de forma que uma face da superfície mais grosseira não contenha muito mais faces (ou represente muito mais área) da superfície original. As superfícies com diferentes granularidades usadas nos trabalhos anteriores eram dizimações simples da superfície original.

Por isso desenvolvemos um algoritmo sistemático para criar várias superfícies mais simples, aumentando gradativamente o nível de complexidade. Este algoritmo é descrito no próximo capítulo.



Figura 1.8: Descrição do algoritmo Bayes Variacional

Capítulo 2

Problema Inverso Abordagem Multi-Grid

2.1 Multi-escala - Downsampling

Neste capítulo vamos introduzir e implementar um novo algoritmo para redimensionamento da superfície cortical, gerando 5 ordens de granularidade diferentes do córtex, mapeando cada face de uma superfície menos granular em 4 novas faces de uma superfície com 4 vezes mais triângulos.

O principal problema é a geometria complicada da superfície do córtex. Esta superfície tem em média 2400 cm^2 [Toro et al., 2008] e dobra em si mesma aproximadamente da mesma forma para todas as pessoas.

Existem evidencias fisiológicas de que a atividade em um determinado instante se concentra em regiões macroscópicas bem definidas, sendo que esta informação já foi usada em trabalhos anteriores para melhorar a estimava das correntes [Friston et al., 2008] e de voxels de fMRI [Amaral et al., 2004]. Porém nós modelamos o córtex com triângulos de diferentes tamanhos, e usamos a solução do algoritmo VB em uma escala com granularidade menor como valor inicial (ou seja, valor do Σ_{α}) para o algoritmo VB em uma escala com granularidade maior.

Seja *o* a ordem da superfície. Se $o \in \mathbb{N}$, a superfície com o menor número de triângulos tem o = 1 e a superfície com o maior número de triângulos $o = \mathcal{O}$. Cada superfície terá um número de triângulos N^o . Primeiro atualizamos o = 1. Queremos gerar uma superfície inicial o mais simples possível, e que forneça uma aproximação inicial do córtex.

Na esfera, o problema de gerar superfícies com menos faces fica mais fácil. Assim, inserimos um poliedro regular gerado analiticamente dentro da esfera que aparece na figura 2.1.



Figura 2.1: Hemisfério esquerdo, da esquerda para direita: dobrado (ou "real"), inflado e esférico. As bordas dos triângulos estão transparentes e as cores dos mesmos representam a curvatura na superfície original.

O poliedro escolhido foi o icosaedro. Este é o passo 0 da figura 2.8. Em seguida, encontramos quais são os vértices da superfície original esférica mais próximos dos vértices do icosaedro. Como a lista de **faces** permanece a mesma em todas as superfícies de uma mesma ordem (dobrada, inflada e esférica), temos agora uma aproximação inicial do córtex na superfície original dobrada. Esta superfície possui ordem 1, sendo $N^1 = 20$. Isto pode ser visto na figura 2.2. a primeira linha possui duas aproximações na superfície dobrada, e segunda linha as mesmas duas aproximações na esfera.


Figura 2.2: Na coluna da esquerda esta representada a configuração inicial do algoritmo RE: os vértices vermelhos são os vértices da superfície esférica (circunscrita transparente na segunda linha) que melhor se aproximam o icosaedro criado analiticamente em seu interior. Na primeira linha estão os mesmos vértices porém na superfície original. As cores das faces representam a soma das áreas dos triângulos da superfície original compostos pelos vértices que ficam acima das faces do icosaedro. Na coluna da direita estão as mesmas representações, porém após 22 iterações do algoritmo RE, após a variância parar de diminuir com pequenas mudanças das posições dos vértices. Os números na legenda significam a média e a variância da área na superfície original que cada face representa, em mm^2 .¹

 $^{^1\}mathrm{Na}$ verdade este valor equivale a aproximadamente 3 vezes o
 valor da área, pois cada triangulo original foi contado 3 vezes



Figura 2.3: Segunda ordem do algoritmo RE com 80 faces. A primeira coluna representa a configuração inicial, dada pela divisão de cada aresta da ordem anterior no meio, adicionando um novo vértice vermelho, enquanto os vértices que vieram da ordem anterior, os azuis. Na coluna da direita estão representadas as configurações finais do algoritmo RE. Os vértices azuis foram mantidos imóveis enquanto os vermelhos sofreram pequenas mudanças de posição até que a variância parasse de diminuir. Novamente os valores das legendas significam a área média na superfície original representada por cada nova face, e entre parenteses a variância, em mm^2 .



Figura 2.4: Terceira ordem.



Figura 2.5: O mesmo da figura 11, porém para a 4^a ordem com 1280 faces. É possível ver que as faces laranjas somem, ou seja, a variância diminuiu. Vemos que isso corresponde a pequenas distorções no grid na superfície esférica.



Figura 2.6: 5^a ordem com 5120 faces. Nesta escala a variância praticamente não diminuiu, indicando que a divisão já ocorreu em uma condição inicial bem homogênea.

A diferença entre as duas colunas da figura 2.2 é a área na superfície original que cada face do icosaedro representa. Diferentes escolhas do icosaedro podem levar a superfícies no córtex muito diferentes. Isto se deve a que o número de pontos nas faces do icosaedro pode variar muito. Precisamos eliminar esta arbitrariedade na escolha da superfície inicial. Para isso introduzimos um algoritmo de balanceamento do número de pontos em cada face, descrito a seguir.

Primeiro identificamos para quais pontos da superfície \mathcal{O} ($N^{\mathcal{O}} = 283148$), o raio da esfera intercepta a face criada por vértices da superfície o ($N^1 = 20$), usando a seguinte fórmula

$$t\mathbf{v}_{k}^{\mathcal{O}} = \mathbf{v}_{1}^{o} + u(\mathbf{v}_{2}^{o} - \mathbf{v}_{1}^{o}) + v(\mathbf{v}_{3}^{o} - \mathbf{v}_{2}^{o}),$$
(2.1)

onde $\mathbf{v}_k^{\mathcal{O}}$ é o ponto da esfera que estamos testando, $\mathbf{v}_p^o, p \in [1\,2\,3]$ são os pontos que definem uma face da superfície o, e $t, u, v \in \mathbb{R}$. Quando o raio do ponto k qualquer $\mathbf{v}_k^{\mathcal{O}}$ corta esta face da superfície que estamos criando

$$t \in [0, 1]$$

$$u, v \in [0, 1], (u + v) \leq 1.$$
 (2.2)

Em seguida no passo 1, para cada ponto que obedece a equação 2.2, a área dos triângulos que pertencem a superfície \mathcal{O} e tem este ponto como vértice, são computadas e somadas. No final, para cada face da nova superfície

$$A_n^o = \sum_{m \in A_n^o} A_m^{\mathcal{O}},\tag{2.3}$$

onde n é uma das faces do icosaedro, e m são faces da superfície original \mathcal{O} . Com $m \in A_n^o$ queremos dizer: as m faces formadas por vértices da superfície \mathcal{O} que interceptam a face n da superfície o. Este valor A_n^o pode ser visualizado como as cores das faces das figuras na primeira linha das figuras 2.2 a 2.6.

