

UNIVERSIDADE DE SÃO PAULO  
FFCLRP - DEPARTAMENTO DE FÍSICA E MATEMÁTICA  
PROGRAMA DE PÓS-GRADUAÇÃO EM FÍSICA APLICADA À  
MEDICINA E BIOLOGIA

Detecção e caracterização da resposta hemodinâmica pelo desenvolvimento de novos métodos de processamento de imagens funcionais por ressonância magnética

Marcio Junior Sturzbecher

Dissertação apresentada à Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto da USP, como parte das exigências para a obtenção do título de Mestre em Ciências, Área: Física Aplicada à Medicina e Biologia.

RIBEIRÃO PRETO – SP  
- 2006 –

# **Livros Grátis**

<http://www.livrosgratis.com.br>

Milhares de livros grátis para download.

UNIVERSIDADE DE SÃO PAULO  
FFCLRP - DEPARTAMENTO DE FÍSICA E MATEMÁTICA  
PROGRAMA DE PÓS-GRADUAÇÃO EM FÍSICA APLICADA À  
MEDICINA E BIOLOGIA

**Detecção e caracterização da resposta hemodinâmica pelo  
desenvolvimento de novos métodos de processamento de  
Imagens Funcionais por Ressonância Magnética**

Dissertação apresentada à Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto da USP, como parte das exigências para a obtenção do título de Mestre em Ciências, Área: Física Aplicada à Medicina e Biologia.

Aluno: Marcio Junior Sturzbecher

Orientador: Prof. Dr. Dráulio Barros de Araújo

RIBEIRÃO PRETO – SP

- 2006 -

## **FICHA CATALOGRÁFICA**

Preparada pela Biblioteca Central  
do Campus Administrativo  
de Ribeirão Preto

**Sturzbecher, Marcio Junior.**

Detecção e caracterização da resposta hemodinâmica pelo desenvolvimento de novos métodos de processamento de Imagens Funcionais por Ressonância Magnética, 2006, 139p.

Dissertação de Mestrado apresentada à Faculdade de Filosofia Ciências e Letras de Ribeirão Preto da Universidade de São Paulo – Área de Concentração: Física Aplicada a Medicina e Biologia.

Orientador: de Araújo, Dráulio Barros.

1. fMRI; 2. Teoria da Informação; 3. Atraso da HRF.

*Dedico esta obra a Ferzita, minha inspiração, por sua compreensão, seu amor, seu carinho, sua presença, sua confiança e incansável apoio durante a elaboração deste trabalho.*

## AGRADECIMENTOS

Ao Prof. Dr. Dráulio Barros de Araújo pela orientação e amizade, que, nos anos de convivência, muito me ensinou, contribuindo para meu crescimento científico e intelectual.

Ao Dr. Walfred Tedeschi pelas discussões, amizade, atenção e apoio científico durante o desenvolvimento do trabalho.

Ao Prof. Dr. Uiraci Pereira da Costa Neves pelas discussões teóricas pertinentes.

Ao Prof. Dr. Bernardo Mello pelo apoio científico depositado neste trabalho.

Aos voluntários que dispuseram de seu precioso tempo para participarem desta pesquisa.

A todos os amigos do Departamento de Física e Matemática que proporcionaram momentos agradáveis e produtivos.

A todos os amigos que conheci em Ribeirão Preto, com os quais vivi momentos de grande crescimento pessoal.

Ao Tiago pela longa amizade compartilhada e pelo apoio que tem me dado.

A FFCLRP, pela oportunidade de realização do curso de mestrado e ao Departamento de Física e Matemática por colocar a disposição o laboratório.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pela concessão da bolsa de mestrado e pelo apoio financeiro para a realização desta pesquisa.

A minha família de Mococa: Dito, Nezita, Nana, Dé, Flá, Luander, que tem me acolhido com muito carinho e proporcionado um novo lar.

E a minha família: meu pai Érico (*In Memoriam*), minha mãe Armide, meu irmão Julio e meu irmão Carlos e família e a Fernanda pelo apoio, pelo incentivo e pelo amor.

*Quero apossar-me do é da coisa. Esses instantes que decorrem no ar que  
respiro: em fogos de artifício eles espocam mudos no espaço.*

**Água Viva de Clarice Lispector**

## RESUMO

Os estímulos evento-relacionados (*event-related functional magnetic resonance imaging* - ER-fMRI) têm se tornado mais popular nos últimos anos. A maioria das análises das séries temporais geradas nesses exames é baseada em um modelo específico da função resposta hemodinâmica, *hemodynamic response function* (HRF). Contudo, a considerável variabilidade da HRF e a baixa relação sinal ruído das imagens, tem dificultado sua caracterização e localização. Neste sentido, recentemente foi proposto um método baseado na computação da entropia de Shannon, que não faz suposição sobre a forma da resposta. Desse modo, no presente trabalho, buscou-se, em um primeiro momento, avaliar alguns parâmetros do método proposto a fim de aumentar sua sensibilidade. Além disso, generalizou-se o método com a entropia de Tsallis, que introduz um novo parâmetro,  $q$ . Foi proposto, também, um novo método baseado no cálculo da distância de Kullback-Leibler (KL). Todos os métodos foram aplicados a dados simulados e dados reais. Curvas ROC (*Receiver Operating Characteristic*) foram usadas para estudar a dependência dos valores de corte e os diferentes parâmetros dos métodos em relação à sensibilidade e especificidade. A entropia de Tsallis apresentou maior poder de detecção nos intervalos de  $-1 < q < 1$  ( $q \neq 0$ ) com dois níveis. O método de KL mostrou uma melhora significativa da sensibilidade em dados simulados e apresentou a habilidade em discriminar regiões funcionais ativas e em repouso, para um estímulo motor. Por fim, foi proposta uma primeira tentativa de identificar simultaneamente duas regiões cerebrais não correlacionadas no tempo e no espaço com os métodos de Tsallis e KL-fMRI em um experimento de estímulo duplo (visual e motor). Além disso, um novo método mostrou-se eficiente, ainda que de forma preliminar, para detectar o atraso temporal de duas HRFs.



## ABSTRACT

Event-related functional magnetic resonance imaging (ER-fMRI) has become more popular in the last years. Most of analysis applied to such exams is based on a specific model of the hemodynamic response function (HRF). However, the variability of HRF and the low signal-to-noise ratio of such exams have imposed interesting challenges to both characterization and localization of the HRF. Recently, it was proposed a method based on computation of Shannon entropy, which makes no assumption about the shape of the response. We herein evaluated some parameters of the previous method in order to increase its sensibility. Moreover, this method was generalized with Tsallis entropy, which introduces a new parameter,  $q$ . Besides, a new method was proposed based on a distance of Kullback-Leibler (KL) measurement. Every method was applied to simulated and real data. Receiver Operating Characteristic curves were used to study the threshold dependence and the different methods parameters related to sensitivity and specificity. Tsallis entropy showed high detection power in the intervals  $-1 < q < 1$  ( $q \neq 0$ ) with two bins. The KL method allowed significant improvement of sensibility in simulated data and reviewed the ability to discriminate regions of functional activity and rest in a motor stimulus. Finally, it was proposed a new technique to identify two brain regions without correlation in time and space through Tsallis and KL-fMRI methods, in a experiment of double stimuli (visual and motor). Furthermore, in a preliminary study, we developed a strategy to detect temporal delay between two HRFs.

## LISTA DE ABREVIATURAS

ANOVA	Acrônimo do idioma inglês que representa: <i>Analysis of Variance</i>
ATP	Acrônimo do idioma inglês que representa: <i>Adenosin Triphosphate</i>
BOLD	Acrônimo do idioma inglês que representa: <i>Blood Oxygen Level Dependent</i>
CBF	Acrônimo do idioma inglês que representa: <i>Cerebral Blood Flux</i>
CBV	Acrônimo do idioma inglês que representa: <i>Cerebral Blood Volume</i>
CC	Coefficiente de Correlação
CMRO <sub>2</sub>	Acrônimo do idioma inglês que representa: <i>cerebral metabolic rate of oxygen</i>
CT	Acrônimo do idioma inglês que representa: <i>Computed Tomography</i>
dHb	Hemoglobina deoxigenada
EEG	Acrônimo do idioma inglês que representa: <i>electroencefalography</i>
EPI	Acrônimo do idioma inglês que representa: <i>Eco-Planar Imaging</i>
ER-fMRI	Acrônimo do idioma inglês que representa: <i>event-related functional MRI</i>
FA	Acrônimo do idioma inglês que representa: <i>Flip Angle</i>
FID	Acrônimo do idioma inglês que representa: <i>Free Induction Decay</i>
fMRI	Acrônimo do idioma inglês que representa: <i>Functional Magnetic Resonance Imaging</i>

FN	Falso Negativo
FNF	Acrônimo do idioma inglês que representa: <i>False Negative Fraction</i>
FP	Falso Positivo
FPF	Acrônimo do idioma inglês que representa: <i>False Positive Fraction</i>
Gd-DTPA	Acrônimo do idioma inglês que representa: <i>Gadolinium-diethylenetriamine pentaacetic acid</i>
GLM	Acrônimo do idioma inglês que representa: <i>General Linear Model</i>
Hb	Hemoglobina oxigenada
HRF	Acrônimo do idioma inglês que representa: <i>hemodynamic response function</i>
ICA	Acrônimo do idioma inglês que representa: <i>Independent Component Analysis</i>
M1	Cortex Motor Primário
MRI	Acrônimo do idioma inglês que representa: <i>Magnetic Resonance Imaging</i>
NMR	Acrônimo do idioma inglês que representa: <i>Nuclear Magnetic Resonance</i>
PET	Acrônimo do idioma inglês que representa: <i>Positron Emission Tomography</i>
RF	Radiofrequência
ROC	Acrônimo do idioma inglês que representa: <i>Receiver Operating Characteristic</i>
SNR	Acrônimo do idioma inglês que representa: <i>Signal-to-noise ratio</i>
TE	Tempo de Eco

TNF	Acrônimo do idioma inglês que representa: <i>True Negative Fraction</i>
TPF	Acrônimo do idioma inglês que representa: <i>True Positive Fraction</i>
TR	Tempo de Repetição
V1	Cortex Visual Primário
VN	Verdadeiro Negativo
VP	Verdadeiro Positivo

# ÍNDICE GERAL

<b>INTRODUÇÃO .....</b>	<b>15</b>
<b>CAPÍTULO I - TEORIA DA INFORMAÇÃO .....</b>	<b>18</b>
1. INTRODUÇÃO .....	18
2. ENTROPIA DE SHANNON .....	19
3. ENTROPIA DE TSALLIS.....	21
4. DISTÂNCIA DE KULLBACK-LEIBLER .....	23
5. EVOLUÇÃO TEMPORAL DAS ENTROPIAS .....	25
<b>CAPÍTULO II – IMAGEM FUNCIONAL POR RESSONÂNCIA MAGNÉTICA.....</b>	<b>28</b>
1. INTRODUÇÃO.....	28
2. SEQÜÊNCIA DE PULSO EPI.....	28
3. ESTUDOS FUNCIONAIS USANDO AGENTES DE CONTRASTE.....	30
4. EFEITO BOLD .....	30
5. EXPERIMENTOS EM fMRI.....	34
6. PARADIGMAS EVENTO-RELACIONADO .....	35
7. PRÉ-PROCESSAMENTO DAS IMAGENS .....	36
7.1 Correção Temporal.....	37
7.2 Correção de Movimento .....	37
7.3 Filtro Temporal e Espacial .....	39
7.4 Remoção de Drift .....	40
8. PÓS-PROCESSAMENTO EM fMRI.....	40
8.1 Teste-t de Student .....	42
8.2 Correlação Cruzada.....	44
8.3 Análise de Variância.....	46

8.4	Modelo Linear Geral .....	48
9.	APRESENTAÇÃO DOS MAPAS ESTATÍSTICOS .....	50
10.	CRONOMETRIA MENTAL .....	51
<b>OBJETIVOS.....</b>		<b>56</b>
<b>CAPÍTULO III – MATERIAL E MÉTODOS.....</b>		<b>57</b>
1.	INTRODUÇÃO.....	57
2.	ENTROPIA DE SHANNON APLICADA À FMRI .....	57
3.	ENTROPIA DE TSALLIS APLICADA À FMRI.....	59
4.	DISTÂNCIA DE KULLBACK-LEIBLER APLICADA À FMRI .....	60
5.	MÉTODO DE DETECÇÃO DE ATRASO DA HRF .....	61
6.	DADOS SIMULADOS .....	64
7.	DADOS REAIS.....	65
7.1	Voluntários.....	65
7.2	Estímulos.....	66
7.3	Aquisição das Imagens .....	66
7.4	Análise dos Dados .....	67
<b>CAPÍTULO IV – RESULTADOS E DISCUSSÃO .....</b>		<b>68</b>
1.	INTRODUÇÃO.....	68
2.	ENTROPIA DE SHANNON E PARADIGMA ER-MOTOR .....	68
3.	ENTROPIA DE TSALLIS.....	71
4.	DISTÂNCIA DE KULLBACK-LEIBLER .....	74
5.	CURVAS ROC.....	75
6.	COMPORTAMENTO DOS MÉTODOS EM RELAÇÃO A SNR.....	80
7.	ANÁLISE DE DADOS REAIS .....	82
8.	ESTÍMULO DUPLO.....	90

<b>CONCLUSÃO E TRABALHOS FUTUROS.....</b>	<b>96</b>
1. CONCLUSÃO.....	96
2. TRABALHOS FUTUROS .....	97
<b>REFERÊNCIAS.....</b>	<b>101</b>
<b>APÊNDICE A - IMAGEM POR RESSONÂNCIA MAGNÉTICA.....</b>	<b>108</b>
<b>APÊNDICE B - CURVAS CARACTERÍSTICAS DE UM TESTE DE DIAGNÓSTICO.....</b>	<b>134</b>

---

## INTRODUÇÃO

As técnicas de neuroimagem funcional, como a tomografia por emissão de pósitrons, *Positron Emission Tomography* (PET), a tomografia por emissão de fóton único, *Single Photon Emission Computer Tomography* (SPECT), a eletroencefalografia (EEG), a magnetoencefalografia (MEG), e a imagem funcional por ressonância magnética, *Functional Magnetic Resonance Imaging* (fMRI) têm progredido sensivelmente nos últimos anos. Dentre essas técnicas, particularmente, a fMRI tem se tornado uma das principais técnicas no mapeamento das funções cerebrais, principalmente, pelo seu caráter não-invasivo e pela boa resolução espaço-temporal [OGAWA *et al.* 1990a,b; LOGOTHETIS, 2002]. Essa técnica tem sido utilizada de forma bastante consistente em processos cognitivos elevados, em aplicações clínicas (dentre as quais se destaca o mapeamento pré-cirúrgico), processos de reorganização cortical e, mais recentemente, estudos em cronometria mental. Há, entretanto, várias questões que ainda exigem refinamentos, em especial no que diz respeito aos métodos de análises dessas imagens.

A aquisição das imagens utiliza técnicas de amostragem rápida, essenciais para captar o sinal com uma resolução temporal relativamente baixa. Contudo, essas técnicas fornecem baixa relação sinal ruído, o que, somado às outras fontes de ruído, dificulta a inspeção visual direta das áreas de atividade cerebral, tornando necessária a utilização de métodos computacionais específicos. Em geral, os métodos clássicos de análise, como, por exemplo, o GLM, modelam a função de resposta hemodinâmica, *Hemodynamic Response Function* (HRF). Porém, diante da variabilidade dessas funções e das dificuldades em se estabelecer o que é estatisticamente significativo na detecção das regiões funcionalmente ativas do cérebro, ainda não há um consenso da comunidade de fMRI sobre qual método utilizar. Nesse sentido, o desenvolvimento de novos métodos de análise soma uma importante contribuição na melhor caracterização desses sinais.

Recentemente, técnicas baseadas na Teoria de Informação mostraram-se consistentes em experiências com fMRI. A análise entrópica fundamenta-se no cálculo da entropia de Shannon dependente do tempo [DE ARAUJO *et al.*, 2003] e não faz nenhuma suposição sobre a forma da resposta he-



modinâmica. Contudo, aspectos de otimização do método não foram inspecionados em detalhes, a fim de avaliar as condições para sua maior sensibilidade. Desse modo, um dos objetivos do presente trabalho é realizar o estudo de alguns parâmetros, procurando identificar os valores apropriados que apresentam maior detecção da HRF.

Além disso, um segundo objetivo vem da generalização desse método, pela utilização de uma nova definição de entropia, proposta por Constantino Tsallis. Essa formulação introduz um novo parâmetro,  $q$ , o que permite realizar um estudo em função dessa nova variável, com um possível maior poder de detecção.

Como terceiro objetivo, propõe-se um segundo método, baseado no cálculo da distância de Kullback-Leibler (KL) das séries temporais de fMRI.

Por fim, os mapas tradicionais de ativação funcional não fornecem informações sobre a dinâmica das conexões entre diferentes regiões cerebrais, envolvidas em uma determinada tarefa. A viabilização desses estudos passa pela necessidade de medidas precisas de propriedades temporais específicas da HRF, em diferentes regiões do cérebro, dentre as quais, o atraso da HRF. Em um primeiro momento, propõe-se uma tentativa de identificar simultaneamente duas regiões cerebrais não correlacionadas no tempo e no espaço com os métodos de Tsallis e KL-fMRI, em um experimento de estímulo duplo: visual e motor. Em um segundo momento, iniciou-se o desenvolvimento de um método de detecção de atraso temporal.

O material apresentado nesta dissertação abrange diversos aspectos teóricos e experimentais. O primeiro capítulo introduz os conceitos fundamentais da teoria da informação utilizados nos métodos de detecção, com ênfase nos conceitos de entropia na sua forma clássica e generalizada, bem como a medida de distância entre duas distribuições de probabilidade.

O capítulo dois é dedicado à técnica de fMRI, em que se descreve aspectos teóricos das aquisições das imagens, do contraste BOLD (*Blood Oxygen Level Dependent*), dos processos envolvidos no pré

---

e pós-processamentos das imagens com métodos clássicos, além de uma descrição da aplicação da f-MRI em estudos de cronometria mental.

Os três últimos capítulos foram destinados à descrição dos métodos experimentais, aos resultados e discussão e, por fim, às conclusões e perspectivas obtidas neste trabalho.

Os apêndices, em um total de dois, dão ao leitor a oportunidade de se aprofundar em alguns aspectos teóricos utilizados ao longo da dissertação. Princípios físicos da Imagem por Ressonância Magnética e princípios da formação de imagem são introduzidos no primeiro apêndice. Os fundamentos teóricos para a construção das curvas ROC (*Receiver Operating Characteristic*) estão apresentados no segundo apêndice.

---

## Capítulo I - Teoria da Informação

### 1. Introdução

A Teoria da Informação (TI) tem sua origem na teoria de probabilidades introduzida por Claude Shannon (1916-2001) com a publicação do artigo “*The Mathematical Theory of Communication*” na revista *Bell System Technical Journal* [SHANNON, 1948]. O objetivo de Shannon era descobrir as leis que regulavam a capacidade de sistemas transmitirem, armazenarem e processarem a informação. Ainda, pretendia definir medidas quantitativas para esses processos. Uma das principais contribuições do modelo foi considerar a comunicação como um problema matemático rigorosamente embasado na estatística. Shannon propôs uma forma de medir a informação contida em um evento probabilístico, baseada na tradicional expressão da entropia de Boltzmann (1896), associando à entropia uma medida de informação.

Em seu modelo de comunicação (fonte-canal-receptor), Shannon estabeleceu um esquema de transmissão de informação. Uma mensagem que parte de uma fonte é codificada e emitida por um transmissor, passa por um canal de comunicação, sofre perturbações por ruídos, e chega ao receptor através de um sistema de decodificação. Nesse modelo, a quantidade de informação transmitida em uma mensagem é função da previsibilidade da mensagem, ou seja, independentemente da mensagem, a quantidade de informação está relacionada à possibilidade da mensagem ocorrer. Se essa probabilidade for pequena, a mensagem contém muita informação, caso contrário, se ela for previsível, então conterá pouca informação. Para medir a quantidade de informação, Shannon criou o conceito de entropia, algo diferente do conceito homônimo em termodinâmica, e definiu a quantidade de informação com base na incerteza, ou dificuldade de previsão dessa mensagem. Desse modo, a noção de entropia está ligada ao grau de desorganização existente na fonte de informação. Quanto maior a desordem, maior o potencial de informação dessa fonte. Uma fonte que responda com uma única e mesma mensagem a toda e qualquer pergunta, não transmite informação, já que não há redução de incerteza.

## 2. Entropia de Shannon

No contexto de Shannon a entropia está relacionada à incerteza de se obter uma informação e a capacidade informativa da fonte. De um modo geral, conceitos de entropia possibilitam uma comparação das propriedades de um sistema em termos numéricos, averiguando como é a sua distribuição de probabilidades.

**Definição:** Seja  $X$  uma variável discreta aleatória com vetor de probabilidade  $(p_1, \dots, p_n)$ . Define-se a entropia de Shannon da variável aleatória  $X$  como:

$$H(X) = -k \sum_{i=1}^n p_i(x) \log_2 \{p_i(x)\}, \quad \dots\dots\dots (1.1)$$

em que  $k$  é uma constante positiva. Vê-se que a entropia de Shannon depende somente da distribuição de  $X$ . Por convenção,  $x \log x$  é igual a zero quando  $x=0$ , dado que  $\lim_{x \rightarrow 0} x \log x = 0$ , para que se preserve a continuidade de  $H$ .

A base da função logaritmo determina a unidade de medida de informação. A rigor, ela pode ser qualquer número maior que 1, sendo usual a utilização da base 2 para sistemas digitais de informação. Ao utilizar essa base, a unidade resultante é a unidade binária (*bit*).

A grandeza  $H(x)$  pode ser interpretada como uma medida da incerteza sobre o valor da variável aleatória  $X$ , antes de observá-la, ou como a quantidade de informação sobre  $X$  que foi adquirida depois de realizar a medida. Em particular, se  $X$  tem algum valor com probabilidade 1 ( $p_i = 1$ ) então nenhuma informação é ganha após a observação de  $X$ . Neste caso, verifica-se, diretamente da definição 1.1, que  $H(X)$  é zero. Note que a entropia  $H(X)$  é sempre não-negativa:  $H(X) \geq 0$ .

A função entropia de Shannon pode ser representada graficamente, no caso de se terem dois resultados possíveis. Sabendo que as probabilidades  $p_1$  e  $p_2$  verificam a relação  $(p_2 = 1 - p_1)$ ,  $H(p_1, p_2)$ , ou seja,  $H(p_1, 1 - p_1)$ , tem a forma:

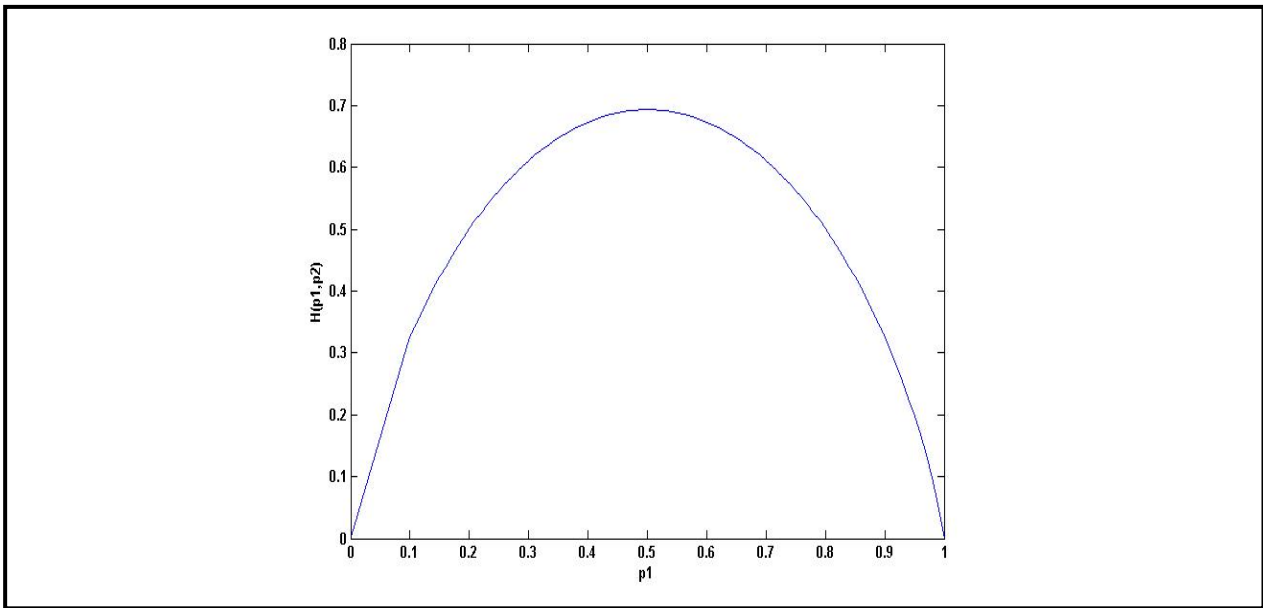


Figura 1.1. Gráfico da função entropia de Shannon  $H(p_1, p_2)$ .

Analisando a entropia de Shannon, verifica-se que ela atinge um máximo quando os estados estão igualmente distribuídos e um mínimo quando os estados têm uma distribuição concentrada. Para exemplificar essa questão, seja:

#### Caso 1: Distribuição Uniforme:

$$p(x) = \begin{cases} p(0) = 1/2 & \text{se } x = 0 \\ p(1) = 1/2 & \text{se } x = 1 \end{cases}.$$

Calculando-se a entropia de Shannon:

$$H(X) = -1/2 \log(1/2) - 1/2 \log(1/2) = -1 \log(1/2) = \log(2).$$

#### Caso 2: Distribuição Binária

$$p(x) = \begin{cases} p(0) = 1 & \text{se } x = 0 \\ p(1) = 1 & \text{se } x = 1 \end{cases},$$

A entropia de Shannon, segue:

$$H(X) = 0 - 1 \log_2(1) = 0.$$

Nesse exemplo, os limites da entropia de  $H(X)$  para as distribuições de probabilidade propostas, confirmam a noção intuitiva de que quanto mais espalhado o sistema se encontra, maior sua entropia e maior é a medida de informação.

Dentro da Teoria da Informação foram formuladas algumas propostas de generalização da entropia, sendo uma das mais conhecidas a de Rényi, de 1961. Ainda, recentemente, uma nova forma de entropia, conhecida por entropia de Tsallis [TSALLIS, 1988], tem atraído bastante interesse. Na sua definição, a entropia clássica de Boltzmann é generalizada, tornando-a conveniente quando aplicada para a caracterização de diferentes sistemas naturais [TEDESCHI *et al.*, 2005; HANSEN *et al.*, 2005; TORRES *et al.*, 2003].

### 3. Entropia de Tsallis

As entropias generalizadas, que podem ser ajustadas para uma medida de informação, são sustentadas pelo seu significado operacional ou por um conjunto de postulados que as caracterize. Em 1988, Constantino Tsallis [TSALLIS, 1988] publica o trabalho “*Possible Generalization of Boltzmann-Gibbs Statistics*” no *Journal of Statistical Physics*, propondo uma nova forma de entropia. Sua expressão fundamenta uma generalização da Mecânica Estatística usual – a Mecânica Estatística Não Extensiva. Ela estende o domínio da aplicabilidade dos procedimentos da Mecânica Estatística usual a sistemas em que a estatística de Boltzmann-Gibbs apresenta sérias dificuldades matemáticas. A expressão proposta, que define a entropia de Tsallis,  $S_q$ , é caracterizada por um parâmetro “ $q$ ”, e tem a forma:

$$S_q = k \cdot \frac{1 - \sum_{i=1}^W p_i^q}{q - 1}, \dots\dots\dots (1.2)$$

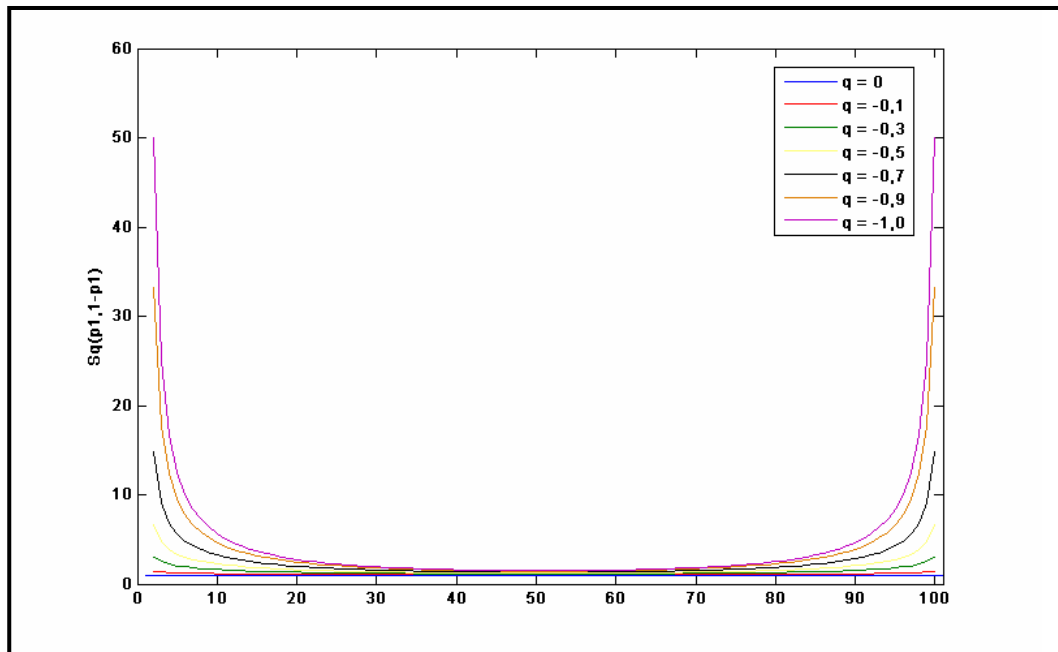
sendo  $k$  uma constante positiva que define a unidade em que a entropia é medida,  $q \in \mathfrak{R}$ , que caracteriza a estatística particular,  $W$  é o número total de microestados e  $p_i$  é o conjunto de probabilidades associado aos estados, sendo  $\sum_{i=1}^W p_i = 1$ . No limite em que  $q \rightarrow 1$ , a equação (1.2) remonta a expressão de Shannon, já que,

$$\lim_{n \rightarrow 0} \frac{x^n - 1}{n} = \ln x,$$

verifica-se que,

$$\begin{aligned} S_1(p_1, \dots, p_n) &= \lim_{q \rightarrow 1} S_q(p_1, \dots, p_n) = \lim_{q \rightarrow 1} k \frac{1 - \sum_i p_i^q}{q - 1} \\ S_1(p_1, \dots, p_n) &= \lim_{q \rightarrow 1} k \frac{\sum_i p_i - \sum_i p_i p_i^{q-1}}{q - 1} = \lim_{q \rightarrow 1} k \frac{\sum_i p_i (1 - p_i^{q-1})}{q - 1} \\ S_1(p_1, \dots, p_n) &= k \sum_i p_i \lim_{q \rightarrow 1} k \frac{1 - p_i^{q-1}}{q - 1} = k \sum_i p_i (-\ln p_i) \\ S_1(p_1, \dots, p_n) &= -k \sum_i p_i \ln p_i \end{aligned}$$

Na figura seguinte estão representados gráficos da entropia de Tsallis de uma distribuição com dois resultados possíveis  $S_q(p_1, p_2)$ , para alguns valores de  $q$ .



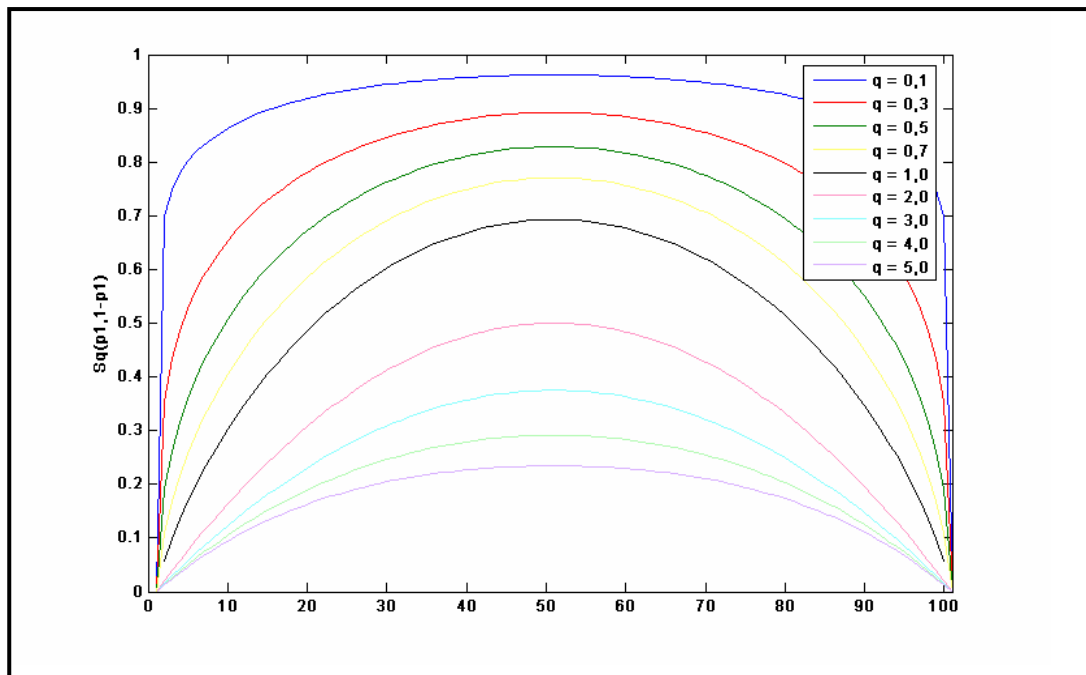


Figura 1.2. Gráficos de  $S_q(p_1, p_2)$ , em um sistema de dois estados com os valores de  $q$  indicados.

A entropia de Tsallis depende de “ $q$ ” e tem uma concavidade bem definida, sendo côncava quando  $q > 0$  e convexa quando  $q < 0$ .

Até agora a medida de informação foi estudada para uma única variável aleatória, a partir de definições de entropia. Contudo, se há mais de uma distribuição de probabilidades, pode-se, a princípio, medir diferenças relativas entre elas. Foi com esse intuito que Solomon Kullback e Richard Leibler em 1951 [COVER & THOMAS, 1991] introduziram o conceito de distância de Kullback-Leibler.

#### 4. Distância de Kullback-Leibler

A distância de Kullback-Leibler (KL), ou Entropia Relativa, Divergência I ou Entropia de Kullback-Leibler, consiste da medida de uma distância entre duas distribuições. A entropia relativa traduz o ganho de informação resultante da substituição da distribuição (de referência ou a *priori*) pela distribuição (a *posteriori*).

**Definição:** A distância KL entre duas distribuições de função probabilidade, discretas,  $p(x)$  e  $q(x)$  é definida por [COVER & THOMAS, 1991]:



$$D(p || q) = \sum_{i=1}^W p_i \log_2 \frac{p_i}{q_i} \dots\dots\dots (1.3)$$

Na definição acima, usamos  $0 \log 0/q = 0$  e  $p \log p/0 = \infty$  de acordo com o limite da função. A entropia relativa é sempre não negativa e é zero se e somente se  $p = q$ . Embora seja usual pensar na entropia relativa como uma medida de “distância” entre duas distribuições estatísticas, esta não pode ser considerada uma distância no sentido usual, uma vez que não apresenta simetria, ou seja,

$$\sum_{i=1}^W p_i \log_2 \frac{p_i}{q_i} \neq \sum_{i=1}^W p_i \log_2 \frac{q_i}{p_i}, \dots\dots\dots (1.4)$$

o que não pode ser considerada uma distância métrica.

À semelhança da generalização da entropia de Shannon, a distância KL também pode ser generalizada e associada a um parâmetro  $q$ . Em 1998, Borland e colaboradores [BORLAND *et al.*, 1998] definiram essa generalização, designada por  $q$ -informação de KL, através da informação própria<sup>1</sup> não extensiva de cada evento por:

$$I_i^q = - \frac{p_i^{1-q} - 1}{1-q} \dots\dots\dots (1.5)$$

Considerando as distribuições discretas de probabilidade  $p(x)$  e  $p'(x)$  de eventos medidos, a variação da informação é dada por:

$$\Delta I_i^q = \frac{1}{1-q} \left( \left( 1 - p_i'^{1-q} \right) - \left( 1 - p_i^{1-q} \right) \right) \dots\dots\dots (1.6)$$

A variação média de informação é obtida através da expressão:

$$D_q(p || p') = \sum_{i=1}^W p_i^q \frac{1}{1-q} \left( p_i^{1-q} - p_i'^{1-q} \right), \dots\dots\dots (1.7)$$

que representa o ganho de informação generalizado de Kullback-Leibler. Pode-se escrever esta equação de outra forma, mediante a definição de  $q$ -logaritmo:

$$\log_q(p) = \frac{p^{1-q} - 1}{1-q} \dots\dots\dots (1.8)$$

---

<sup>1</sup> Informação própria é o ganho de informação associada à ocorrência de um evento

Obtêm-se a forma:

$$D_q(p || p') = \sum_{i=1}^W p_i \log_q \frac{p_i}{p'_i} \dots\dots\dots (1.9)$$

A função generalizada da entropia relativa anula-se quando as distribuições  $p'(x)$  e  $p(x)$  são iguais ou quando  $q=0$ .

### 5. Evolução Temporal das Entropias

Levando em consideração que a entropia calculada sobre todo o sinal,  $s(t)$ , fornece pouca informação sobre sua evolução temporal, é interessante calcular uma entropia dependente do tempo. Para tanto, a entropia pode ser dividida em intervalos de tempo por meio de uma janela deslizante [CAPURRO *et al.*, 1999].

Seja  $s(t)$  a série temporal formada por um conjunto discreto de amplitudes amostradas no tempo (conjunto S):

$$S = \{s(t_k), k = 1, \dots, k\} \dots\dots\dots (1.10)$$

A janela deslizante W,

$$W(m; w, \Delta) = \{s(t_k), k = 1 + m\Delta, \dots, w + m\Delta\}, m = 0, 1, 2, \dots, M \dots\dots\dots (1.11)$$

divide o curso temporal do sinal em intervalos de tempo, sobre os quais a entropia pode ser calculada. Da expressão,  $m$  são os valores possíveis de posicionamento da janela,  $w \in \mathbb{N}$  é a largura da janela, e  $\Delta$  é o fator de deslocamento da janela de centro em  $\ell$ . Note que  $w, m$  e  $\Delta$  devem ser escolhidos de modo a obedecer às condições:  $w \leq k$  e  $(k - w) / \Delta \in \mathbb{N}$ . O centro da janela é localizado em  $t_k = l = w/2 + m\Delta$ .

Considerando o conjunto de amplitudes S e o conceito de janelas deslizantes, W, é possível definir os níveis acessíveis ao sistema de duas maneiras: método global e método local.

• **Método global**

Nesse método, calcula-se o máximo e o mínimo da amplitude de todo o sinal  $s(t)$  e, em seguida, obtém-se os níveis acessíveis ao sistema através da partição do sinal em  $L$  intervalos. Matematicamente tem-se:

$$s_0 < s_1 < s_2 < \dots < s_M, \quad \dots\dots\dots (1.12)$$

logo,

$$s_0 \equiv \min[S] = \min_k \{s(t_k), k = 1, \dots, K\}, \quad \dots\dots\dots (1.13)$$

$$s_M \equiv \max[S] = \max_k \{s(t_k), k = 1, \dots, K\}, \quad \dots\dots\dots (1.14)$$

Desse modo, define-se os intervalos de amplitude,  $I_\ell$ , como:

$$I_\ell = [s_{l-1}, s_l), \quad \dots\dots\dots (1.15)$$

sendo que  $\ell$  varia de 1 até o número total de níveis  $L$ . Logo, a probabilidade de ocupação de cada nível, no interior da janela, pode ser calculada segundo a equação:

$$p_W(i) = \frac{1}{n} \sum_{j \in W} \delta_{iN(j)}, \quad \dots\dots\dots (1.16)$$

em que  $n$  é o número de pontos na janela e  $\delta_{iN(j)}$  é igual a 1 se e somente se o índice  $i$  do nível em questão for igual ao nível do ponto  $j$ , de outro modo  $\delta_{iN(j)} = 0$ .

Por fim, dada a distribuição de probabilidades, em cada janela, calcula-se as entropias, atribuindo esse valor calculado ao ponto referente ao centro da janela.

• **Método local**

No caso do método local, os níveis acessíveis ao sistema são definidos a partir dos máximos e mínimos pertencentes às janelas  $W$ . Matematicamente tem-se:

$$j_0 < j_1 < j_2 < \dots < j_M, \quad \dots\dots\dots (1.17)$$

em que:

$$j_0 \equiv \min[W] = \min_k \{j(t_k), k = 1 + m\Delta, \dots, w + m\Delta\}, \dots\dots\dots (1.18)$$

$$j_M \equiv \max[W] = \max_k \{j(t_k), k = 1 + m\Delta, \dots, w + m\Delta\}, \dots\dots\dots (1.19)$$

Deste modo, os intervalos de amplitude estão definidos no interior da janela, da seguinte maneira:

$$I_\ell = [j_{\ell-1}, j_\ell), \dots\dots\dots (1.20)$$

em que  $\ell$  varia de 1 até o número total de níveis  $L$ . Novamente, calcula-se a probabilidade de ocupação de cada nível no interior da janela segundo a equação 1.16.

---

## Capítulo II – Imagem Funcional por Ressonância Magnética

### 1. Introdução

A imagem por ressonância magnética, *magnetic resonance imaging* (MRI), está baseada no fenômeno de ressonância magnética nuclear [Apêndice A] e é capaz de produzir imagens do corpo humano com excelente contraste entre tecidos moles, permitindo distinguir tecidos saudáveis dos anormais, p.ex., tumores.

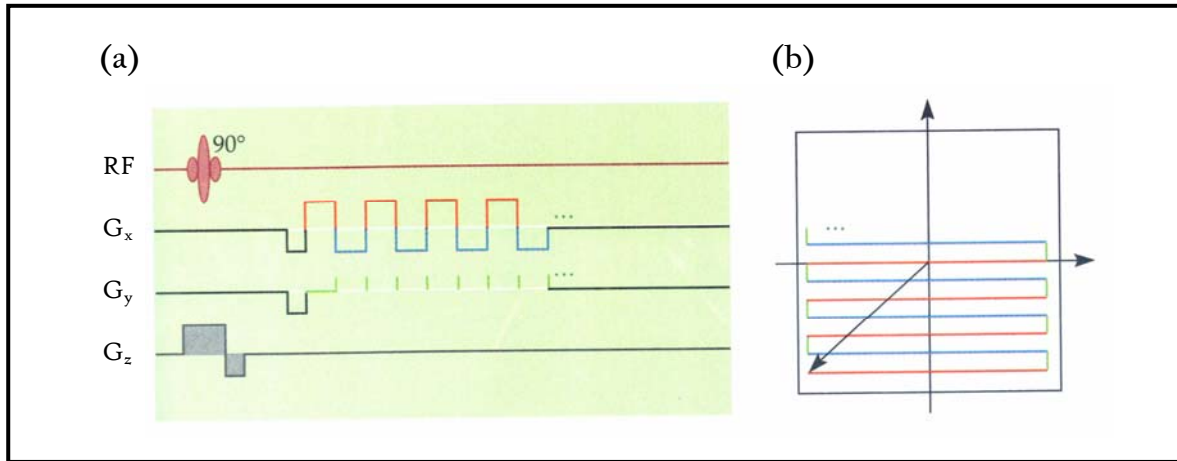
O aperfeiçoamento da técnica de MRI, com o desenvolvimento de seqüências de pulso específicas, do tipo eco planar, *Echo Planar Imaging* (EPI), e descobertas de processos fisiológicos que produzem contraste endógeno, abriu a possibilidade de estudos das funções cerebrais em humanos a partir dessa técnica. A Imagem Funcional por Ressonância Magnética, *Functional Magnetic Resonance Imaging* (fMRI), como assim é chamada, tem obtido um papel de destaque, dentre todas as técnicas de mapeamento funcional, por seu caráter não-invasivo, e a prevalência de tomógrafos de ressonância em instituições hospitalares. Além disso, essa técnica tem oferecido uma melhora substancial da resolução espaço-temporal comparado a qualquer outro método não-invasivo [LOGOTHETIS, 2002].

Nesta seção serão descritos os aspectos básicos da técnica de fMRI, como a aquisição das imagens, o mecanismo de contraste endógeno utilizado, os desenhos experimentais, e as técnicas de pré-processamento e pós-processamento. Além disso, serão apresentados aspectos recentes da sua utilização como ferramenta para o estudo da cronometria mental.

### 2. Seqüência de Pulso EPI

O mapeamento das funções cerebrais requer a aquisição de imagens com resolução temporal alta. Nesse sentido, seqüências de pulsos rápidas têm sido desenvolvidas para adquirir um número grande de imagens dentro de um curto período de tempo. Os métodos convencionais preenchem o espaço-k linha por linha, necessitando de várias excitações separadas para reconstruir uma imagem. Em 1977, Peter Mansfield propôs o método EPI, no qual todo o espaço-k é preenchido com um único

pulso RF, seguindo uma trajetória zig-zag no espaço recíproco [MANSFIELD, 1977]. A figura 2.1 ilustra um diagrama com a seqüência de pulso EPI, e sua respectiva trajetória no espaço k.



**Figura 2.1.** Seqüência de pulso EPI do tipo Gradiente Eco (a) e a trajetória no espaço-k (b). Observe que as direções dos gradientes são alteradas rapidamente no tempo para permitir a trajetória no espaço-k. *Figura modificada de Huettel et al. 2004.*

Juntamente com o pulso RF de  $90^\circ$ , o gradiente de seleção de fatia ( $G_z$ ) é aplicado. Na seqüência, há um pulso de gradiente negativo de codificação de freqüência e outro de fase para posicionar a coordenada do espaço-k no seu canto inferior esquerdo. Nesse momento, um gradiente positivo  $G_x$  é aplicado e uma linha na direção  $k_x$  é adquirida. O processo continua pela aplicação de um pequeno gradiente positivo  $G_y$ , que move as coordenadas do espaço-k uma linha acima, e um pulso negativo de  $G_x$  dirige as coordenadas para trás, varrendo a segunda linha, e assim sucessivamente, até o completo preenchimento do espaço-k. Na EPI, toda informação necessária para reconstrução da imagem é obtida em um simples período TR. Um problema central desse processo é que a janela de aquisição de dados é limitada pela magnetização transversal ( $T_2$  e  $T_2^*$ ), que pode variar de 50 – 100 ms. A espessura da fatia pode ser ajustada alternando a intensidade do gradiente de campo  $G_z$ , e a posição central da fatia pode ser ajustada alternando a freqüência do pulso de excitação.

Enquanto seqüências de pulsos convencionais têm resolução temporal de alguns segundos, a EPI alcança a ordem de alguns milissegundos, chegando a 20 ms. Contudo, a melhora na resolução temporal resulta em uma imagem com baixa resolução espacial [HUETTEL, *et al.*, 2004].

Embora haja outras estratégias para a aquisição rápida das imagens, a consolidação de novas tecnologias para produção de gradientes tem tornado a EPI o método de escolha nos estudos de fMRI.

### 3. Estudos Funcionais Usando Agentes de Contraste

Uma estratégia para o aumento do contraste nas imagens por ressonância magnética é o uso de agentes de contraste exógenos, como o Gd-DTPA (ácido diethylenetriaminepentacético). Essas substâncias, altamente paramagnéticas, podem ser injetadas na corrente sanguínea, alterando o contraste local, pelo encurtamento dos tempos de relaxação. Esses agentes de contraste como são conhecidos, têm grande importância em imagens clínicas, especialmente na detecção de tecidos patológicos, incluindo tumores cerebrais.

Os primeiros estudos de fMRI em humanos foram idealizados pelo uso de agentes de contraste exógenos, por Belliveau e colegas, em 1991 [BELLIVEAU *et al.*, 1991]. Eles mediram o aumento de contraste no lobo occipital pelo mapeamento de Gd-DTPA durante estimulação visual. Por um lado o uso de agentes de contraste promove uma enorme mudança do contraste local, entretanto, o método demanda a injeção endovenosa dessa substância. Desse modo, pouco tempo depois, outro método foi concebido. Nesse, não há necessidade de utilização de um agente exógeno, e está baseado em um processo fisiológico há muito conhecido.

### 4. Efeito BOLD

O cérebro humano contém aproximadamente  $10^{11}$  neurônios que realizam todos os processos cognitivos, sensoriais e motores do corpo. A transferência de informação no cérebro ocorre por propagação de impulsos nervosos de uma célula nervosa a outra, pela liberação de neurotransmissores em sinapses e suas subseqüentes interações com receptores específicos nos neurônios alvos. As interações entre os neurotransmissores pré-sinápticos e receptores levam a mudanças do fluxo de corrente na membrana, despolarizando a membrana pós-sináptica, gerando o potencial de ação.

As mudanças metabólicas nos neurônios e células gliais, que acompanham a liberação de neurotransmissores, requerem energia, que é, em grande parte, usada nas sinapses. Como a produção de energia no cérebro normal depende do metabolismo oxidativo (produção de Adenosina Trifosfato, *Adenosin Triphosphate* - ATP), a maior demanda local de oxigênio ocorre pelo aumento da atividade sináptica. O consumo de ATP requer um suprimento contínuo de glicose e oxigênio, que é regulado, em parte, pelo fluxo sanguíneo cerebral, *Cerebral Blood Flow* (CBF), uma vez que o cérebro não tem uma reserva de oxigênio armazenada localmente. Nesse sentido, Ângelo Mosso, em 1881, demonstrou a correlação entre a demanda de energia e o CBF [MOSSO (1881 **apud** LOGOTHETIS, 2002)]. Alguns anos mais tarde, em 1890, Roy e Sherrington demonstram que a estimulação do cérebro causa um aumento local de fluxo sanguíneo [ROY & SHERRINGTON (1890 **apud** LOGOTHETIS, 2002)].

Por outro lado, em 1936, Linus Pauling e seu estudante Charles Coryell, descobriram que a molécula de hemoglobina tem propriedades magnéticas que dependem do estado da oxigenação dos grupos heme. A hemoglobina oxigenada (Hb) é diamagnética, ou seja, tem momento magnético nulo. Em contraste, a deoxi-hemoglobina (dHb) é paramagnética, tendo um momento magnético diferente de zero [HUETTEL *et al.* 2004]. Sabe-se, ainda, que a alteração do campo magnético local de um tecido altera os tempos de relaxação. Assim, mudanças na susceptibilidade magnética local mediante mudanças de concentração da dHb levam à alteração no sinal MR. Essa predição foi verificada experimentalmente nos estudos realizados por Thulborn e colegas, que demonstraram, mediante um tubo de teste com sangue, que a taxa de relaxação  $T_2$  depende da proporção de hemoglobina oxigenada [THULBORN *et al.* (1982 **apud** HUETTEL *et al.* 2004)]. O fato de que a oxigenação do sangue tem um efeito mensurável no sinal de MR dos tecidos vizinhos foi descoberto por Ogawa e colegas, em um estudo com roedores anestesiados [OGAWA, *et al.*, 1990a]. Eles encontraram que o sinal MR em torno das veias diminuiu quando a concentração de oxigênio do ar inspirado foi reduzida, e que o efeito foi revertido quando essa concentração retornou a valores normais. Além disso, eles observaram que a redução de sinal não ocorria exclusivamente no sangue, mas afetava tecidos em torno dos vasos, sugerindo uma



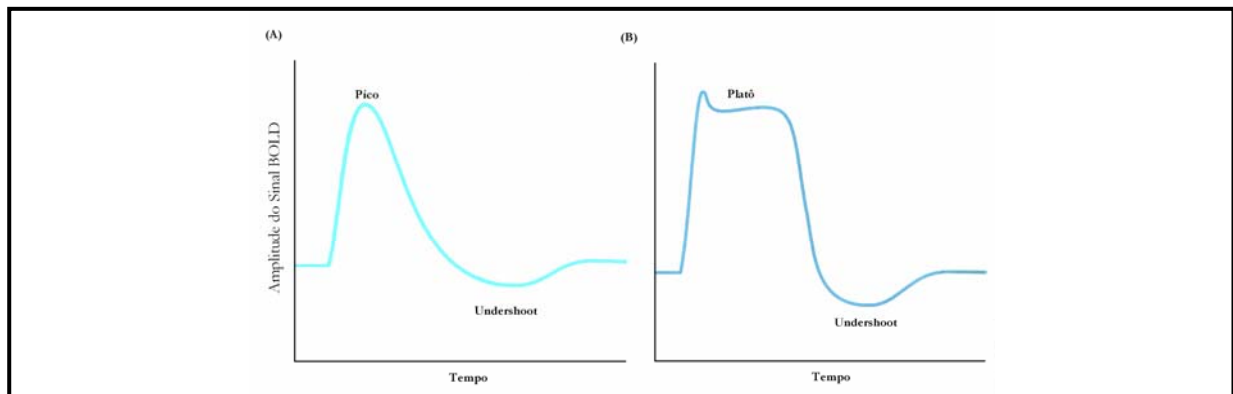
redução de  $T_2^*$  nos espaços intra e extravascular, similar a mudanças de susceptibilidade causadas pelos agentes de contraste.

Para verificar se a alteração do contraste observado resultava da demanda metabólica de oxigênio, Ogawa e colaboradores mediram as alterações de sinal enquanto manipulavam gases inalados em ratos anestesiados, comparando níveis baixos e altos de anestesia. O contraste foi muito maior em baixos níveis de anestesia em relação a altos níveis, indicando que a demanda metabólica de oxigênio era um pré-requisito para a observação do contraste [OGAWA, *et al.*, 1990b].

Pouco tempo depois, Kwong e colaboradores demonstraram que era possível observar o mesmo efeito em humanos [KWONG, *et al.*, 1992]. Esse mecanismo de contraste passou, então, a se chamar BOLD, *Blood Oxygenation Level Dependent*, e é utilizado largamente nos estudos atuais de fMRI. Em resumo, o contraste BOLD depende da quantidade total de dHb presente em uma dada região cerebral, que, por sua vez, depende do balanço entre consumo e fornecimento de oxigênio. Pela previsão de Ogawa e colaboradores seria esperado que o aumento da atividade neuronal resultasse em um aumento do consumo de oxigênio, aumentando a dHb no sangue e, conseqüentemente, diminuindo o sinal de MR. Não obstante, o oposto é observado: durante o aumento da atividade neural há uma diminuição da dHb e aumento do sinal de MR. Estudos efetuados por Fox e Raichle ajudaram a explicar essa observação [FOX & RAICHLER, 1986]. A disparidade entre a utilização de oxigênio e a demanda significa que mais oxigênio é fornecido para uma região cerebral, que a quantidade consumida. O contraste BOLD que acompanha a atividade neural ocorre não porque a Hb aumenta o sinal MR, mas porque ela substitui a dHb, suprimindo sua influência na diminuição da intensidade do sinal MR. De fato, eles observaram que durante a atividade neural há um leve crescimento da extração de oxigênio do sangue, em detrimento de uma grande elevação do fluxo sangüíneo cerebral, carregando com ele mais Hb.

A evolução temporal do sinal BOLD é conhecida como função resposta hemodinâmica, *hemodynamic response function* (HRF). A figura 2.2 ilustra a HRF em uma região cerebral. A forma da HRF varia com as propriedades do estímulo evocado e com a atividade neural subjacente. A altura do patamar está associada ao grau de atividade neural. Se ocorrer aumento da taxa de atividade neural, podemos esperar

um aumento da amplitude da HRF e, da mesma maneira, se aumentarmos a duração da atividade neural, a largura da HRF também poderá aumentar. A relação exata entre os eventos neurais que provocam a HRF e sua forma ainda não estão bem claros. Sabe-se, no entanto, que as respostas neurais corticais, que ocorrem dentro de aproximadamente 10 ms, só promovem as primeiras mudanças observáveis da HRF em torno de 1 ou 2 segundos depois do estímulo.



**Figura 2.2. Representação esquemática da resposta hemodinâmica BOLD. (a) Função resposta hemodinâmica para um evento de curta duração e (b) para um evento de longa duração. Modificada de HUETTEL et al. 2004.**

Geralmente, a HRF para uma área de atividade neural específica, como, por exemplo, a área visual, é composta por sete fases distintas: (i) antes da apresentação do estímulo, o sinal encontra-se em uma linha de base; (ii) em seguida, geralmente logo após a apresentação do estímulo, a resposta decresce por um breve intervalo de tempo. Esse padrão foi atribuído à extração de oxigênio inicial antes do aumento do fluxo sanguíneo e aumento transiente na concentração da dHb, mas esse efeito não é sempre visto, sendo observado com mais freqüência em estímulo visual [ROSEN *et al.*, 1998]; (iii) após a depressão do sinal, ele inicia uma etapa de derivada positiva (BOLD positivo anterior), atingindo um máximo, usualmente após 7-10 segundos do estímulo. Este máximo é conhecido como o pico da HRF (Figura 2.2-A). (iv) o sinal permanece em um platô por poucos segundos. Se a atividade neural é estendida através de um bloco de eventos no tempo, o pico pode ser estendido dentro de um platô (Figura 2.2-B); (v) em seguida, ele passa a decrescer, formando um efeito de BOLD positivo posterior; (vi) esse decréscimo é seguido por uma nova depressão do sinal, negativo com respeito à linha de base, conhecido como *undershoot*, uma vez que, seguindo o término da atividade neural, o fluxo sanguíneo diminui mais rapidamente que o volume sanguíneo. Se o volume permanece acima do nível da linha de base,

enquanto o fluxo estava na linha de base, então uma maior quantidade de dHb estará presente, fazendo com que o sinal de MR seja levado a valores abaixo da linha de base. (vii) como o volume sanguíneo retorna lentamente para níveis normais, o sinal fMRI aumentará, gradativamente, voltando para a linha de base.

Embora o modelo descrito anteriormente seja bem aceito para um grande número de processos cerebrais, a forma da resposta hemodinâmica ainda não é totalmente compreendida, e depende de várias características específicas, como a região do cérebro, a idade do paciente, e alterações patofisiológicas [PINEIRO *et al.*, 2002]. Desse modo, nos últimos anos, o interesse em explorar a natureza fisiológica do sinal BOLD e sua dependência na duração e no tipo de estímulo neural vem crescendo enormemente [KIM, 2003].

## 5. Experimentos em fMRI

O arranjo experimental de exames por fMRI deve estar adequado para responder às hipóteses inicialmente formuladas e buscar minimizar o tempo do exame. A aquisição dessas é acompanhada por uma série de tarefas realizadas pelos sujeitos, com intuito de encontrar áreas de atividade cerebral desencadeadas pela tarefa. O desenho de como apresentar essas tarefas é denominado de paradigma, que devem conter, no mínimo, duas condições experimentais. Os paradigmas podem ser divididos esquematicamente em três grupos principais:

1. Paradigmas em bloco: alternam longos períodos de uma tarefa com períodos igualmente longos de outra tarefa, que pode ser o estado de repouso;
2. Paradigmas contínuos ou paramétricos: os estímulos são apresentados continuamente, em geral de forma aleatória;
3. Paradigmas evento-relacionado, *event-related functional MRI* (ER-fMRI): tarefas curtas intercaladas com longos períodos de repouso.

Aqui, descreveremos em mais detalhes apenas o paradigma ER, uma vez que este será utilizado neste trabalho.

## 6. Paradigmas Evento-Relacionado

Atualmente, os paradigmas em bloco são os mais utilizados para exames de fMRI, em especial sua utilização clínica. Contudo, o uso de paradigmas ER-fMRI tem sido cada vez mais utilizado, dada algumas de suas características, que lhe conferem maior flexibilidade. A suposição central dos paradigmas ER-fMRI é que a atividade neural de interesse ocorrerá dentro de intervalos curtos e discretos. Desse modo, estímulos são apresentados por um curto período de tempo, geralmente da ordem de 1 a 2 segundos, seguido por um período relativamente longo de repouso, de aproximadamente 20 segundos, o que dá origem a uma HRF similar ao demonstrado na figura 2.3.

Um dos primeiros estudos realizados para comparar paradigmas ER-fMRI com paradigmas em bloco foi relatado por Buckner e colegas em 1996. Neste estudo, sujeitos eram instruídos a completar mentalmente as iniciais de palavras que lhes eram apresentadas. Quatro desses sujeitos foram testados com um TR de 2 s, e outros dois com um TR de 1 s. Além de encontrar respostas em ambos paradigmas, eles anteciparam que o paradigma ER-fMRI poderia fornecer informações adicionais sobre aspectos temporais do sinal BOLD. De fato, em um dos dois sujeitos, sobre o qual os dados foram amostrados com TR de 1 s, foi observada uma clara diferença na latência da HRF entre o córtex occipital e o pré-frontal [HUETTEL *et al.*, 2004]. A princípio, com esse paradigma, não apenas podemos observar quais áreas cerebrais estão sendo ativadas pela tarefa, como, também, em qual ordem temporal essas regiões se associam.

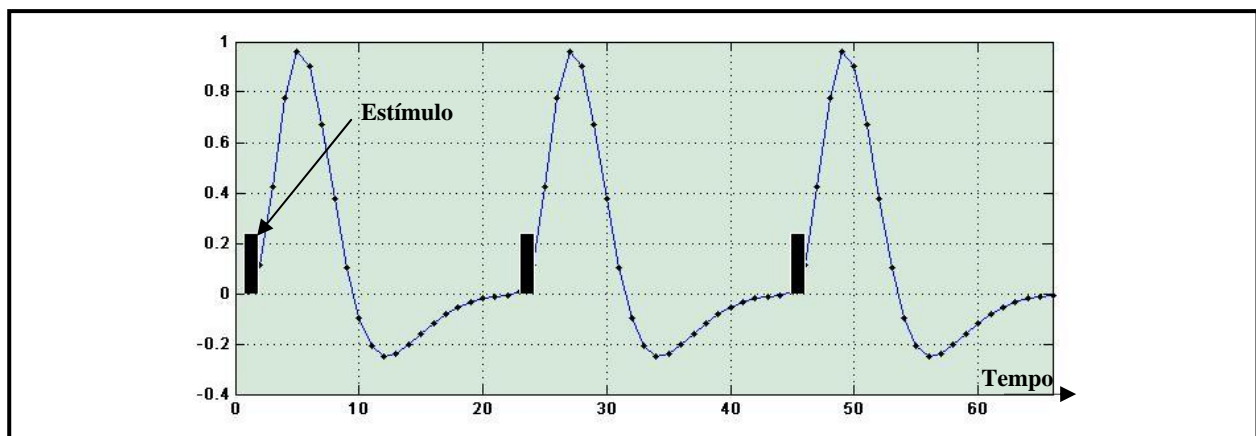


Figura 2.3. Série temporal simulada de um voxel em um paradigma ER-fMRI.

Em resumo, o paradigma ER-fMRI mede mudanças transientes da atividade cerebral associada a estímulos curtos. Sua resolução temporal é maior que estudos em bloco, podendo atingir frequências de amostragem de algumas dezenas de milisegundos, dependendo do campo magnético. Várias são as vantagens e desvantagens desse paradigma. Podemos citar, dentre as mais importantes:

Vantagens	Desvantagens
<p>Bom poder de estimação da forma da HRF;</p> <p>Estratégias de análise mais flexíveis;</p> <p>Resolução temporal melhorada;</p> <p>Possibilidade de misturar diferentes processos cognitivos dentro de um único paradigma;</p> <p>Possibilidade de separação de regiões corticais envolvidas em diferentes tarefas, como, por exemplo, estudos de armazenamento e recuperação de memória.</p> <p>Possibilidade de inferir sobre o fluxo de informação neural.</p>	<p>Pode ter o poder de detecção reduzido pela baixa Relação sinal-ruído;</p> <p>Sensível a erros na HRF preditora.</p>

Uma vez adquiridas, as imagens devem passar por uma série de estratégias de pré-processamento, dada a baixa relação sinal-ruído, *Signal to Noise Ratio* (SNR), dos exames de fMRI e a presença de artefatos.

## 7. Pré-Processamento das Imagens

Todos os dados de fMRI sofrem de inexatidão espacial e temporal causados por várias fontes de ruído. Dentre elas, podemos citar: ruído térmico intrínseco no sujeito e na eletrônica dos tomógrafos; ruídos do sistema associados às imperfeições no hardware; artefatos resultantes do movimento involuntário da cabeça dos sujeitos, respiração, batimento cardíaco e outros processos fisiológicos; dentre outras. Se não corrigido, esta variabilidade pode reduzir (ou até eliminar) o poder de detecção de áreas cerebrais envolvidas em um experimento.

Nesta seção será discutido uma série de procedimentos computacionais de pré-processamentos que operam sobre os dados de fMRI, visando condicionar os dados para maximizar a sensibilidade da técnica. Dentre essas etapas, estão: correção temporal entre fatias, a correção de movimento, aplicação de filtros espaciais e temporais e remoção de *drift*.

## 7.1 Correção Temporal

Nos exames de fMRI, cada fatia é adquirida em um tempo diferente dentro do TR, ou seja, as fatias que compõem um volume cerebral encontram-se ligeiramente defasadas no tempo. Esse fato pode ser problemático, em especial sobre paradigmas ER-fMRI, já que influencia a correspondência entre os dados observados e hipóteses experimentais. A correção desse erro é feita pelo alinhamento temporal das fatias, o que, de regra, utiliza técnicas de interpolação temporal. Ela usa informação dos pontos temporais vizinhos para estimar a amplitude do sinal MR no início do TR.

Há muitas estratégias de interpolação, mas nenhuma pode recuperar perfeitamente informações perdidas entre as fatias. Sua exatidão depende de dois fatores: variabilidade nos dados experimentais e taxa de amostragem. Se os dados variam mais lentamente, quando comparados à taxa de amostragem, a interpolação será mais efetiva e também quando os dados são adquiridos com um TR curto (i.e., 1 s).

## 7.2 Correção de Movimento

O movimento da cabeça constitui uma das maiores fontes de erros nos estudos de fMRI e podem ser divididos em intrínsecos e extrínsecos. O primeiro é decorrente de flutuações internas provenientes de movimentos dos tecidos intracranianos e de processos fisiológicos como respiração e batimentos cardíacos. O segundo é decorrente de movimentos involuntários da cabeça do indivíduo.

Quando há movimento da cabeça durante um experimento, algumas imagens podem ser adquiridas com o cérebro em posições diferentes, resultando em padrões falsos de atividade, isto é, regiões que no momento anterior apresentavam baixo contraste passam a apresentar brilho mais intenso.

Desse modo, estratégias de correção têm o objetivo de ajustar espacialmente as séries temporais das imagens, de modo que todos os voxels se encontrem na mesma posição, em todas as imagens.

No processo de alinhamento, sucessivos volumes em uma série temporal são co-registrados com um volume de referência, utilizando transformação de corpo rígido, de modo que um pode ser sobreposto exatamente sobre o outro por uma combinação de três translações (ao longo dos eixos x, y, z) e três rotações (em torno dos planos x-y, x-z e y-z).

Para determinar a quantidade de movimento, os algoritmos computacionais identificam um conjunto de parâmetros de translação e rotação que fornecem a melhor combinação entre o volume de referência e o corrigido, como por exemplo, a primeira imagem adquirida na sessão (Figura 2.4). Em seguida, uma medida de similaridade, ou função de custo, é realizada a fim de identificar a qualidade do ajuste. Uma vez que grandes diferenças são muito mais problemáticas do que pequenas diferenças, uma função de custo mais comum é a soma do quadrado das diferenças de intensidade dos voxels. Uma vez determinado um conjunto de parâmetros de realinhamento, os dados originais são re-amostrados por interpolação espacial.

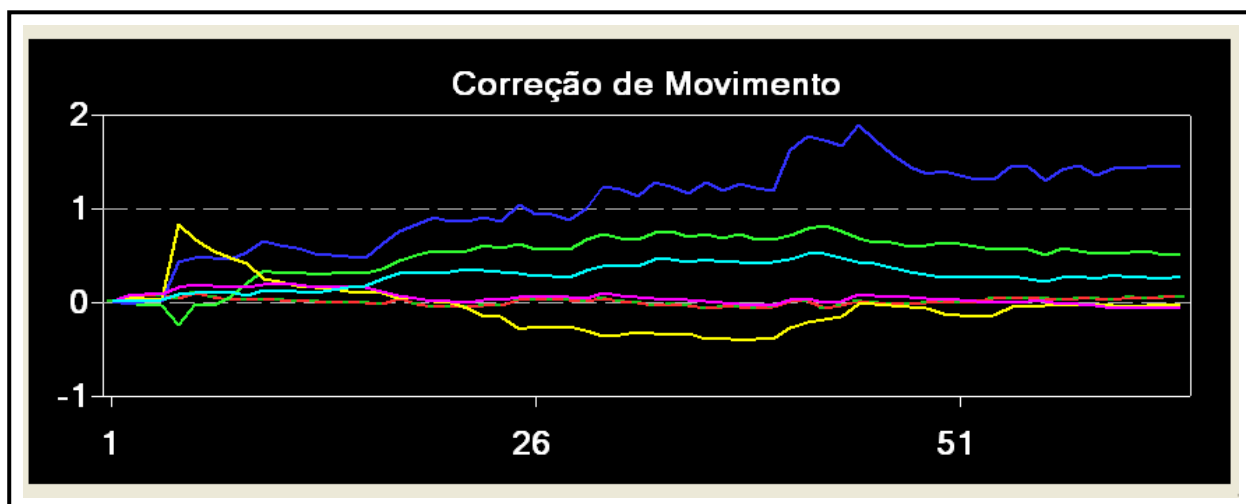


Figura 2.4. Gráfico que demonstra os movimentos da cabeça de translação e rotação para uma simples sessão de fMRI.

Uma outra etapa de pré-processamento é a utilização de filtros. Na fMRI, os filtros são usados para remover variações indesejáveis provocadas pelas fontes de ruído, enquanto o sinal de interesse é preservado. Abordaremos dois típicos filtros utilizados no pré-processamento de fMRI: filtros temporal e espacial.

### 7.3 Filtros Temporal e Espacial

O uso de filtros temporais podem aumentar substancialmente a qualidade dos dados de fMRI melhorando a SNR. Para melhor descrever a aplicação de um filtro temporal, consideremos uma série temporal de um único vóxel, no domínio do tempo. Aplicando-se a Transformada de Fourier (TF), os dados podem ser convertidos para o domínio da frequência. A escala de frequência que está presente em um sinal depende da taxa de amostragem que, em fMRI, está associada ao valor TR. Como é de conhecimento, a frequência máxima que pode ser identificada, ou frequência de Nyquist, é igual à metade da taxa de amostragem. Por exemplo, se a taxa de amostragem for de 0.5 Hz (TR de 2000 ms), nenhuma frequência do sinal poderá ser maior que 0.25 Hz, de maneira a evitar o fenômeno *aliasing*. Devido à limitação de Nyquist, nós devemos amostrar o cérebro em duas vezes a taxa do fenômeno de interesse.

São três os filtros temporais, que atenuam seletivamente as frequências nos sinais: filtro passa-banda, que atenua uma gama ou banda de frequências; filtro passa-baixa, que deixa frequências baixas passarem enquanto atenua altas frequências; e os filtros passa-alta, que atenuam apenas baixas frequências. Desse modo, a utilização dos filtros temporais depende da variabilidade que se quer eliminar. Em um experimento de fMRI, típicas taxas de batimento cardíaco podem variar, mas encontram-se entre 1.0 e 1.5 Hz. A taxa de respiração é lenta, aproximadamente 0.2 a 0.3 Hz. Para uma comparação, um experimento típico pode apresentar blocos alternados de 12 s de estímulo e 12 s de repouso, adquiridos com uma taxa de amostragem de 0.04 Hz. Um filtro passa baixa que exclui frequências acima de 0.2 Hz pode remover oscilações fisiológicas sem reduzir significativamente a habilidade de detectar o efeito da tarefa de interesse. Entretanto, se um experimento usa um paradigma ER-fMRI muito rápido,



a tarefa e a respiração podem ter frequências similares, e a utilização de filtros pode ser extremamente difícil.

Por outro lado, a aplicação de filtros espaciais aumenta a SNR e melhora as características de apresentação dos resultados mediante a suavização, embora leve a uma perda de resolução espacial. Geralmente, as análises de fMRI fazem uso de filtros do tipo Gaussiano, que tem a forma de uma distribuição normal e consiste na convolução da imagem com funções do tipo:

$$f(x, y, z) = \exp\left\{-\left[\frac{x^2}{2s_x^2} + \frac{y^2}{2s_y^2} + \frac{z^2}{2s_z^2}\right]\right\}, \quad \dots\dots\dots (2.1)$$

em que  $x$ ,  $y$  e  $z$  correspondem às posições geométricas das imagens, e  $s_x$ ,  $s_y$  e  $s_z$  são os desvios padrão da função, definida pelo usuário. As características do filtro são bem definidas pelo parâmetro largura a meia altura, ou FWHM (*Full-Width-Half-Maximum*), que determina a forma de atuação do filtro sobre a imagem, sendo que as larguras típicas utilizadas em fMRI estão entre 2 a 10 mm FWHM.

## 7.4 Remoção de *Drift*

Em muitos casos, o sinal BOLD observado tem a característica de derivar linearmente ao longo do tempo, o que reduz a significância estatística dos padrões de atividade. Este artefato é conhecido como *baseline drift* e pode resultar de muitas fontes diferentes como a instabilidade instrumental, ruído fisiológico e movimento do sujeito [SMITH *et al.*, 1999], e é removido por vários métodos como os lineares [BANDETTINI *et al.*, 1993; LOWE & RUSSELL, 1999], polinomiais [MATTAY *et al.*, 1996], filtros passa alta [JEZZARD & CLARE, 2001] dentre outros.

## 8. Pós-processamento em fMRI

Conforme visto anteriormente, o efeito BOLD tem uma baixa SNR, promovendo uma diferença de contraste muito pequena, tornando necessária a manipulação estatística do conjunto de ima-

gens adquiridas por métodos específicos. Ainda, um exame típico é formado por um conjunto de imagens, adquiridas sequencialmente no tempo, durante as condições experimentais de ativação e repouso, formando, então, um volume de imagens 3-D (Figura 2.5). É possível, então, associar uma série temporal a cada voxel da imagem, em que cada elemento da série tem o valor do voxel das imagens funcionais adquiridas no tempo. São essas as séries temporais submetidas aos métodos estatísticos, capazes de localizar a resposta hemodinâmica e, então, distinguir os voxels ativos dos não ativos.

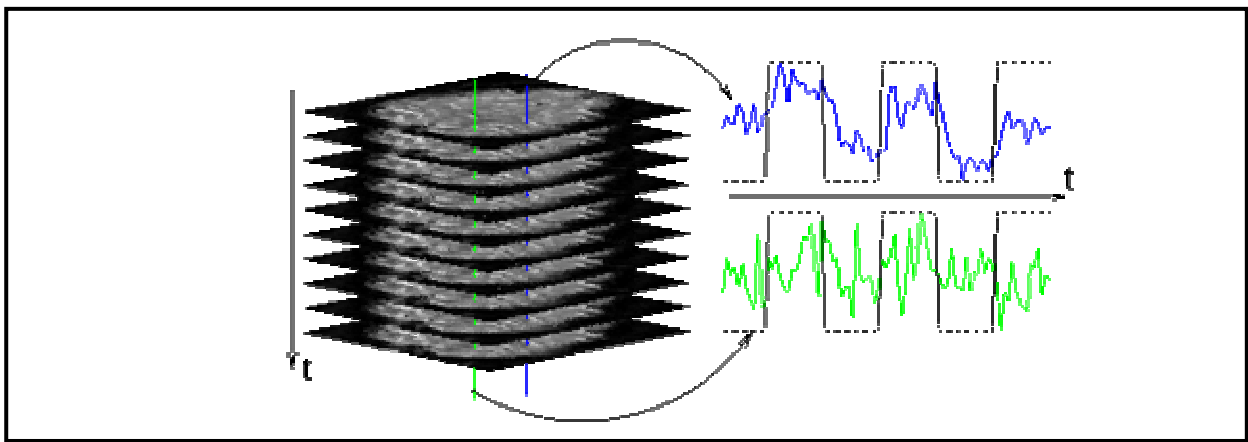


Figura 2.5. *Diagrama* representativo da série temporal formadas por imagens funcionais. No detalhe apresenta-se o curso temporal de um voxel ativo (linha azul) e outro não ativo (linha verde), quando da utilização de um paradigma de bloco.

Portanto, a análise estatística é uma etapa crítica do mapeamento cerebral por técnica de f-MRI. Muitos métodos têm sido propostos e dependem, em grande parte, do tipo de fenômeno que se pretende observar. Dentre esses métodos, podemos citar o modelo linear geral, *General Linear Model* (GLM) [FRISTON *et al.*, 1995a], a correlação cruzada [HUETTEL *et al.*, 2004], o teste-t de Student [HUETTEL *et al.*, 2004], a Análise de Componentes Independentes, *Independent Component Analysis* (ICA) [MCKEOWN *et al.* 1998], *Kolmogorov-Smirnov test* [HUETTEL *et al.*, 2004], *Bayesian approach* [AMARAL *et al.*, 2004], dentre outros.

Nessa sessão, serão apresentados alguns dos métodos acima mencionados, em que, embora possuam características comuns, todas essas abordagens diferem em sua suposição e objetivos:

- Elas expressam a significância como a probabilidade do resultado ocorrer sobre a hipótese nula ( $H_0$ );

- os mapas de cores que formam a base das imagens de fMRI expressam esta probabilidade.
- voxels cujos níveis de probabilidade estão abaixo de um valor de corte, *threshold*, são chamados de significantes, enquanto voxels cuja probabilidade está acima deste valor não tem significância estatística.

Assim como não há um único paradigma experimental apropriado para todas as questões a serem pesquisadas, também não há uma única abordagem de análise que pode ser usada para todos os experimentos. O método estatístico correto para qualquer experimento depende do paradigma utilizado no experimento, das hipóteses a serem testadas, e do tipo de erro que se quer minimizar.

### 8.1 Teste-t de Student

Um dos parâmetros mais usados para gerar mapas estatísticos em fMRI é o Teste-t de Student. Em fMRI, a técnica calcula uma significância estatística para cada série temporal dos voxels, possibilitando determinar se os voxels estão ativos ou não.

Para conduzir esse teste, consideram-se dois subconjuntos, formados pelos estados de ativação e de repouso das séries temporais. Nesse contexto, o teste-t procura determinar se há diferença significativa entre a média destes dois estados [atividade ( $\bar{X}'_{on'}$ ) e repouso ( $\bar{X}'_{off'}$ )], em todos os voxels da imagem de tal sorte que o valor t pode ser calculado por [McPHERSON, 1990; SPIEGEL, 1972]:

$$t = \frac{\bar{X}'_{on'} - \bar{X}'_{off'}}{S_{\bar{X}'_{on'} - \bar{X}'_{off'}}} \dots\dots\dots (2.2)$$

em que

$$S_{\bar{X}'_{on'} - \bar{X}'_{off'}} = \sqrt{\frac{S_p^2}{n'_{on'}} + \frac{S_p^2}{n'_{off'}}} \dots\dots\dots (2.3)$$

sendo que  $n'_{on'}$  e  $n'_{off'}$  correspondem o número de pontos dos subconjuntos nos estados ativado e repouso. Finalmente  $S_p^2$  é a variância amostral combinada (*pooled variance*) dada por:

$$S_p^2 = \frac{\sum_{i=1}^{n'_{on'}} (X'_{on'} - \bar{X}'_{on'})^2 + \sum_{i=1}^{n'_{off'}} (X'_{off'} - \bar{X}'_{off'})^2}{n'_{on'} + n'_{off'} - 2} \dots\dots\dots (2.4)$$

O resultado desse teste pode então ser convertido para um valor de probabilidade, valor-p (*p-value*), que representa o menor nível de significância para o qual  $H_0$  pode ser rejeitado. Esse parâmetro está baseado nos graus de liberdade, *degrees of freedom* (df), dado por  $n'_{on'} + n'_{off'} - 2$  e no tipo de teste de hipótese (bi-caudal).

Em um teste de hipótese podem ocorrer dois tipos de erro: erro do tipo I e erro do tipo II. Se  $H_0$  é a hipótese nula que será testada e  $H_1$  a hipótese alternativa, o primeiro consiste em rejeitar  $H_0$ , sendo  $H_0$  verdadeira, enquanto o segundo consiste em aceitar  $H_0$  sendo  $H_0$  falsa. Nesse sentido, um valor de probabilidade  $\alpha$  do erro tipo I pode ser definido, e é denominado o nível de significância do teste. Para determinar se  $H_0$  pode ser rejeitada, o valor-p é comparado com a probabilidade  $\alpha$ . No contexto da fMRI,  $H_0 = \mu'_{on'} - \mu'_{off'} = 0$ . Isso significa que voxels ativos têm grandes valores de t e baixos valores de p, e voxels não ativos possuem baixos valores de t e grandes valores de p. Para determinar se um voxel está ativo ( $H_1$ ) ou não está ativo ( $H_0$ ) a seguinte comparação é realizada:

$$\begin{cases} \text{se } p \leq \alpha, \text{ então rejeita - se } H_0 \text{ para o nível } \alpha \\ \text{se } p > \alpha, \text{ então aceita - se } H_0 \text{ para o nível } \alpha \end{cases}$$

Essa técnica tem a vantagem de não usar qualquer informação sobre a forma da resposta hemodinâmica. Entretanto, a escolha do valor  $\alpha$  (*threshold*) é fundamental para a construção dos mapas de ativação. Um critério convencional é utilizar  $\alpha=0,05$  para um único voxel, que pode levar a um número muito grande de falsos positivos. Não obstante, há algumas propostas para a escolha mais criteriosa no controle de falsos positivos, como a *Random Field Theory* [ADLER, 1981], *False Discovery Rate* [BENJAMINI & HOCHBERG, 1995; GENOVESE *et al.*, 2001] e a Correção de Bonferroni [HUETTEL *et al.*, 2004]. Dentre esses métodos, o mais usado, é a correção de Bonferroni. Se há k testes a serem exe-

cutados, a correção de Bonferroni substitui o nível de significância alfa com um nível de alfa/k para cada teste. Por exemplo, em fMRI, poderá existir um total de 10.000 voxels a serem analisados. Para garantir que todos os voxels tenham um mesmo nível de significância, aplica-se a correção de Bonferroni, no qual o valor  $\alpha$  escolhido deve ser dividido pelo número de voxels, ou  $\alpha = 0,000005$ .

Outro problema na aplicação do teste-t consiste em decidir quais pontos temporais pertencem às condições experimentais de atividade e repouso, uma vez que a HRF não segue precisamente o estímulo.

## 8.2 Correlação Cruzada

Enquanto o teste-t não usa qualquer informação sobre a forma da HRF, a análise estatística por correlação cruzada é realizada incluindo-se a HRF esperada. Proposta pela primeira vez em 1993 [BANDETTINI *et al.*, 1993], a análise de correlação cruzada procura quantificar a correspondência entre uma HRF modelo e os dados experimentais observados.

É muito simples conduzir uma análise por correlação sobre os dados de fMRI. Primeiro, identifica-se as épocas<sup>2</sup> dos dados experimentais. Em seguida, escolhe-se um modelo apropriado para a HRF com o mesmo número de pontos de uma época de dados experimentais. Segundo, calcula-se, para cada voxel, a covariância entre o curso temporal dos dados experimentais e o modelo da HRF, indicado pela seguinte equação:

$$r = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})^2} \sqrt{\sum_i (y_i - \bar{y})^2}} \dots\dots\dots (2.5)$$

em que x e y correspondem às funções do sinal experimental e referência, respectivamente. As variáveis  $\bar{x}$  e  $\bar{y}$  representam as médias de ambas as funções.