Isto é feito porque o algoritmo RE impõe o vínculo de que cada nova face represente aproximadamente a mesma área na superfície com a granularidade original. Pois, após o algoritmo VB convergir, a corrente estimada para uma face não irá conter muito mais informação da superfície original do que outras, já que esta informação será dividida entre todas as faces da superfície com maior granularidade (e é representada pelos parâmetros α_n como veremos na próxima seção).

Em seguida no passo 2, nós fazemos pequenas movimentações dos vértices da superfície o (da ordem de 1 mm) de modo a minimizar a variância da área da superfície original que fica acima de cada face, procurando afastar os vértices de faces com A_n^o grande e aproximar de faces com A_n^o pequeno. Para escolher a direção e o tamanho do passo, inicializamos ε com a distância média entre os vértices da superfície \mathcal{O} . Depois, para cada vértice \mathbf{v}_a^o da nova superfície o, um versor é criado apontando na direção das medianas de cada uma das N^a faces que ele faz parte (ou seja, todos os triângulos que \mathbf{v}_a^o é um vértice).



Figura 2.7: Representação da N^a faces que um determinado vértice \mathbf{v}_a^o qualquer faz parte. \mathbf{v}_b^o e \mathbf{v}_c^o representam dois possíveis outros vértices que formam uma das faces, e as setas indicam os versores representados por $\frac{\mathbf{v}_b^o + \mathbf{v}_c^o - 2\mathbf{v}_a^o}{|\mathbf{v}_b^o + \mathbf{v}_c^o - 2\mathbf{v}_a^o|}$ na fórmula abaixo 2.4.

Por fim, a posição do vértice \mathbf{v}_a^o é atualizada da seguinte forma

$$(\mathbf{v}_a^o)_{new} = \mathbf{v}_a^o + \varepsilon \sum_{a\hat{b}c \in N^a} (A^o_{a\hat{b}c} - \langle A^o \rangle_j) \frac{\mathbf{v}_b^o + \mathbf{v}_c^o - 2\mathbf{v}_a^o}{|\mathbf{v}_b^o + \mathbf{v}_c^o - 2\mathbf{v}_a^o|},\tag{2.4}$$

onde $(\mathbf{v}_a^o)_{new}$ é a nova posição do vértice \mathbf{v}_a^o , A_{abc}^o é a área da face formada pelos vértice \mathbf{v}_a^o , \mathbf{v}_b^o e \mathbf{v}_b^o , e $abc \in N^a$ significa todos os conjuntos de 3 vértices a, b e c que formam as faces contidas em N^a , sendo que a permanece fixo, e $\langle \rangle_j$ é a média sobre todas as j áreas desta superfície o, onde $j \in [1..N^o]$. Depois que todos os pontos recebem esta atualização, o algoritmo retorna para o passo 1. Quando a variância

$$\sigma_A^2 = <(A^o)^2 >_j - _j^2,$$

diminui em relação a iteração anterior, apenas aceita-se a mudança. Quando ela aumenta, o passo ε é dividido por dois, e o passo 2 é executado novamente. Quando a variância se apresenta aproximadamente estável após uma mudança de posição dos vértices no passo 2, estes são **fixados** (representados nas figuras 2.3 a 2.6 pela cor azul), permanecendo o mesmo durante todo o resto do algoritmo.

Caso já tenhamos a superfície com a granularidade (quantidade de triângulos) desejada, podemos encerrar o algoritmo ¹. Porém normalmente vamos ter que seguir ao menos uma vez para o passo 3, onde dividimos as faces do poliedro anterior em 4 faces menores, criando *um novo vértice para cada aresta*. Estes vértices são representados pela cor vermelha (figuras 2.3 a 2.6). Em seguida, atualizamos o = o + 1 e seguimos de volta para o passo 1. Após o cálculo das áreas A_n^o , apenas os novos vértices vermelhos são movidos no passo 2, se afastando de faces com muita área original acima e se aproximando de faces com pouca área. Os passo 1 e 2 são repetidos novamente até a variância convergir. Quando isso acontece, estes novos vértices vermelhos se tornam azuis e são fixados, e então voltamos para o passo 3. Este processo é repetido quantas vezes for necessário, até que o número de faces desejado seja atingido e enquanto o número de faces $N_o < N_O$. O algoritmo RE esta esquematizado na figura 2.8.

Resolvemos parar após a 5^a iteração, quando temos 10240 faces no total (ambos os hemisférios, 5120 cada um), com uma distância inter-vértice de aproximadamente 5.30 (2.01) mm, pois trabalhos anteriores mostraram que seriam necessários mais sensores para conseguir uma melhor resolução com o algoritmo VB [Nummenmaa, 2008].

Esta superfície de ordem 5 irá representar a nossa *superfície real*, ou seja, vamos colocar dipolos ativos nesta superfície, calcular o campo resolvendo o problema direto e usar apenas este campo (corrompido com ruido) para inferir quais fontes geraram o mesmo. Depois podemos comparar com as fontes *reais*. Isto sera melhor descrito na próxima seção.

Após o algoritmo convergir, cada face de cada ordem terá um dipolo associado. A posição do dipolo é definida como a média das m medianas das faces da superfície \mathcal{O} representadas pela face n da superfície o (sendo que pelo nosso critério $o \in [1, 2, 3, 4, 5]$)

¹Nós estamos no momento pensando em implementar um algoritmo semelhante com a granularidade final $(N^1 = 20)$ para localização de fontes de equipamentos portáteis de 16 eletrodos, por exemplo.



Figura 2.8: Esquematização do algoritmo RE. A representação desejada escolhida foi 10240 faces, o que equivale a repetir o passo 3 quatro vezes.

$$\mathbf{r}_{n}^{o} = \frac{\sum_{m \in A_{n}^{o}} \mathbf{r}_{m}^{\mathcal{O}}}{N_{m}^{\mathcal{O}}},\tag{2.5}$$

onde *n* representa uma face na superfície *o*, \mathbf{r}_n^o a posição de um dipolo representando a face *n* da superfície *o*, \mathbf{r}_m^o uma mediana de um triangulo da superfície \mathcal{O} e $N_{on}^{\mathcal{O}}$ representa o numero triângulos formados por vértices da superfície \mathcal{O} que interceptam a face *n* da superfície *o*.