A covariância pode ser positiva ou negativa, e expressa o grau com o qual a medida do sinal segue a função modelo. Se a covariância é positiva, valores positivos dos dados experimentais significam grandes valores positivos da HRF modelo, e valores negativos dos dados experimentais significam

---

<sup>2</sup> Uma época corresponde a uma janela de pontos temporais durante o período de estímulo e repouso.

valores negativos da HRF modelo, indicando uma sintonia de fase entre os dois sinais. Por outro lado, se a covariância é negativa, há uma diferença de fase importante entre as duas funções. Normaliza-se a covariância pela divisão do produto do desvio padrão das duas épocas. O resultado é o coeficiente de correlação, ou valor-*r*, que pode ter valores em uma escala de 1 até -1. O valor 1 representa uma correlação positiva perfeita, enquanto que -1 refere-se à correlação negativa perfeita. Uma correlação igual a 0 indica a não existência de correlação entre as duas funções.

Nos experimentos de fMRI, a hipótese nula ( $H_0$ ) é  $r = 0$ , necessária para determinar se o voxel está ativo ( $H_1$ ) ou não está ativo ( $H_0$ ). O coeficiente de correlação é então usado como um parâmetro estatístico para mapeamento cerebral, e o valor de corte sobre *r* é escolhido para selecionar voxels que estão ativos.

Para qualquer valor de *r*, há um valor de *t* correspondente, que permite associar um nível de significância alfa. A equação 2.6 exprime matematicamente essa relação com N-2 graus de liberdade [W. H. PRESS *et al.*, 1992]:

$$t = \frac{r\sqrt{N-2}}{\sqrt{1-r^2}} \dots\dots\dots (2.6)$$

Na maioria dos métodos de análise, como é o caso da correlação cruzada, há a necessidade de escolher uma função de referência, definida a priori. Desse modo a escolha de uma função de referência adequada é fundamental. Em geral, funções do tipo quadrada (*Box-car function*) são as mais utilizadas nos paradigmas em bloco. Já nos paradigmas ER-fMRI a HRF pode ser modelada por funções gamma amortecida [LANGE *et al.*, 1997], do tipo:

$$h(t) = \frac{\left(\frac{t}{\tau}\right)^{(n-1)} e^{-(t/\tau)}}{\tau(n-1)!} \dots\dots\dots (2.7)$$

em que *t* é o tempo, *n* e  $\tau$  são parâmetros variáveis na representação da HRF.

Dessa forma, uma desvantagem da técnica de correlação cruzada é a necessidade de se criar um modelo para a HRF, uma vez que suas características podem diferir entre sujeitos, regiões cerebrais, estímulo, etc. [HANDWERKER *et al.*, 2004, SOLTYSIK *et al.*, 2004]. Além disso, o método de corre-

lação cruzada, assim como outros métodos convencionais tem a premissa de que a relação entre a resposta e o estímulo é linear [VAZQUEZ *et al.* 1998]. Em outras palavras, quanto maior a duração ou a quantidade de estímulos, maior será a intensidade do sinal BOLD. Os primeiros experimentos [BOYTON *et al.*, 1996] apontavam para essa relação linear, entretanto, artigos publicados recentemente têm mostrado não linearidade para vários tipos de estímulos [SOLTYSIK *et al.*, 2004], cuja origem é desconhecida.

Nesse sentido, uma estratégia interessante foi proposta por Clare e colaboradores: o método de análise da variância (ANOVA). Ele procura localizar áreas cerebrais ativas sem fazer suposição sobre a forma da HRF [CLARE *et al.*, 1999].

### 8.3 Análise de Variância

O método consiste basicamente do cálculo da razão entre a variância da média do conjunto de dados e a variância do conjunto de dados para cada pixel da imagem. Considerando que cada ponto da série temporal é independente, a razão dos pixels com amplitudes randômicas (não ativação) será de  $1/n$ , em que  $n$  é o número de ciclos. Os pixels ativos, entretanto, terão uma razão significativamente alta, uma vez que a variância da média do conjunto de dados com a resposta BOLD será grande.

Se  $X_{ij}$  é a intensidade do pixel medido, a série temporal pode ser representada como:

		Experimento						
		1	2	...	j	...	n	
	$\tau$	$X_{11}$	$X_{12}$	...	$X_{1j}$	...	$X_{1n}$	$\bar{X}_1$
	$2\tau$	$X_{21}$	$X_{22}$	...	$X_{2j}$	...	$X_{2n}$	$\bar{X}_2$
Tempo	...	...	...	...	...	...	...	...
	$i\tau$	$X_{i1}$	$X_{i2}$	...	$X_{ij}$	...	$X_{in}$	$\bar{X}_i$
	$m\tau$	$X_{m1}$	$X_{m2}$	...	$X_{mj}$	...	$X_{mn}$	$\bar{X}_m$
								$\bar{X}$

em que  $i$  representa o ponto temporal após o estímulo e  $j$  um experimento. Com  $n$  experimentos e  $m$  pontos temporais de intervalo  $\tau$  por experimento. A hipótese nula é que não há diferença significativa entre as variâncias. Isso pode ser testado pela comparação de duas estimativas da variância da popula-

ção,  $\sigma^2$ , baseado sobre a variação das medidas do mesmo ponto temporal, e baseado sobre a variação entre os pontos temporais. A variância de qualquer ponto temporal pode ser calculada como:

$$s_i^2 = \sum_{j=1}^n \frac{(X_{ij} - \bar{X}_i)^2}{n-1}, \quad \dots\dots\dots (2.8)$$

e, então, a variância média dos pontos temporais é dado por

$$\hat{\sigma}_W^2 = \sum_{i=1}^m \frac{s_i^2}{m} = \sum_{i=1}^m \sum_{j=1}^n \frac{(X_{ij} - \bar{X}_i)^2}{m(n-1)}, \quad \dots\dots\dots (2.9)$$

baseado sobre  $m(n - 1)$  graus de liberdade. A variância da média dos pontos temporais é dada por

$$s_{\bar{X}}^2 = \sum_{i=1}^m \frac{(\bar{X}_i - \bar{X})^2}{m-1}. \quad \dots\dots\dots (2.10)$$

E, desde que

$$s_{\bar{X}}^2 = \frac{\sigma^2}{n}, \quad \dots\dots\dots (2.11)$$

$\sigma^2$  pode ser estimado de:

$$\hat{\sigma}_B^2 = n \cdot s_{\bar{X}}^2 = n \cdot \sum_{i=1}^m \frac{(\bar{X}_i - \bar{X})^2}{m-1}, \quad \dots\dots\dots (2.12)$$

no qual está baseada sobre  $m-1$  graus de liberdade. Sobre a hipótese nula,  $\hat{\sigma}_B^2$  e  $\hat{\sigma}_W^2$  são estimativas independentes da variância da população  $\sigma^2$ . Isto significa que a razão

$$F = \frac{\hat{\sigma}_B^2}{\hat{\sigma}_W^2}, \quad \dots\dots\dots (2.13)$$

terá uma distribuição F de Snedecor com  $m - 1$  e  $m(n - 1)$  graus de liberdade. Se existir qualquer mudança no sinal ligado ao estímulo o valor  $\hat{\sigma}_B^2$  será grande e  $H_0$  pode ser rejeitada.

Nesse método, a equação 2.13 forma um mapa paramétrico de teste estatístico F, no qual o valor-p pode ser obtido usando tabelas estatísticas. Calculado o valor de F procura-se o valor crítico,  $F_{(\alpha, m(n-1), m-1)}$ , correspondente. Se  $F < F_{(\alpha, m(n-1), m-1)}$ , a hipótese  $H_0$  não será rejeitada (significa



que a probabilidade de  $F$  é maior que o nível de significância adotado), concluindo que as amostras têm variâncias iguais. Por outro lado, se  $F \geq F_{(\alpha, m(n-1), m-1)}$  rejeitamos  $H_0$ , concluindo que as variâncias comparadas são diferentes.

### 8.4 Modelo Linear Geral

O modelo linear geral, *general linear model* (GLM), consiste de uma grande classe de modelos estatísticos que relaciona uma resposta às combinações lineares de variáveis de um “preditor”. É uma ferramenta estatística robusta e flexível a qual incorpora um grande número de testes estatísticos tradicionais (Teste-t, Correlação, ANOVA, etc.). O GLM, introduzido por Nelder e Wedderburn em 1972 [NELDER & WEDDERBURN, 1972], foi amplamente estudado e utilizado em estudos funcionais de PET e posteriormente estendido a fMRI [FRISTON *et al.*, 1995a].

A essência do GLM é descrever a variável dependente ( $Y$ ) como uma combinação linear de vários fatores modelo ( $X_i$ ) e um termo de erro ( $\epsilon$ ):

$$Y = a_0 + a_1 X_1 + a_2 X_2 + \dots + a_n X_n + \epsilon \quad \dots\dots\dots (2.14)$$

Os parâmetros de peso ( $a_i$ ) são as amplitudes e indicam quanto cada fator contribui para os dados, enquanto o termo  $a_0$  reflete a contribuição de todos os fatores que se mantêm constantes durante o experimento. O GLM considera que os erros são independentes e com distribuição normal [ $N(0, \sigma^2)$ ].

No contexto da fMRI,  $Y$  é a variação da série temporal dos dados observados em um voxel. Na equação do modelo linear, apenas os dados experimentais são conhecidos, pois os fatores modelos representam componentes hipotéticos dos dados. Dado um conjunto específico de fatores modelo e os dados experimentais, é possível calcular qual a melhor combinação de pesos que minimiza o termo de erro. O termo de erro mínimo, após resolver o modelo linear, é conhecido como erro residual. Quando há apenas uma variável dependente, a equação 2.14 é conhecida como um modelo de regressão múltipla univariada. Mas a mesma equação pode ser estendida para incluir um grande número de variáveis dependentes, como os vários pontos temporais de um estudo fMRI, por meio do GLM.

No GLM, os dados experimentais estão representados como uma matriz bidimensional constituída de  $n$  pontos temporais e  $V$  voxels. A estrutura espacial dos dados fMRI não são usados, uma vez que valores dos pesos dos parâmetros e o termo de erro são calculados independentemente para todos os voxels. A *design matrix*, consiste de fatores modelo que representam contribuições possíveis em pre-dizer a forma da resposta BOLD para um dado padrão de estímulo. Ela consiste de  $M$  fatores modelos. A matriz parâmetro contém  $n$  linhas, tal que cada célula indica a amplitude de um dos fatores modelo para um dado voxel. Finalmente, o termo de erro é um vetor com  $n$  linhas. Em alguns sistemas de notação, a *design matrix* é denotada por  $X$  e os parâmetros da matriz são denotados por  $\beta$ . Logo o GLM tem a forma matricial:

$$\begin{pmatrix} Y_1 \\ \vdots \\ Y_n \end{pmatrix} = \begin{pmatrix} X_{11} & \cdots & X_{n1} \\ \vdots & \ddots & \vdots \\ X_{1V} & \cdots & X_{nV} \end{pmatrix} \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{pmatrix}.$$

De uma forma mais sintética:

$$Y = X\beta + \varepsilon. \tag{2.15}$$

Matematicamente, o melhor ajuste para os parâmetros de peso é dado da seguinte maneira:

$$\hat{\beta} = (X^T X)^{-1} X^T Y. \tag{2.16}$$

Para testar a significância do fator modelo em um dado voxel, a amplitude do parâmetro associado é dividida pelo erro residual. A significância é baseada em quão bem os dados experimentais ajustam uma resposta hemodinâmica modelo.

Novamente, as funções respostas hemodinâmicas podem ser modeladas de diversas maneiras. Por exemplo, a função gamma, descrita anteriormente, e o modelo proposto por Friston em 1998 [GLOVER, 1999], que consiste em um modelo fisiológico mais plausível, no qual a diferença de duas funções gamma é computada:

$$h(t) = \left(\frac{t}{d}\right)^a \exp\left(-\frac{t-d}{b}\right) - c \left(\frac{t}{d'}\right)^{a'} \exp\left(-\frac{t-d'}{b'}\right) \tag{2.17}$$

sendo  $d = ab$  é o tempo para o pico,  $d' = a'b'$  é o tempo para o *undershoot*, com  $a = 6$ ,  $a' = 12$  e  $b = b' = 0.9$  s, e  $c = 0,35$ .

Desse modo, assim como na análise de correlação cruzada, a necessidade de informar a forma da HRF constitui a principal desvantagem neste método, pois se esta não for precisamente modelada, o método perderá sensibilidade de detecção.

## 9. Apresentação dos Mapas Estatísticos

Os mapas estatísticos resultantes das diferentes técnicas de análise são apresentados na forma de mapas de cores. Cada voxel que compõe a imagem recebe uma cor que varia de acordo com a sua significância estatística. Em um possível padrão, voxels cuja significância estatística é elevada recebem uma coloração mais próxima do vermelho, enquanto que aqueles de significância estatística desprezível têm coloração azulada. Um exemplo típico de um exame funcional de regiões primárias motoras é apresentado na figura 2.6-A. Nela observamos uma espécie de aglomerado de pontos avermelhados, indicando uma alta probabilidade de ser a região ativa por uma tarefa motora.

Em um segundo momento, os mapas de ativação podem ser apresentados em superposição a imagens que tragam informações anatômicas. A primeira possibilidade é atribuir uma coloração específica, na própria imagem que gerou os mapas estatísticos (p.ex. EPI). Na figura 2.6-B, um desses exemplos é apresentado. O mapa gerado em um paradigma motor é superposto à sua respectiva imagem EPI. Note que apenas regiões mais significativas são apresentadas, nesse caso regiões cujos coeficientes de correlação foram superiores a 0.75.

Infelizmente, as imagens EPI têm baixa resolução espacial e pouco contraste anatômico. Logo, é possível, sobrepor os resultados sobre imagens anatômicas de boa resolução, como as ponderadas em  $T_1$ .

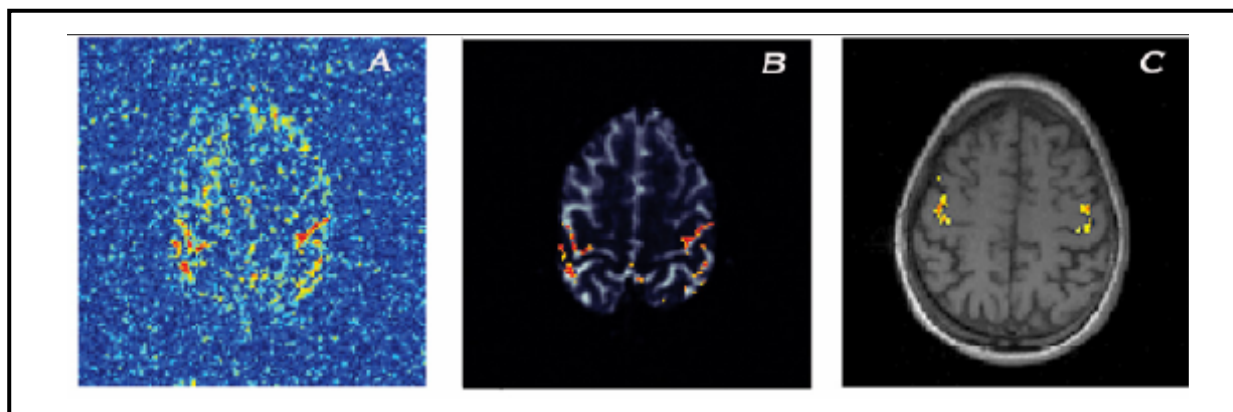


Figura 2.6. Possíveis apresentações de exames de fMRI (a) Mapa estatístico de um experimento com paradigma motor, *analisado* por correlação cruzada. (b) Superposição dos mapas estatísticos à respectiva imagem EPI de um paradigma motor com valor de corte de 0.75. (c) Sobreposição dos resultados a um plano axial de imagem ponderada em  $T_1$ . *Imagem retirada DE ARAÚJO 2002.*

## 10. Cronometria Mental

Uma das possíveis aplicações da fMRI, que tem despertado bastante interesse nos últimos anos, é a cronometria mental. Pesquisas sobre percepção e cognição humana têm sido relacionadas à análise de eventos mentais dentro de estágios de processamento hierárquicos, em que os aspectos temporais destes processamentos são denominados cronometria mental [POSNER (1978 *apud* Menon et al., 1998)]. A cronometria mental é a tentativa de decompor uma tarefa primária cognitiva, ou de percepção, dentro de uma seqüência de estágios de processamento sobre a base dos tempos de respostas medidos [FORMISANO & GOEBEL, 2003].

Inicialmente, estudos de cronometria mental foram baseados exclusivamente sobre análises do tempo de resposta comportamental ou tempo de reação, *Reaction Time* (RT), como uma função da condição da tarefa [DONDEERS (1969 *apud* Formisano & Goebel, 2003)]. Recentemente, a informação do RT comportamental tem sido complementada com medidas invasivas [GEORGOPOULOS & PELLIZZER, 1995], e não invasivas da atividade cerebral, através da eletroencefalografia (EEG) por potenciais evocados, *evoked potencial* (EP) [HILLYARD (1993 *apud* FORMISANO *et al.*, 2002)] e magnetoencefalografia (MEG) [HARI *et al.*, 2000].

Dada a necessidade de boa resolução temporal, as técnicas de EEG e MEG têm sido largamente aplicadas em humanos e, particularmente, usadas para relacionar seqüências de eventos neurais às seqüências de estágios cognitivos na execução de uma tarefa. De fato, o registro de EEG e MEG de respostas relacionadas a eventos reflete mudanças na atividade neuronal sobre uma escala temporal de milisegundos. Contudo, devido a limitações intrínsecas, esses métodos fornecem informações com baixa resolução espacial.

Por outro lado, a fMRI, que fornece informações espaciais mais precisas, conta com mudanças relativamente lentas da resposta hemodinâmica cerebral. Mesmo assim, os recentes avanços da ER-fMRI têm possibilitado mapear a atividade cerebral usando paradigmas similares aos usados com EP, capazes de medir diferenças de tempo entre ativação de diferentes substratos neurais no cérebro humano. Logo, o paradigma ER-fMRI pode ser utilizado para traçar a seqüência de ativação cortical entre regiões cerebrais durante uma tarefa [ROSEN *et al.* 1998]. Resultados recentes [MENON *et al.*, 1999] apontam para a possibilidade de a fMRI traçar eventos seqüenciais com uma resolução temporal de ~100-200 milisegundos. Contudo, a resolução temporal de fMRI é influenciada pela amostragem temporal das respostas hemodinâmicas, caracterizada pelo tempo de repetição (TR), e que depende do equipamento de MRI. A diminuição do valor de TR resulta em uma melhor estimativa da HRF, melhorando a avaliação das suas propriedades temporais, já que a HRF será descrita com um maior número de pontos temporais.

A resolução espaço-temporal da fMRI também está limitada pelas características da resposta hemodinâmica. Recentes estudos mostram que o entendimento sobre a natureza do acoplamento entre a resposta hemodinâmica e o domínio neural permanece vago [LOGOTHETIS, 2002]. Muitos fatores biofísicos não neuronais podem afetar diferentemente o atraso e outros parâmetros das HRFs em diferentes áreas cerebrais, assim dificultando uma interpretação neural das observações com a fMRI. Apesar de incertezas fisiológicas, os resultados de Menon e colegas [MENON *et al.*, 1998], juntamente com outras evidências experimentais [THIERRY *et al.*, 1999; BELLGOWAN *et al.*, 2002; FORMISANO *et al.*, 2002], sustentam a idéia de que estudos de cronometria mental podem ser realizados por fMRI.

Thierry e colaboradores [THIERRY et al. 1999] traçaram a seqüência de ativação cerebral de campos receptivos de linguagem (córtex auditivo e área de Wernicke) para campos expressivos de fala (área de Broca) e regiões de resposta motora (giro pré-central).

Bellgowan e co-autores [BELLGOWAN et al., 2002] estudaram a dinâmica do sistema neural que está envolvida na decisão léxica, por fMRI. Eles mostraram que a estimativa do atraso e largura das HRFs combinadas com uma tarefa apropriada, podem ser usadas para diferenciar rotas neurais no processamento de palavras.

Em um recente estudo, Formisano e colaboradores [FORMISANO et al., 2002] estudaram o papel da diferenciação funcional entre várias áreas corticais relacionadas à tarefa do relógio mental – uma tarefa complexa de imagens mentais visuo-espacial. Neste paradigma, sujeitos eram solicitados a imaginar pares de relógios analógicos, e deviam decidir qual dos dois formava o maior ângulo. A análise resolvida no tempo detectou uma seqüência de ativação cortical da percepção auditiva para a resposta motora com seqüência de ativação do córtex parietal posterior, *posterior parietal cortex* (PPC), esquerdo e direito, sugerindo que estas regiões realizam distintas funções na tarefa de formação de imagens mentais.

Em resumo, a cronometria mental por fMRI está baseada na hipótese de que as seqüências temporais de eventos neurais localizados espacialmente são refletidos em seqüências de respostas BOLD localizados espacialmente [MENON, *et al.*, 1999]. Contudo, esses padrões espaço-temporais da ativação hemodinâmica devem ser detectados usando aquisição e ferramentas de análises que preservem a estrutura temporal dos dados.

De uma perspectiva metodológica, a possibilidade de realizar estudos de cronometria mental usando ER-fMRI depende crucialmente da capacidade de se obter medidas precisas de outras propriedades temporais específicas da HRF, em diferentes regiões do cérebro. Nesse sentido, o atraso temporal para o início da resposta hemodinâmica tem sido usado para resolver etapas do processamento cerebral, e descrever a cronometria mental (Figura 2.7). Mais que isso, via de regra as HRFs estão contaminadas por ruído, tornando ainda mais difícil essa estimativa de atraso.

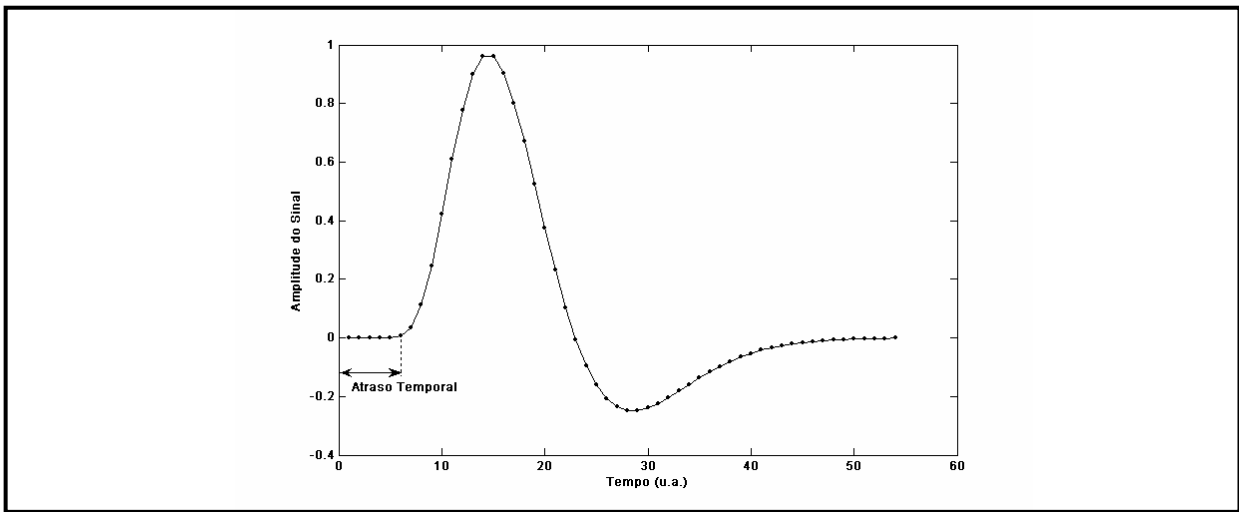


Figura 2.7: Ilustração do parâmetro de atraso temporal para uma função resposta hemodinâmica.

Cuidados adicionais devem ser tomados na interpretação das medidas de atraso entre regiões cerebrais diferentes, bem como na manipulação das séries temporais. No pré-processamento das séries temporais, por exemplo, o realinhamento temporal entre as fatias, assim como a eliminação de *drifts* lineares e não-lineares, deve ser realizado de modo ligeiramente diferente.

Alguns métodos têm sido propostos para estimar o atraso e outros parâmetros das HRFs [SAAD *et al.*, 2001; FORMISANO *et al.*, 2002; HENSON *et al.*, 2001; MIEZIN *et al.*, 2000; FRISTON *et al.*, 1998; PURDON *et al.*, 2001; LIAO *et al.*, 2002]. Por exemplo, a distribuição espacial dos atrasos entre regiões cerebrais pode ser obtida usando a transformada de Hilbert e a correlação cruzada entre os modelos de HRF atrasados e as séries temporais de fMRI [SAAD *et al.*, 2001]. Este método pode consumir muito tempo se diferentes estímulos têm diferentes atrasos. Dentro do contexto do GLM, mapas de latência podem ser obtidos incluindo, na *design matrix*, não apenas um modelo da resposta esperada mas também sua derivada temporal [FRISTON *et al.*, 1998] ou modelos da resposta esperada com atrasos [FORMISANO *et al.*, 2002]. Ainda no contexto do GLM, outras duas propostas recentes fornecem métodos estatísticos gerais e robustos para estimar e fazer inferências sobre latências da HRF [HENSON *et al.*, 2001, LIAO *et al.*, 2002].

Todos os métodos mencionados contam com modelos fixos e predefinidos da HRF, embora a forma precisa da resposta BOLD para uma dada região seja usualmente desconhecida. Como conse-

qüência, para um dado modelo, a análise pode resultar em uma sensibilidade diferente em uma ou outra região do cérebro, dependendo de quão bem a representação do modelo específico da HRF foi realizado.

Uma abordagem complementar no mapeamento do padrão espacial da ativação cerebral é o uso do método de ICA (*Independent Component Analysis*) [MCKEOWN *et al.* 1998]. Este método tem a vantagem de não fazer hipótese alguma sobre o curso temporal das HRFs e pode ser usado para obter uma descrição da seqüência do padrão espacial da ativação cerebral com ER-fMRI em estudos de cronometria mental baseado na fMRI [CALHOUN *et al.*, 2001]. Contudo, além de dispendioso, do ponto de vista computacional, a análise de componentes independentes requer a escolha, a posteriori, das HRFs que indicam a resposta correta. Essa escolha é feita pelo experimentador.



## OBJETIVOS

A presente dissertação tem como objetivos principais:

- Realizar um estudo de parâmetros (número de níveis e de épocas), sobre o método de entropia de Shannon, com intuito de encontrar os valores que apresentam maior detecção de sinal de fMRI;
- Generalizar o método de Shannon, com a entropia de Tsallis. Essa formulação introduz um novo parâmetro,  $q$ , permitindo, então, realizar um estudo capaz de estabelecer intervalos de valores  $q$  com maior poder de detecção da HRF;
- Desenvolver um novo método baseado no cálculo da distância de Kullback-Leibler das séries temporais de paradigmas ER-fMRI. Da mesma forma, neste método realizamos o estudo de parâmetros do número de épocas e número de níveis;
- Comparar os métodos clássicos com os métodos baseados na Teoria de Informação em termos da sensibilidade e especificidade;
- Estudo de detecção de atraso temporal das Funções Respostas Hemodinâmicas no contexto da cronometria mental.

---

## Capítulo III – Material e Métodos

### 1. Introdução

O estudo de parâmetros nos métodos entrópicos foi realizado com dados simulados e a análise com as curvas ROC [Apêndice B]. Esse tipo de procedimento permite um controle sobre os dados de entrada no sistema de detecção definindo exatamente o posicionamento dos voxels ativos e dos não ativos, o que não é possível em dados reais de fMRI.

Testes sobre um novo método de análise devem seguir alguns passos bem determinados para a comprovação de sua funcionalidade e sua caracterização. No caso específico de novos métodos aplicados à análise de dados de fMRI, uma forma de avaliar a qualidade dos resultados vem da comparação da resposta encontrada com as áreas funcionais esperadas. Adicionalmente, para a comprovação da eficiência do método é importante saber como esse se relaciona com os métodos clássicos de análise, o que é realizado com o auxílio de dados simulados.

### 2. Entropia de Shannon aplicada à fMRI

Em 2003, de Araujo e colaboradores, propuseram um novo método para análise de séries temporais ER-fMRI, baseado em um aspecto específico da teoria de informação: o cálculo da entropia de um sinal usando a formulação de Shannon [DE ARAUJO *et al.*, 2003].

Para computar a entropia de Shannon dependente no tempo, pode-se definir estados acessíveis do sistema como a amplitude do sinal dentro de um intervalo específico (janela), centrado em um instante  $t$  [CAPURRO *et al.*, 1999]. De forma mais detalhada, suponha que  $s(t)$  denota a resposta BOLD ao longo de uma época. Além disso, considere que a resposta BOLD de um único voxel, localizado dentro de uma área ativa do cérebro, exibe uma série temporal semelhante à apresentada na figura 3.1. Cada ponto corresponde a um único valor do voxel em função do tempo,  $s(t)$ , para uma época. Durante períodos de ativação, um aumento da resposta BOLD é observado por aproximadamente 10 pontos temporais, e em seguida, retornam à linha de base.

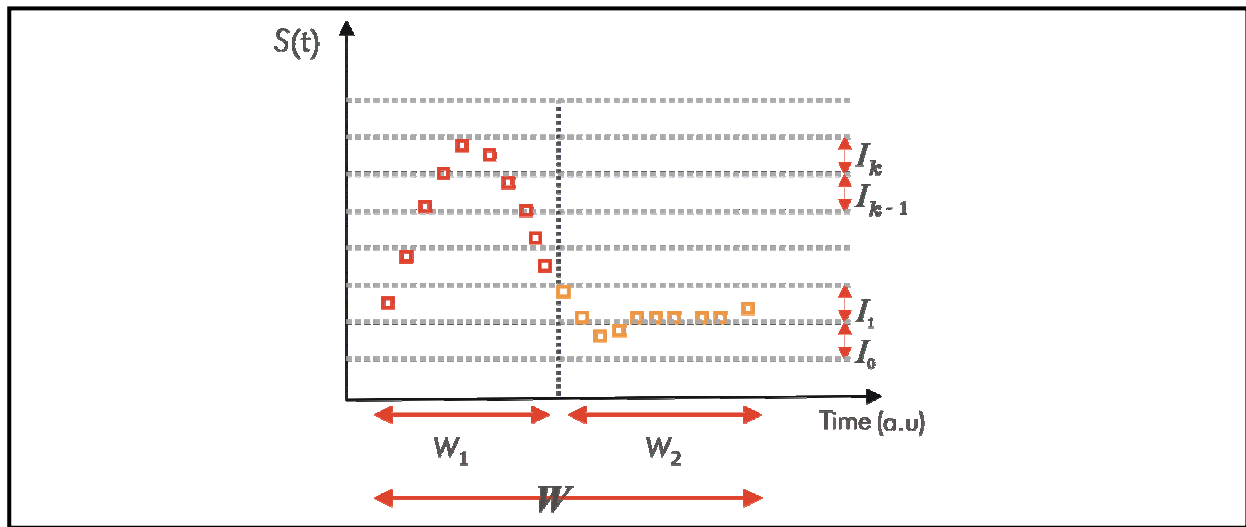


Figura 3.1: Diagrama descritivo da série temporal do sinal BOLD. As probabilidades podem ser obtidas sobre cada janela separadamente ( $W_1$  e  $W_2$ ) e para cada nível  $I_k$ . Imagem retirada de de Araújo *et al.* 2003.

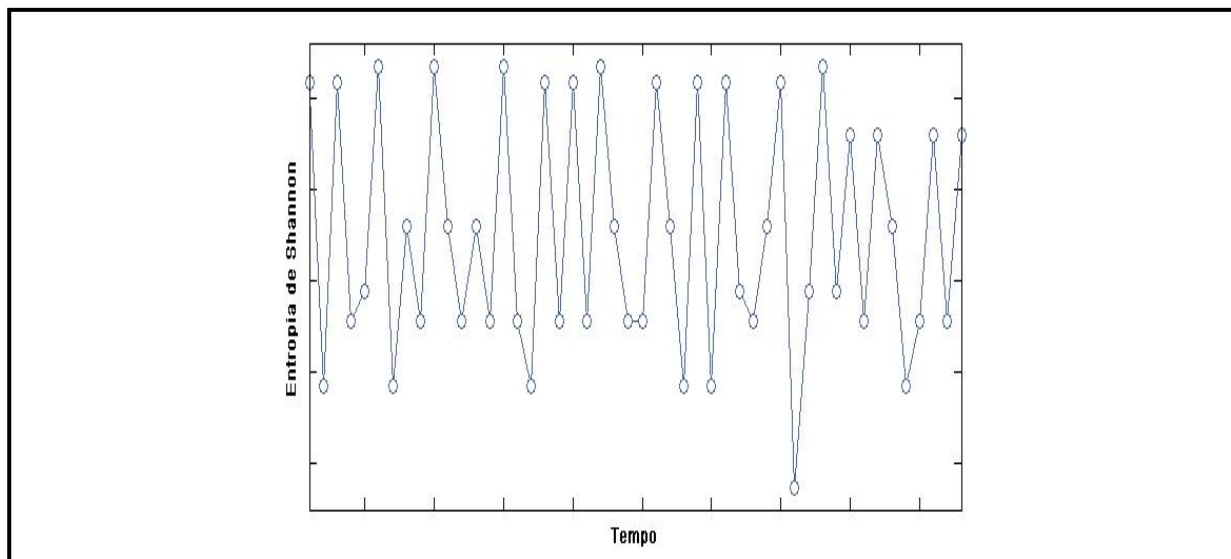
Para calcular a entropia de Shannon de um sinal deve-se, inicialmente, dividir a HRF em duas janelas ( $W_1$  e  $W_2$ ). A primeira está relacionada ao aumento do sinal, enquanto a segunda corresponde principalmente a valores da linha de base. Segundo, a janela ( $W$ ), que compreende uma época inteira, é dividida em um número discreto ( $k$ ) de níveis de amplitude do sinal ( $I_k$ ). Essa divisão é realizada calculando-se o máximo e o mínimo da HRF, dentro de cada época (Método Local). Para calcular a densidade de probabilidade de um nível de amplitude específico  $I_k$  que contém parte do sinal, aplica-se a seguinte relação matemática:

$$p^m(I_k) = \frac{\text{n.º valores de } s(t_k) \in W_m \text{ no intervalo } I_k}{\text{n.º de valores de } s(t_k) \in W_m} \dots\dots\dots (3.1)$$

De posse dessa informação, é possível obter os valores da entropia de Shannon para cada uma das janelas independentemente, de acordo com a expressão 1.1.

O método é baseado na computação independente da entropia de Shannon para os segmentos ( $W_1$  e  $W_2$ ) separadamente, de modo que cada conjunto de pontos,  $s(t_k) \in W_1$ , tem um valor de entropia. Intuitivamente, a primeira porção do sinal, na qual o estímulo está presente, deve fornecer altos valores de entropia, superiores à contida em  $W_2$ , por se tratar de um sistema mais desorganizado. Em regiões onde existir o predomínio de atividade em resposta a uma determinada tarefa, haverá um pa-

drão que descreve algo semelhante a uma função do tipo “dente de serra” (ver figura 3.2). O algoritmo é concluído aplicando-se uma correlação cruzada entre uma função dente de serra simulada e os valores de evolução temporal das entropias de Shannon calculadas. É possível então gerar um mapa estatístico com valores de coeficientes de correlação que se referem às áreas cerebrais com maior chance de estarem ativas.



**Figura 3.2: Evolução Temporal da entropia de Shannon calculada em cada janela (W1 e W2). Períodos com entropia alta correspondem regiões do sinal dominadas pelo período de crescimento do sinal BOLD, enquanto valores de baixa entropia são relacionados à linha de base.**

### 3. Entropia de Tsallis aplicada à fMRI

Este método é uma generalização do método de entropia de Shannon dependente do tempo. Todas as etapas, até o cálculo da densidade de probabilidade de um nível de amplitude específico,  $I_k$ , são realizadas da mesma maneira que no método de Shannon. Contudo, em seguida, calcula-se a entropia de Tsallis, ao invés da de Shannon, para cada uma das janelas independentemente, através da expressão 1.2.

Finalmente, aplica-se a correlação cruzada entre uma função dente de serra e a série que descreve a evolução temporal da entropia de Tsallis. Do mesmo modo, os coeficientes de correlação resultantes geram um mapa estatístico que se refere à atividade.

#### 4. Distância de Kullback-Leibler aplicada à fMRI

Nesta sessão, propõe-se um novo método para a análise de séries temporais ER-fMRI baseado em um aspecto específico da Teoria de Informação: a distância entre duas distribuições de probabilidades de um sinal usando a distância KL [COVER & THOMAS, 1991], em que nenhuma hipótese é feita sobre a forma da HRF.

O método consiste da computação da distância KL derivada das séries temporais geradas em paradigmas ER-fMRI. Considere que  $s(t)$  denota a resposta BOLD de um único voxel, de uma área funcional ativa (Figura 3.1). Durante os períodos de ativação (10 pontos temporais) há um aumento da resposta BOLD e durante períodos de repouso os pontos temporais retornam para a linha de base. Para calcular a distância KL dependente no tempo, definimos estados acessíveis ao sistema dentro de uma janela deslizante ( $W$ ), que corresponde a uma época inteira, centrado em um instante  $t$  (Método Local). Inicialmente divide-se  $W$  em duas janelas ( $W_1$  e  $W_2$ ), do mesmo modo que foi feito anteriormente. Calculado o valor máximo e o mínimo da época, definem-se os níveis acessíveis do sistema pela divisão da janela ( $W$ ) em um número discreto de *bins* ( $L$ ) de amplitude do sinal ( $I_k$ ).

Para computar a distância KL, calculam-se duas densidades de probabilidades de um nível de amplitude específico ( $I_k$ ) que contém parte do sinal:

$$p^m(I_k) = \frac{\text{n.º valores de } s(t_k) \in W_m \text{ no intervalo } I_k}{\text{n.º de valores de } s(t_k) \in W_m}, \dots\dots\dots (3.2)$$

e,

$$q^m(I_k) = \frac{\text{n.º valores de } s(t_k) \in W_{m+1} \text{ no intervalo } I_k}{\text{n.º de valores de } s(t_k) \in W_{m+1}} \dots\dots\dots (3.3)$$

a primeira  $p^m(I_k)$  é calculada na janela  $W_1$  e a segunda,  $q^m(I_k)$ , é calculada sobre a janela  $W_2$ .

De posse dessas informações é possível obter os valores das distâncias KL para cada uma das janelas  $W$ , em que  $W = W_1 + W_2$ , independentemente [COVER & THOMAS, 1991]:

$$D_S^m(p||q) = \sum_{k=1}^L p_k^m(I_k) \log_2 \frac{p_k^m(I_k)}{q_k^m(I_k)}; \dots\dots\dots (3.4)$$

Finalmente o método é concluído realizando-se a medida das distâncias KL da série temporal:

$$D_S(p||q) = \frac{1}{m} \sum_{k=1}^m D_{S_k} \dots\dots\dots (3.5)$$

com  $m = 1, 2, 3, \dots, n^\circ$  de épocas. Intuitivamente, um voxel que apresenta a resposta BOLD tem em sua primeira porção valores de densidade de probabilidades mais uniformemente distribuídos, muito diferente das encontradas em  $W_2$ , cujas densidades de probabilidades são mais localizadas. Dessa forma, as distâncias KL serão grandes em regiões onde haja predomínio de atividade. Por outro lado, voxels em regiões não-ativas tem distribuições de probabilidade semelhantes em ambas as janelas, apresentando distâncias menores.

Portanto, para cada série temporal associa-se uma distância KL. É possível, então, gerar um mapa estatístico que, após um valor de corte, pode ser apresentado superposto às imagens EPI ou a imagens de boa resolução anatômica (MPR). Para o valor de corte, utilizou-se uma estimativa da distância, aquela que fosse maior que dois desvios padrões ( $D > 2\sigma$ ).