Além disso a orientação do dipolo é fixada como o versor que aponta na direção da soma de todas as normais das faces formadas por estes pontos

$$\hat{\mathbf{q}}_{n}^{o} = \frac{1}{N_{on}^{\mathcal{O}}} \frac{\sum_{m \in A_{n}^{o}} \hat{\mathbf{q}}_{m}^{\mathcal{O}}}{||\sum_{m \in A_{n}^{o}} \hat{\mathbf{q}}_{m}^{\mathcal{O}}||}.$$
(2.6)

Deste modo temos uma maneira sistemática de realizar a redução dimensional da superfície original, e representar uma superfície muito granular com poucos dipolos, aumentando gradualmente a quantidade de dipolos e mapeando cada 4 novos dipolos em 1 dipolo de uma superfície mais grossa. Abaixo nas figuras 2.9 a 2.12 temos uma sequência de imagens do córtex em diferentes escalas, passando de ordem 1 para 2, 2 para 3, 3 para 4 e 4 para 5 respectivamente. Cada imagem representa uma execução do passo 3 da figura 2.8. As ordens 1, 2, 3, 4 e 5 possuem 40, 160, 640, 2560 e 10240 faces, respectivamente. A imagem da esquerda de cada figura sempre possui uma cor aleatória associada a cada face, sendo que esta cor é mantida na imagem da direita, e é associada a cada uma das 4 novas faces em que a face anterior foi dividida.



Figura 2.9: Visualização das divisões das faces realizadas no passo 3 do algoritmo RE. Ordem 1 para Ordem 2.



Figura 2.10: Visualização das divisões das faces realizadas no passo 3 do algoritmo RE. Ordem 2 para Ordem 3.



Figura 2.11: Visualização das divisões das faces realizadas no passo 3 do algoritmo RE. Ordem 3 para Ordem 4.



Figura 2.12: Visualização das divisões das faces realizadas no passo 3 do algoritmo RE. Ordem 4 para Ordem 5.

2.2 O problema direto para cada uma das 5 superfícies

Agora que temos 5 superfícies com diferentes números de faces, uma posição e uma direção de um dipolo associado a cada uma destas faces, podemos calcular o *Lead Field* de cada uma destas superfície

$$\mathbf{V} = \mathbf{G}^o \mathbf{J}^o + \xi, \tag{2.7}$$

onde $o \in [1, 5]$, **V** é uma matriz M x T com o potencial medido em cada sensor em cada instante de tempo, com M sendo o número de sensores e T o número de medidas, **G**^o é uma matriz M x N^o com o Lead Field de ordem o, com N^o sendo o número de dipolos (e de faces) da ordem o, e **J**^o é uma matriz N^o x T com a intensidade de cada dipolo, que esta localizado em \mathbf{r}_n^o com orientação $\hat{\mathbf{q}}_n^o$, onde $n \in [1, N^o]$.

As matrizes G^o foram geradas usando o programa Fieldtrip, conforme discutido anteriormente. Nós escolhemos um modelo de 3 esferas para as simulações com EEG. O potencial elétrico gerado por estes dipolos na superfície da cabeça é obtido simplesmente multiplicando o *Lead Field* pelas intensidades associadas aos dipolos. Durante todas as simulações, escolhemos M = 128 sensores.

Abaixo na figura 2.13 temos uma interpolação do potencial gerado por 2 dipolos em duas faces aleatórias da superfície de ordem o = 5. As faces da superfícies estão transparentes para facilitar a visualização e as arestas estão verdes. Os dipolos são representados por setas pretas.



Figura 2.13: Potencial elétrico gerado por dois dipolos, em V. Os eixos X e Y estão em mm. O campo visualizado em volta é uma interpolação dos valores conhecidos nos vértices do polígono desenhado acima de tudo (estes vértices são os sensores).

2.3 Multi-escala - Bayes Variacional

Utilizando o algoritmo RE, nós fizemos uma modificação no algoritmo VB, de forma a considerar no Prior de J a informação resultante da solução do problema em um escala com resolução menor. Este é o algoritmo Multigrid ou MG.

Conforme discutido anteriormente, uma intensa pesquisa sobre os valores iniciais de α_{0n} e $\gamma_{0n\alpha}$ foi realizada no trabalho de [Nummenmaa, 2008]. Ele verificou que o modelo é bastante sensível as variáveis iniciais, sendo que estas dão acesso a diferentes partes da massa de probabilidade durante a execução do algoritmo VB.

Em [Amaral et al., 2004] foi proposto um método Multi-escala para lidar com a caracterização da distribuição a priori em fMRI. Dado que a natureza do problema de fMRI e M/EEG é diferente precisamos repensar o uso de tal idéia. A atividade em fMRI esta representada por um escalar enquanto que em M/EEG as densidades de corrente são vetores. Após várias tentativas infrutíferas de tentar definir graus de liberdade vetoriais renormalizados, percebemos que isso afortunadamente não é necessário. Trabalhamos com o escalar α , o inverso da variância da distribuição a priori de correntes. Esta é a extensão mais natural e simples do método de multigrid fMRI para correntes de M/EEG, que descrevemos a seguir com mais detalhes.

Decidimos distribuir a informação da probabilidade de existência de uma corrente em uma face de uma superfície com menos triângulos em faces de uma superfície com mais triângulos (4 faces). A melhor maneira de propagar informação de uma superfície com menos faces para uma nova com mais faces era utilizar o valor de $\hat{\alpha}_{n_{new}}$ estimado na última iteração do algoritmo VB em uma superfície com menos triângulos como o valor de α_{0n} para as 4 novas superfícies em que esta foi dividida. Desta forma regiões com mais variância terão um peso maior no Prior da corrente naquela região, e regiões com menos variância terão um peso menor. Podemos verificar isso na figura 2.14



Figura 2.14: A esquerda dois plots do prior de α para dois valores de $\alpha_0 \text{ com } \gamma_0 = 10$, a direita dois plots do prior de J para dois α diferentes (não normalizadas)