### 5. Método de Detecção de Atraso da HRF

Durante uma experiência de fMRI, o estímulo é repetido  $N_S$  vezes em um período  $T_S$  com intervalo  $\Delta t$  entre as aquisições das imagens. Os dados experimentais são descritos pela série temporal  $f_n \equiv f(n\Delta t)$  com  $N_P = T_S / \Delta t$ , e  $1 \leq n \leq N_S N_P$ . Em que  $N_P$  é o número de pontos temporais de uma época. Assumido que os dados são compostos pelo sinal  $s_n$  e um ruído  $\eta_n$  com desvio padrão  $\sigma_{\eta,n}$ , a série temporal pode ser escrita na forma:

$$f_n = s_n + \eta_n \cdot \dots\dots\dots (3.6)$$

Pretende-se investigar a latência para o início do sinal BOLD gerado pelo estímulo. Em um primeiro momento, o sinal pode ser estimado pela média dos  $N_S$  ciclos:

$$S_n = \frac{1}{N_S} \sum_{m=1}^{N_S} f_{n+mN_P} = s_n + \xi_n \dots\dots\dots (3.7)$$

Outra informação relevante fornecida pelo sinal é a sua derivada  $f'_n$  que pode ser estimada por:

$$f(t)'_n = \frac{f_n(t + \Delta t) - f_n(t)}{\Delta t} = s'_n + \eta'_n, \dots\dots\dots (3.8)$$

com  $\Delta t$  igual a 2 e  $t = 1 \dots N_S N_P$ . Calcula-se a média da derivada  $f'_n$  dos  $N_S$  ciclos:

$$S'_n = \frac{1}{N_S} \sum_{m=1}^{N_S} f'_{n+mN_P} = s'_n + \xi'_n \dots\dots\dots (3.9)$$

Se cada conjunto  $\eta_{n+mN_P}$ , para um dado  $n$ , é independente e identicamente distribuído (IID), o teorema do Limite Central, *central limit theorem* (CLT) pode ser aplicado. Essa condição é alcançada se as flutuações individuais não ultrapassam tempos maiores que  $T_s$ , e o ruído é uniforme. De acordo com o CLT, o desvio padrão de  $S_n$  é  $\sigma_{S_n} = \sigma_{f_n} / \sqrt{N_S}$ . Ele será estimado por:

$$\sigma_{S_n}^2 = \frac{1}{N_S(N_S - 1)} \sum_{m=1}^{N_S} [f_{n+mN_P} - S_n]^2 \dots\dots\dots (3.10)$$

Da mesma maneira, o desvio padrão de  $S'_n$  é  $\sigma'_{S_n} = \sigma'_{f'_n} / \sqrt{N_S}$  e é estimado por:

$$\sigma'^2_{S_n} = \frac{1}{N_S(N_S - 1)} \sum_{m=1}^{N_S} [f'_{n+mN_P} - S'_n]^2 \dots\dots\dots (3.11)$$

Sabe-se que  $S_n / \sigma_{S_n}^2$  e  $S'_n / \sigma'^2_{S_n}$  seguem a distribuição de Student com  $N_S-1$  graus de liberdade que será denotado por  $P_{N_S-1}(t)$  e  $P'_{N_S-1}(t')$  respectivamente. No contexto de um teste de hipótese, isso é descrito pelo valor de probabilidade (valor-p):

$$p_{S_n} = 2 \int_t^\infty P_{N_S-1}(t) dt \dots\dots\dots (3.12)$$

e

$$p_{S'_n} = 2 \int_t^\infty P'_{N_{S-1}}(t') dt' \dots\dots\dots (3.13)$$

Sendo  $p_{S_n}$  e  $p_{S'_n}$  o valor-p do sinal e da derivada, respectivamente e o valor 2 multiplicativo aparece uma vez que o teste é bi-caudal. Com esses valores, a probabilidade combinada pode ser calculada da seguinte maneira:

$$p_{S_n;S'_n} = \prod_{n=S_n}^{S'_n} p_n \dots\dots\dots (3.14)$$

$p_{S_n;S'_n}$  é a probabilidade de um ruído ocorrer com módulo maior do que  $t$ , quando  $s_n = 0$ . Em outras palavras, define-se um valor  $\alpha_I$ , denominado limiar individual, tal que se  $p_{S_n;S'_n} < \alpha_I$ , um sinal está presente. Se, na realidade, não há sinal,  $\alpha_I$  é a probabilidade de assumir, erroneamente, que um sinal está presente.

Com o intuito de decidir se um aumento contínuo do sinal está presente, aplicam-se os seguintes critérios:

1. Procura-se pelo maior número de pontos consecutivos da série temporal tal que  $p_{S_n;S'_n} < \alpha_I$ , para todos os pontos;
2. Considera-se um aumento contínuo do sinal, em um cluster delimitado por  $N_1 \dots N_n$ , se  $p_{N_1 \dots N_n} < \alpha_C$ , onde:

$$p_{N_1 \dots N_n} = \prod_{m=N_1}^{N_n} (p_{S_n;S'_n})_m \dots\dots\dots (3.15)$$

em que  $\alpha_C$  é o limiar contíguo. Uma vez satisfeito esses critérios, em uma dada série temporal, utiliza-se o primeiro ponto desse conjunto contíguo  $N_1$  para definir  $t = N_1 \Delta t$ , como sendo o tempo de início do sinal, *signal starting time* (SST), ou seja, o início do sinal BOLD.



## 6. Dados Simulados

A avaliação do desempenho de um método passa pela estimativa da sua sensibilidade e especificidade. Na maioria das técnicas isso é realizado por comparação com outros métodos, considerados padrões – *Gold Standard*. Entretanto, diante da dificuldade de se obter um teste padrão para a fMRI, é rotineiro se estudar a sensibilidade e especificidade por meio de dados simulados. Para tanto, neste trabalho, foram gerados artificialmente conjuntos de dados 3-D, de 64x64 voxels com séries temporais de HRFs repetidas no tempo, de 22 pontos temporais cada época, seguindo àquela proposta por Friston (2.17).

Da coluna de 1 até 32 estão representados voxels com padrão de ativação, ou seja, contendo um sinal BOLD, e a coluna 32 até 64 só contém ruído, descrito por um padrão Gaussiano uniforme. O ruído foi adicionado a esses traçados com a SNR variando entre -15 dB e 15 dB, tendo sido calculada pela seguinte expressão:

$$\text{SNR} = 10 \log \left[ \frac{\sigma_S^2}{\sigma_R^2} \right], \quad \dots\dots\dots (3.16)$$

em que  $\sigma_S^2$  e  $\sigma_R^2$  são a variância do sinal e do ruído, respectivamente. Na figura 1.4 são apresentadas as séries temporais de um único voxel, que contém uma HRF simulada, para diferentes valores de SNR.

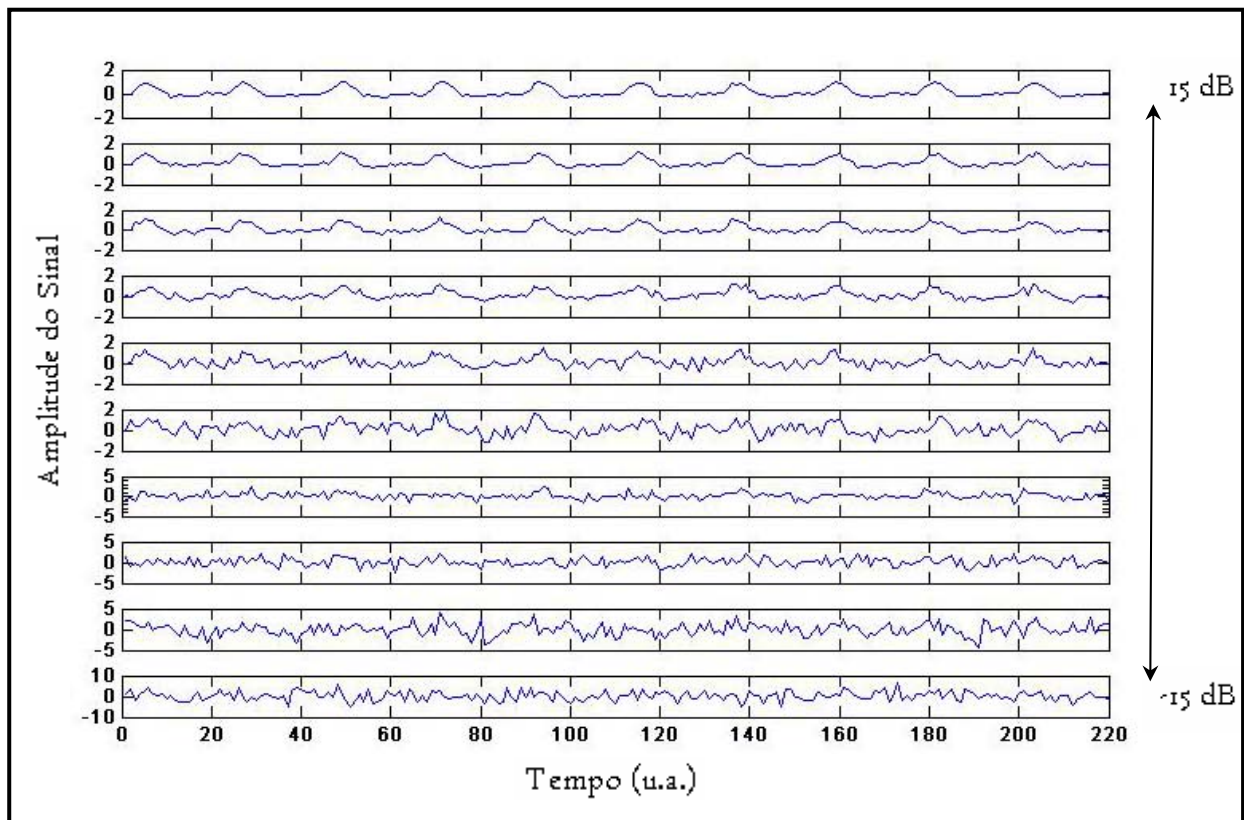


Figura 1.4.: Uma resposta hemodinâmica simulada com SNRs variando de 15 dB a -15 dB

A performance de detecção foi avaliada pelo cálculo da área abaixo da curva ROC [Apêndice B], gerada após o processo de detecção. O desempenho dos métodos propostos neste trabalho foi comparado aos resultados obtidos pelos métodos convencionais (correlação cruzada e teste-t), quando aplicados ao mesmo conjunto de dados simulados.

## 7. Dados Reais

### 7.1 Voluntários

Os métodos também foram testados sobre dois conjuntos de dados reais. No primeiro conjunto, uma tarefa motora foi realizada por 9 voluntários assintomáticos, destros. No segundo conjunto, de estímulo duplo, 6 voluntários participaram do estudo, todos destros, adultos (5 homens e uma mulher: média das idades 30,2 anos).

## 7.2 Estímulos

O primeiro experimento foi desenhado com um paradigma ER, que consistiu de um estímulo motor (*finger-tapping*) da mão direita, com duração de 3 segundos, seguido por um período de repouso de 20 segundos. O experimento contém 24 épocas. Cinco fatias axiais foram amostradas 14 vezes no tempo, ao longo da duração de cada época, três durante os estados de atividade, e outras onze durante os períodos de repouso.

No segundo experimento, dois estímulos foram apresentados, um em seguida do outro: o primeiro visual e o segundo motor. Para o estímulo visual, foi apresentado um retângulo oscilante, tipo tabuleiro (cinza e preto), com uma taxa de cintilação de 2 Hz, durante 3 segundos. Logo em seguida, era solicitado aos voluntários que realizassem uma tarefa motora unilateral (*finger tapping* da mão direita), que durava também 3 segundos. Após a exposição aos estímulos, os voluntários permaneciam em repouso por um período de 24 segundos. O TR possibilitou o registro de seis imagens durante os estados de atividade, e outras doze durante o repouso, resultando um total de 18 imagens para cada época. Cada estudo era formado por um total de 10 épocas.

Os estímulos foram desenvolvidos em um programa específico: Presentation® 0.60 (Neuro-behavioral Systems), e projetados por um canhão de projeção (Infocus Systems, modelo LitePro 210) sobre uma tela translúcida que estava localizada aos pés do voluntário. O estímulo era visualizado pelo voluntário, através de um espelho acoplado à bobina de cabeça.

## 7.3 Aquisição das Imagens

Todas as imagens foram adquiridas em um scanner de 1.5 T (*Siemens, Magnetom Vision*) com bobina de quadratura transmissora/receptora de cabeça e polarização circular comercialmente disponível (em atividade regular no HCFMRP – Serviço de Radiodiagnóstico).

No primeiro estudo, cinco fatias axiais, foram posicionadas sobre porções do córtex motor primário. Os parâmetros das imagens foram: TR = 1,68 s; TE = 118 ms; matriz = 128 x 128, FOV =

210 mm; dimensão do voxel = 1,64 x 1,64 x 4,00 mm; e número total de 336 imagens por fatia, dividida em 24 épocas de 14 imagens cada.

No segundo estudo, foram adquiridos 185 volumes, correspondendo às 10 épocas de 18 imagens cada. As cinco primeiras imagens foram descartadas para permitir o efeito de saturação de T1. Cada volume continha 16 fatias axiais, cujos parâmetros são: TR = 1,68 s; TE = 66 ms; matriz = 64 x 64, FOV = 220 mm; *flip angle* (FA) = 90°; espessura da fatia = 5 mm.

Em ambos os estudos, imagens com alta resolução espacial do tipo MPR foram adquiridas (*Multipplanar Reconstruction*), para a eventual necessidade de sobreposição dos mapas estatísticos. Essas imagens têm os seguintes parâmetros de aquisição: TR = 9.7 ms, TE = 4.0 ms, FA = 12°, FOV 256 mm, matriz 256 x 256 largura da fatia = 1 mm e número de fatias = 156.

## 7.4 Análise dos Dados

O pré-processamento das imagens foi realizado no programa Brain Voyager<sup>TM</sup> QX (Versão 1.3.8; Brain Innovation, Maastricht, Holanda). Nesta etapa foi realizada a correção de movimento utilizando interpolação tri-linear, a correção do tempo entre fatia, a suavização espacial (através de um Filtro Gaussiano com FWHM de 4 mm) e também a filtragem temporal por um filtro passa alta 3 ciclos/s.

As análises dos dados simulados e pós-processamento das imagens reais, através dos métodos descritos anteriormente, foram realizadas utilizando algoritmos implementados em MATLAB 6.5 (Mathworks, Sherborn, MA).

---

## Capítulo IV – Resultados e Discussão

### 1. Introdução

Neste capítulo, são apresentados os resultados dos estudos conduzidos com os métodos de entropia de Shannon, entropia de Tsallis e KL-fMRI sobre dados simulados e dados reais. Estes métodos são comparados aos métodos clássicos, como a correlação e teste-t de Student. O estudo da Especificidade e Sensibilidade foi realizado com auxílio das curvas ROC [Apêndice B] em dados simulados. Além disso, foi calculada a área abaixo da curva ROC, pelo método de trapézio, que traz informação sobre o poder discriminante em um teste de diagnóstico.

### 2. Entropia de Shannon e Paradigma ER-Motor

No estudo realizado anteriormente, que fazia uso da entropia de Shannon, os números de níveis e de épocas foram mantidos constantes, e iguais a 3 e 24, respectivamente. Desse modo, para testar variantes do método, realizou-se um estudo para avaliar sua eficiência sob diferentes condições do número de parâmetros.

O objetivo principal desse estudo é encontrar os parâmetros que melhor respondam à detecção de voxels ativos em dados simulados. Para tanto, calculou-se a área abaixo da curva ROC, em um intervalo de 2 até 5 níveis em função do número de épocas, que variou de 5 até 35 épocas. O resultado desse processamento é apresentado na figura 4.1.

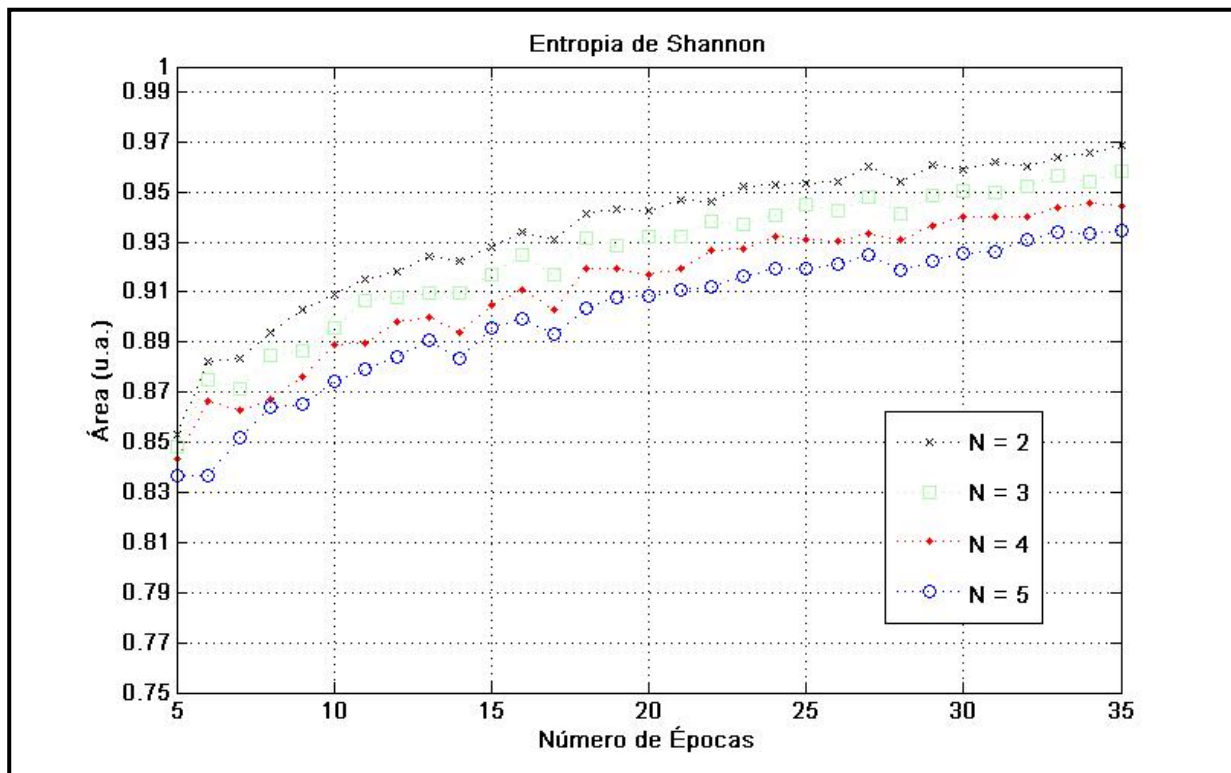


Figura 4.1. Gráfico do comportamento da área ROC em função do número de épocas e do número de níveis, no método de Shannon.

A curva com maior área abaixo da curva ROC representa o maior poder discriminante do teste. Logo, conforme figura 4.1, conclui-se que o método de Shannon apresentará melhor desempenho se a probabilidade de distribuição for calculada com dois níveis.

Aumentando o número de níveis, verifica-se um aumento da entropia de Shannon correspondente a linha de base do sinal BOLD. Isso pode ser verificado na figura 4.2, em que é mostrada a evolução temporal da entropia de Shannon de um voxel simulado com SNR = 15 dB. Em voxels ativos, com baixa SNR, a entropia na janela W2 tenderá a ser maior quanto maior o número de níveis, diminuindo a correlação com a função dente de serra e, então, diminuindo o poder de detecção do sinal. Observa-se também que a área abaixo da curva cresce monotonicamente com o aumento do número de épocas. Esse comportamento já era esperado, uma vez que aumentando o número de épocas, aumentamos o tamanho da amostra das entropias melhorando, assim, sua distribuição.

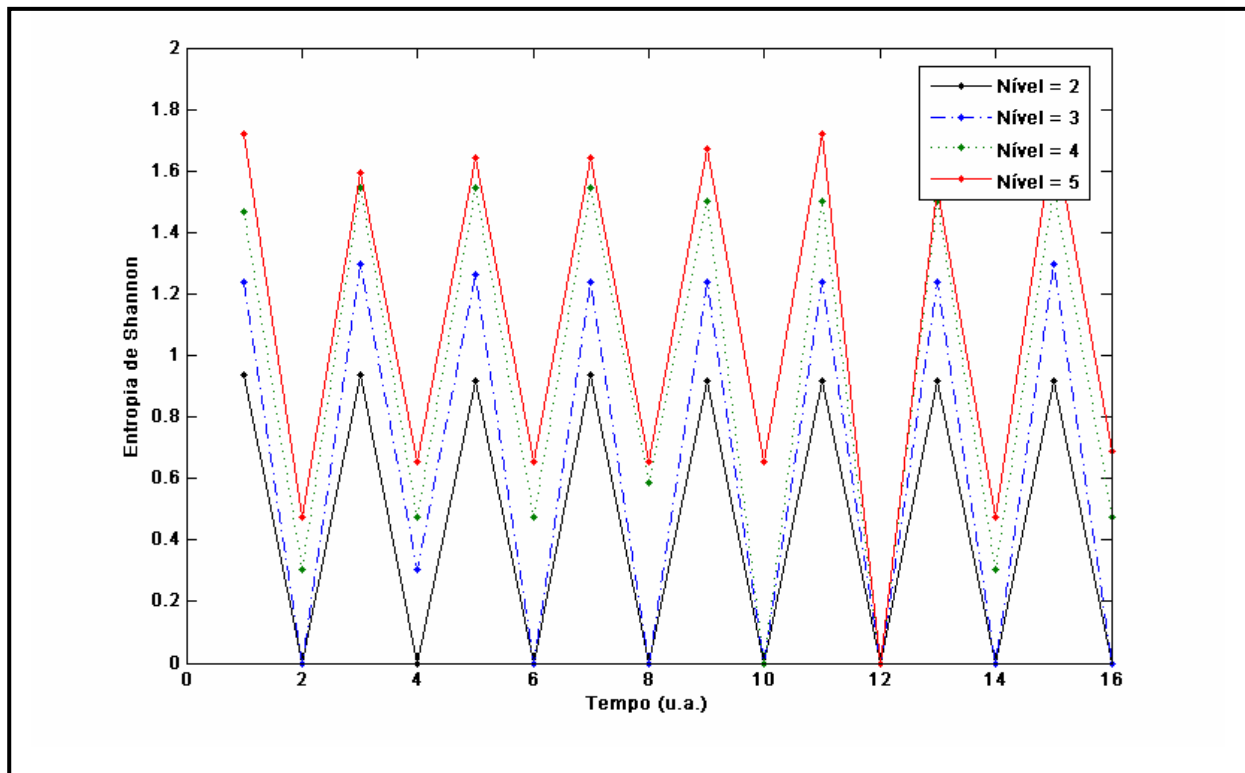


Figura 4.2: Comparação entre as evoluções temporais da entropia de Shannon quando processadas com números de níveis igual a 2, 3, 4 e 5.

Em mais detalhes, a tabela 4.1 mostra os valores das áreas ROC, em função do número de épocas, para dois níveis. De acordo com a tabela, não há um critério de seleção para um número de épocas apropriado, uma vez que o poder de detecção cresce com o aumento do número de épocas. É importante notar que, do ponto de vista prático, quanto maior o número de épocas, mais longa a duração dos exames.

Geralmente os métodos clássicos de análise são aplicados sobre estudos com poucas séries temporais. Aqui, considerou-se um estudo com 5 épocas. Os métodos de correlação e teste-t apresentaram áreas ROC de 0,991 e 0,971 (Tabela 4.2), respectivamente, para dados simulados com 5 épocas. Observa-se, então, que o método de Shannon exige um número maior de épocas se comparado com métodos clássicos. Contudo, sendo que não há grande variação na área ROC depois da 23ª época, e está próxima das áreas encontradas nos métodos clássicos, adotou-se esse número de épocas para a comparação entre os métodos.

Tabela 4.1 – Áreas Roc calculadas para um intervalo de 5 até 35 épocas utilizando dois níveis.

Número de Épocas	Área	Número de Épocas	Área
5	0.8532	21	0.9465
6	0.8822	22	0.9464
7	0.8838	23	0.9523
8	0.8939	24	0.9531
9	0.9032	25	0.9537
10	0.9089	26	0.9539
11	0.9151	27	0.9600
12	0.9180	28	0.9539
13	0.9245	29	0.9605
14	0.9222	30	0.9587
15	0.9282	31	0.9618
16	0.9337	32	0.9600
17	0.9308	33	0.9638
18	0.9411	34	0.9654
19	0.9434	35	0.9686
20	0.9428		

### 3. Entropia de Tsallis

Além dos parâmetros estudados no método de Shannon, a entropia de Tsallis permite a introdução de um novo parâmetro:  $q$ . Neste estudo, também foi calculada a área abaixo da curva ROC em um intervalo de 2 a 5 níveis e 5 até 35 épocas, para valores de  $q$  variando de -4 a 4. Varrendo as três regiões distintas, ou seja,  $q < 0$ ,  $0 < q < 1$  e  $q > 1$ . Os gráficos que resultaram desse processamento são apresentados na figura 4.3.



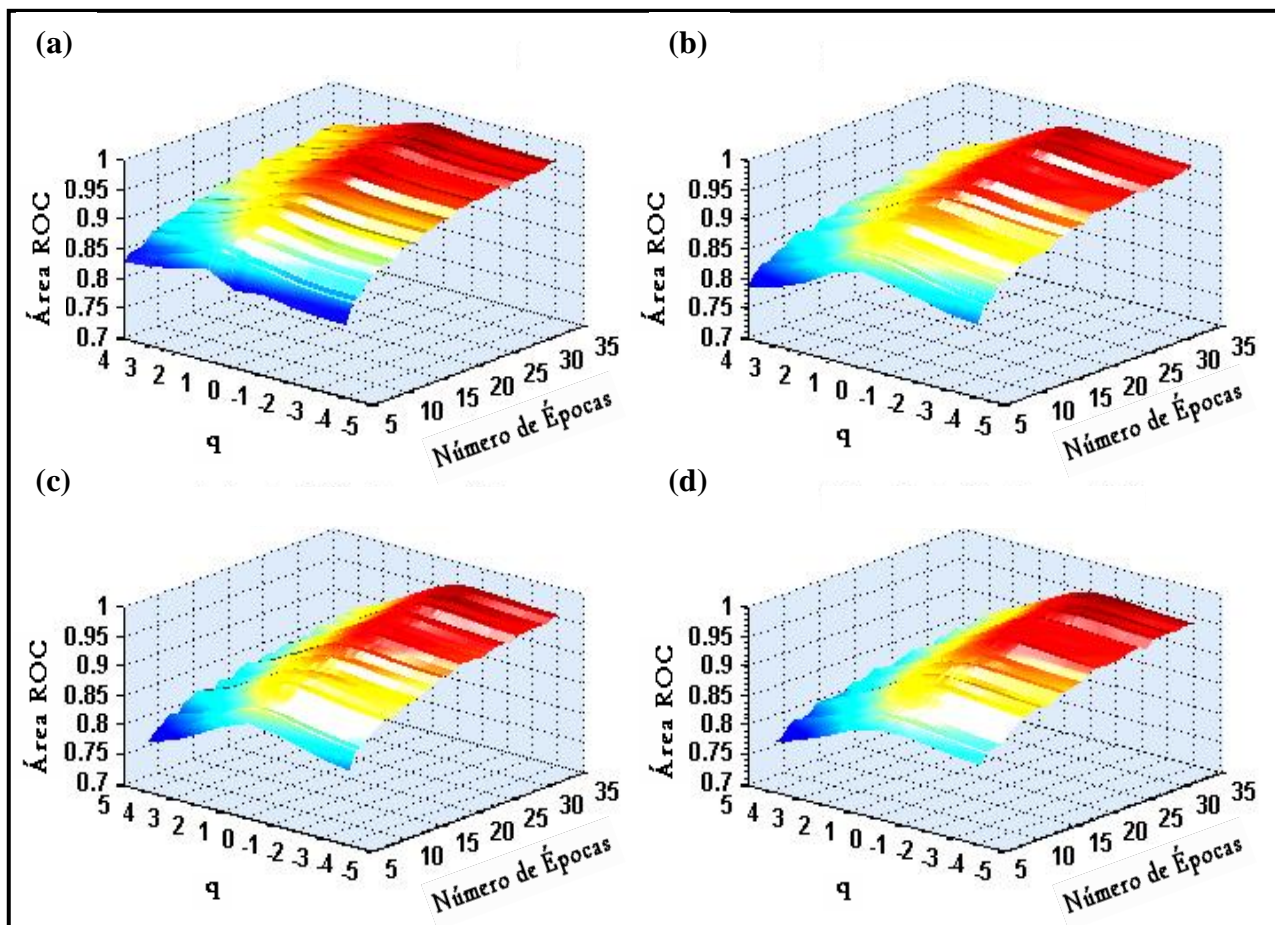


Figura 4.3. Gráfico do comportamento da área ROC em função do número de épocas, número de níveis e parâmetro  $q$  no método de entropia de Tsallis (a) número de níveis = 2, (b) número de níveis = 3, (c) número de níveis = 4 e (d) número de níveis = 5.

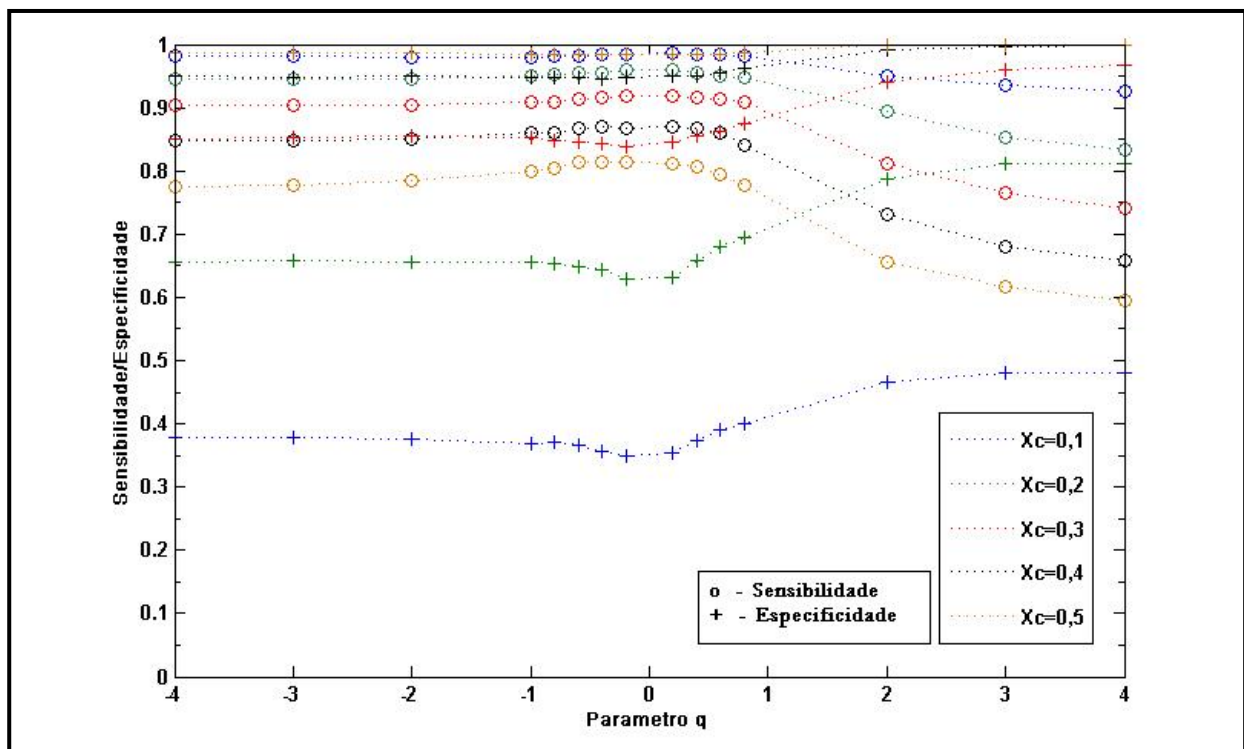
Os gráficos da figura 4.3 mostram que a área ROC diminui com o aumento do número de níveis e que apresenta maior poder discriminante quando a probabilidade de distribuição é calculada com dois níveis. Além disso, similar ao método de Shannon, a área cresce monotonicamente com o aumento do número de épocas, mas, neste caso, a entropia é dependente do parâmetro  $q$ .

Dentre a faixa dos valores de  $q$  testados, observa-se que há um pico no intervalo de  $-1 < q < 1$  ( $q \neq 0$ ), na qual a entropia de Tsallis apresenta os melhores resultados. A qualidade da detecção diminui para valores de  $q < -1$  e  $q > 1$ , restringindo, assim, o intervalo de valores de  $q$  que podem ser usados. Note que esses valores são correlacionados com a curva teórica (Figura 1.2).

É importante uma inspeção mais cautelosa dos padrões de sensibilidade e especificidade do método proposto. A figura 4.4 apresenta essa dependência, em função de  $q$ , extraída das curvas ROC

com 23 épocas e dois níveis. Observa-se uma dependência do parâmetro  $q$  e o limiar  $X_c$ , com respeito à sensibilidade e à especificidade. Entretanto, a faixa de valores de  $q$  que apresenta maior relação sensibilidade e especificidade localiza-se entre  $-1 < q < 1$  ( $q \neq 0$ ).

Dependendo do objetivo do estudo de fMRI, as aproximações devem ser conservadoras, e enfatizar a redução dos falsos positivos (maior especificidade). Isso significa que o valor de corte ( $X_c$ ) deve ser alto para garantir o máximo de especificidade possível. Desse modo, conforme gráfico da figura 4.4, um  $X_c > 0,9$  garantiria o máximo de especificidade, no intervalo de  $q$  otimizado. Por outro lado quanto maior o  $X_c$  maior a taxa de falsos negativos. Logo, dentro da faixa de  $q$  estabelecida, e com um valor de  $X_c > 0,80$  é garantida uma maior sensibilidade. Pode-se, ainda, optar por um valor de corte menos restrito, por exemplo,  $X_c > 0,5$ . Nesses termos, o método fornecerá, ainda, uma especificidade razoável com um aumento da sensibilidade considerável.



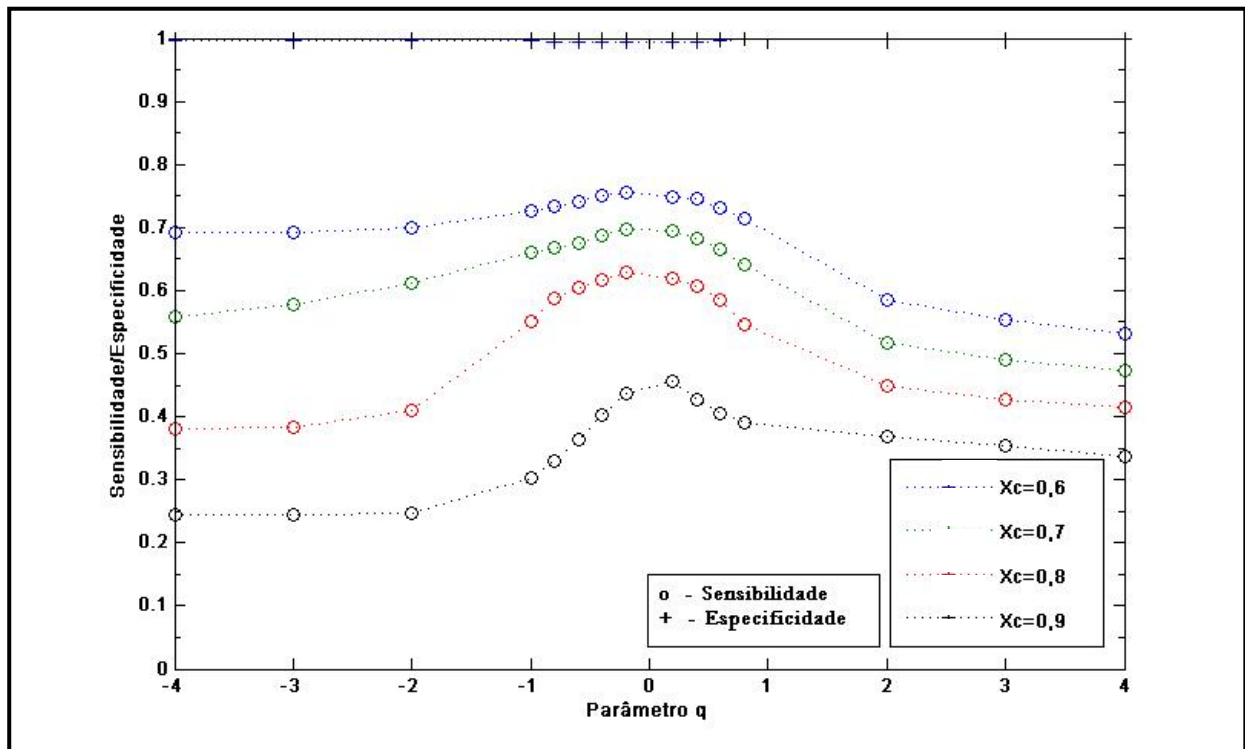


Figura 4.4: Dependência do parâmetro  $q$  de Tsallis com o valor de corte  $X_c$ . Acima, a dependência da sensibilidade/especificidade, em função do parâmetro  $q$ , para valores de  $X_c$  variando entre 0,1 e 0,5. Abaixo, a mesma dependência, para  $X_c$  entre 0,6 e 0,9.

#### 4. Distância de Kullback-Leibler

Usando a mesma estratégia anterior, aplicou-se o método de KL-fMRI sobre imagens funcionais 3-D artificiais para encontrar os números de nível e números de época mais adequados. O resultado desse estudo é apresentado na figura 4.5. Nessa figura, apresentamos a dependência da área sob a curva ROC em função do número de épocas, para diferentes números de níveis.

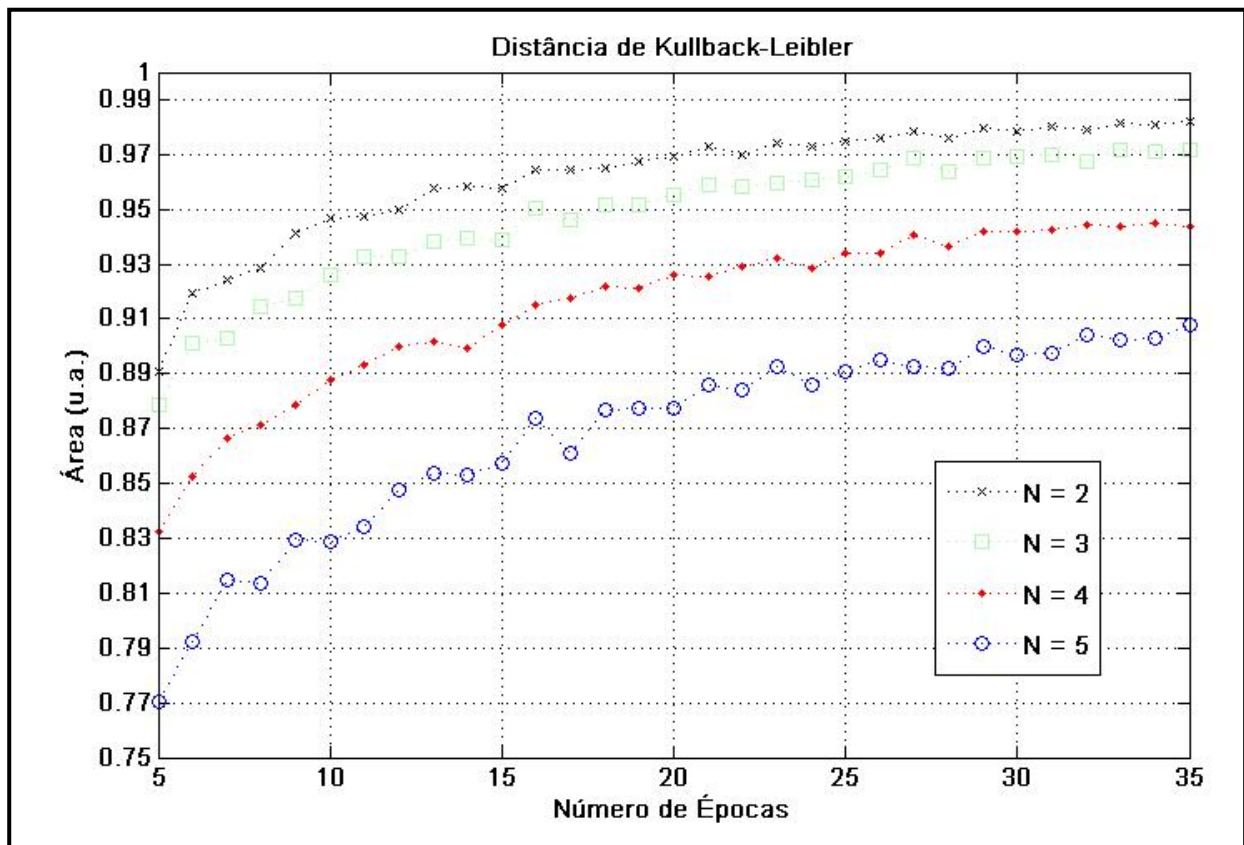


Figura 4.5: Gráfico do comportamento da área sob as curvas ROC em função do número de épocas e número de níveis, no método de KL-fMRI, aplicado a dados simulados.

Conforme o gráfico da figura 4.5, o método KL-fMRI apresentará maior poder de detecção no mapeamento de áreas ativas se a probabilidade de distribuição for calculada com dois níveis. Novamente, observa-se que a área ROC cresce monotonicamente com o aumento do número de épocas. Contudo, a área ROC apresenta maior intensidade se comparada com o método de Shannon e Tsallis.

O método KL-fMRI apresenta vantagem sobre os métodos de Shannon e Tsallis, por não precisar calcular a correlação cruzada, reduzindo assim as etapas de processamento necessárias para a análise e o tempo computacional.

## 5. Curvas ROC

Uma vez testadas as condições ideais de funcionamento sobre todos os métodos, passamos a avaliar seu poder de detecção e comparar com os métodos clássicos. Para tanto, todos eles foram apli-

cados sobre as matrizes simuladas, sob diferentes condições de sinal ruído. As fatias simuladas tinham as colunas 1-32 com ruído e as colunas 32-64 ativação. Ruído gaussiano foi adicionado com SNR diminuindo de 15 dB a -15 dB, linha a linha. Os resultados estão mostrados na figura 4.6. No eixo y apresentam-se os valores descritivos de cada método. Para a entropia de Tsallis utilizou-se o valor de  $q=0,20$  (Figura 4.6-d).

A análise de correlação cruzada (Figura 4.6-a) apresenta uma sensível diminuição nos valores descritivos dos voxels ativos com o aumento do ruído. O teste-t de Student (Figura 4.6-b) também apresenta diminuição dos níveis descritivos dos voxels ativos com a diminuição da SNR, porém, em alguns casos há um aumento dos valores de t nos voxels inativos. Métodos determinísticos, como a correlação cruzada, tendem a gerar falsos negativos em casos de baixa SNR, por outro lado métodos estatísticos tendem a gerar falsos positivos que podem contribuir erroneamente na determinação de regiões ativas [MULLER *et al.*, 2002]. Assim como o método de Shannon (Fig. 4.6-c), a entropia de Tsallis (Fig. 4.6-d), aparentemente, apresenta uma diminuição menos acentuada nos valores descritivos dos voxels ativos com a diminuição da SNR, evidenciando a possibilidade de maior sensibilidade para valores de corte mais restritos. Do mesmo modo, o mapa estatístico processado com o método KL-fMRI (Fig. 4.6-e) mostra uma extensa região de voxels ativos com altos valores descritivos, diminuindo lentamente com o aumento do ruído. Por outro lado, há um aumento bastante acentuado nos valores descritivos dos voxels inativos, o que pode levar a um aumento na taxa de falsos positivos com um limiar menos restrito.