Podemos ver que os lugares com α **pequeno** permitem que a corrente assuma mais valores do que um α **grande**, pois este parâmetro é o inverso da variância. Assim qualquer variação na corrente nessa região tem uma grande importância na maximização da entropia. Já regiões que desde a maior granularidade se mostram silenciosas, influenciarão menos neste processo. Desta forma temos uma forma de determinar a relevância dos dados a partir dos próprios dados, ou uma Determinação Automático de Relevância (Automatic Relevance Determination ARD [Neal, 1997]). Para o algoritmo VB original usamos em todas as simulações $\alpha_{0n} = 1$ e $\gamma_{0n\alpha} = 0.1$. Para o algoritmo MG, usamos estes valores apenas na primeira superfície, com 40 faces. Depois, aumentamos a ordem da superfície e usamos o valor de $\hat{\alpha}_{n_{new}}$ atribuído a cada uma das 40 faces como α_{0n} para as novas 160 faces, distribuídas conforme o mapa na figura 2.9. Antes de distribuir porém, nós "espalhamos" o valor de $\hat{\alpha}_{n_{new}}$ entre os primeiros vizinhos, pois o sinal provavelmente esta próximo ao dipolo associado àquela face, mas pode estar na face vizinha. Assim, antes de distribuir o $\hat{\alpha}_{n_{new}}$ para o α_{0n} das novas faces, nós atualizamos o mesmo da seguinte forma

$$\bar{\hat{\alpha}}_{n_{new}} = \hat{\alpha}_{n_{new}} + \frac{1}{N_{vn}} \sum_{m \in V(n)} \hat{\alpha}_{m_{new}}, \qquad (2.8)$$

onde $m \in V(n)$ quer dizer todas as faces vizinhas a face n. Desta forma ocorre um "espalhamento" do α nas proximidades de onde a variância é alta, pois a localização é aproximada e o sinal verdadeiro pode estar nas vizinhanças e não na face onde a escala mais grosseira convergiu. Esta informação é apenas um prior para a próxima iteração do algoritmo. Como é possível ver na figura 2.22, muitas vezes ela muda durante a convergência do algoritmo naquela escala. Na próxima seção mostraremos resultados da simulação de 2 dipolos.

2.4 Exemplo de Simulação

Nas figuras a seguir 2 dipolos estão ativos na superfície de ordem 5, a 55 mm de distancia um do outro, com os primeiros vizinhos também ativos porém com intensidade mais baixa. O potencial foi calculado usando o modelo direto e apenas esta informação foi utilizada para estimar as fontes. O potencial foi corrompido com ruído com Noise-to-Signal Ratio NSR igual a 0.1. A figuras 2.15, 2.17 e 2.16 mostram as soluções, ou seja, a corrente estimada. As demais figuras mostram a valor do inverso de α_n (a variância) para o algoritmo VB e para cada uma das iterações do algoritmo MG.

Esta simulação é uma das 100 utilizadas na comparação mais detalhada feita no próximo capítulo. Ela foi escolhida por exemplificar uma característica muito interessante do algoritmo MG. Sabidamente as fontes mais profundas são estimadas como correntes espalhadas na superfície. Outras distribuições a priori foram utilizadas para levar em conta este efeito (e.g. [Pascual-marqui, 1999]). Porém como podemos ver na simulação a seguir, a identificação de uma fonte localizada em um sulco foi realizada com sucesso pelo método MG, enquanto o algoritmo original VB espalhou o sinal em algumas fontes na superfície do córtex. Isto aconteceu em pelo menos 10 das 100 simulações em posições aleatórias realizadas, sugerindo que o método é robusto com relação a profundidade das fontes.



Figura 2.15: Visualização das correntes de um exemplo que geraram o sinal utilizado para a inferência. 26 faces possuem corrente diferente de zero.



Figura 2.16: Visualização das correntes estimadas pelo método VB. Podemos ver que as 3 faces visíveis, que representam 90% do espectro de corrente, estão espalhadas longe das posições onde realmente exite corrente.



Figura 2.17: Visualização das correntes estimadas pelo método MG. Podemos ver que as 21 faces visíveis, que representam 68% do espectro de corrente, estão localizadas próximas à região com atividade real.



Figura 2.18: Visualização da variância inicial e final usada no algoritmo VB.



Figura 2.19: Visualização da variância inicial e final usada na primeira iteração do algoritmo MG. Podemos ver que desde a escala mais baixa a localização da corrente já é coerente com a posição real das fontes.



Figura 2.20: Visualização da variância inicial e final usada na segunda iteração do algoritmo MG. Nesta escala já ocorre uma divisão entre as duas fontes.



Figura 2.21: Visualização da variância inicial e final usada na terceira iteração do algoritmo MG. Ocorre uma grande concentração em torno das fontes reais.



Figura 2.22: Visualização da variância inicial e final usada na quarta iteração do algoritmo MG. A convergência na coluna da direita mostra a correção de localização que ocorreu principalmente na fonte da parte superior, enquanto a da parte inferior quase não sofre mudanças. O interessante é que a correção na parte superior foi da ordem de uma ou duas faces da escala anterior.



Figura 2.23: Visualização da variância inicial e final usada na última iteração do algoritmo MG. Podemos ver claramente que a variância do Prior se distribui uniformemente em torno da fonte real, mostrando como a informação trazida da escala anterior foi modificada corretamente pelo método de ARD.

Capítulo 3 Resultados

Todas as simulações a seguir foram feitas com 2 dipolos ativos, em posições aleatórias, separados por uma distância entre 55 e 65 mm. A intensidade de cada dipolo foi espalhada para os primeiros vizinhos, com intensidade 1/3 para os 3 vizinhos que compartilham uma aresta e 1/9 para os 9 vizinhos que compartilham um vértice. Ainda, cada simulação tem T = 51 ms, sendo que a evolução temporal segue sin $\frac{t\pi}{T}$, atingindo o máximo em 26 ms, e diminuindo lentamente até sumir em 51 ms. A análise das correntes é feita no instante 26 ms.