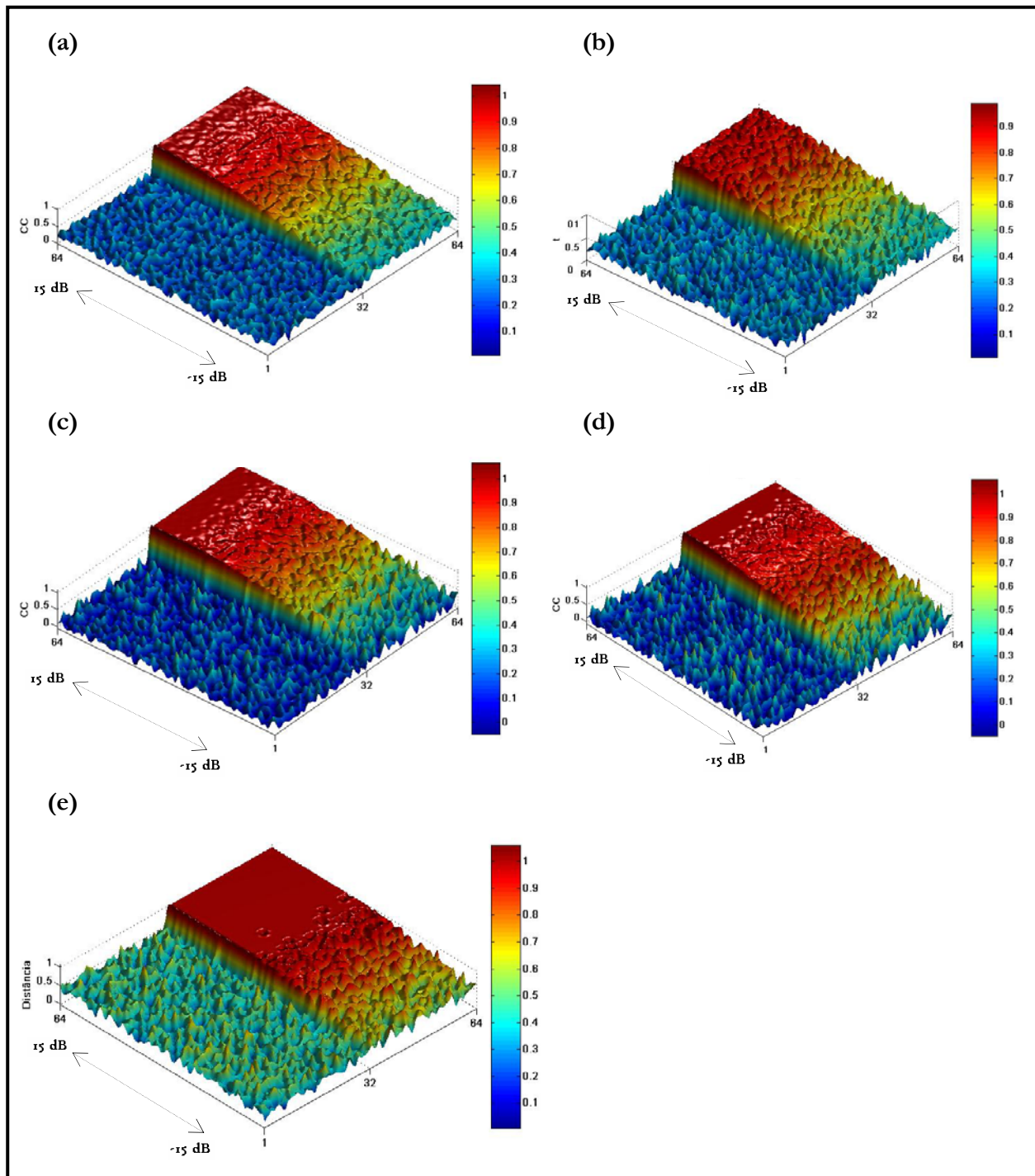


Figura 4.6. Resultados da análise funcional para dados simulados. SNR diminui de -15 até 15 dB. Metade da direita: presença de sinal simulado BOLD, metade da esquerda: somente ruído (a) resultado para análise de correlação cruzada, (b) resultado para teste-t de Student, (c) entropia de Shannon, (d) entropia de Tsallis com  $q = 0,20$  e (e) KL-fMRI.

Uma comparação quantitativa foi realizada baseada na análise das curvas ROC. As curvas foram construídas com o mesmo número de pontos, e  $X_c$  variando de 0 até 1, de maneira a permitir estu-

dos em uma região de interesse. Na figura 4.7 são apresentadas as curvas ROC para os resultados das simulações da figura 4.6. Quanto mais próximo a curva ROC se encontra do canto superior esquerdo do gráfico, maior o poder de detecção.

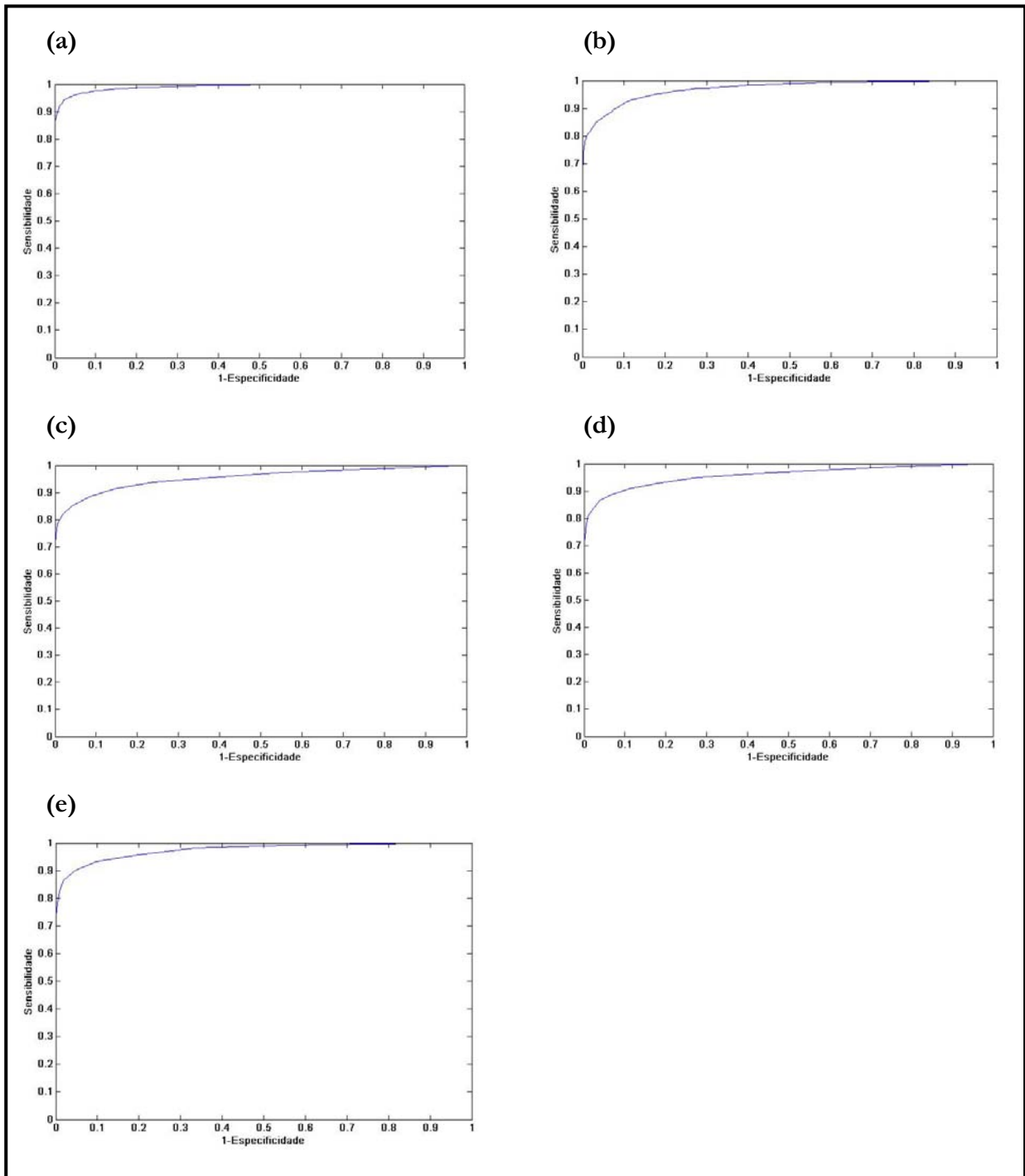


Figura 4.7. Curvas ROC para os métodos aplicados nos conjuntos de dados simulados: (a) análise de correlação, (b) teste-t de Student (c) entropia de Shannon, (d) entropia de Tsallis com  $q=0,20$  e (e) KL-fMRI.

Dado a semelhança entre os gráficos, a análise quantitativa e comparativa entre os métodos foi realizada pelo cálculo dos índices TPF, TNF, FPF e FNF, no ponto ótimo de operação ( $p_0$ ) e área abaixo da curva ROC em cada caso. Os índices TPF e TNF indicam, respectivamente, a sensibilidade e a especificidade do método, ou seja, encontrar sinais em regiões onde apenas sinal esta presente ou encontrar ruído em regiões onde somente ruído esta presente. O ponto ótimo de operação consiste no valor de corte que apresenta maior relação entre a sensibilidade e especificidade. A tabela 4.2 mostra os resultados para cada método processado sobre as dados simulados.

**Tabela 4.2 – Índices quantitativos da curva ROC para o ponto ótimo de operação dos métodos clássicos e entrópicos.**

Método	Área ROC	$X_{c_{p_0}}$	Valor Estatístico	Nível de Significância	TPF <sub>p<sub>0</sub></sub>	TNF <sub>p<sub>0</sub></sub>	FPF <sub>p<sub>0</sub></sub>	FNF <sub>p<sub>0</sub></sub>
Correlação Cruzada	0.991	0,397	CC = 0,20	$p < 0,0362$	0,942	0,978	0,022	0,058
Teste-T Student	0.971	0,382	$t = 2,56$	$p < 0,0115$	0,893	0,926	0,074	0,107
Entropia de Shannon	0.953	0,361	CC = 0,361	$p < 0,0138$	0,863	0,941	0,059	0,137
Entropia Tsallis $q=0,2$	0.959	0,422	CC = 0,422	$p < 0,00349$	0,865	0,956	0,044	0,135
KL-fMRI	0.974	0,583	D = 63,30		0,891	0,958	0,042	0,109

Pela tabela 4.2, verifica-se a eficiência do método de correlação cruzada. Por outro lado, o método de entropia de Shannon apresentou maior especificidade em relação ao teste-t de Student com um valor-p menos estrito ( $p < 0,0138$ ). Entretanto a sensibilidade do teste-t de Student foi superior ao encontrado pelo método de entropia de Shannon. Da mesma maneira, o método de entropia de Tsallis, apresentou maior especificidade e menor sensibilidade em relação ao método de teste-t de Student, porém com um valor-p mais estrito ( $p < 0,00349$ ). O método KL-fMRI apresentou especificidade muito próxima ao método de entropia de Tsallis, porém, mostrou maior sensibilidade equivalente ao Teste-t de Student. Cabe notar que a eficiência da correlação cruzada (CC) vem, em grande parte, do fato de que a HRF simulada é a mesma usada de referência pela CC.

A tabela 4.3 apresenta valores de TPF, TNF, FPF e FNF para uma região de  $X > 0,80$ , valor esse estrito, para avaliar a maior especificidade do método. Observa-se que nenhum método apresentou falso positivo. Entretanto, os métodos de Entropia de Tsallis e KL-fMRI fornecem sensibilidades mai-



ores quando comparados com os métodos clássicos, dados que os valores-p são menos estritos ( $p < 2,6 \times 10^{-11}$ ). Além disso, nestas condições, o método KL-fMRI, apresentou sensibilidade superior sobre todos os métodos estudados.

**Tabela 4.3 – Índices quantitativos da curva ROC para um valor de corte estrito ( $X_c > 0,80$ ) dos métodos clássicos e entrópicos.**

Método	Estatística	Valor-p	TPF	TNF	FPF	FNF
Correlação Cruzada	CC = 0,726	$2,9 \times 10^{-19}$	0.492	1,00	0,00	0,508
Teste-T Student	t = 10,15	$2,1 \times 10^{-17}$	0.335	1,00	0,00	0,665
Entropia de Shannon	CC = 0,80	$2,6 \times 10^{-11}$	0.515	1,00	0,00	0,485
Entropia Tsallis	CC = 0,80	$2,6 \times 10^{-11}$	0.620	1,00	0,00	0,380
KL-fMRI	D = 82,4		0.697	1,00	0,00	0,303

## 6. Comportamento dos Métodos em Relação a SNR

Em fMRI, a SNR exerce grande influencia na detecção do sinal BOLD. Entretanto, utilizando dados simulados, pretende-se avaliar o comportamento dos métodos clássicos e dos métodos de entropia de Shannon, Tsallis e KL-fMRI, nas condições ideais de funcionamento, quando a SNR diminui. Neste sentido, imagens compostas com SNRs de -15 dB até 15 dB, com variação de 1 dB para cada imagem, foram geradas de acordo com a descrição anterior. Valores de corte ( $X_c$ ) de 0 a 1 foram aplicados a estas imagens para estimar as taxas correspondentes de falsos positivos e verdadeiros positivos. Finalmente, a área abaixo da curva ROC foi estimada.

Os resultados são apresentados na Figura 4.8. As figuras 4.8-(a)-(e) apresentam o relacionamento entre voxels classificados como verdadeiros positivos e falsos positivos em uma simulação, para diferentes SNRs. Nas simulações, há 2048 voxels contendo um sinal BOLD mais ruído, e 2048 voxels somente com ruído. Em todos os métodos observa-se que a diminuição da SNR aumenta consideravelmente o número de falsos positivos, em regiões de pequena probabilidade de falsos positivos (especificidade próxima a um). Como resultado, as SNRs das imagens devem ser grandes suficientes para diminuir o número de falsos positivos em relação ao número de verdadeiros positivos.

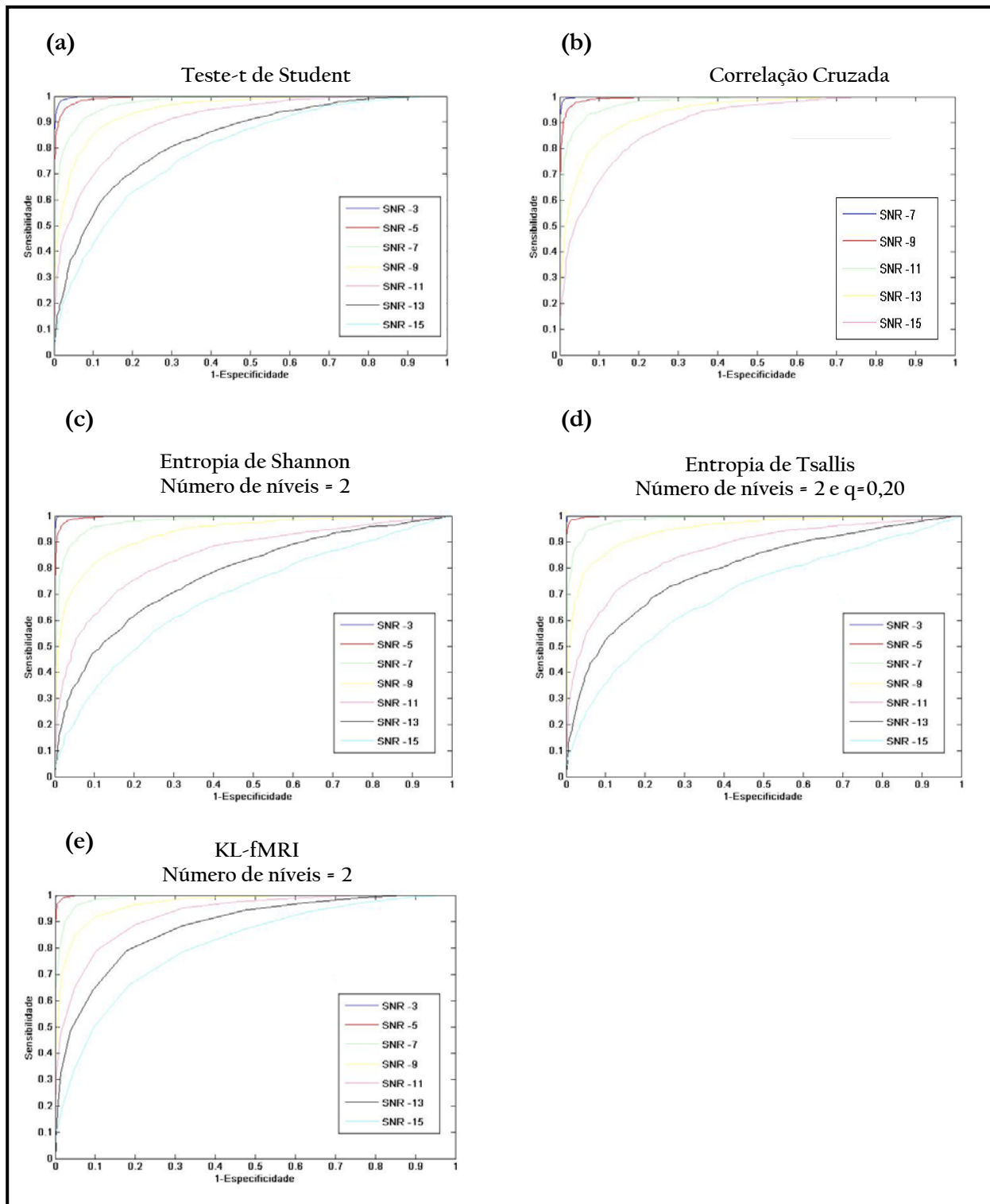


Figura 4.8. Ilustração da relação entre as medidas de SNR e curvas ROC usando dados simulados. (a)-(e) Curvas ROC para diferentes SNRs entre -15 dB a 15 dB para os métodos estudados. Os pontos das curvas foram calculados aplicando-se valores de corte para a correspondente simulação, e o número de pontos de verdadeiros positivos e falsos positivos foi avaliado.

A figura 4.9 apresenta a relação entre a área abaixo das curvas ROC (Fig. 4.8), e a SNR das imagens processadas com os métodos clássicos e entrópicos. Conforme SNR aumenta, a área ROC tende para seu valor máximo em todos os métodos.

Comparativamente, a análise de correlação cruzada teve o melhor desempenho dentre os métodos analisados. Por outro lado, esse método supõe que a HRF pode ser padronizada [VAZQUEZ *et al.*, 1998]. Ainda mais neste caso, em que os dados foram simulados a partir de uma função conhecida, e que é a mesma utilizada de referência para cálculo da correlação cruzada. Daí boa parte da eficiência do método neste caso. Contudo, o método de KL-fMRI teve melhor desempenho sobre os métodos de entropia de Shannon, Tsallis e teste-t de Student, quando a SNR dos dados era diminuída.

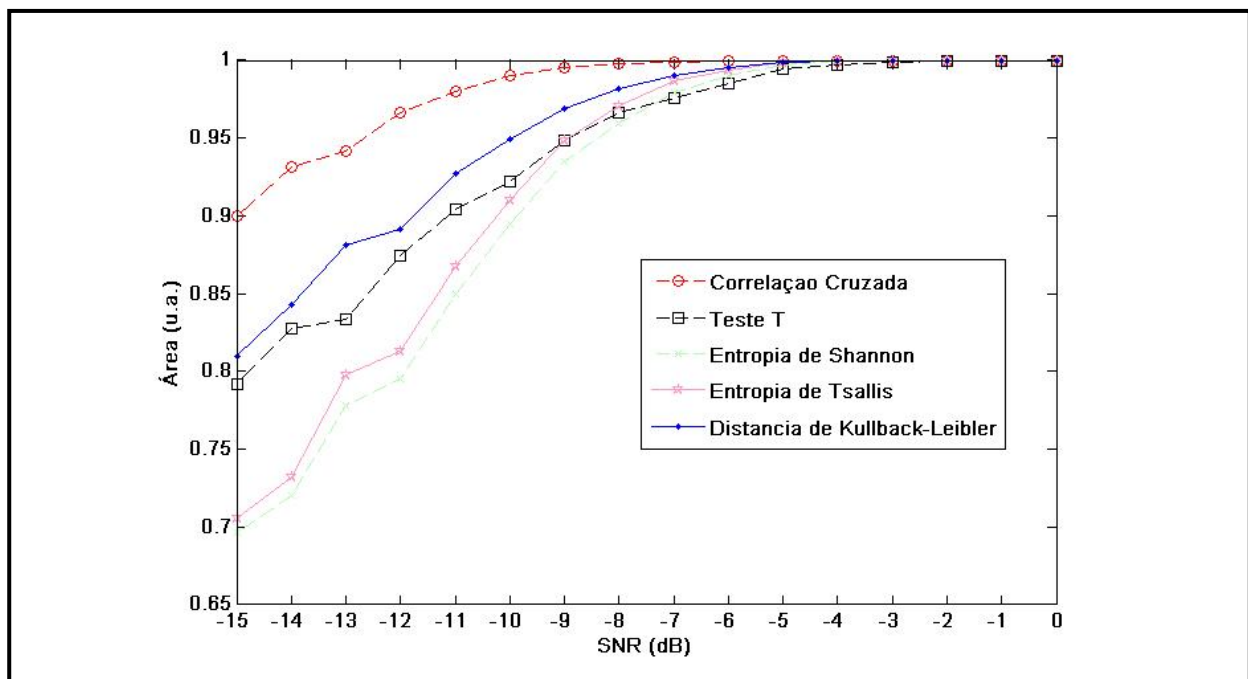


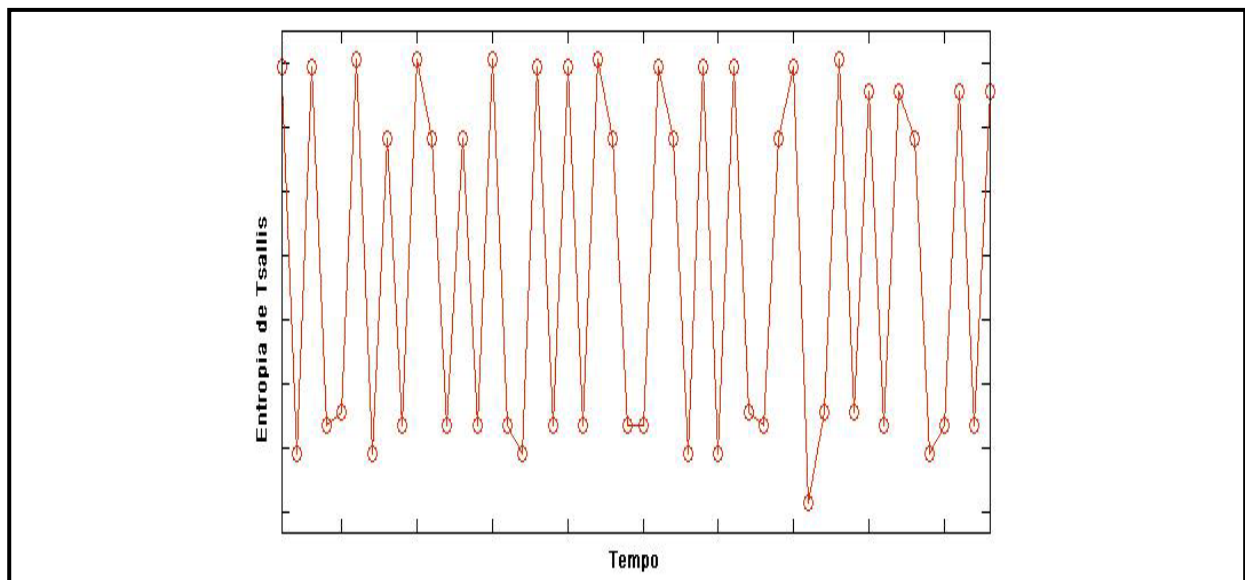
Figura 4.9: Gráfico comparativo. Área abaixo da Curva ROC para cada SNR da imagem. Entropia de Shannon; Entropia de Tsallis com  $q=0,20$  e KL-fMRI.

## 7. Análise de Dados Reais

Partindo da análise simulada foi possível caracterizar os métodos e saber qual a região de maior eficiência de cada um, sob diferentes condições. Desse modo, essa sessão tem como objetivo testar os métodos de entropia de Tsallis e KL-fMRI em dados reais de fMRI. Os métodos foram testados nos

dois paradigmas descritos anteriormente, ambos ER-fMRI: uma tarefa motora e uma tarefa dupla, motora e visual.

A evolução temporal da entropia de Tsallis, calculada para cada voxel da série temporal de imagens, é apresentada na figura 4.10. Como era de se esperar, a entropia tem um comportamento diferente em cada uma das duas metades. O gráfico resultante tem a característica de uma onda dente de serra, para os voxels que foram ativados, padrão que não se repete para voxels que são constituídos apenas de ruído.



**Figura 4.10:** Evolução Temporal da entropia de Tsallis ( $q=0,2$ ) calculada em cada janela ( $W_1$  e  $W_2$ ), Períodos com entropia alta correspondem a regiões do sinal dominadas pelo BOLD, enquanto valores de baixa entropia são relacionados a estados de repouso.

A figura 4.11 mostra os mapas estatísticos das imagens funcionais reais, originados de uma tarefa motora bilateral das mãos, computados pelos métodos de correlação cruzada (Fig. 4.11-a e b), teste-t de Student (Fig. 4.11-c), entropia de Shannon com  $N = 2$  (Fig. 4.11-d), entropia de Tsallis com  $N = 2$  e  $q = 0,20$  (Fig. 4.11-e) e KL-fMRI com  $N = 2$  (Fig. 4.11-f). A análise de correlação foi realizada com duas HRF simuladas: a função gamma e a função de Friston.

A figura 4.12 apresenta os mapas funcionais obtidos por todos os métodos, após aplicação de limiares (*thresholds*) e sobrepostos às respectivas imagens EPIs. Os limiares foram escolhidos de modo a que somente áreas estatisticamente significativas fossem mostradas.

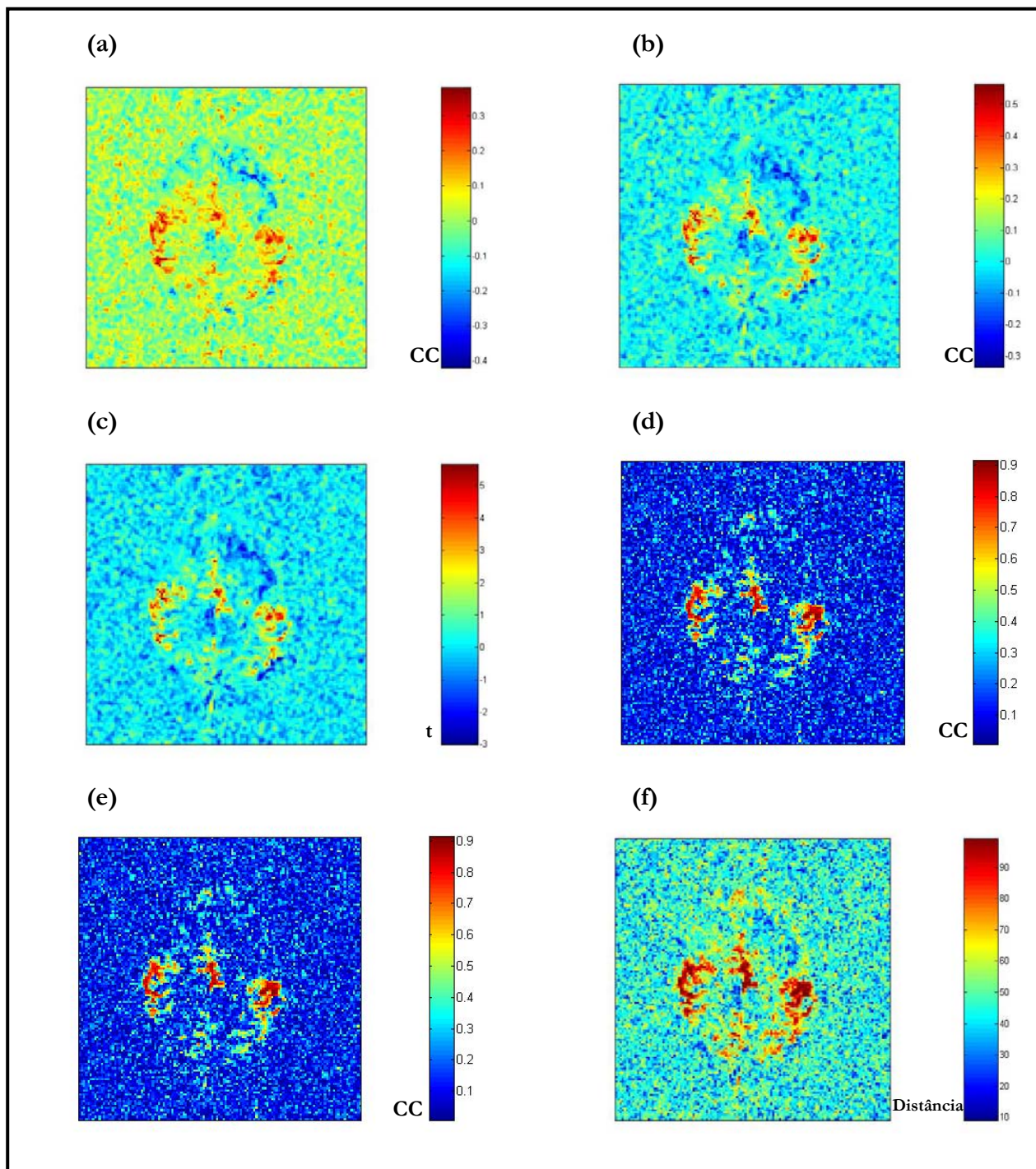


Figura 4.11: Mapas funcionais obtidos pelos métodos (a) análise de correção cruzada, usando a HRF de Friston, (b) análise de correlação usando uma gamma amortecida (c) teste-t de Student, (d) entropia de Shannon, (e) entropia de Tsallis e (f) KL-fMRI.

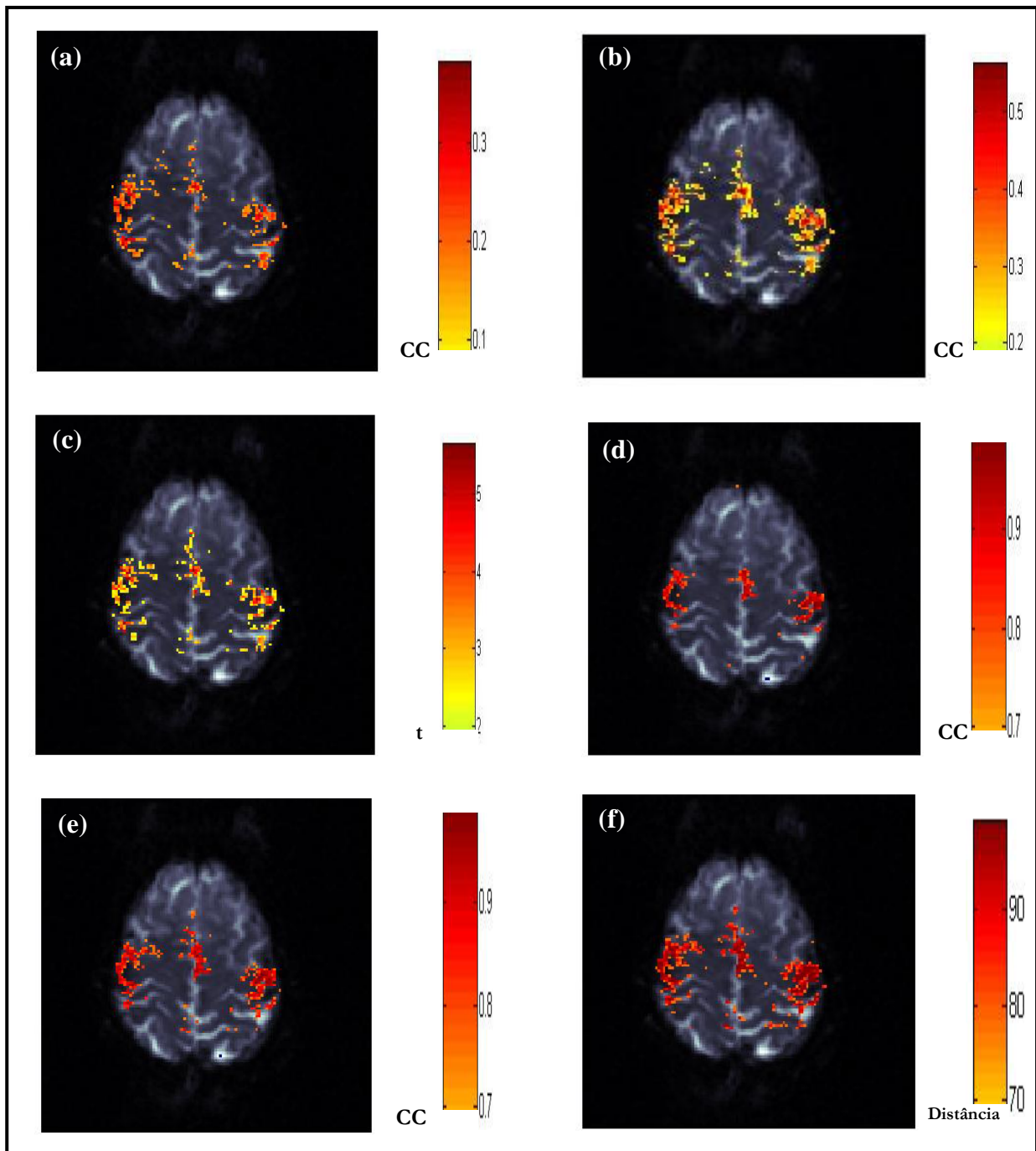


Figura 4.12: Mapas funcionais sobrepostos a imagens EPIs obtidos pelos métodos (a) análise de correção cruzada com a HRF Friston ( $CC > 0,16$ ;  $p < 0,00327$ ), (b) análise de correção cruzada com uma HRF gamma ( $CC > 0,21$ ;  $p < 0,000147$ ), (c) teste-t de Student ( $t > 2,22$ ;  $p < 0,027$ ), (d) entropia de Shannon ( $CC > 0,7$ ;  $p < 0,00000003$ ), (e) entropia de Tsallis ( $CC > 0,7$ ;  $p = 0,00000003$ ) e (f) KL-fMRI ( $(D(2\sigma) > 73.3125)$ ).

Todos os métodos revelam padrões de atividade sobre porções laterais do giro pré-central. Regiões motoras suplementares também apresentam significância estatística, implicando que, em uma primeira análise, os métodos entrópicos são eficientes. Embora todos os métodos sejam capazes de

localizar atividade sobre áreas correspondentes à tarefa motora, observa-se que as análises clássicas, análise de correlação e teste-t de Student, apresentaram baixa significância estatística se comparados aos métodos de Shannon, Tsallis e KL-fMRI.

Uma inspeção das séries temporais identificou a presença de artefatos do tipo “*drift*”, que, como já discutido anteriormente, pode diminuir a sensibilidade dos métodos na detecção dos sinais. Na figura 4.13-a, apresenta-se um exemplo dessa constatação, sobre uma série temporal de um voxel ativo da imagem real (Fig. 4.12).

A remoção da flutuação foi realizada pela subtração entre os pontos temporais da serie temporal e uma reta linear, ajustada por mínimos quadrados, linha azul da figura 4.13-a. A figura 4.13-a, mostra a série temporal de um voxel ativo da imagem real com o ajuste linear, linha azul, e o respectivo histograma do sinal, que consiste na distribuição do sinal nos estados ativos e repouso. A figura 4.13-b mostra a série temporal do voxel e o histograma da série temporal com a remoção do *drift*. Note que a série temporal com a remoção do *drift* linear, ainda apresenta artefatos *drifts* de outras ordens polinômiais. Contudo, embora ainda haja *drift*, a separação dos histogramas foi melhorada.

Para melhor testar a influência desse artefato sobre os resultados obtidos, a partir dos vários métodos, uma HRF foi simulada utilizando a equação de Friston, contendo 10 épocas de 14 pontos temporais cada, SNR = -5 dB e com flutuação linear. O sinal simulado foi processado por todos os métodos, objetivando verificar a variação da significância estatística quando o sinal estava com e sem a flutuação linear. A figura 4.14 mostra o sinal simulado antes e depois da retirada do *drift* linear. A tabela 4.4 apresenta os resultados dos métodos estatísticos sobre os dados simulados com e sem o *drift* linear. Como pode ser verificado, os métodos clássicos têm seus níveis descritivos muito influenciados pela flutuação linear, uma vez que seus valores melhoraram significativamente após a remoção da flutuação linear. Por outro lado, os métodos de entropia de Shannon, entropia de Tsallis e KL-fMRI, não sofreram nenhuma influência. Este fato pode ser explicado pela utilização da janela deslizante, em que todo o processamento é realizado localmente, ou seja, em cada época. A remoção do *drift* linear é importante

nos métodos clássicos, uma vez que permite utilizar um valor de corte mais estrito, diminuindo a probabilidade de falsos positivos.

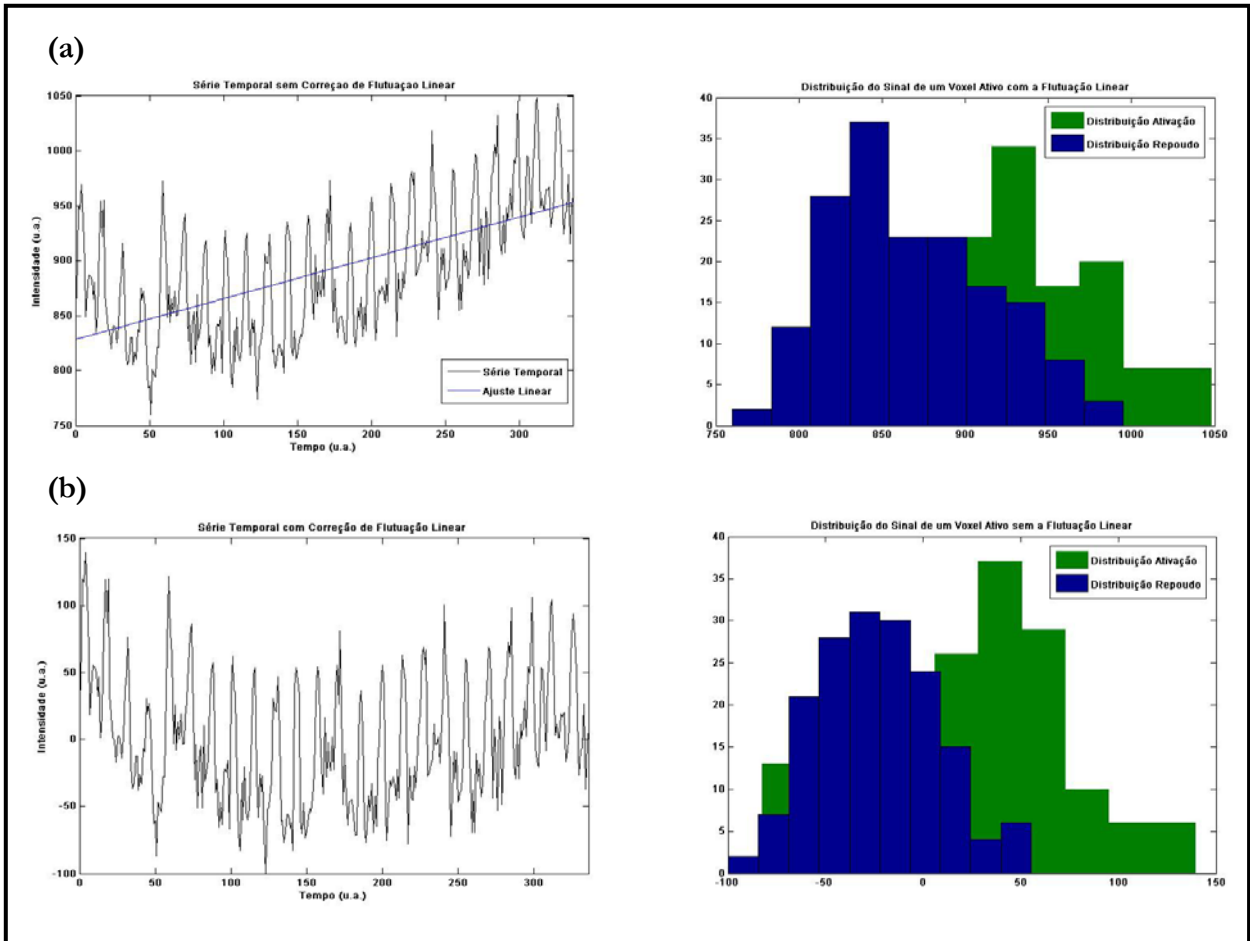


Figura 4.13: Série temporal de um voxel em uma imagem real. (a) Série temporal com ajuste linear e respectivo histograma com *drift* (b) Série temporal com eliminação de flutuação linear e histograma da distribuição do sinal.

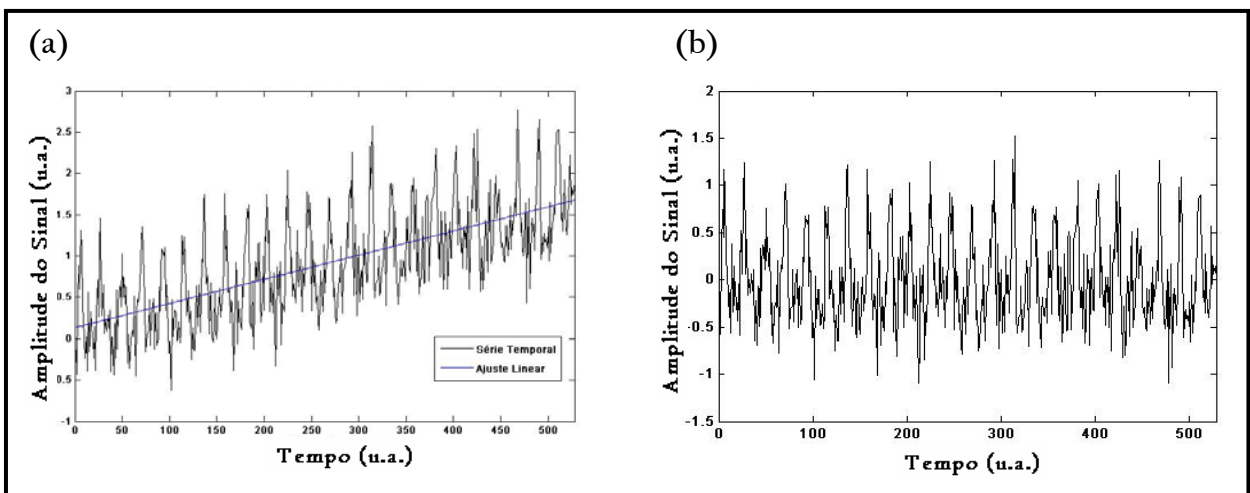


Figura 4.14: Série temporal simulada de um voxel com 10 épocas de 14 pontos temporais cada, com SNR = -5 dB e flutuação linear. (a) Série temporal com ajuste linear (b) Série temporal com eliminação de flutuação linear.