Para verificar a qualidade dos resultados, criamos um parâmetro de ordem ρ que compara a corrente estimada e a corrente real. Se considerarmos as correntes como vetores de 10240 dimensões (para a superfície de ordem 5 com 10240 faces que estamos utilizando), este parâmetro é o logaritmo do cosseno entre o valor absoluto (Norma L^1) da corrente real J_{real} e o valor absoluto da corrente estimada J_{ψ} , onde ψ é o algoritmo testado, que pode ser VB (Bayes Variacional Original), MG (Bayes Variacional Multi Grid) ou MNE (Estimativa de Norma Mínima)

$$\rho^{\psi} = \ln \frac{|J_{real}| \cdot |J_{\psi}|}{\|J_{real}\|^2 \|J_{\psi}\|^2}.$$
(3.1)

Porém este parâmetro compara apenas os algoritmos face a face, ou seja, se o sinal estimado estiver na face ao lado de onde estava o sinal real, e o sinal real for estritamente zero nesta face, o valor somado à ρ^{ψ} para esta face no algoritmo testado será zero. Isso não é bom, pois acetar a face vizinha ao foco de corrente é uma boa estimativa. Por isso introduzimos um novo grau de liberdade s, e calculamos o ρ^{ψ} para as médias de cada face com os seus vizinhos mais próximos. Assim

$$\rho_s^{\psi} = \ln \frac{|J_{real}| \cdot \langle |J_{\psi}| \rangle_s}{\|J_{real}\|^2\| \langle |J_{\psi}| \rangle_s \|^2},\tag{3.2}$$

onde $\langle \rangle_s$ significa que cada face do vetor $\langle |J_{\psi}| \rangle_s$ é a média desta face com os : s = 1primeiros vizinhos, s = 2 primeiros e segundos vizinhos, e assim sucessivamente.

O parâmetro ρ_s^{ψ} foi estimado para cada simulação, e os gráficos mostram a média deste parâmetro entre todas as simulações. Ainda, a área média para os *s* primeiros vizinhos foi estimada e é apresentada no eixo das abcissas no lugar de *s*. Assim é possivel saber a área sobre qual a média foi feita.

O algoritmo MNE usa apenas a Equação 1.3.3 para estimar a corrente, sem atribuir um prior para o parâmetro α e tentar estimar um α ótimo usando inferência Bayesiana. O α usado no algoritmo MNE é o mesmo α_{0n} usando no algoritmo VB e na primeira superfície do algoritmo MG, ou seja, 1.

3.1 Diferentes superfícies iniciais

Primeiro analisamos o comportamento do algoritmo MG partindo de diferentes condições iniciais. Fizemos 100 simulações como descritas acima utilizando como primeira superfície inicial para o algoritmo MG a superfície de ordem 1. Depois, repetimos as mesmas 100 simulações porém começando o algoritmo MG da superfície de ordem 2. Isso equivale a não realizar a iteração na superfície 1 (a figura 2.19), mas na a superfície 2 (lado esquerdo da figura 2.20) com as mesmas constantes usadas no algoritmo VB ($\alpha_{0n} = 1 \ e \ \gamma_{0n\alpha} = 0.1$). A seguir, realizamos o mesmo partindo da superfície 3, e em seguida apenas da superfície 4. Assim verificamos de que forma o comportamento do algoritmo MG se aproxima do algoritmo original VB, que usa como condição inicial um valor contante para todos α_n (como pode ser visto na figura 2.18).

Neste primeira simulação, todos os sinais foram corrompidos com ruído com Noise-to-Signal Ratio (NSR) igual a 0.1.

Além do parâmetro ρ_s^{ψ} para $s \in [1\,2\,3\,4]$, geramos também dois histogramas: um com a distância entre o dipolo mais intenso estimado (VB, MG e MNE) e o dipolo mais intenso que gerou o sinal (Real) e outro com a distância entre o segundo dipolo mais intenso estimado (VB, MG e MNE) e o segundo dipolo mais intenso que gerou o sinal (Real).



Figura 3.1: Comparação dos ângulos entre o vetor corrente real e as estimativas dos diferentes algoritmos : Bayes Variacional (VB), Norma Mínima (MNE), Multigrid começando da superfície de ordem 1 (MG 1), ordem 2 (MG 2), ordem 3 (MG 3) e ordem 4 (MG 4).

Podemos ver que ρ assume o seu valor mais próximo da corrente real quando iniciamos o algoritmo MG na superfície 3. Este valor é ligeiramente melhor que começando na superfície 4, e sempre melhor que o algoritmo VB original. Como vemos a seguir, este resultado se repete quando analisamos o histograma com a distância entre os dipolos mais fortes na corrente real e na corrente estimada, sendo que número de zeros (significando que a estimativa acertou a face onde o sinal original era mais forte) é ligeiramente melhor quando iniciamos a partir da superfície 3, e sempre melhor que o VB.



Figura 3.2: Histograma com as distâncias entre a localização do dipolo mais forte estimado e o dipolo mais forte real para o algoritmo começando da superfície 1 (MG 1) e da superfície 2 (MG 2).



Distância entre dipolos mais fortes : início na superfície 3

Figura 3.3: Histograma com as distâncias entre a localização do dipolo mais forte estimado e o dipolo mais forte real para o algoritmo começando da superfície 3 (MG 3) e da superfície 3 (MG 3).

Temos também um pico em torno de 60 mm, já que todas as simulações possuem dois focos com corrente, sempre com 55 a 65 mm de distância um do outro. Podemos ver por outro lado nos histogramas dos segundos dipolos mais intensos que o mínimo de erro ocorre quando inicializamos o algoritmo MG da superfície 1



Figura 3.4: Histograma com as distâncias entre a localização do segundo dipolo mais forte estimado e o segundo dipolo mais forte real para o algoritmo começando da superfície 1 (MG 1) e da superfície 2 (MG 2).



Figura 3.5: Histograma com as distâncias entre a localização do segundo dipolo mais forte estimado e o segundo dipolo mais forte real para o algoritmo começando da superfície 3 (MG 3) e da superfície 4 (MG 4).

3.2 Diferentes intensidades de ruído

Nesta simulação corrompemos os dados com diferentes intensidades de ruído e verificamos o comportamento dos algoritmos. Primeiro escolhemos 5 simulações onde ambos os algoritmos se comportaram bem, acertando a localização de ambas as fontes com boa precisão quando o sinal foi corrompido com ruído pequeno (NSR 0.1). O ruído, como descrito na seção 1.3.1, é gaussiano. Assim a média do ruído é nula e a variância é estimada de forma que o NSR seja igual à um valor desejado. Para calcular o NSR nós usamos

$$NSR = \frac{M \sigma_n^2}{V'V},\tag{3.3}$$

onde M é o número de sensores, V é o sinal e σ_n^2 é a variância do ruído somado ao sinal.

Em seguida, aumentamos o ruído mantendo as mesmas simulações, começando com NSR = 0.1 e subindo até NSR = 2.1. Para analisar o resultado fizemos um gráfico da média de ρ_1^{ψ} e outro da média de ρ_2^{ψ} para estas 5 simulações em função do NSR figuras (3.6 e 3.7).

Além disso, no lugar do histograma feito na seção anterior fizemos apenas a média da distância em mm da face com a corrente real mais forte e a face com a corrente estimada mais forte. E da mesma forma, fizemos um gráfico da distância em mm da face com a segunda corrente real mais forte para a face com a segunda corrente estimada mais forte (figuras 3.8 e 3.9).

Ainda, nestas novas simulações fizemos a média da distância dos Falso Positivos para os dois focos de corrente gerados. Para estimar os falsos positivos, seja N o número de faces ativas na corrente real. Todas as N correntes mais intensas estimadas foram consideradas como *ativas* e as demais como *inativas*. Estes gráficos estão nas figuras (3.10 e 3.11).

Por fim, ainda utilizando este conceito, estimamos uma curva ROC da sensibilidade

$$sensibilidade = \frac{Positivos Verdadeiros}{Positivos Verdadeiros + Falsos Negativos},$$
(3.4)

por 1 - especificidade, onde

$$especificidade = \frac{Negativos Verdadeiros}{Negativos Verdadeiros + Falsos Positivos}.$$
(3.5)

Estas imagens estão nas figuras 3.12 e 3.13.

Para verificar melhor as diferenças encontradas na simulação anterior, realizamos todas as simulações começando da superfície 1 e da superfície 3 no algoritmo MG, para comparar a performance das duas inicializações quando aumentamos o ruído. Podemos perceber que todas as
curvas (talvez com exceção do dipolo mais forte, que é equivalente) são melhores para o algoritmo começando da superfície de ordem 1, com a menor resolução disponível. Isso mostra que o algoritmo é mais resistente a ruído se usamos todas as superfícies disponíveis, principalmente com relação a localização dos dipolos menos intensos.

Não achamos necessário testar o algoritmo MNE nesta etapa, pois estamos mais interessados em comparar os algoritmos VB e MG.



Figura 3.6: Parâmetro ρ_1 para os algoritmos VB, MG 1 e MG 3. Podemos ver como o algoritmo começando da superfície com menor resolução possível é mais resistente ao aumento do ruído.



Figura 3.7: Parâmetro ρ_2 para os algoritmos VB, MG 1 e MG 3. Podemos ver como o algoritmo começando da superfície com menor resolução possível é mais resistente ao aumento do ruído.



Figura 3.8: A distância entre o dipolo mais intenso estimado pelo algoritmo VB e o dipolo mais intenso real é a curva azul (quadrados). A distância entre o dipolo mais intenso estimado pelos algoritmos MG 1 e MG 3 e o dipolo mais intenso real são as curvas vermelhas (círculos) com linha sólida e tracejada, respectivamente. Podemos ver que a distância aumenta de maneira compatível entre os dois algoritmos.



Figura 3.9: A distância entre o segundo dipolo mais intenso estimado pelo algoritmo VB e o segundo dipolo mais intenso real é a curva azul (quadrados). A distância entre o segundo dipolo mais intenso estimado pelos algoritmos MG 1 e MG 3 e o dipolo mais intenso real são as curvas vermelhas (círculos) com linha sólida e tracejada, respectivamente. Podemos ver que os algoritmos MG são constantemente melhores que o algoritmo VB original, enquanto o MG 1 é mais robusto a ruído que o MG 3.



Figura 3.10: Erro entre os falsos positivos e o primeiro foco de corrente: média da Distância Euclidiana em mm entre o máximo do primeiro foco de corrente e os falsos positivos. VB é o algoritmo Bayes Variacional original, MG 1 é o algoritmo Multigrid começando da superfície 1 e MG 3 é o algoritmo Multigrid começando da superfície 3. Podemos ver claramente como o algoritmo MG é mais robusto ao ruído.



Figura 3.11: Erro entre os falsos positivos e o segundo foco de corrente: média da Distância Euclidiana em mm entre os falsos positivos e o máximo do segundo foco de corrente. VB é o algoritmo Bayes Variacional original, MG 1 é o algoritmo Multigrid começando da superfície 1 e MG 3 é o algoritmo Multigrid começando da superfície 3. Podemos ver claramente como o algoritmo MG é mais robusto ao ruído.



Figura 3.12: Pontos no espaço ROC (Sensibilidade por 1 - Especificidade) para diferentes níveis de ruído. VB é o algoritmo Bayes Variacional original e MG 1 é o algoritmo Multigrid começando da superfície 1.



Figura 3.13: Pontos no espaço ROC (Sensibilidade por 1 - Especificidade) para diferentes níveis de ruído. VB é o algoritmo Bayes Variacional original e MG 3 é o algoritmo Multigrid começando da superfície 3.

Capítulo 4

Conclusão

Neste trabalho estudamos a estimativa da localização de fontes de M/EEG. O Problema Direto foi estudado por ser indispensável para qualquer trabalho de estimativas de fontes, destacando que ele pode ser resolvido analiticamente para aproximações esféricas da cabeça ou numericamente para modelos reais de cabeça que usam imagens de sMRI.

Estudamos também o modelo de *Norma Mínima* proposto por [Hämäläinen and Ilmoniemi, 1994], e a sua generalização usando idéias de campo médio proposto por [Sato et al., 2004].

Introduzimos um novo algoritmo que utiliza informação Multi-escala a priori para estimar as fontes de corrente chamado Multigrid ou MG, sendo que foi necessário desenvolver um algoritmo de Re-escala ou RE das superfícies corticais extraídas das imagens de sMRI, que não são apenas dizimações simples como normalmente é feito nos algoritmos convencionais.

O algoritmo MG se mostrou superior ao algoritmo VB original em geral. Em particular ressaltamos que é notadamente melhor na identificação de fontes profundas, i.e. localizadas em sulcos. Estudamos também o desempenho quando a informação foi corrompida com ruído. O algoritmo MG se mostrou robusto, tolerando melhor que o VB o aumento do NSR. Ainda, o algoritmo MG identificou a face onde a corrente real era mais forte como a face com a corrente mais forte o dobro de vezes que o algoritmo original VB durante as 100 simulações realizadas (como podemos ver na figura 3.2).

Pretendemos no futuro melhorar o algoritmo de RE incluindo um *simulated annealing* no passo 2 da figura 2.8 de modo a tentar minimizar a variância das áreas e reduzir ainda mais o viés introduzido pela escolha da posição inicial do icosaedro. Neste mesmo espírito, queremos impor um novo vínculo na convergência do algoritmo RE de forma a fixar uma região com caráter funcional conhecido em cada uma das faces da escala de menor resolução, semelhante ao trabalho de [Ou et al., 2010]

Queremos também estudar o comportamento do algoritmo MG quando removemos eletrodos, de modo a verificar a aplicação deste em análises clínicas com poucos eletrodos de baixa impedância distribuídos em uma pequena região da cabeça.