**Tabela 4.4 – Índices Estatísticos e Níveis descritivos dos métodos estatísticos antes e depois da remoção da flutuação linear.**

Método	Índices estatísticos – com Drift Linear	Nível descritivo – com Drift Linear	Índices estatísticos – sem Drift Linear	Nível descritivo – sem Drift Linear
Correlação Cruzada	CC = 0.2564	p = 0,00223	CC = 0.3793	p = 0,00000381
Teste-t de Student	t = 1.3428	p = 0,182	t = 2.4575	p = 0,0152
Entropia de Shannon	CC = 0.4577	p = 0,00107	CC = 0.4577	p = 0,00107
Entropia de Tsallis	CC = 0.5672	p = 0,0000263	CC = 0.5672	p = 0,0000263
KL-fMRI	D(2 $\sigma$ ) = 80		D(2 $\sigma$ ) = 80	

A figura 4.15 apresenta alguns mapas estatísticos após a remoção do *drift* linear. Pode-se observar que é possível utilizar valores de cortes mais estritos, aumentando a significância estatística nos métodos clássicos. Por outro lado, os mapas funcionais gerados pelos métodos de entropia de Shannon, Tsallis e KL-fMRI não apresentaram mudanças aparentes.

O sinal BOLD é uma função complexa que reflete a atividade dos sistemas hemodinâmicos [MILLER *et al.*, 2001; VAZQUEZ *et al.*, 1998]. Além disso, a resposta BOLD segue algumas fases distintas de acordo com o estímulo utilizado caracterizando a HRF. As respostas hemodinâmicas podem, em alguns casos, conter uma depressão anterior ao aumento do sinal [ROSEN *et al.*, 1998]. Essa situação, mais comum em estímulos visuais, não pôde ser constatada neste estudo. Contudo, um aumento notável do sinal após o estímulo foi observado regularmente. Os métodos clássicos, como a análise de correlação cruzada, pressupõem que a resposta BOLD pode ser padronizada. Contudo, conforme dito anteriormente, vários estudos têm mostrado que a HRF apresenta variabilidade sobre inúmeros aspectos [SOLTYSIK *et al.*, 2004, HANDWERKER *et al.* 2003, MIEZIN *et al.*, 2000] podendo afetar a significância estatística nesses métodos. Desse modo, os métodos aqui propostos, baseados na Teoria da Informação, buscam uma alternativa às abordagens clássicas, em que não há inferência com relação às HRFs. A idéia central é localizar regiões funcionais sem o conhecimento prévio da forma da HRF. Além disso, com dados simulados, nós mostramos que o mapeamento baseado na teoria de informação não é sensível ao ruído de flutuação linear.

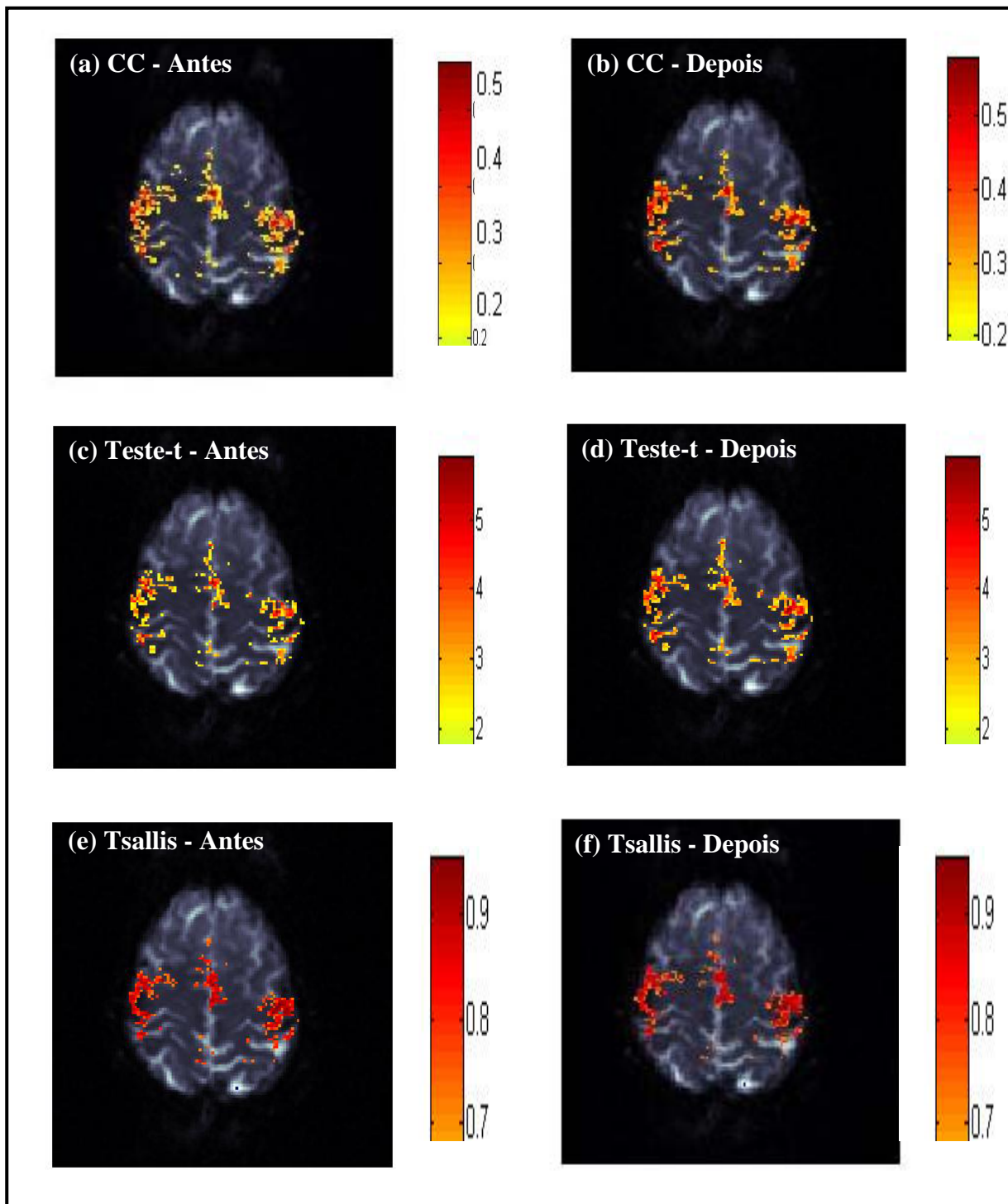


Figura 4.15: Mapas funcionais comparativos, sobrepostos a imagens EPIs, obtidos antes e depois das remoções do *Drift* linear. Análise de correção cruzada com Equação Gamma (a) antes ( $CC > 0,21$ ;  $p < 0,000147$ ) e (b) depois ( $CC > 0,26$ ;  $p < 0,00000136$ ). Teste-t de Student (c) antes ( $t > 2.22$ ;  $p < 0,027$ ) e (d) ( $t > 2.46$ ;  $p < 0,0144$ ). Entropia de Tsallis (e) Antes ( $CC > 0,7$ ;  $p < 0,00000003$ ) e (f) depois ( $CC > 0,7$ ;  $p < 0,00000003$ ).

É importante, ainda, observar que os parâmetros da janela, sobre a qual os valores de entropia ou distância são calculados, foram otimizados a fim de minimizar o tempo total da experiência e maxi-

mizar a sensibilidade e especificidade. No paradigma motor utilizado, a resposta atinge o máximo, em média, após 10 segundos após o estímulo. Assim, a janela deve ser escolhida de modo a conter toda a parte positiva da função da HRF. No entanto, o método continua válido no caso de respostas hemodinâmicas que não sigam o padrão exato observado nos casos de áreas motoras. Neste caso, a única exigência é que o tamanho da janela escolhida seja suficientemente longo a fim de conter a maior parte do complexo positivo da HRF. Nesse caso, o cálculo da entropia também fornecerá respostas semelhantes, haja vista que a primeira metade da janela fornecerá valores entrópicos maiores que a segunda metade, assim como a entropia relativa também fornecerá maiores distâncias já que a primeira metade da janela fornecerá uma distribuição mais uniforme comparado com a segunda.

## 8. Estímulo Duplo

Enquanto os métodos predominantes na análise de dados fMRI têm sido desenvolvidos e aplicados para identificar regiões do cérebro em resposta a uma tarefa específica, métodos que forneçam informações sobre as interações funcionais entre as áreas cerebrais envolvidas são cruciais, bem como para o entendimento da sua dinâmica. Além do mais, métodos de localização de regiões cerebrais que não levam em consideração a forma e atrasos das HRFs são de grande relevância no contexto da cronometria mental. Neste sentido, o objetivo deste estudo é testar e adaptar um experimento ER-fMRI nos métodos de entropia de Tsallis e KL-fMRI propostos, em uma tentativa de identificar simultaneamente duas regiões cerebrais não correlacionadas no tempo e no espaço. Nesse sentido, o desenho do experimento foi realizado de modo que a janela  $W_1$  escolhida envolvesse toda a parte positiva da função da HRF. Logo, estes métodos foram aplicados a dados de ER-fMRI, adquiridos durante a performance de uma tarefa visual e motora.

Na figura 4.16 observa-se que os métodos detectaram clusters de voxels, estatisticamente significativos, no córtex visual e motor, primários.

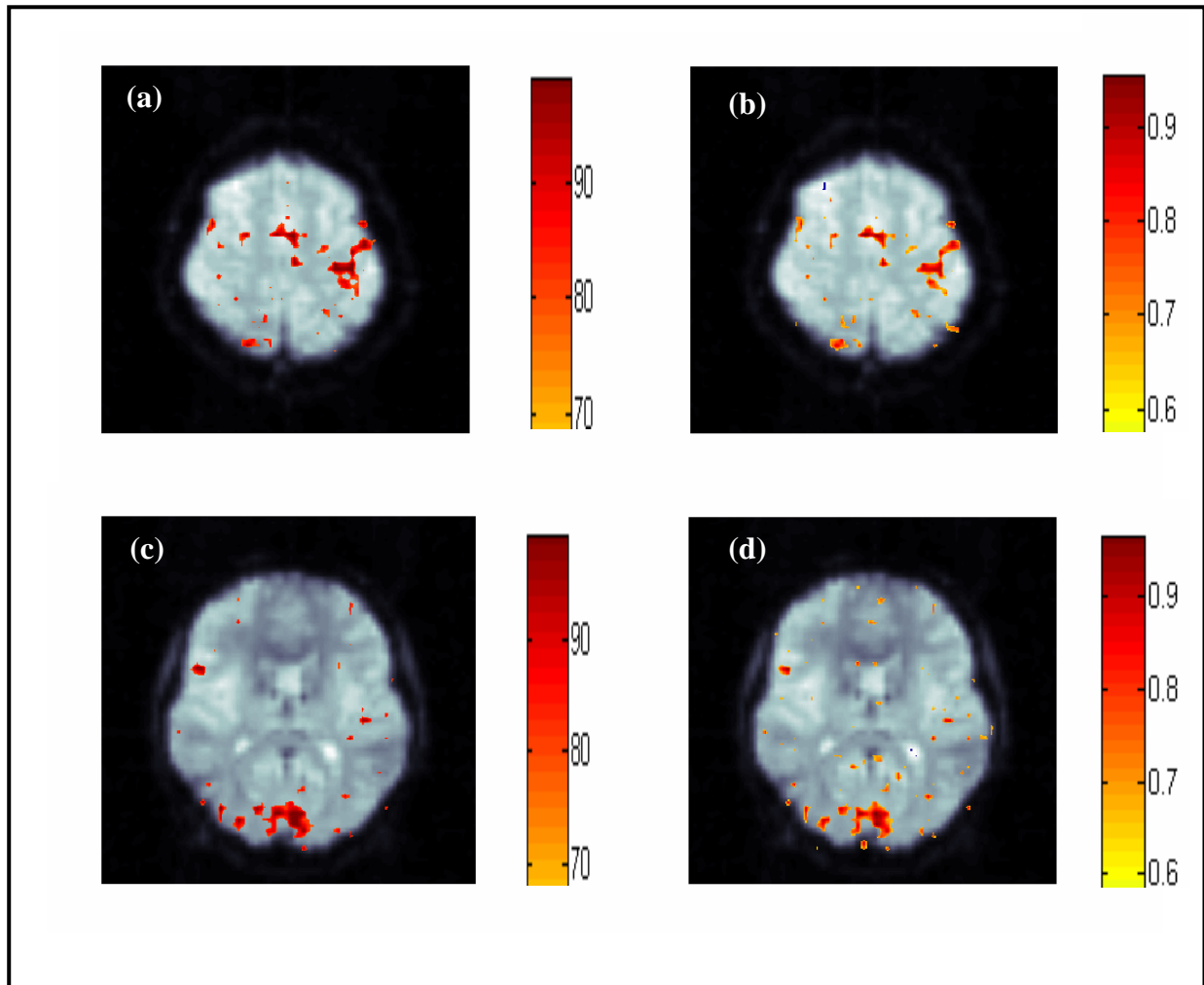


Figura 4.16: Mapas sobrepostos as imagens EPIs obtidos pelos métodos KL-fMRI e entropia de Tsallis com  $q=0,20$ . Ativação do córtex motor pelos métodos de (a) KL-fMRI com  $D(2\sigma) > 77.4534$  e (b) Entropia de Tsallis com  $CC = 0,64$  ( $p < 0,000222$ ) e ativação do córtex visual pelos métodos de (c) KL-fMRI com  $KL - D(2\sigma) > 0.7926$  e (d) Entropia de Tsallis com  $CC = 0,64$  ( $p < 0,000222$ ).

As HRFs das duas áreas são apresentadas na figura 4.17. A primeira curva (linha vermelha) provém de um voxel localizado na região motora, enquanto a segunda curva (linha azul) refere-se a um voxel da região visual. Como pode ser observado há uma evidente diferença de fase entre elas.

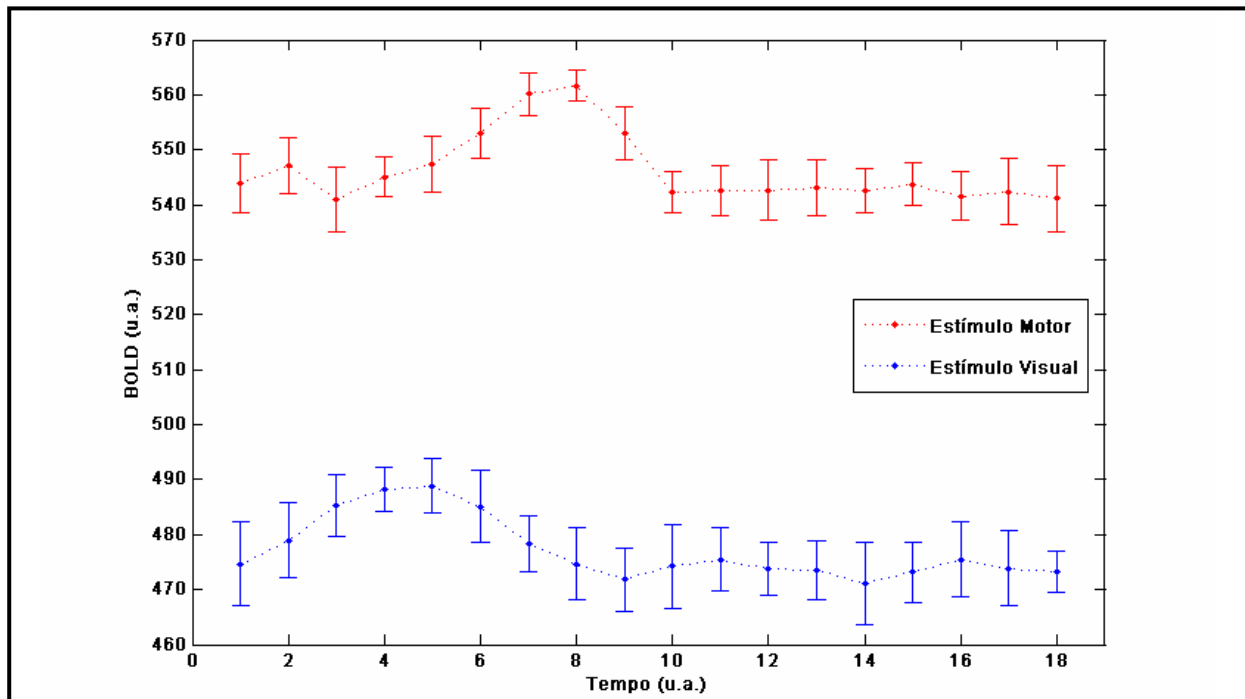


Figura 4.17 – Média das Funções Respostas Hemodinâmicas para cada região (motora e visual) detectada pelos métodos propostos de um único sujeito. O eixo x representa o tempo (TR=2s) e o y representa a mudança de sinal BOLD.

Uma vez determinado que o método é capaz de detectar respostas provenientes de duas regiões cerebrais distintas, com um atraso de fase entre elas, é interessante o desenvolvimento de um método capaz de caracterizar aspectos temporais específicos das HRFs, que pode ser aplicado no âmbito da cronometria mental. Nesse sentido, em um segundo momento, estudos preliminares de um novo método para detecção do atraso temporal da HRF foram realizados sobre dados simulados, em que nenhuma inferência sobre o atraso e a forma da HRF é realizada. Como já discutido, os achados dispõem de inúmeros métodos para a medida do atraso das respostas HRF [MENON *et al.*, 1998, LIAO *et al.*, 2002; SAAD *et al.* 2003]. Uma vez que a HRF tem sido apresentada com alta variabilidade entre sujeitos e regiões [AGUIRE *et al.*, 1998; HANDWERKER *et al.*, 2004], a estimativa do atraso pode ser dependente de quão bem o modelo do ajuste da HRF é realizado nos dados. Além disso, esses métodos dependem de um modelo do paradigma, e eles não podem ser usados para medir latências em condições onde nenhuma informação temporal específica está disponível, como períodos de atraso durante

uma tarefa de memória [POSTLE *et al.* (2000 *apud* SUN *et al.*, 2005)], ou períodos de repouso [BISWAL *et al.* (1995 *apud* SUN *et al.*, 2005)].

## 5. Dados Simulados para a Detecção do Atraso da HRF

Para testar o método de detecção de atraso, simulou-se uma série temporal composta por 10 repetições de uma função gamma [LANGE & ZEGER] com 26 pontos temporais. À série foi adicionado um atraso temporal igual a 7 (u.a.) e um ruído gaussiano com SNR de -5 dB.

Para essa simulação,  $\alpha_I = 0,01$  e  $\alpha_C = 0,0001$ . A figura 4.18 mostra os resultados do processamento da série temporal, juntamente com os parâmetros utilizados para a detecção do atraso temporal. O método encontrou um atraso de 7(u.a.), conforme o sinal simulado, demonstrando seu potencial para a detecção de atrasos temporais da HRF.

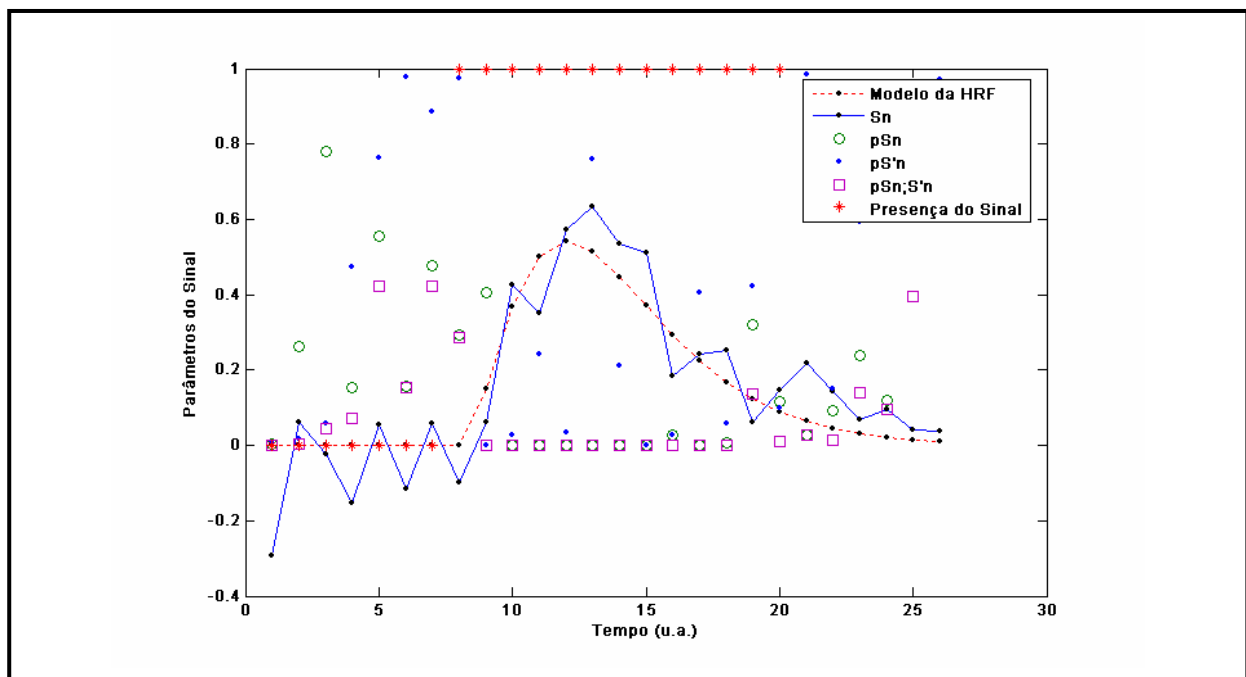


Figura 4.18: Resultado do processamento de uma série temporal com 10 épocas de uma função gamma como modelo da HRF. Além dos diversos parâmetros utilizados para encontrar o atraso da HRF, pode-se observar os pontos da presença de sinal, evidenciando um atraso temporal igual a 7 (u.a.).

Contudo, o método apresenta uma dependência com o limiar individual ( $\alpha_I$ ), contíguo ( $\alpha_C$ ) e a SNR. Neste sentido, foi realizado um estudo com dados simulados para verificar essa dependência e

procurar encontrar os valores de limiar individual e contíguo, para diferentes SNRs, quando o erro era minimizado.

Foi gerada uma série temporal utilizando uma função gamma como modelo da HRF. A série foi simulada com 10 épocas, de 25 pontos temporais cada e atrasada de 6 (u.a.). Ruídos gaussianos de média zero foram adicionados, com SNRs de 15, 5, -5, -15 dB. Foi testado um intervalo de 0,5-0,0001 para o limiar individual e contínuo. Para cada situação, eram geradas 10.000 séries temporais e calculado o desvio padrão com relação ao atraso simulado.

Na figura 4.19, apresenta-se os resultados da simulação. Os gráficos representam o comportamento do erro-padrão do atraso simulado em relação aos limiares individual e contíguo. Verifica-se, pois, a dependência de ambos os limiares estudados com relação à SNR. Para valores altos de SNR (15 e 5 dB), verificou-se que os menores valores de limiar contíguo e individual do intervalo estudado tiveram o menor erro. Por outro lado, para valores baixos de SNR (-5 e -15 dB), o inverso ocorre, isto é, o erro é maior para os limiares nesse intervalo.

Apesar dessa dependência, o método proposto mostrou-se capaz de detectar o atraso temporal em dados simulados. É importante notar que esses dados são simulados tendo como base valores característicos de uma HRF, inclusive no que diz respeito à SNR. De todo modo, a perspectiva é de generalizar o método para retirar tal dependência e, assim, torná-lo tão amplo quanto possível no que diz respeito à forma da HRF e valores de SNR.

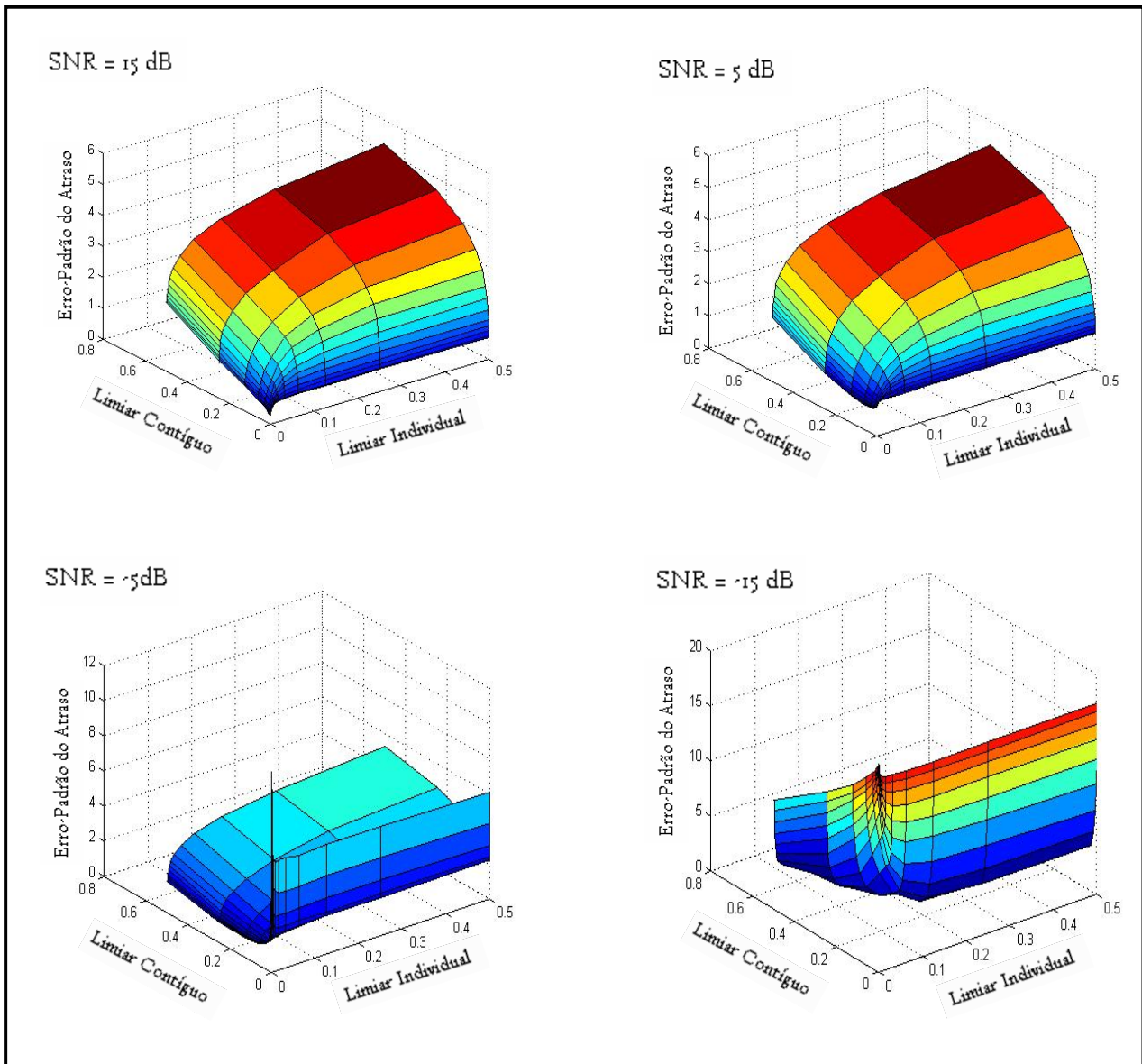


Figura 4.19: Erro-Padrão do Atraso temporal de 6(u.a.) sobre os atrasos temporais encontrados pelo método proposto, quando o limiar Contíguo e Individual variavam de 0,5-0,0001 para algumas SNRs. Cada Erro-Padrão do Atraso foi calculado sobre 10.000 séries temporais simuladas.



---

## Conclusão e Trabalhos Futuros

### 1. Conclusão

A teoria da informação fornece uma ferramenta capaz de quantificar informações, quando apenas a distribuição de probabilidade é conhecida. Como medida não-extensiva de informação, a entropia de Tsallis foi particularmente útil para a localização de padrões de sinais ER-fMRI. Apresentou-se, ainda, um novo método, baseado na distância KL, para análise de ER-fMRI. O método apresentou alta especificidade e sensibilidade na localização de regiões ativas em dados simulados de fMRI. Assim como nos métodos entrópicos, este método não requer o conhecimento a priori da resposta hemodinâmica. Contudo, mais estudos devem ser feitos para a implementação final de um teste de hipóteses.

Com as curvas ROC foi possível caracterizar a sensibilidade e especificidade quando o número de níveis e o número de época variavam, sob diferentes condições de SNR. Os métodos de Shannon, Tsallis e KL-fMRI apresentaram melhor poder de detecção quando computados com dois níveis. Além disso, verificou-se que a sensibilidade e especificidade aumentavam quando um número maior de épocas era amostrado, o que já era esperado, pois, sabe-se que quanto maior a amostra, melhor a representação da população e, então, maior significância estatística. Contudo, o aumento no número das épocas aumenta o tempo dos experimentos em fMRI.

As curvas ROC também permitiram estabelecer o intervalo dos valores  $q$ , da entropia de Tsallis, que apresentavam maior especificidade e sensibilidade. A região ótima restringiu-se ao intervalo  $-1,0 < q < 1,0$  ( $q \neq 0$ ). Entretanto, para a análise de fMRI é possível ser mais restritivo usando  $q > 1,0$ . Por outro lado, os resultados são geralmente menos restritivos para  $q < -1,0$ .

Comparativamente, dentre os métodos propostos de detecção da HRF, o método de KL-fMRI apresentou maior sensibilidade e especificidade em relação ao método de Shannon e Tsallis. Ainda, o método de Tsallis, dentro do intervalo de  $q$  otimizado, apresentou uma melhora no poder de detecção em relação ao método de Shannon. Comparando os métodos clássicos com os métodos propostos, verificou-se que o método de análise de correlação foi superior a todos os outros métodos em da-

dos simulados, o que não se confirmou em dados reais já que é um método que depende da forma da HRF.

Os resultados da aplicação dos métodos de Tsallis e KL-fMRI sobre dados reais, demonstraram a possibilidade de encontrar regiões de funções primárias em um experimento motor com um voluntário assintomático. Demonstrou-se, também, com dados simulados, que o mapeamento das funções cerebrais baseado na teoria de informação não é sensível ao ruído de flutuação linear, o que não ocorreu com os métodos clássicos. Esse resultado foi confirmado nas imagens reais.

Por fim, em um experimento de estímulo duplo, mostrou-se que os métodos de KL-fMRI e entropia de Tsallis localizaram, simultaneamente, duas regiões cerebrais não correlacionadas no tempo e no espaço. Foi proposto, por fim, um novo método na detecção de atrasos temporais das HRFs. Ele foi capaz de encontrar atrasos em dados simulados, porém, alguns parâmetros do método dependem da SNR. Logo, é necessário realizar mais estudos para retirar essa dependência.

## 2. Trabalhos Futuros

A falta de conhecimento da distribuição da distância KL não permitiu a inferência de significância estatística dentro de um teste de hipóteses. Uma possibilidade para resolver esse problema, é obter uma distribuição cumulativa aleatória da distância KL mediante simulações de Monte Carlo e, então, aplicar o teste de Kolmogorov-Smirnov (KS) para verificar se a distribuição é conhecida. Nesse teste, se houver concordância entre as distribuições teórica e da amostra, a hipótese de que a amostra foi extraída de uma população com uma distribuição específica é aceita. Além disso, há a possibilidade de se estudar uma variação do método aqui proposto, em que a distância KL pode ser computada uma só vez em toda a série temporal. Isto pode ser feito pela soma das probabilidades  $p(x)$  e  $q(x)$  de cada nível, calculadas com o método local, ao longo da série temporal. Esta proposta aumentaria o número de pontos temporais melhorando a inferência da distância KL.

Para melhor visualização, os métodos de Shannon, Tsallis e KL-fMRI poderiam ser implementados no ambiente do Brain Voyager, através do desenvolvimento de *plugins* em ambiente C++.

A utilização do método de atraso temporal apresenta uma dependência dos limiares contíguo ( $\alpha_C$ ) e individual ( $\alpha_I$ ) em relação à SNR. Logo, a aplicação do método de detecção de atraso temporal sobre o mapa de ativação gerado pelos métodos de Tsallis e KL-fMRI no experimento de duplo estímulo, passa pela necessidade de ajustes para que não haja mais essa dependência. Igualmente, a análise de erro deve ser implementada.

Os métodos propostos, baseados na teoria de informação, mostraram-se bastante promissores no mapeamento de regiões cognitivas primárias, vislumbrando possíveis aplicações em problemas mais complexos como a localização da zona irritativa de pacientes com epilepsia.

O termo epilepsia refere-se a um conjunto de doenças crônicas caracterizadas por crises paroxísticas (não-contínuas), espontâneas, e originadas pela hipersincronia da atividade elétrica cerebral focal, ou generalizada. Na maioria dos pacientes, a epilepsia pode ser tratada com medicamentos, que tem o objetivo típico de diminuir a excitabilidade neuronal, e cuja eficiência depende do tipo e causa da epilepsia. O tratamento medicamentoso das epilepsias é, em última análise, um tratamento sintomático, isto é, visa primariamente o controle das crises epilépticas. Nesse sentido, a epilepsia pode ser controlada em cerca de 70% dos casos. Entretanto, a medicação é ineficiente no restante dos pacientes. Estes pacientes possuem a chamada epilepsia refratária ao tratamento medicamentoso e são denominados farmacologicamente resistentes. Nesses casos, o tratamento cirúrgico tem sido um importante e efetivo meio de controlar as crises, tendo uma eficiência em torno de 80% a 90% [AI-HONG *et al.*, 2005]. Contudo, antes de submeter o paciente a um procedimento cirúrgico é importante que se faça uma avaliação pré-cirúrgica, que consiste em combinar diferentes informações clínicas, com objetivo de definir a possibilidade de remoção da zona epileptogênica, bem como avaliar os riscos no pós-operatório. Essas técnicas podem ser referidas como invasivas, por exemplo, quando eletrodos profundos são necessários ou não-invasivas, i.e., com mínima interferência física ao paciente. De regra, ferramentas não-

invasivas têm sido utilizadas para investigar pacientes com epilepsia refratária, quais sejam: registros de vídeo-EEG e a MRI.

A EEG tem sido utilizada no contexto de desordens epiléticas desde sua descoberta. Esses estudos são críticos nos diagnósticos de epilepsia, na classificação das desordens epiléticas e na localização das fontes de atividade epiléticas. Entretanto, a localização do EEG é limitada pela sua baixa resolução espacial. Assim, para melhorar a caracterização diagnóstica, e localização da zona epileptogênica, o EEG pode ser associado à aquisição de dados comportamentais, gerados por um aparelho de vídeo, configurando o vídeo-EEG [GOTMAN, 1999].

Estudos de neuroimagem, como por MRI, têm contribuído em muito para a definição dessas diferentes regiões. A MRI permite a identificação de alterações estruturais com excelente resolução espacial, identificando lesões atróficas, displásicas ou neoplásicas. Ainda, métodos quantitativos podem ser usados para aumentar a capacidade de a MRI distinguir tecidos normais de anormais. Por exemplo, a medida de volume de estruturas mesiais pode ajudar a encontrar atrofia [CENDES, 1997] e a análise de textura automatizada pode apontar para malformações em regiões específicas [ANTEL, 2003].

Além da avaliação estrutural, nos últimos anos, o aperfeiçoamento da fMRI tem exercido um importante papel para a avaliação de funções cognitivas e avaliação pré-cirúrgica de pacientes com epilepsia [DETRE, 2004]. A variabilidade da HRF é mais evidente em tecidos patológicos, como regiões epileptogênicas cerebrais [BÉNAR *et al.*, 2002]. Nesse sentido, os métodos baseados na teoria de informação têm a vantagem de mapear regiões ativas sem a necessidade de adotar nenhuma função de referência definitiva.

No contexto da epilepsia, a fMRI pode medir mudanças hemodinâmicas locais associadas à estimulação de descargas epileptiformes interictais com sensibilidade alta o suficiente para localizar a zona irritativa em pacientes individuais. O recente desenvolvimento da aquisição combinada de EEG e fMRI criou a oportunidade de aliar a resolução espacial da fMRI com a resolução temporal da EEG, na localização da crise. Registrando o EEG no interior do magneto, pode-se, a princípio, promover a aquisição de sinais efetivos de fMRI, provocados pelas descargas (*spikes*) epileptiformes interictais, para i-

identificar áreas cerebrais envolvidas na geração de *spikes*. A localização de ativação pela fMRI pode ser comparada com o foco de *spikes* no EEG e anormalidades estruturais de MRI. Esta técnica ajuda a definir a localização do foco epileptogênico e é importante para o melhor entendimento das patogêneses da epilepsia [LI *et al.*, 2000; AGHAKHANI *et al.*, 2004; GOLDMAN R. I. *et al.*, 2000; BAGSHAW *et al.*, 2004a].

Poucos estudos têm sido realizados na aquisição simultânea de EEG e fMRI. Os trabalhos desenvolvidos com essa técnica diferem da maioria das aplicações de fMRI, pois os estímulos são gerados internamente, aleatórios quanto ao aparecimento temporal, não muito freqüentes, e de limitada duração (de uma fração de 1s até 10 s). Isto faz o estudo de eventos epileptiformes interictais extremamente desafiante em termos do número e tipo de eventos que podem ser investigados em um paciente. Ainda, um dos principais problemas da técnica de EEG-fMRI diz respeito aos métodos de detecção da HRF decorrentes das descargas epileptiformes interictais. Nesses estudos, é sempre mais difícil a localização precisa do sinal BOLD envolvido no evento eletrofisiológico, uma vez que o padrão da HRF não é bem caracterizado [BENAR *et al.*, 2002]. Pesquisas têm sido realizadas com intuito de aperfeiçoar esses modelos [BAGSHAW *et al.*, 2004b]. Desse modo, é possível aplicar as análises aqui propostas, com o intuito de serem utilizadas em dados provenientes da aquisição simultânea de fMRI e eletroencefalografia (EEG), para o estudo das respostas hemodinâmicas de descargas interictais em pacientes com epilepsia.

## Referências

- ADLER, R. (1981). *The Geometry of Random Field*. Wiley, New York.
- AGHAKHANI Y, BAGSHAW AP, BENAR CG et al. (2004) fMRI activation during spike and wave discharges in idiopathic generalized epilepsy. *Brain*; 127:1127-1144.
- AGUIRE GK, ZARAHN E., D'ESPOSITO M. The variability of human, BOLD hemodynamic responses. *NeuroImage*, 8: 360-369.
- AI-HONG YU, KUN-CHENG LI, CHANG-FU PIAO, HONG-LI LI. (2005) Application of functional MRI in epilepsy, *Chin Med J*.;118(12):1022-1027.
- ANTEL S.B., COLLINGS D.L., BERNASCONI N., ANDERMANN F., SHINGHAL R., KEARNEY R.E., ARNOLD D.L. (2003), Bernasconi A. Automated detection of focal cortical dysplasia lesions using computational models of their MRI characteristics and textures analysis. *Neuroimage*, vol. 19, no. 4, 1748-1759.
- BAGSHAW A. P., AGHAKHANI Y., BENAR C. G., KOBAYASHI E., HAWCO C., DUBEAU F., PIKE G. B., AND GOTMAN J. (2004a), "EEG-fMRI of focal epileptic *spikes*: Analysis with multiple haemodynamic functions and comparison with gadolinium-enhanced MR angiograms," *Human Brain Mapping*, vol. 22, pp. 179-192.
- BAGSHAW A. P., HAWCO C., BÉNAR CG., KOBAYASHI E., AGHAKHANI Y., DUBEAU F., PIKE G.B., AND GOTMAN J. (2004b), Analysis of the EEG-fMRI response to prolonged bursts of interictal epileptiform activity, *NeuroImage* 24,1099-1112.
- BANDETTINI P.A., JESMANOWICZ A., WONG E.C., HYDE J.S. (1993). Processing strategies for time-course data sets in functional MRI of human brain, *Magn. Reson. Med.*, vol. 30: 161-173
- BELLIVEAU JW, KENNEDY DN JR, MCKINSTRY RC, BUCHBINDER BR, WEISSKOFF RM, COHEN MS, VEVEA JM, BRADY TJ, ROSEN BR. (1991) Functional mapping of the human visual cortex by magnetic resonance imaging. *Science* 254:716-719
- BENAR C. G., GROSS D. W., WANG Y. H., PETRE V., PIKE B., DUBEAU F. O., AND GOTMAN J. (2002), The BOLD response to interictal epileptiform discharges, *Neuroimage*, vol. 17, pp. 1182-1192.
- BENJAMINI, Y, HOCHBERG, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J.R. Stat. Soc.* 57, 289-300.