Por fim, gostaríamos de estender o trabalho de [Amaral et al., 2004] junto com este, fazendo uma abordagem Multigrid para a fusão de dados de fMRI com M/EEG.

Bibliografia

- [Amaral et al., 2004] Amaral, R., Rabbani, S. R., and Caticha, N. (2004). Multigrid priors for a Bayesian approach to fMRI. *NeuroImage*, 23:654 – 662.
- [Baillet et al., 2001] Baillet, S., Mosher, J. C., and Leahy, R. M. (2001). Electromagnetic Brain Mapping. *IEEE Signal Processing Magazine*, (November).
- [Bear et al., 2006] Bear, M. F., Connors, B. W., and Paradiso, M. A. (2006). Neuroscience: Exploring the Brain, volume 5. Lippincott Williams & Wilkins.
- [Belliveau et al., 1991] Belliveau, J. W., Kennedy, D. N., McKinstry, R. C., Buchbinder, B. R., Weisskoff, R. M., Cohen, M. S., and Vevea, J. M. (1991). Functional Mapping of the Human Visual Cortex by Magnetic Resonance Imaging. *Science*, 254.
- [Berger, 1929] Berger, H. (1929). Uber das elektrenkephalogramm des menschen. (on the electroencephalogram in man). Arch. Psychiatr. Nervenkr., 87:527 – 543.
- [Bishop, 2007] Bishop, C. M. (2007). Pattern Recognition and Machine Learning (Information Science and Statistics). Springer, 1 edition.
- [Bossavit, 1988] Bossavit, A. (1988). Whitney Forms : a class of finite elements for threedimensional computations in electromagnetism. *IEE Proceedings*, 135.
- [Cohen, 1968] Cohen, D. (1968). Magnetoencephalography: evidence of magnetic fields produced by alpha-rhythm currents. *Science*, 161(843):784–6.
- [Cuffin, 1995] Cuffin, B. N. (1995). A method for localizing EEG sources in realistic head models. IEEE transactions on bio-medical engineering, 42(1):68–71.

- [Dale and Sereno, 1993] Dale, A. M. and Sereno, M. I. (1993). Improved Localization of Cortical Activity by Combining EEG and MEG with MRI Cortical Surface Reconstruction: A Linear Approach. Journal of Cognitive Neuroscience, 76.
- [Fischl et al., 1999a] Fischl, B., Sereno, M. I., and Dale, A. M. (1999a). Cortical surface-based analysis. I. Segmentation and Surface Reconstruction. *NeuroImage*, 9(2):195–207.
- [Fischl et al., 1999b] Fischl, B., Sereno, M. I., and Dale, a. M. (1999b). Cortical surfacebased analysis. II: Inflation, flattening, and a surface-based coordinate system. *NeuroImage*, 9(2):195–207.
- [Friston et al., 2008] Friston, K., Harrison, L., Daunizeau, J., Kiebel, S., Phillips, C., Trujillobarreto, N., Henson, R., Flandin, G., and Mattout, J. (2008). Multiple sparse priors for the M / EEG inverse problem. *NeuroImage*, 39:1104 – 1120.
- [Friston, 2009] Friston, K. J. (2009). Modalities, modes, and models in functional neuroimaging. Science (New York, N.Y.), 326(5951):399–403.
- [Geselowitz, 1967] Geselowitz, D. (1967). On Bioelectric Potentials in an Inhomogeneous Volume Conductor. *Biophysical Journal*, 7(1):1–11.
- [Gramfort et al., 2010] Gramfort, A., Papadopoulo, T., Olivi, E., and Clerc, M. (2010). Open-MEEG: opensource software for quasistatic bioelectromagnetics. *Biomedical engineering online*, 9:45.
- [Hämäläinen and Ilmoniemi, 1994] Hämäläinen, M. and Ilmoniemi, R. (1994). Interpreting magnetic fields of the brain: minimum norm estimates. *Medical and Biological Engineering and Computing*, 32:35–42. 10.1007/BF02512476.
- [Jaynes, 1994] Jaynes, E. T. (1994). Probability Theory.
- [Kim et al., 1997] Kim, S. G., Richter, W., and UAurbil, K. (1997). Limitations of temporal resolution in functional MRI. *Magnetic resonance in medicine*, 37(4):631–636.
- [Lin et al., 2006] Lin, F.-H., Belliveau, J. W., Dale, A. M., and Hämäläinen, M. S. (2006). Distributed current estimates using cortical orientation constraints. *Human brain mapping*, 27(1):1–13.

[Mann,] Mann, M. D. http://www.unmc.edu/physiology/Mann/mann13.html.

- [Mosher and Leahy, 1998] Mosher, J. C. and Leahy, R. M. (1998). Recursive MUSIC: a framework for EEG and MEG source localization. *IEEE transactions on bio-medical engineering*, 45(11):1342–54.
- [Mosher et al., 1999] Mosher, J. C., Leahy, R. M., and Lewis, P. S. (1999). EEG and MEG: forward solutions for inverse methods. *IEEE transactions on bio-medical engineering*, 46(3):245–59.
- [Mosher et al., 1992] Mosher, J. C., Lewis, P. S., and Leahy, R. M. (1992). Multiple dipole modeling and localization from spatio-temporal MEG data. *IEEE transactions on bio-medical* engineering, 39(6):541–57.
- [Neal, 1997] Neal, R. M. (1997). Bayesian Learning for Neural Networks (Lecture Notes in Statistical Vol. 118). Journal of the American Statistical Association, 92(438):791.
- [Nummenmaa, 2008] Nummenmaa, A. (2008). *Hierarchical bayesian aspects of distributed neu*romagnetic source models. PhD thesis, Helsinki University of Technology.
- [Nunez and Srinivasan, 2006] Nunez, P. L. and Srinivasan, R. (2006). Electric Fields of the Brain: The Neurophysics of EEG, 2nd Edition. Oxford University Press.
- [Okada, 1981] Okada, Y. (1981). Neurogenesis of evoked magnetic fields, pages 399–408. de Gruyter, Berlin.
- [Oostendorp and van Oosterom, 1989] Oostendorp, T. F. and van Oosterom, a. (1989). Source parameter estimation in inhomogeneous volume conductors of arbitrary shape. *IEEE transactions on bio-medical engineering*, 36(3):382–91.
- [Opper and Winther, 2001] Opper, M. and Winther, O. (2001). From Naive Mean Field Theory to the TAP Equations. Advanced mean field methods theory and pratice, pages 1–19.
- [Ou et al., 2010] Ou, W., Nummenmaa, A., Ahveninen, J., Belliveau, J. W., Hämäläinen, M. S., and Golland, P. (2010). NeuroImage Multimodal functional imaging using fMRI-informed regional EEG / MEG source estimation. *NeuroImage*, 52(1):97–108.
- [Parisi, 1988] Parisi, G. (1988). Statistical Field Theory. Addison Wesley, 1 edition.

- [Pascual-marqui, 1999] Pascual-marqui, R. D. (1999). Review of Methods for Solving the EEG Inverse Problem. INTERNATIONAL JOURNAL OF BIOELECTROMAGNETISM, 1(1):75– 86.
- [Phillips, 2001] Phillips, C. (2001). Source estimation in EEG. Docteur en Sciences Appliquees.
- [Plonsey, 1969] Plonsey, R. (1969). Bioelectric Phenomena. McGraw-Hill.
- [Rockel et al., 1980] Rockel, A. J., Hiorns, R. W., and Powell, T. P. (1980). The basic uniformity in structure of the neocortex. *Brain: A journal of neurology*, 103(2):221–244.
- [Sato and Yoshioka, 2008] Sato, M.-a. and Yoshioka, T. (2008). Hierarchical Bayesian Inference of Brain Activity. *Lecture Notes in Computer Science*, 4984:576–585.
- [Sato et al., 2004] Sato, M.-a., Yoshioka, T., Kajihara, S., Toyama, K., Goda, N., Doya, K., and Kawato, M. (2004). Hierarchical Bayesian estimation for MEG inverse problem. *Science And Technology*, 23:806 – 826.
- [Schwartz et al., 1989] Schwartz, E., Shaw, a., and Wolfson, E. (1989). A numerical solution to the generalized mapmaker's problem: flattening nonconvex polyhedral surfaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(9):1005–1008.
- [Tanzer, 2006] Tanzer, I. O. (2006). Numerical Modeling in Electro- and Magnetoencephalography. PhD thesis, Helsinki University of Technology.
- [Toro et al., 2008] Toro, R., Perron, M., Pike, B., Richer, L., Veillette, S., Pausova, Z., and Paus, T. (2008). Brain size and folding of the human cerebral cortex. *Cerebral cortex (New York,* N.Y. : 1991), 18(10):2352–7.
- [Wikimedia, a] Wikimedia. http://en.wikipedia.org/wiki/File:Synapse_Illustration_unlabeled.svg.
- [Wikimedia, b] Wikimedia. http://en.wikipedia.org/wiki/File:Neuronno_labels.png.
- [Wikswo, 1983] Wikswo, J. P. (1983). Cellular action currents / edited by samuel j. williamson ... [et al.]. Biomagnetism, An Interdisciplinary Approach, pages 173–207.
- [Wipf and Nagarajan, 2009] Wipf, D. and Nagarajan, S. (2009). A unified Bayesian framework for MEG / EEG source imaging. *NeuroImage*, 44(3):947–966.

[Zhang, 1995] Zhang, Z. (1995). A fast method to compute surface potentials generated by dipoles within multilayer anisotropic spheres. *Physics in medicine and biology*, 40(3):335–49.

Arquivos em Anexo

Encaminhamos em anexo um CD com o resultado de duas simulações. Ele contém figuras do Matlab que permitem uma visualização 3D de imagens semelhantes àquelas apresentadas no final do capítulo 2, em duas simulações diferentes: Uma com fontes mais superficiais onde ambos os algorítmos obtiveram resultados satisfatórios e outra com fontes mais internas onde o algorítmo MG foi capaz de identificar ambos os focos de corrente, ao contrário do algorítmo original VB que identificou apenas 1. Este comportamento foi observado diversas vezes nas dezenas de simulações que realizamos e por isso estas simulações foram escolhidas. O CD possui a seguinte estrutura de diretórios:

```
Matlab

Corrente em Fontes Mais Internas

Inflada

Corrente em Fontes Mais Superficiais

Inflada

Original

Videos

Corrente em Fontes Mais Internas

Corrente em Fontes Mais Superficiais
```

Os diretórios *Inflada* e *Original* contém as figuras do Matlab com as simulações representadas na superfície Inflada e Original, respectivamente. Cada diretório contém os seguintes arquivos:

- 1. Real.img : Contém as faces com a corrente real
- 2. MG.img : Contém a corrente estimada pelo método Multigrid ou MG
- 3. VB.img : Contém a corrente estimada pelo método Bayes Variacional ou VB
- 4. MNE.img : Contém a corrente estimada pelo método de Norma Mínima ou MNE
- 5. VarMG[n][e].img: Contém a variância ou inverso de α para cada face de cada superfície [n] do algorítmo MG, onde [n] = 1, 2, 3, 4 e 5. [e] = I é variância inicial naquela superfície e [e] = F é a variância quando o algorítmo convergiu.
- VarVB5[e].img: Contém a variância ou inverso de α para cada face do algorítmo VB, no início ([e] = I) e no final ([e] = F) do algorítmo.

cada uma das 4 imagens que representam a corrente estimada e real tem os seguintes controles:



Figura 1: Controles nas figuras do Matlab.

Para abrir as imagens : basta abrir o Matlab e mudar o diretório corrente para o diretório "Matlab" contido no CD, ou adiciona-lo no Path. No linux, em geral:

cd /media/TeseMG/Matlab/

ou

```
path(path, '/media/TeseMG/Matlab/');
```

Depois disso, qualquer imagem funcionará normalmente, bastando abri-la pelo menu do Matlab ou usando o comando open. Por exemplo

open('/media/TeseMG/Matlab/Corrente em Fontes Mais Internas/Original/MG.fig');

Além das imagens do matlab foram gerados filmes com as superfícies originais de ambas as simulações, caso o Matlab não esteja disponível. Neste caso os controles acima não estão disponíveis e as variâncias iniciais e finais estão unidas em um mesmo filme:

- 1. Real.avi : Contém as faces com a corrente real
- 2. MG.avi : Contém a corrente estimada pelo método Multigrid ou MG
- 3. VB.avi : Contém a corrente estimada pelo método Bayes Variacional ou VB
- 4. MNE.avi : Contém a corrente estimada pelo método de Norma Mínima ou MNE
- 5. VarMG[n].avi: Contém a variância ou inverso de α para cada face de cada superfície [n] do algorítmo MG, onde [n] = 1, 2, 3, 4e5.
- 6. VarVB5.avi: Contém a variância ou inverso de α para cada face do algorítmo VB.