- BORLAND L., PLASTINO A., TSALLIS C., ( 1998). Information gain within nonextensive thermostatics, *Journal of Mathematical Physics*, vol. 39, no 12: 6490-6501
- BOYNTON G.M., ENGEL, S.A., GLOVER, G.H., HEEGER, D.J., 1996. Linear systems analysis of functional magnetic resonance imaging in human V1. *J. Neurosci.* 16, 4207-4221.
- BURDETTE JH, MINOSHIMA S, BORGHT TV, TRAN DD, KUHL DE (1996) Alzheimer Disease: Improved Visual Interpretation of PET Images by using Three-Dimensional Stereotaxic Surface Projections, *Radiology*, 198: 837-843
- CALLOUN VD, ADALI T, MCGINTY VB, PEKAR JJ, WATSON TD, PEARLSON GD. (2001) fMRI activation in a visual-perception task: network of áreas detected using the general linear model and independent components analysis. *NeuroImage*, 14:1080-1088.
- CAPURRO A., DIAMBRA L., LORENZO D., MACADAR O., MATRIN M.T., MOSTACCIO C., PLASTINO A., PÉREZ J., ROFMAN E., TORRES, M.E., VELLUTI J. (1999). Human brain dynamics: the analysis of EEG signals with Tsallis information measure, *Physica A* 265: 235-254
- CENDES F., CARAMANOS Z., ANDERMANN F., DUBEAU F., ARNOLD D.L. (1997) Próton magnetic resonance spectroscopic imaging and magnetic resonance imaging volumetry in the lateralization of temporal lobe epilepsy: a series of 100 patients. *Ann. Neurol.*, vol. 42, no. 5, 737-746.
- CLARE S, HUMBERSTONE M, HYKIN J, BLUMHARDT LD, BOWTELL R, MORRIS P. (1999). Detecting activations in event-related fMRI using analysis of variance. *Magn Reson Med.* 42(6):1117-22.
- CONSTABLE, R.T., SKUDLARSKI, P., GORE JC, (1995). An ROC Approach for evaluating functional brain MR imaging and postprocessing protocols. *Magn. Reson. Med.* 34(1), 57-64
- COVER T.M. & THOMAS J.A. (1991). *Elements of Information Theory*. (John Wiley & Sons, New York).
- DAMADIAN R. (1971). Tumor detection by nuclear magnetic resonance imaging. *Science*, 171: 1151-1153
- DE ARAUJO DB, TEDESCHI W, SANTOS AC, ELIAS J JR, NEVES UP, BAFFA O. (2003). Shannon entropy applied to the analysis of event-related fMRI time series. *Neuroimage*. 2003 Sep;20(1):311-7.

- DE ARAÚJO, D.B. (2002) Sobre neuroimagens funcionais por magnetoencefalografia e ressonância magnética: novos métodos e aplicações. Tese (Doutorado). Universidade de São Paulo. FFCLRP.
- DETRE JA. (2004), fMRI: applications in epilepsy. *Epilepsia*; 45:26-31.
- FORMISANO E, LINDEN DE, DI SALLE F, TROJANO L, ESPOSITO F, SACK AT, GROSSI D, ZANELLA FE, GOEBEL R. (2002) Tracking the mind's image in the brain I: time-resolved fMRI during visuospatial mental imagery. *Neuron*. 3;35(1):185-94.
- FORMISANO E.; GOEBEL R. Tracking cognitive processes with functional MRI mental chronometry. *Curr Opin Neurobiol*. 2003 Apr;13(2):174-81.
- FOX P.T & RAICHLE, M.E. (1986). Focal physiological uncoupling of cerebral blood flow and oxidative metabolism during somatosensory stimulation in human subjects. *Proc. Natl. Acad. Sci. U.S.A.*, 83(4): 1140-1144
- FRISTON K.J., HOLMES A.P., WORSLEY K.J, POLINE J.P, FRITH C.D AND DRACKOWIAK, R.S.J. (1995a). Statistical parametric maps in functional imaging: a general linear approach. *Hum. Brain Mapp*. 2, 189-210.
- FRISTON K.J., FLETCHER P, JOSEPHS O, HOLMES AP., RUGG MD, TURNER R. (1998). Event-Related: characterizing differential responses. *Neuroimage*, 7:30-40.
- GENOVESE C.R, LAZAR N.A., NICHOLS T. (2002). Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *Neuroimage* 15(4): 870-8.
- GEORGOPOUTOS AP, PELLIZZER G (1995). The mental and the neural: psychological and neural studies of mental rotation and memory scanning. *Neuropsychologia* 33: 1531-1547
- GLOVER G.H., Deconvolution of Impulse Response in Event-Related BOLD fMRI. *NeuroImage*, 9:416-429, 1999.
- GOLDMAN R. I., STERN J. M., ENGEL J., AND COHEN M. S. (2000), "Acquiring simultaneous EEG and functional MRI," *Clinical Neurophysiology*, vol. 111, pp. 1974-1980.
- GOTMAN J. (1999). Automatic detection of seizures and spikes. *J. Clin. Neurophys.*, vol. 16, no. 2, 130-140
- HAHN EL. Spin echoes. *Phys Rev* 1950; 80:580-594.



- HANDWERKER DA, OLLINGER JM, D'ESPOSITO M. (2004). Variation of BOLD hemodynamic responses across subjects and brain regions and their effects on statistical analyses. *Neuroimage* 21(4):1639-51.
- HANSEN S.H., EGLI D., HOLLENSTEIN L., SALZMANN C. (2005). Dark matter distribution function from non-extensive statistical mechanics, *New Astronomy* 10, 379.
- HARI R., LEVANEN S., RAIJ T. (2000). Timing of human cortical functions during cognition: role of MEG. *Trends Cogn Sci* 4: 455-462.
- HENSON RN, PRICE CJ, RUGG MD, TURNER R, FRISTON KJ. (2002) Detecting latency differences in event-related BOLD responses: application to words versus nonwords and initial versus repeated face presentations. *NeuroImage*, 15:83-97.
- HUETTEL S.A., SONG A.W., MCCARTHY G. *Functional Magnetic Resonance Imaging*, 1<sup>a</sup> ed. Sunderland: Sinauer Associates, 2004.
- INSTITUTE FOR BIOMEDICAL ENGINEERING. ETH and University of Zurich. Disponível em <http://www.mr.ethz.ch/>. Acesso em: 16/01/2006.
- JEZZARD P. & CLARE S. (2001) *Principles of nuclear magnetic resonance and MRI*. Em *Functional MRI. Na Introduction to Methods* (edited by Jezzard P, Matthews P.M., & Smith S.M.). Oxford University Press
- KIM, SG. (2003). Progress in understanding functional Imaging signals. *PNAS*, 100:3550-3552.
- KWONG K.K., BELLIVEAU J.W., CHESLER D.A., GOLDBERG I.E., WEISSKOFF R.M., PONCELET B.P., KENNEDY D.N., HOPPEL B.E., COHEN M.S., TURNER R. (1992) Dynamic magnetic resonance imaging of human brain activity during primary sensory stimulation. *Proc. Natl. Acad. U.S.A.* 89, 5675.
- LANGE N. & ZEGER S.L. (1997). Non-linear Fourier time series analysis for human brain mapping by functional magnetic resonance imaging (with discussion). *Applied Statistics*, 46, 1-29.
- LAUTERBUR PC. Image formation by induced local interactions: Examples employing nuclear magnetic resonance. *Nature* 1973; 242-190.
- LI EZ, WENG XC, HAN Y (2000) Asymetry of brain functional activation: fMRI study under language and music stimulation. *Chin Med J*, 113:154-158.

- LIAO CH., WORSLEW KJ., POLINE JB., ALTON JA., DUNCAN GH., EVANS AC. (2002). Estimating the delay of the fMRI response. *Neuroimage* 16: 593-606
- LOGOTHETIS, N.K.(2002). The neural basis of the blood-oxygen-level-dependent functional magnetic resonance imaging signal. *Philos. Trans. R. Soc. Lond B. Biol. Sci.* 357 (1424): 1003-1037.
- LOWE M.J A. & RUSSELL D. P. (1999). Treatment of baseline drifts in fMRI time series analysis, *J. Comput. Assist. Tomogr.*, vol. 23, no. 3: 463-473, 1999.
- MANSFIELD P. Multi-planar image-formation using NMR spin echoes. *Journal of Physics C-Solid State Physics* 1977; 10:L55-L58.
- MATTAY V.S., FRANK J.A., SANTHA A.K., PEKAR J.J., DUYN, J.H., MCLAUGHLIN A.C., WEINBERGER D.R. (1996). Whole-brain functional mapping with isotropic MR imaging. *Radiology* 201(2):399-404.
- MCKEOWN MJ, MAKEIG S., BROWN GG, JUNG TP, KINDERMANN SS., BELL AJ., SEJNOWSKI TJ (1998). Analysis of fMRI data by blind separation into independent spatial components. *Hum Brain Mapp* 6: 160-188
- McPHERSON G. (1990) *Statistic in Scientific Investigation: Its Basis, Application, and Interpretation.* [S.l.]: Springer-Verlag.
- MENON RS, LUKNOWSKY DC, GATI JS: (1998). Mental chronometry using latency-resolved functional MRI. *Proc Natl. Acad USA*, 95: 10902-10907.
- MENON RS., KIM SG (1999) Spatial and temporal limits in cognitive neuroimaging with fMRI. *Trends Cogn Sci* 3:207-216.
- METZ, C.E. (1986) *Statistical Analysis of ROC Data in Evaluating Diagnostic Performance. Multiple Regression Analysis: Applications in the Health Sciences*, number 13, edited by Donald E. Herbert and Raymond H. Myers. 365-384. American Institute of Physics.
- MIEZIN FM, MACCOTTA L, OLLINGER JM, PETERSEN SE, BUCKNER RL. (2000) Characterizing the hemodynamic response: effects of presentation rate, sampling procedure, and the possibility of ordering brain activity based on relative timing. *Neuroimage*. 11(6):735-59.
- MILLER KL, LUH W., LIU TT, MARTINEZ A, OBATA T, WONG EC, FRANK LR, BUXTON RB. (2001). Nonlinear temporal dynamics of the cerebral blood flow response. *Hum Brain Mapp* 13:1-12.

- MILLER, K.L. Novel Methods for Steady-State Neuroimaging. Tese (Doutorado). Stanford University. 2004.
- MULLER H., KRAFT E., LUDOLPH A., ERNE S., (2002). New methods in fMRI analysis, Hierarchical cluster analysis for improved signal-to-noise ratio compared to standard techniques, *IEEE Eng. Med. Biol. Mag.* 21(5): 134-142
- NELDER J.A. & WEDDERBURN R.W.M. (1972). Generalized Linear Models. *Journal of the Royal Statistical Society, A*, v. 135: 384-1586.
- OGAWA S., LEE, T.M., KAY, A.R. & TANK, D.W. (1990b). Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *Proc. Natl. Acad. Sci. USA.* 87(24): 9868-72
- OGAWA, S., LEE, T., NAYAK, A.S. & GLYNN, P. (1990a) Oxygenation-sensitive contrast in magnetic resonance image of rodent brain at high magnetic fields. *Magn. Resonance Med.* 14(1): 68-78
- PARKER J., et al. (1995). Classification of Ductal Carcinoma in Situ by Image Analysis of Classification from Digital Mammograms, *The British Journal of Radiology* 68: 150-159
- PINEIRO R, PENDLEBURY S, JOHANSEN-BERG H, MATTHEUWS PM. (2003). Altered hemodynamic responses in patients after subcortical stroke measured by Functional MRI. *Stroke*, 33:103-109
- PURDON PL, SOLO V., WEISSKOFF RM., BROWN EN (2001). Locally regularized spatiotemporal modeling and model comparison for functional MRI. *Neuroimage* 14: 912-9123
- ROSEN BR, BUCKNER R.L., Dale A.M. (1998). Event-related functional MRI: Past, present, and future. *Proc. Natl. Acad. Sci. USA* 95, 773-780
- SAAD, Z.S., DEYOE, E.A., ROPELLA, K.M., (2003). Estimation of fMRI response delays. *NeuroImage* 18, 494-504.
- SAAD, Z.S., ROPELLA K.M., COX, R.W., DEYOE E.A. (2001). Analysis and Use of fMRI Response Delays. *Human Brain Mapp* 13: 74-93
- SHANNON, C.E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, 27, 379-423

- SMITH A.M., LEWIS B.K., RUTTMANN U.E., YE F.Q., SINNWELL T.M., YANG Y., DUYN J. H., FRANK J.A. (1999). Investigation of low frequency drift in fMRI signal, *NeuroImage* 9(5), 526-533
- SMITH, H. & Raballo, F. A non-mathematical approach to basic MRI, Medical Physics Publ., Madison, 1989.
- SOLTYSIK D.A., PECK K.K., WHITE, K.D., CROSSON B., BRIGGS R. W. (2004). Comparison of hemodynamic response nonlinearity across primary cortical areas. *NeuroImage* 22, 1117-1127
- SPIEGEL M.R. (1972). Estatística [S.l.]: McGraw-Hill
- SUN F.T, MILLER L.M, D'ESPOSITO M. (2005). Measuring temporal dynamics of functional networks using phase spectrum of fMRI data. *Neuroimage* 28:227-237.
- TEDESCHI W., MÜLLER HP., DE ARAUJO D.B, SANTONS A.C., NEVES U.P.C., ERNÈ S.N., BAFFA O. (2005). Generalized mutual information tests applied to fMRI analysis. *Physica A* 352: 629-644
- THIERRY G., BOULANOUAR K., KHERIF F., RANJEVA JP., DEMONET JF. (1999) Temporal sorting of neural components underlying phonological processing. *NeuroReport* 10: 2599-2603.
- TORRES M.E., ANIMO M.M., SCHLOTTHAUER G. (2003) Automatic detection of slight parameter changes associated to complex biomedical signal using multiresolution q-entropy,. *Medical Eng. and Physics* 25, 859.
- TSALLIS C. (1988). "Possible Generalization of Boltzmann-Gibbs Statistics", *Journal of Statistical Physics*, vol. 52, pp. 479-487
- VAZQUEZ AL., NOLL DC. (1998). Nonlinear aspects of the BOLD response in functional MRI. *Neuroimage*;7:108-18.
- W. H. Press et al. Numerical Recipes. 2nd. ed. [S.l.]: Cambridge University Press, 1992.
- WIKIPEDIA A Enciclopédia LIVRE. Disponível em <http://www.wikipedia.com/>. Acesso em: 07/12/2005.

---

## APÊNDICE A - Imagem por Ressonância Magnética

### 1. Uma Breve História da MRI

São várias as descobertas científicas que culminaram na imagem por ressonância magnética, *Magnetic Resonance Imaging* (MRI). Estudos realizados pelo físico austríaco Wolfgang Pauli em 1924 trouxeram a idéia de que o núcleo atômico tem propriedades magnéticas. Em 1936, depois de estudos desenvolvidos por Stern e Gerlach no núcleo atômico, o físico americano Isidor Rabi descobre o fenômeno de ressonância magnética, sendo agraciado com Prêmio Nobel de 1944, pela primeira demonstração do efeito de ressonância magnética nuclear. No final da Segunda Guerra Mundial, ano de 1945, dois físicos, Felix Bloch em Stanford e Edward Purcell no MIT/Harvard, realizam investigações independentes sobre ressonância magnética nuclear, *Nuclear Magnetic Resonance* (NMR) na matéria, detectando o efeito, bem como processos de relaxação associados. Pelas descobertas, Purcell e Bloch receberam o Prêmio Nobel de física em 1952.

A primeira aplicação desta nova técnica foi a análise química de amostras. A NMR mostrou-se muito útil no estudo da composição química de substâncias homogêneas através de espectroscopia por NMR.

Em 1960, medidas de NMR revelaram diferenças entre moléculas de água, dependendo se estas estavam ou não em tecidos biológicos. O médico americano Raymond Damadian supôs que estas diferenças poderiam ser observadas em células cancerosas em relação às não cancerosas, podendo ser um excelente método de detecção desses tecidos. Desse modo, em 1971, através de suas pesquisas com ratos, ele mostrou que os tempos de relaxação de tecidos cancerosos eram muito mais longos do que os saudáveis [DAMADIAN, 1971].

Um ano depois, em 1972, o físico-químico americano Paul Lauterbur, realiza os primeiros testes em um espectrômetro NMR, na tentativa de adquirir informações espaciais da amostra. No ano seguinte publica a primeira imagem por ressonância magnética na revista *Nature* [LAUTERBUR, 1973]. Na mesma época, o físico inglês Peter Mansfield desenvolve métodos de formação de imagens por

NMR, baseados na aplicação de campos magnéticos de variação espacial [gradientes espaciais (G)]. Por suas contribuições para a formação de imagens utilizando a NMR, Lauterbur e Mansfield receberam o Prêmio Nobel de Fisiologia e Medicina em 2003.

Em 1977, um tomógrafo foi construído por Damadian e colegas (conhecido como O Indomável) pela FONAR Corporation, quando produziram a imagem do corpo humano, cuja aquisição demorou 4 horas (ver figura 1.1).

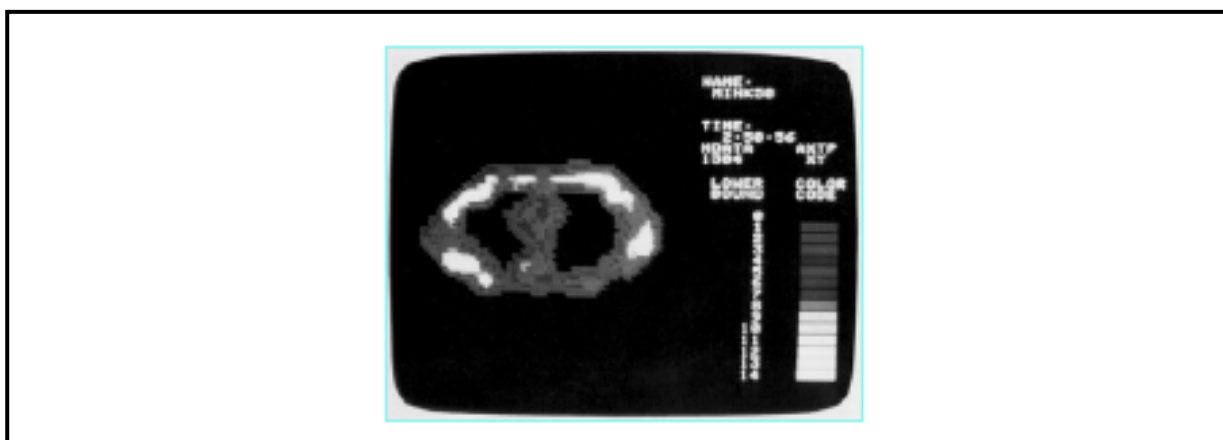


Figura 1.1. Imagem de ressonância magnética do corpo humano, produzida pelo pós-doutorando Larry Minkoff do grupo do Prof. Damadian.

Dada a importância da NMR para a produção de MRI, abordaremos o tema em mais detalhes, a seguir.

## 2 Ressonância Magnética Nuclear

Os átomos, que compõem toda matéria, consistem, dentro de um modelo clássico, de um núcleo central e de elétrons em sua órbita. O núcleo contém núcleons que são subdivididos em prótons e nêutrons; enquanto os prótons têm carga positiva, os elétrons carga negativa e os nêutrons carga nula.

Pelas leis do eletromagnetismo, uma partícula carregada em movimento gera um campo magnético. Um elétron girando em torno do seu núcleo induz um campo magnético, comportando-se como um pequeno magneto, que é representado pelo momento magnético  $\vec{\mu}$ , medido em  $\text{Am}^2$ . Da mesma maneira, podemos associar um momento magnético aos nucleons. Quando o número de massa é

ímpar, ou seja, quando prótons ou nêutrons estão desemparelhados, existe um momento magnético nuclear permanente, e só nesses casos observa-se a NMR. Dentre os núcleos mais utilizados para a aplicação clínica encontra-se o hidrogênio ( $^1\text{H}$ ). Ele contém apenas um próton (número atômico e de massa  $1^3$ ) e é usado por ser muito abundante no corpo humano, apresentar sensibilidade quando um campo magnético externo é aplicado, além de ter momento magnético relativamente intenso. A tabela a seguir mostra a abundância natural, *spin* nuclear e frequência de ressonância de alguns núcleos de interesse biológico.

**Tabela 1.1 – Alguns núcleos de importância biológica com suas respectivas propriedades de interesse na NMR. Observe que a frequência de ressonância do  $^1\text{H}$  é a mais alta, devido seu momento magnético. Tabela modificada de Jezzard 2001 pg. 69.**

Núcleo	Abundância Natural (%)	<i>Spin</i>	Frequência
$^1\text{H}$ (Hidrogênio)	99,9	1/2	42,577 MHz
$^{13}\text{C}$ (Carbono)	1,1	1/2	10,708 MHz
$^{14}\text{N}$ (Nitrogênio-14)	99,63	1	3,078 MHz
$^{17}\text{O}$ (Oxigênio)	0,037	5/2	5,774 MHz

Na ausência de um campo magnético aplicado, os momentos magnéticos dos núcleos de hidrogênio têm uma orientação aleatória. Contudo, quando colocados na presença de um campo magnético estático externo, seus momentos magnéticos se alinham a este campo, fazendo aparecer um momento magnético resultante na direção do campo aplicado. O efeito do campo magnético estático sobre a orientação dos *spins* dos prótons, bem como sua manipulação pela aplicação de campos magnéticos variáveis no tempo e no espaço, pode ser explicado por dois modelos: um clássico, outro quântico. Aqui, trataremos os dois em mais detalhes, a seguir.

<sup>3</sup> O número atômico ( $Z$ ) é a soma dos prótons no núcleo e o número de massa ( $A$ ) é a soma dos prótons e nêutrons no núcleo.

## 2.1 Modelo Clássico

Como o próton tem massa, sua rotação produz um momento angular  $\vec{J}$ . No modelo clássico, o momento angular total de uma partícula carregada de massa  $m$ , percorrendo uma órbita circular de raio  $r$ , velocidade linear  $v$  está relacionado com o momento linear ( $p$ ) dado por:

$$\vec{J} = \vec{r} \times \vec{p} = \vec{r} \times (m\vec{v}) = m(\vec{r} \times \vec{v}). \quad \dots\dots\dots (1.1)$$

A taxa temporal de variação de  $\vec{J}$  ( $d\vec{J}/dt$ ) é dada por

$$\frac{d\vec{J}}{dt} = \frac{d(\vec{r} \times \vec{p})}{dt} \implies \frac{d\vec{J}}{dt} = \frac{d\vec{r}}{dt} \times \vec{p} + \vec{r} \times \frac{d\vec{p}}{dt}. \quad \dots\dots\dots (1.2)$$

Entretanto,

$$\frac{d\vec{r}}{dt} \times \vec{p} = \vec{v} \times \vec{p} = m\vec{v} \times \vec{v} = 0, \quad \dots\dots\dots (1.3)$$

e pela segunda lei de Newton de movimento temos:

$$\vec{r} \times \frac{d\vec{p}}{dt} = \vec{r} \times \vec{F} = \vec{\tau} \implies \frac{d\vec{J}}{dt} = \vec{\tau}. \quad \dots\dots\dots (1.4)$$

Ou seja, quanto maior o torque ( $\vec{\tau}$ ), mais rápida é a precessão.

Um próton com momento magnético  $\vec{\mu}$  em um campo magnético aplicado  $\vec{B}_0$ , sofre uma interação entre o seu momento magnético e o campo magnético aplicado, experimentando um torque na tentativa de alinhar  $\vec{\mu}$  e  $\vec{B}_0$ , cuja expressão é dada por:

$$\vec{\tau} = \vec{\mu} \times \vec{B}_0. \quad \dots\dots\dots (1.5)$$

Este torque, que causa o movimento de precessão dos prótons em torno do campo, é igual à taxa de variação do seu momento angular. Classicamente, essa variação é:



$$\frac{d\vec{J}}{dt} = \vec{\tau} = \vec{\mu} \times \vec{B}_0. \dots\dots\dots (1.6)$$

Assim, se um próton é colocado em um campo magnético ele experimenta um torque devido à interação entre o momento magnético e o campo do magneto. Quanto maior o valor do campo magnético  $\vec{B}_0$ , mais rápido é a taxa de precessão. Cada núcleo tem uma precisa relação entre o momento angular intrínseco  $\vec{J}$ , e seu momento magnético  $\vec{\mu}$ . Esta relação é estabelecida matematicamente por:

$$\vec{\mu} = \gamma \vec{J}. \dots\dots\dots (1.7)$$

A equação 1.7 pode então ser reescrita como:

$$\frac{d\vec{\mu}}{dt} = \gamma \frac{d\vec{J}}{dt} = \gamma(\vec{\mu} \times \vec{B}_0). \dots\dots\dots (1.8)$$

Essa equação nos diz que a taxa que o núcleo precessa em um campo magnético depende da magnitude de  $\vec{\mu}$  e de  $\vec{B}_0$ .

A soma vetorial de um ensemble de momentos magnéticos por unidade de volume resulta em uma medida macroscópica denominada magnetização  $\vec{M}$ . A magnetização  $\vec{M}$  na presença de um campo magnético  $\vec{B}_0$  exibirá um movimento de precessão, cuja taxa de variação da magnetização será governada pela seguinte equação:

$$\frac{d\vec{M}}{dt} = \vec{M} \times \gamma \vec{B}_0. \dots\dots\dots (1.9)$$

A velocidade angular que o vetor magnetização descreve é constante e tem módulo igual à frequência de Larmor. O ângulo entre  $\vec{M}$  e  $\vec{B}_0$  também é constante.

Para medirmos características físicas do sistema, como os tempos de relaxação, devemos alterar a orientação da magnetização, aplicando um segundo campo magnético  $B_1$ , circularmente polarizado no plano x-y e perpendicular a  $\vec{B}_0$  com frequência de oscilação  $\omega_1$ .

$$B_1 = B_1 [\cos(\omega_1 t) i - \text{sen}(\omega_1 t) j]. \dots\dots\dots (1.10)$$

Na presença do campo magnético  $B_1$ , a equação 1.9 pode ser rescrita como:

$$\frac{d\vec{M}}{dt} = \vec{M} \times \gamma(\vec{B}_0 + \vec{B}_1). \quad \dots\dots\dots(1.11)$$

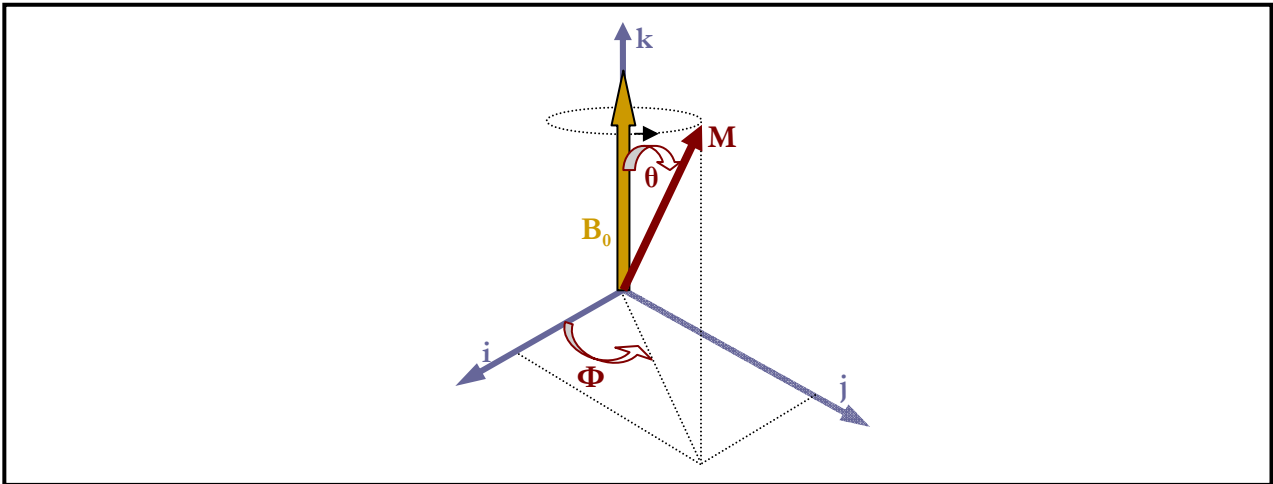


Figura 1.2. Sistema de coordenadas de precessão do vetor magnetização quando um campo magnético estático  $\vec{B}_0$  é aplicado.

De acordo com a figura 1.2 podemos escrever os vetores  $\vec{M}$  e  $\vec{B}_0$  da seguinte maneira:

$$\vec{M} = M[\text{sen}(\theta)\text{cos}(\phi)\hat{i} + \text{sen}(\theta)\text{sen}(\phi)\hat{j} + \text{cos}(\theta)\hat{k}], \quad \dots\dots\dots(1.12)$$

$$\vec{B} = B_0\hat{k}. \quad \dots\dots\dots(1.13)$$

Substituindo a equação de 1.12 em 1.11 obtemos a taxa de variação temporal de  $\theta$ :

$$\frac{d\theta}{dt} = -\gamma B_1 \text{sen}(\omega_1 t + \theta). \quad \dots\dots\dots(1.14)$$

Essa variação será máxima quando  $\text{sen}(\omega_1 t + \theta) = 1$ , logo:

$$\frac{d\theta}{dt} = -\gamma B_1 \text{ ou } \theta = \gamma B_1 \Delta t. \quad \dots\dots\dots(1.15)$$

Observa-se que o ângulo  $\theta$  depende da intensidade do campo  $B_1$  circularmente polarizado e do tempo de aplicação  $\Delta t$ . Como esse tempo geralmente é muito curto, a aplicação deste campo magnético é denominada de pulso de radiofrequência. Num primeiro momento a magnetização encontra-se no equilíbrio térmico, realizando a precessão em torno do campo  $\vec{B}_0$ . A aplicação de um pulso de ra-

diófreqüência altera o ângulo  $\theta$ , provocando a rotação do vetor magnetização. Se  $\theta = \pi/2$ , o vetor magnetização será girado para o plano x-y em 90°, sendo denominado pulso de 90°.

## 2.1.1 Processos de Relaxação

Classicamente, os prótons de  $^1\text{H}$  na presença de um campo magnético estático  $B_0\hat{k}$  e no equilíbrio térmico, exibem um movimento de precessão do vetor magnetização ( $M_0$ ) nessa direção. A aplicação do pulso de RF no plano x-y modifica a orientação de  $M_0$  que traçará uma espiral em torno de  $B_0\hat{k}$  modificando o ângulo com o eixo z, e como resultado um movimento de precessão em torno de  $B_1$ .

O sinal MR criado após o pulso de excitação decai no tempo, por um fenômeno chamado de relaxação nuclear. O sistema buscará, naturalmente, o estado de menor energia através da troca de energia entre os *spins* e sua vizinhança<sup>4</sup>. São dois os mecanismos básicos que contribuem para a perda do sinal MR: tempo de relaxação  $T_1$ , também denominado relaxação *spin-rede* ou relaxação longitudinal e tempo de relaxação  $T_2$ , chamado relaxação transversal ou *spin-spin*.

### 2.1.1.1 Relaxação Longitudinal ( $T_1$ )

Do ponto de vista clássico, a magnetização transversal gradualmente diminui, aumentando a longitudinal. A recuperação da magnetização longitudinal segue uma forma exponencial (Figura 1.3 (a)):

$$M_z = M_0 \left( 1 - e^{-t/T_1} \right). \quad \dots\dots\dots (1.16)$$

$M_z$  é a componente da magnetização na projeção do eixo z,  $M_0$  é a magnitude do vetor magnetização no equilíbrio térmico;  $t$  é o tempo decorrido após o pulso de rádio-freqüência e  $T_1$  é a constante de tempo na função exponencial. Quando  $t$  é igual a  $T_1$  tem-se o seguinte valor:

<sup>4</sup> A vizinhança no estado sólido representa o movimento vibracional da rede cristalina, já no líquido o movimento browniano de átomos e moléculas.

$$M_0 \left(1 - e^{-1}\right) = M_0 \left(1 - \frac{1}{2,7}\right) = 0,63M_0 \quad \dots\dots\dots (1.17)$$

Dessa forma, podemos definir o tempo de relaxação  $T_1$  como o tempo necessário para que 63% da magnetização seja recuperada na direção z.

Em adição ao forte campo magnético estático criado pelo magneto, todos os prótons do corpo também experimentam pequenos campos magnéticos locais, produzidos pelos prótons, que formam a vizinhança. O movimento térmico das moléculas, que inclui movimento linear, rotacional e vibracional, permite que esse campo magnético local flutue. Esta flutuação é multidirecional e consiste de muitas frequências. Dipolos magnéticos com frequência baixa ou alta produzirão pouca influencia sobre os tempos de relaxação  $T_1$ . Aquelas componentes do campo magnético local que tem frequência de Larmor e uma orientação apropriada podem girar os prótons de uma direção paralela para uma antiparalela, ou vice versa. Ainda, o tempo de relaxação  $T_1$  é diferente nos tecidos e depende das flutuações de campos locais.

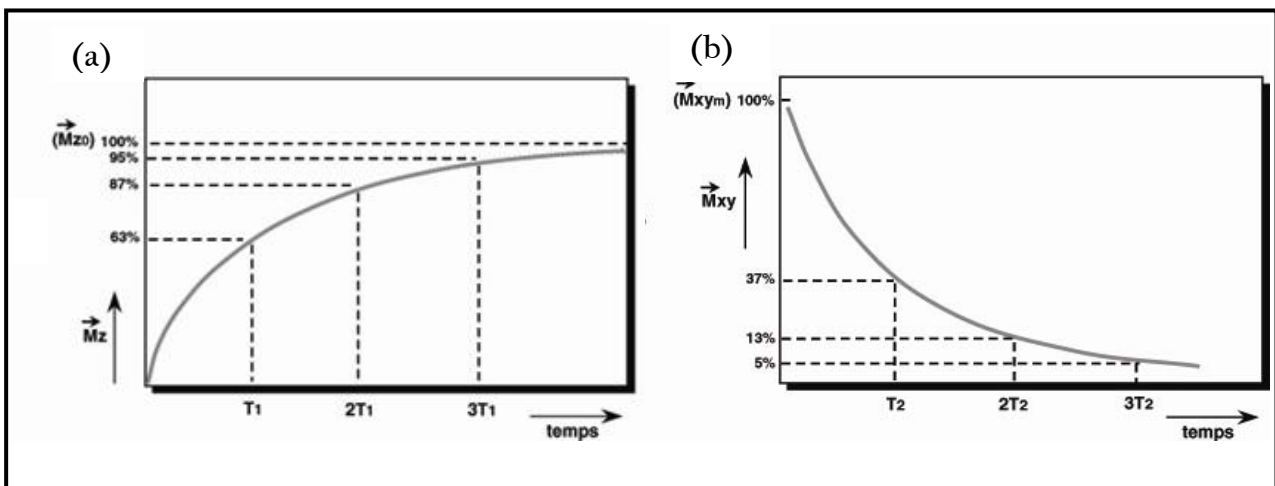


Figura 1.3. Comportamento do vetor magnetização em função do tempo. (a) Decaimento exponencial de  $M_z$  na direção z. (b) Decaimento exponencial do vetor magnetização  $M_{x-y}$  no plano x-y.

### 2.1.1.2 Relaxação Transversal ( $T_2$ e $T_2^*$ )

Depois que a magnetização é girada para o plano transversal pelo pulso RF, os momentos de dipolo da amostra passam a realizar precessão no plano x-y, com mesma fase. Com o tempo, a coerência de fase entre os momentos de dipolo é gradualmente perdida (ver figura 1.4).

Em geral, há duas causas para a relaxação transversal, uma intrínseca e outra extrínseca. A causa intrínseca vem da interação *spin-spin*: quando muitos *spins* estão excitados, há perda de coerência devido aos efeitos de um sobre o outro. A diminuição da magnetização no plano x-y pela perda de coerência de fase ocorre devido a variações randômicas na intensidade do campo magnético local que induz uma alteração na freqüência de precessão. Lembrando que a freqüência de Larmor é  $\omega = \gamma B$ , então se há flutuação de campo local, haverá alteração em  $\omega$ .

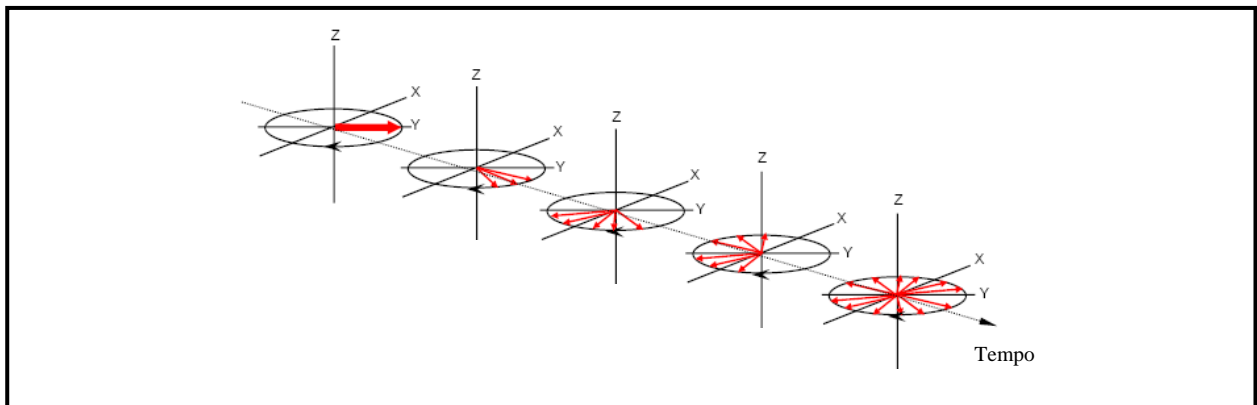


Figura 1.4. Após o pulso RF de 90°, há perda de coerência oriunda de flutuações de campos magnéticos locais, resultando no processo de relaxação T<sub>2</sub>

Não obstante, flutuações de baixa freqüência têm efeito predominante na perda da magnetização M<sub>x-y</sub>. Entretanto, flutuações em torno da freqüência de Larmor também provocam encurtamento de T<sub>2</sub>. Motivo pelo qual T<sub>2</sub> é sempre menor ou igual a T<sub>1</sub>. A perda de sinal deste mecanismo interno é chamado decaimento T<sub>2</sub> e é caracterizado pela constante conhecida como T<sub>2</sub>. Matematicamente, a componente M<sub>x-y</sub> pode ser medida pela seguinte equação:

$$M_{xy} = M_0 e^{-t/T_2} \tag{1.18}$$

Um gráfico desse decaimento é apresentado na figura 1.3 (b). Quando  $t = T_2$ ,  $M_{xy} = 0,37M_0$ . Podemos então definir T<sub>2</sub> como sendo o tempo necessário para a magnetização transversal atingir 37% da magnetização do estado de equilíbrio térmico.

Uma outra fonte, extrínseca contribui para a diminuição de M<sub>x-y</sub>: a inhomogeneidade do campo magnético estático. Uma vez que cada *spin* gira em uma freqüência proporcional ao campo magnético

externo, variações no campo em diferentes locais permitem diferentes frequências de precessão, também levando a perda de coerência. Os efeitos combinados de interação *spin-spin* e a inhomogeneidade de campo leva a uma perda de sinal conhecido como decaimento  $T_2^*$ . Logo a taxa de decaimento  $1/T_2^*$  é:

$$\frac{1}{T_2^*} = \frac{1}{T_2} + \frac{1}{T_{2inh}} \tag{1.19}$$

Observe que o decaimento de  $T_2^*$  é sempre mais rápido do que  $T_2$ , uma vez que inclui um fator adicional da não homogeneidade de campo. Em geral  $T_2^* < T_2 \leq T_1$ .

### 2.1.2 Recepção do Sinal FID

Mostramos que o vetor magnetização do sistema de *spin* pode ser alterado aplicando-se um pulso de excitação eletromagnética com uma bobina transmissora. Para medir o sinal dessa magnetização, após a excitação, precisamos de outra bobina receptora. O sinal adquirido pelas bobinas receptoras é governado pela Lei de Indução de Faraday. Depois que a magnetização da amostra está no plano transversal, o sistema de *spin* tem precessão na frequência de Larmor e causa um fluxo magnético ( $\phi$ ) variável no tempo através da bobina receptora. Esta mudança de fluxo  $d\phi/dt$  induz uma força eletromotriz (fem):

$$fem = -\frac{d\phi}{dt} \tag{1.20}$$

O fluxo magnético gerado pela amostra e que penetra na área da bobina é:

$$\phi = \int_v \bar{B}_1(M(t)dv), \tag{1.21}$$

sendo  $\bar{B}_1$  o campo por unidade de corrente e  $M(t)$  a magnetização criada pela amostra. Substituindo a equação 1.21 em 1.20 verificamos que:

$$fem = -i\omega_0 \int_v \bar{B}_1(M(t)dv) \tag{1.22}$$

Observamos que a força eletromotriz induzida oscila na frequência de Larmor e, lembrando que ambos  $M$  e  $\omega_0$  são proporcionais ao campo magnético  $B_0$ , então a fem é proporcional a  $B_0^2$ . Isso significa que o sinal MR recebido pela bobina detectora aumenta com o quadrado do campo magnético principal. Segundo a equação 1.22 apenas a componente da magnetização no plano x-y gera um sinal detectável na bobina receptora.

O sinal detectado apresenta um padrão de decaimento da magnetização (ver figura 1.4) no plano x-y provocado pelos processos de relaxação  $T_1$  e  $T_2^*$ . Este padrão é conhecido como FID (*Free Induction Decay*).

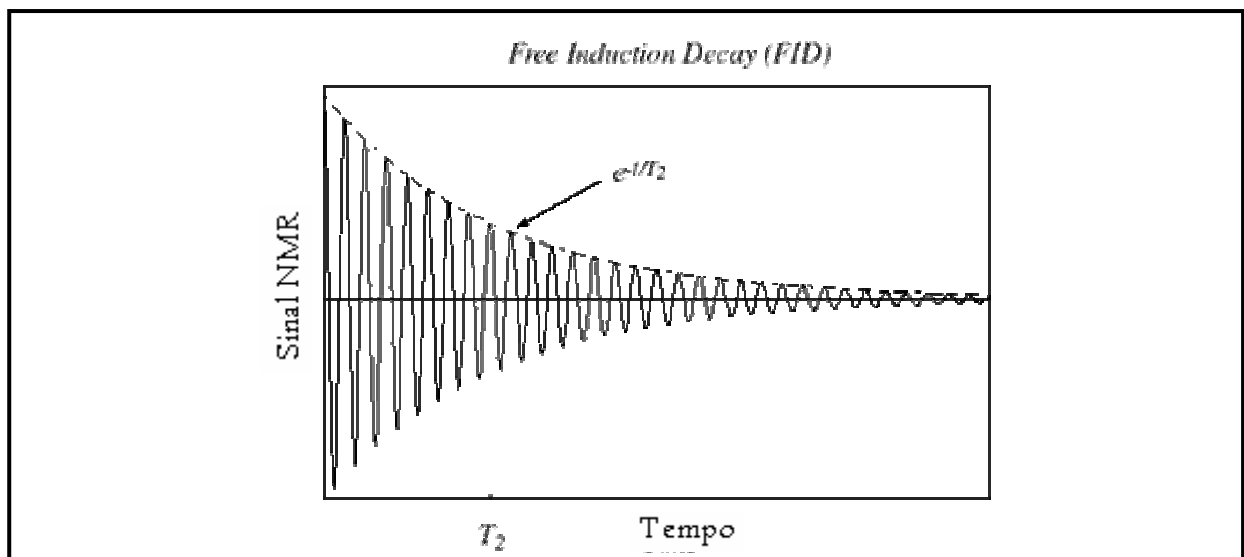


Figura 1.4. Gráfico do sinal detectado por bobinas receptoras no plano x-y. Conhecido como FID, reflete o efeito de  $T_2^*$  sobre a amplitude.

## 2.2 Formação da Imagem, Contraste e Seqüências de Pulso em MRI

Até aqui estudados os princípios básicos da formação do sinal de ressonância magnética que é de extrema importância para a formação da imagem. Esses sinais analógicos são convertidos por um digitalizador para posteriormente formar as imagens digitais e possibilitar sua visualização e manipulação.

Uma imagem (do latim *imago*) digital é essencialmente uma distribuição de propriedades físicas do objeto. É formada por um número finito de elementos chamados *pixel*<sup>5</sup> que tem uma determinada intensidade. O conjunto de *pixels* está distribuído bidimensionalmente (em colunas e linhas) que formam a matriz. Na prática, os sinais que determinam a intensidade dos *pixels* vêm de um volume do tecido, definido por um *pixel* em duas dimensões e a espessura da fatia em uma terceira dimensão. Este volume é chamado de voxel.

O contraste é outro parâmetro importante na imagem, pois reflete as diferenças das intensidades dos *pixels*. Ao valor do *pixel* pode ser atribuída uma escala de cinza, formada por um grande espectro de representações de tonalidades entre branco, cinza e o preto, resultando no brilho de imagem. É interessante notar que se o conversor analógico-digital só tiver 8 bits, ele só poderá atribuir valores entre 0 e 255, se o conversor tiver 10 bits, ele poderá armazenar valores entre 0 e 1024. Em resumo, dependendo do número de bits, o conversor poderá fornecer  $2^{\text{número de bits}}$  diferentes tonalidades.

O objetivo da MRI é a formação de uma imagem, ou seja, um mapa que descreva a distribuição espacial das propriedades relacionadas aos *spins* da amostra. Estas propriedades podem refletir, por exemplo, a densidade dos *spins*, tempos de relaxação  $T_1$  ou  $T_2$  dos tecidos no qual os *spins* se encontram. Por exemplo, imagens ponderadas em  $T_1$  descrevem a distribuição espacial dos valores  $T_1$ , em que voxels com valores curtos de  $T_1$  são brilhantes e voxels com longos  $T_1$  são escuros.

## 2.2.1 Princípios Básicos de Formação da MRI

Nas seções anteriores descrevemos dois tipos de campos magnéticos que são importantes na geração do sinal MR: o campo magnético estático  $B_0$  e o pulso de rádio-freqüência,  $B_1$ . Agora, introduzimos um terceiro tipo de campo magnético, o gradiente espacial,  $G$ , o qual altera a freqüência de precessão dos *spins* dependendo de sua localização espacial, promovendo informação da sua exata localização. Matematicamente, considere um gradiente espacial aplicado da direção  $x$ :

---

<sup>5</sup> *Pixel* é o menor ponto que forma uma imagem digital, sendo que o conjunto de milhares de *pixels* forma a imagem inteira. O nome *Pixel* vem da aglutinação de *Picture Cel*, célula de imagem. [WIKIPEDIA, 2005]



$$B_z = (B_0 + G_x) \hat{k}. \quad \dots\dots\dots (1.23)$$

$G_x = \frac{\partial B_z}{\partial x}$  designa a variação do campo magnético ao longo de x, fazendo com que  $B_z$  cresça na direção de x. Generalizando-se, a intensidade do campo magnético total,  $B_z$ , experimentado pelo sistema de *spins* em uma dada localização espacial (x, y, z) e tempo (t) consiste em uma combinação linear do campo estático e gradientes de campos variando no tempo em direções específicas:

$$B_z = (B_0 + \vec{G} \cdot \vec{r}). \quad \dots\dots\dots (1.24)$$

Como consequência da aplicação do gradiente de campo, a frequência de Larmor é diferente espacialmente:

$$\omega(\vec{r}) = \gamma(B_0 + \vec{G} \cdot \vec{r}) \quad \dots\dots\dots (1.25)$$

Na presença dos gradientes G, a intensidade de  $B_z$  varia dependendo da localização espacial permitindo, assim, selecionar as porções (x, y, z) que tem frequência única ( $\omega_{x,y,z}$ ). Isso significa que na aplicação do gradiente, qualquer plano perpendicular à direção do gradiente terá sua própria intensidade de campo magnético. Os prótons em qualquer plano terão em comum, única frequência de Larmor, diferente dos prótons em qualquer outro plano perpendicular a direção do gradiente.

Os gradientes  $G_z$ ,  $G_x$  e  $G_y$  são respectivamente gradientes de seleção de fatia, codificação de frequência e codificação de fase. Na presença do gradiente de campo  $G_z$ , o sistema de *spins* realizará um movimento de precessão com frequência característica. Se esse volume de tecido é exposto a um pulso de RF, apenas os prótons com frequência de Larmor igual à frequência de pulso RF serão excitados, possibilitando a seleção de fatia. Entretanto a seleção de fatia pode ser realizada não apenas na direção z. Se desejamos fatias sagitais ou coronais, os gradientes  $G_x$  ou  $G_y$  podem ser usados.

Finalmente, para selecionar pontos específicos do espaço, aplicam-se os gradientes nas outras direções. Se a aplicação do gradiente para a seleção de fatia for  $G_z$ , os gradientes na direção x e y são utilizados para selecionar as linhas e os pontos do espaço.

Para adquirirmos as imagens de toda uma região no espaço, o processo deve ser seqüencial. Primeiro selecionamos o plano pela aplicação de um dos gradientes e aplicamos um pulso de RF de forma a girar a magnetização para o plano transversal. Aplicamos, sucessivamente, gradientes de campo nas duas direções selecionando toda região da fatia.

Essas seqüências são conhecidas como seqüências de pulso, geralmente representadas na forma de diagramas, que resume a evolução temporal da aplicação dos pulsos de RF e dos diferentes gradientes de campo.

### 2.2.2 Seqüência de Pulso Eco de *Spins*

Para um entendimento mais claro sobre o mapeamento espacial dos sinais medidos pela NMR, abordaremos a seqüência eco de *spin* (*Spin Echo*) desenvolvido por Hahn [HAHN, 1950]. A figura 1.4 mostra o diagrama desta seqüência de pulso.

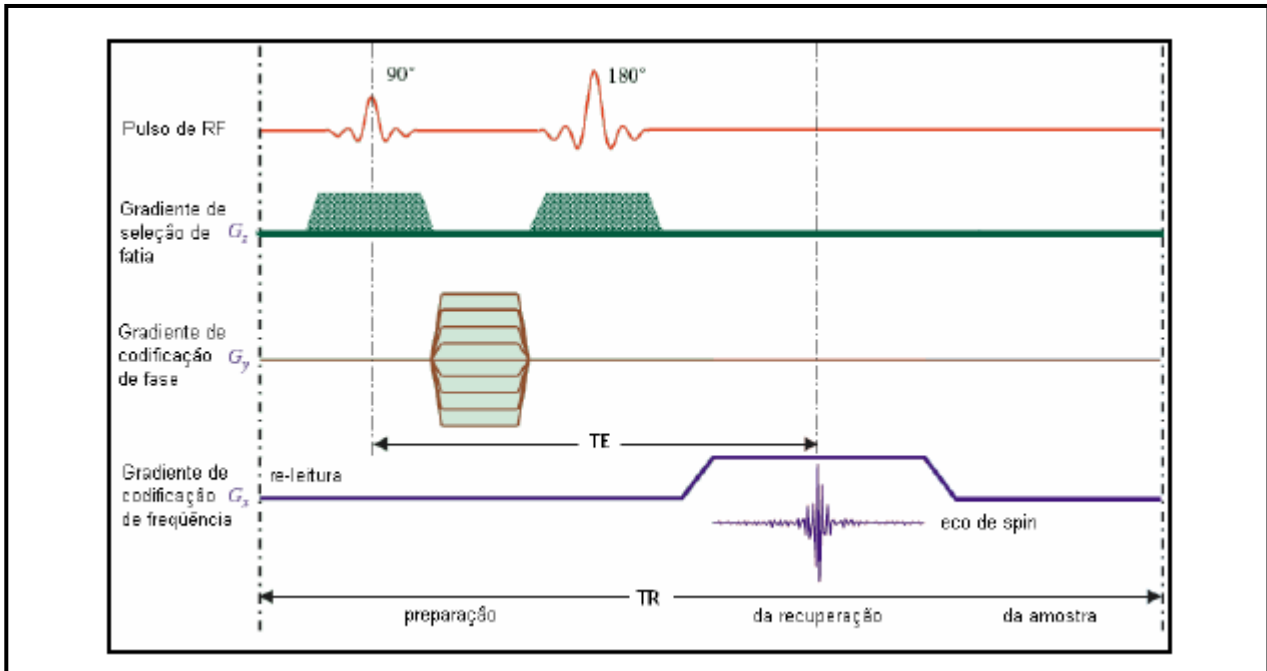


Figura 1.5. Essa seqüência permite a separação dos tempos de relaxação  $T_2$  e  $T_2^*$ , que consiste na aplicação de dois pulsos de RF: um de  $90^\circ$  e outro de  $180^\circ$ , e gradientes de Campo  $G_x$ ,  $G_y$  e  $G_z$ . *Figura modificada de Logothetis 2002.*

Acompanhando a figura 1.5 podemos verificar que em um primeiro momento, o gradiente de seleção de fatia é aplicado ( $G_z$ ) simultaneamente com o pulso RF de  $90^\circ$  que gira o vetor magnetização

para o plano x-y (Figura 1.6(a)). Após um curto período de tempo, os *spins* começam a perder coerência de fase (Figura 1.6(b)), por conta das flutuações de campo magnético local, realizando precessão na frequência de Larmor. Após a retirada do pulso RF, aplica-se o gradiente de codificação de fase ( $G_y$ ) que permite a codificação da posição espacial do sinal ao longo da dimensão y por diferentes ângulos de fase. Nós desejamos perdurar a coerência de fase porque ela fornece um melhor sinal. Um tempo  $\tau$  após a aplicação do pulso de  $90^\circ$ , um outro pulso de  $180^\circ$  (Figura 1.6(c) e (d)) é aplicado combinado ao gradiente de seleção de fatia, invertendo as posições dos dipolos magnéticos no plano transversal. Portanto, os *spins* que estão girando com velocidade mais rápida passam para uma posição mais atrasada em relação aos que estavam girando com velocidade inferior, e esses passam para uma posição mais adiantada (Figura 1.6(e)). Depois de um tempo  $2\tau$ , ocorre o realinhamento dos *spins* produzindo um eco de *spin* (Figura 1.6(f)). Durante a aquisição dos dados, o gradiente  $G_x$  é ligado, mudando a frequência dos *spins* e possibilitando a localização espacial pela frequência.

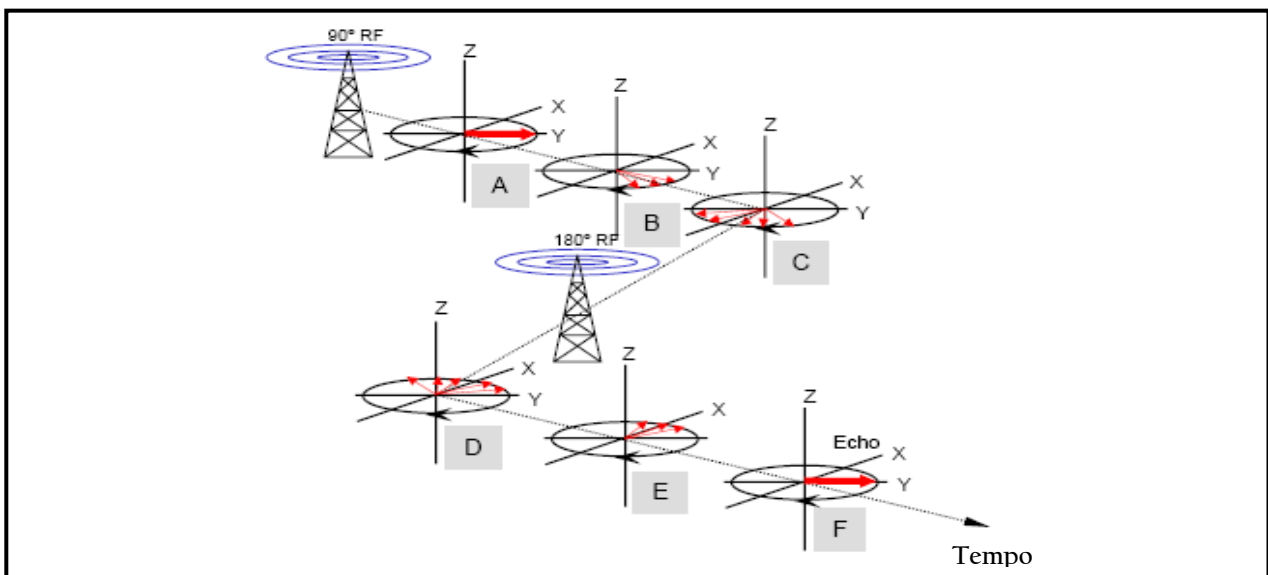


Figura 1.6. Esquema ilustrativo da posição do vetor magnetização em uma seqüência de pulso do tipo eco de *spin*.

Dentre os parâmetros das seqüências de pulso que permitem a manipulação de contraste estão o TR, TE e FA. O tempo de Repetição (TR) consiste no tempo decorrido entre a aplicação de dois pulsos de  $90^\circ$  sucessivos. O TR determina o grau de relaxamento  $T_1$  que pode ocorrer entre o término de um pulso RF ( $\pi/2$ ) e a aplicação do pulso seguinte. O tempo entre o pulso RF ( $90^\circ$ ) e o pico máxi-

mo do sinal induzido é denominado tempo de eco. O TE determina o grau de declínio de magnetização transversal (relaxação  $T_2$ ) que pode ocorrer antes da leitura do sinal. O FA ou *Flip Angle* refere-se à quantidade de graus que o vetor magnetização é girado da sua posição de equilíbrio, podendo estar entre 0-180°. Nas seqüências de eco de *spin* o FA é 90° e 180°, o TR pode estar entre 100-3000 ms e o TE entre 5-250 ms.

### 2.2.3 Mecanismos de Contraste

A MRI é extremamente versátil, se compararmos a outros métodos de neuroimagem sendo excelente na discriminação dos tecidos moles. Nesta técnica, podemos ter diferentes mecanismos de contraste: estático e dinâmico. Os estáticos são sensíveis ao tipo, número e propriedades de relaxação do núcleo atômico dentro de um voxel. Contrastes típicos incluem densidade de prótons, tempos de relaxação  $T_1$ ,  $T_2$  e  $T_2^*$ . Esses contrastes são usualmente utilizados para determinar a anatomia cerebral nos experimentos de MRI. Os contrastes dinâmicos são sensíveis ao movimento do núcleo atômico. Contrastes dinâmicos típicos fornecem informações sobre características dinâmicas dos prótons, como o fluxo sanguíneo, difusão da água, ou irrigação capilar.

Uma distinção adicional pode ser extraída dependendo se o contraste depende das propriedades intrínsecas do tecido biológico (contraste endógeno) ou da presença de substâncias externas que foram introduzidas no corpo (contraste exógeno).

As características físicas que se deseja medir e visualizar na imagem estão intimamente ligadas à seqüência de pulso adotada. Há um número grande de seqüências de pulso disponíveis que permitem medidas de características diferentes dos tecidos. Não obstante, a imagem deve apresentar *pixels* com maior ou menor brilho que depende de parâmetros intrínsecos do tecido bem como de parâmetros de máquina.

A intensidade do sinal ou brilho do *pixel* é dependente desses parâmetros, e variam de acordo com a seqüência de pulso. Novamente, o sinal MR medido depende do grau de magnetização transversal detectado. Essa magnetização depende de fatores como os tempos de relaxação  $T_1$ ,  $T_2$ , densidade de

prótons e parâmetros de pulso como TR e TE já mencionados. Matematicamente a magnetização transversal pode ser escrita como

$$M_{xy}(t) = M_0 \left(1 - e^{-TR/T_1}\right) e^{-TE/T_2} . \quad \dots\dots\dots (1.26)$$

Na tabela 1.2 apresentamos alguns valores de constantes de tempo T<sub>1</sub> e T<sub>2</sub> em um campo magnético de 1.5T.

**Tabela 1.2 – Valores para as constantes de tempo T<sub>1</sub> e T<sub>2</sub> em um campo de 1.5 T. Tabela modificada de HUETTEL 2004 pg. 101.**

	Matéria Cinza	Matéria Branca	Fluído Cérebro Espinhal
T <sub>1</sub>	900 ms	600 ms	4000 ms
T <sub>2</sub>	100 ms	80 ms	2000 ms

Para dois tecidos A e B, o contraste entre eles, C<sub>AB</sub>, é simplesmente a diferença entre o sinal MR associado a cada um deles:

$$C_{AB} = M_{0A} \left(1 - e^{-TR/T_{1A}}\right) e^{-TE/T_{2A}} - M_{0B} \left(1 - e^{-TR/T_{1B}}\right) e^{-TE/T_{2B}} \quad \dots\dots\dots (1.27)$$

Os termos M<sub>0A</sub> e M<sub>0B</sub> correspondem às magnetizações originais dos tecidos A e B, os termos T<sub>1A</sub> e T<sub>1B</sub> são os tempos de relaxação longitudinal de A e B, e T<sub>2A</sub> e T<sub>2B</sub> são os valores T<sub>2</sub> de A e B.

### 2.2.3.1 Contraste por Densidade de Prótons

Como o nome implica as imagens ponderadas por densidade de prótons (DP), fornecem um contraste baseado no número de prótons em um voxel, que difere entre tecidos. Para maximizar o contraste por densidade de prótons, utilizam-se seqüências de pulso que minimizam o contraste de T<sub>1</sub> e T<sub>2</sub>. Para maximizar a ponderação por densidade de prótons, são utilizados TR longo e TE curto nas seqüências de pulso. Um TR maior do que T<sub>1</sub> e um TE menor que T<sub>2</sub> satisfaz esse critério. Os voxels da imagem com alta densidade (i.e., fluido cérebro espinhal) são brilhantes enquanto áreas com baixas

densidades (i.e., ar, matéria branca) escuras. Um exemplo de imagem ponderada por densidade de prótons pode ser vista na figura 1.7 (a).

### 2.2.3.2 Contraste por $T_1$

Quando a intensidade do sinal dos voxels da imagem depende predominantemente de valores  $T_1$  do tecido, dizemos que a imagem é ponderada em  $T_1$ . Usada na obtenção de contraste estrutural para imagens anatômicas do cérebro, as imagens ponderadas em  $T_1$  necessitam de um TR intermediário e TE curto. Imagens com TR muito longo não dependem de  $T_1$ . Por outro lado, com TRs muito curtos não há tempo para uma recuperação de  $M_z$  e nenhum sinal será registrado. Além disso, há necessidade de se minimizar efeitos de  $T_2$ . As seqüências de pulso devem ter um TE curto comparado com  $T_2$ . Quando o TE é muito menor que  $T_2$  o termo  $e^{-TE/T_2}$  da equação 1.41 torna-se aproximadamente igual a 1. Então a equação depende apenas de TR e não de TE, a equação reduz para:

$$C_{AB} = M_{0A} \left(1 - e^{-TR/T_{1A}}\right) - M_{0B} \left(1 - e^{-TR/T_{1B}}\right). \quad \dots\dots\dots (1.28)$$

Em uma imagem ponderada em  $T_1$ , tecidos com um curto valor de  $T_1$  produzirão um eco intenso e aparecerão com maior brilho na imagem. Em adição, tecidos com alta densidade de prótons também aparecerão brilhantes (ver figura 1.7 (b)).

### 2.2.3.3 Contraste por $T_2$

Nas imagens ponderadas em  $T_2$ , a quantidade de perda de sinal depende do tempo entre a excitação e a aquisição de dados, ou TE. Novamente, uma combinação ótima dos parâmetros TR e TE existe, objetivando maximizar o contraste  $T_2$ . Se o TE for curto ou a aquisição dos dados for realizada logo após a excitação, pouca magnetização transversal terá sido perdida e não haverá contraste em  $T_2$ . Por outro lado, se o TE for muito longo, então toda a magnetização transversal será perdida e também

não haverá contraste em  $T_2$ . Mas se a aquisição dos dados for realizada em um TE intermediário, na ordem de  $T_2$ , a magnetização transversal pode ser maximizada.

Para obter contraste em  $T_2$ , nós precisamos minimizar efeitos da relaxação  $T_2$  aplicando um TR longo, permitindo a completa recuperação da magnetização longitudinal. Quando o TR for muito maior do que  $T_1$ , o termo  $e^{-TR/T_1}$  da equação 1.28 aproxima-se de 0 e pode ser eliminado. A equação passa a ter alta dependência do TE:

$$C_{AB} = M_{0A}e^{-TE/T_{2A}} - M_{0B}e^{-TE/T_{2B}} . \quad \dots\dots\dots (1.29)$$

Nas imagens ponderadas em  $T_2$ , regiões com  $T_2$  mais longos apresentarão um brilho mais intenso do que aquelas com  $T_2$  mais curto. O CSF tem um valor de  $T_2$  longo devido à alta concentração de água, a matéria cinzenta tem um TE intermediário pelo seu rico suplemento sanguíneo, e a matéria branca tem um valor de  $T_2$  muito curto. A imagem resultante será brilhante em regiões de fluido como os ventrículos, um brilho intermediário na matéria cinza, e escuro na matéria branca (ver figura 1.7 (c)).

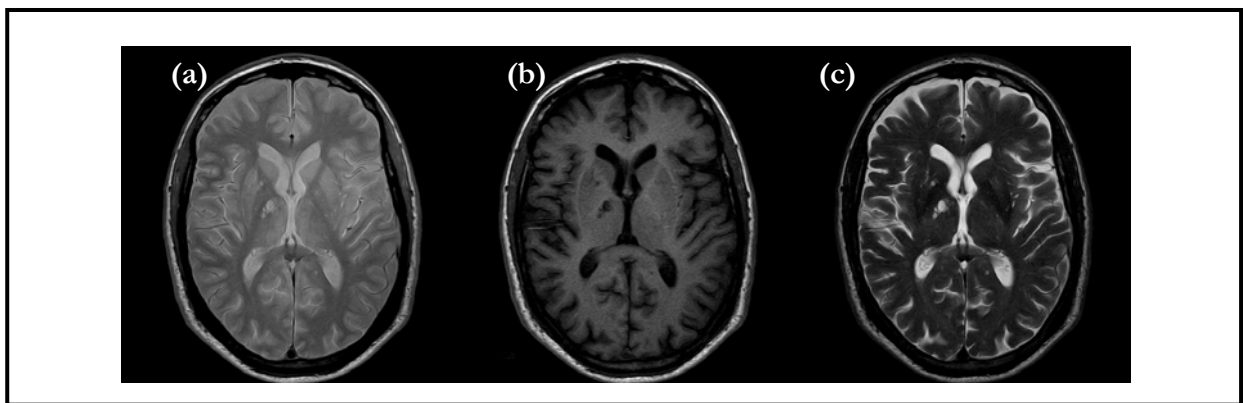


Figura 1.7. Imagens cerebrais no plano axial de um tomógrafo Philips de 3 T. (a) Imagem ponderada em DP, (b) Imagem ponderada em  $T_1$  e (c) imagem ponderada em  $T_2$ . Figura modificada de *Institute for Biomedical Engineering*.

Apresentamos na tabela 1.3, o resumo dos tipos de ponderação discutidos com respectivos parâmetros e efeitos na imagem.

Tabela 1.3 – Valores dos parâmetros TR e TE para as ponderações de imagem e respectivos efeitos na tonalidade. *Tabela modificada de SMITH 1989.*

Tipo de Imagem Ponderada em:	Tamanho relativo de:		Efeitos no brilho da imagem dos tecidos com:		
	TR	TE	Densidade de Próton	T <sub>1</sub> longo	T <sub>2</sub> longo
Densidade de Próton	Longo	Curto	+	(-)	(+)
T <sub>1</sub>	Curto	Curto	+	-	(+)
T <sub>2</sub>	Longo	Longo	+	(-)	+

+: indica que a porção do tecido aparecerá brilhante na imagem;

-: indica que a porção do tecido aparecerá escura na imagem;

(+) (-) indica uma possível menor dependência no brilho da imagem.

### 2.2.4 Reconstrução da Imagem

O método mais utilizado nos equipamentos modernos é o método que utiliza a transformada de Fourier para a reconstrução da imagem. A transformada de Fourier de um sinal consiste na decomposição do sinal periódico S(t) em um conjunto de senos e cossenos de diferentes frequências e amplitudes. A Série de Fourier desta função periódica foi proposta por Jean Baptiste Fourier (1786-1830) [DE ARAÚJO, 2002], e expressa matematicamente por:

$$S(t_j) = a_0 + \sum_{k=1}^{\infty} [a_k \cos(\omega_0 t_j) + b_k \text{sen}(\omega_0 t_j)], \quad \dots\dots\dots (1.30)$$

de modo que os coeficientes a<sub>k</sub> e b<sub>k</sub> podem ser encontrados pelas seguintes equações:

$$a_k = \frac{1}{N} \sum_{j=0}^{N-1} y_j \cos\left(\frac{2\pi jk}{N}\right) \quad \dots\dots\dots (1.31)$$

e

$$b_k = \frac{1}{N} \sum_{j=0}^{N-1} y_j \text{sen}\left(\frac{2\pi jk}{N}\right) \quad \dots\dots\dots (1.32)$$

A série permite o estudo das componentes de frequência, de tal sorte que podemos diferenciar padrões de frequência e fase dos sinais de NMR proveniente de diferentes partes da amostra. Dado um sinal é possível especificar os coeficientes dos pares de senos e cossenos que podem ser usados para gerar o



objeto. Logo, o objeto da imagem de Fourier é a medida dos coeficientes de Fourier cuja imagem pode ser reconstruída através deles. Para a reconstrução de uma imagem bidimensional, o método precisa de um conjunto completo de  $N \times M$  (onde  $N \times M$  é o tamanho da matriz da imagem, i.e.  $128 \times 128$ ) coeficientes de Fourier. Os coeficientes de Fourier são medidos pelo uso apropriado de gradientes de campo magnético seguidos de um pulso de excitação, onde a representação dos gradientes pode ser feita no espaço-k.

### 2.2.5 Espaço-k

O sinal FID pode ser obtido pela excitação, seletiva ou não, por meio de um pulso de radio-freqüência. A contribuição de um sinal que é induzido em uma bobina receptora pelo volume de um elemento da amostra tem magnitude igual ao número de densidade dos *spins* na posição  $(x, y)$  multiplicado pelo tamanho de um elemento de *pixel*  $dx dy$ . A contribuição do sinal na posição  $(x, y)$  pode, entretanto, ser escrito como o somatório de uma parte real e imaginária:

$$dS(x, y, t) = \rho(x, y) \{ \cos [2\pi(\gamma B_0 + \gamma G_x x + \gamma G_y y) t] + i \text{sen} [2\pi(\gamma B_0 + \gamma G_x x + \gamma G_y y) t] \} dx dy \dots\dots\dots (1.33)$$

em que  $\rho(x, y)$  é a densidade de *spins* na localização  $(x, y)$  e  $dS(x, y)$  é a contribuição do sinal dos *spins*.

O sinal que o scanner armazena é a soma da contribuição de todos os pontos em um plano  $(x, y)$  resultado em um sinal  $S(t)$  dado por:

$$dS(x, y, t) = \iint \rho(x, y) \{ \cos [2\pi(\gamma B_0 + \gamma G_x x + \gamma G_y y) t] + i \text{sen} [2\pi(\gamma B_0 + \gamma G_x x + \gamma G_y y) t] \} dx dy \dots\dots\dots (1.34)$$

Esta equação mostra que o sinal total de uma fatia depende da magnetização em todas as localizações  $(x,y)$ . Contudo, essa equação é de difícil visualização e solução, na presente forma. Simplificamos a equação acima, introduzindo dois termos do espaço recíproco. Para gradientes constantes estes termos podem ser escritos como:

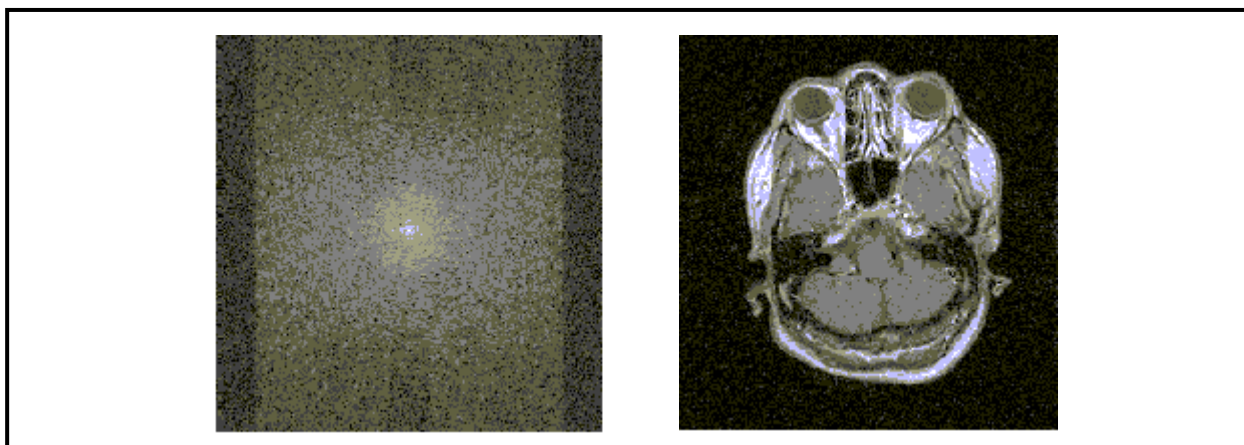
$$k_x(t) = 2\pi\gamma G_x t. \quad \dots\dots\dots (1.35)$$

Substituindo na equação 1.34 os termos do espaço k, a relação fica simplificada da seguinte forma:

$$k_y(t) = 2\pi\gamma G_y t. \quad \dots\dots\dots (1.36)$$

A equação acima pode ser pensada como coeficientes de Fourier das densidades de *spin*,  $\rho$ , e que a densidade de *spin* pode ser reconstruída da medida dos coeficientes de Fourier via transformada de Fourier. Um mapa de  $\rho(x, y)$  é precisamente o que pretendemos medir. A equação 1.36 sugere que uma transformada de Fourier pode converter os dados do espaço-k em uma imagem, um processo conhecido como reconstrução de imagem.

As equações do par de Fourier relacionados ao sinal MR e a densidade  $\rho$  são mostradas abaixo. A figura 1.8 apresenta uma representação esquemática das equações do par de Fourier.



**Figura 1.8.** A figura da esquerda apresenta valores dos coeficientes de Fourier (S) de uma imagem como função das coordenadas  $k_x$  e  $k_y$ . A figura da direita apresenta valores da densidade de *spin*  $\rho$  como uma função de  $x$  e  $y$ .

Os eixos de coordenadas  $k_x$  e  $k_y$ , conhecidos como gradientes de codificação de frequência e de fase respectivamente, definem o espaço-k. A seqüência de aplicação dos gradientes determina as diferentes trajetórias no espaço recíproco. Três trajetórias com aplicações específicas são apresentadas na figura 1.9. A trajetória da Transformada Rápida de Fourier 2D (2DFT) é a mais simples e mais utilizada trajetória na MRI. Um dos benefícios desta trajetória é a facilidade que uma imagem pode ser reconstruída, por outro lado à necessidade de um grande número de excitações torna o processo lento.

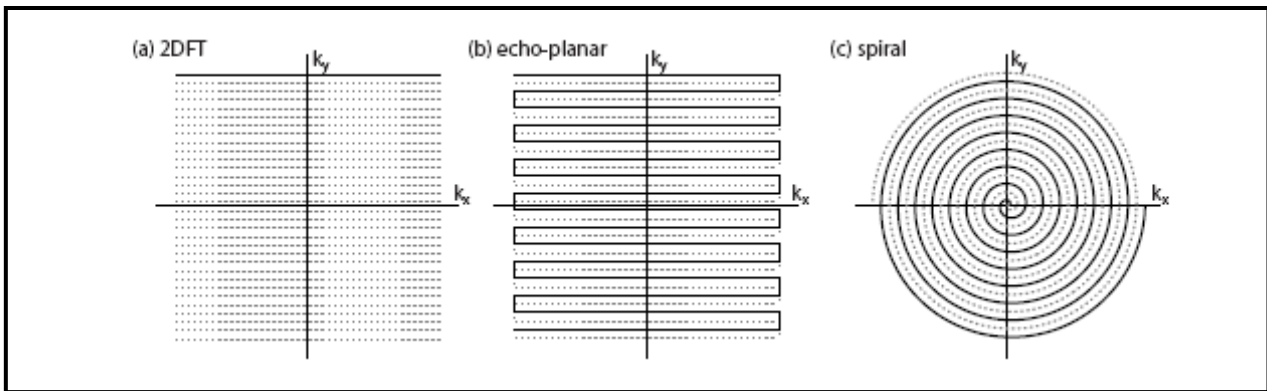


Figura 1.9: Trajetórias do espaço-k mais comuns: (a) 2DFT; (b) *echo-planar imaging* (EPI) e (c) espiral. *Figura modificada de MILLER 2004.*

Uma série de trajetórias alternativas foram propostas para permitir a aquisição de imagens mais rápidas. A trajetória *echo-planar* (EPI) foi desenvolvida por Mansfield [MANSFIELD, 1977] e tem a vantagem de coletar uma matriz bidimensional em um único pulso de RF. Uma outra trajetória rápida desenvolvida mais recentemente é a imagem espiral.

### 2.3 Modelo Quântico

A propriedade do momento magnético que o núcleo exibe pode ser descrito quanticamente, onde os prótons acoplados aos nêutrons por forças nucleares, possuem momento angular de *spin*  $\vec{J}$  que está relacionado ao momento magnético  $\vec{\mu}$  da seguinte maneira:

$$\vec{\mu} = \left( \frac{q\hbar}{2m} \right) \vec{J} \text{ ou } \vec{\mu} = \gamma \vec{J} \dots\dots\dots (1.37)$$

em que  $q$  é a carga da partícula,  $m$  é a massa,  $\hbar$  a constante de Planck e  $\gamma$  é a constante ou razão giro-magnética, que traduz a razão entre o momento magnético e o momento angular dos núcleos. O momento angular pode ainda ser associado a um operador de *spin*  $\vec{I}$ ,

$$\vec{J} = \hbar \vec{I} \dots\dots\dots (1.38)$$

com valores de projeção discretizados e determinados pelo número quântico  $m_I$  que pode assumir:  $\hbar, (1-1)\hbar, \dots, -\hbar$  projeções. Em um campo magnético, um núcleo de número de *spin*  $I$  ocupará  $(2I + 1)$

níveis discretos de energia. O  $^1\text{H}$ , com número quântico de *spin*  $\vec{I} = 1/2$ , apresenta duas orientações relativas possíveis  $m_I = +1/2, -1/2$ . Por outro lado, a interação que ocorre entre o campo magnético aplicado e os momentos magnéticos dos *spins* é regida pelo Hamiltoniano (que deve ser escrito na forma de um operador):

$$H = -\vec{\mu} \cdot \vec{B}. \quad \dots\dots\dots (1.39)$$

Com a equação 1.37 e 1.38, o momento magnético pode ser escrito como:

$$\vec{\mu} = \gamma \hbar \vec{I}. \quad \dots\dots\dots (1.40)$$

E considerando o campo magnético aplicado  $\vec{B}_0$  na direção z, o Hamiltoniano tem a forma:

$$H = -\gamma \hbar \vec{B}_0 I_z. \quad \dots\dots\dots (1.41)$$

Logo, podemos mostrar os níveis de energia Zeeman nuclear pelos autovalores  $E_m$ :

$$E_m = -\gamma \hbar B_0 m_I. \quad \dots\dots\dots (1.42)$$

O sistema de *spin* do  $^1\text{H}$  com número quântico de *spin*  $\vec{I} = 1/2$ , permite dois níveis de energia

$$E_{\pm} = \pm \frac{1}{2} \gamma \hbar B_0 \quad \dots\dots\dots (1.43)$$

Em outras palavras, o próton pode ocupar um estado paralelo, de mais baixa energia, ou anti-paralelo, de mais alta energia. A diferença de energia entre esses níveis é dada por:

$$\Delta E = \gamma \hbar B_0. \quad \dots\dots\dots (1.44)$$

Ainda, sabe-se que a radiação eletromagnética é composta de pacotes discretos de energia chamados fótons, cuja energia é determinada pela frequência da radiação:

$$E = h\nu. \quad \dots\dots\dots (1.45)$$

Se irradiamos prótons em um campo magnético estático com radiação eletromagnética, podemos induzir alguns prótons a absorver energia da radiação e fazê-los passar do estado de mais baixa energia para mais alta energia. Isto somente ocorre, entretanto, se uma condição muito especial, for satisfeita: a energia da radiação deve ser igual à diferença de energia entre os dois estados:

$$\Delta E = h\nu. \quad \dots\dots\dots (1.46)$$

Desse modo, a transição entre dois estados quânticos envolve a absorção de um fóton com energia igual à  $\Delta E$ , ou a emissão de um fóton com energia igual à  $\Delta E$ . Igualando as equações 1.45 e 1.46 podemos verificar que

$$\hbar\gamma B_0 = h\nu \text{ ou } \nu = \frac{\gamma}{2\pi} B_0. \dots\dots\dots (1.47)$$

Essa equação é uma equação fundamental na NMR, denominada Equação de Larmor. Conhecendo a constante giromagnética, podemos calcular a frequência de ressonância para um dado campo magnético aplicado.  $\gamma / 2\pi$  tem unidade de frequência por campo magnético (MHz/T), e para o  $^1\text{H}$  é igual a 42,577 MHz/T. Outros núcleos ativos em RM têm razões giromagnéticas diferentes, tendo portando frequências de precessão diferentes.

Em NMR, o sinal medido tem origem em um ensemble de núcleos. Quando um ensemble de prótons é colocado em um campo magnético estático, aqueles paralelos estarão em número maior do que os antiparalelos. No equilíbrio termodinâmico, a razão entre o número de prótons paralelo ( $N_{\text{paralelo}}$ ) e o número de prótons antiparalelo ( $N_{\text{antiparalelo}}$ ), por unidade de massa do tecido, é governada pela equação de Boltzmann:

$$\frac{N_{\text{paralelo}}}{N_{\text{antiparalelo}}} = e^{-\Delta E/kT}, \dots\dots\dots (1.48)$$

em que  $k$  é a constante de Boltzmann e  $T$  é a temperatura absoluta medida em Kelvin (K). O excesso de momentos magnéticos alinhados ao campo, resulta em uma medida macroscópica denominada magnetização  $M$ , ou seja, é a soma vetorial dos momentos magnéticos individuais por unidade de volume:

$$\vec{M} = \sum_{i=1}^N \vec{\mu}_i. \dots\dots\dots (1.49)$$

Partindo da equação de Boltzmann e a relação de energia, podemos mostrar que:

$$M \approx \left( \frac{\hbar^2 \gamma^2}{4k} \right) \left( \frac{B_0 N_p}{T} \right). \dots\dots\dots (1.50)$$

$N_p$  é o número de prótons por unidade de massa do tecido. Isso significa que a magnetização aumenta com o aumento do campo magnético, com o aumento da densidade de prótons e diminui com o aumento da temperatura. Não obstante, a diferença da população de prótons paralelos em relação aos antiparalelos não é muito grande. Na ausência de agitação térmica todos os momentos magnéticos dos núcleos alinham-se com o campo magnético e a amostra encontra-se no estado de menor energia. Entretanto, em temperaturas fisiológicas há uma grande agitação térmica dos núcleos e um conseqüente desequilíbrio entre o número de *spins* paralelos e antiparalelos. O excesso de prótons alinhados com um campo de 0.5 T é apenas 3 ppm (partes por milhão), em um sistema de 1.0 T há 6 ppm e em um sistema de 1.5 T há 9 ppm. [SMITH *et al.*, 1989].

---

## APÊNDICE B - Curvas Características de um Teste de Diagnóstico

### 1. Introdução

A análise das Curvas Características de Operação de um Receptor, *Receiver Operating Characteristic* (ROC), teve origem na teoria de decisão estatística e foi desenvolvida nos anos 1950 para avaliar a detecção de sinais em radares e na psicologia sensorial [METZ, C.E., 1986]. A aplicação da análise ROC na avaliação de diagnósticos médicos começou a ser utilizada nos anos 1970' e, subsequentemente, passou a ser aplicada com sucesso no diagnóstico de imagens médicas [PARKER J., *et al.*, 1995; BURDETTE *et al.*, 1996]. Somente em 1995, as curvas ROC foram introduzidas na análise de fMRI [CONSTABLE, 1995].

### 2. Teoria de Detecção de Sinal

Na teoria de detecção de sinal, o observador tem a tarefa de identificar sinais e ruídos. Desse modo, detectar um sinal que de fato existe consiste em detectar um **verdadeiro positivo**, *true positive* (TP). Da mesma forma, detectar ruído em uma região em que só há ruído significa detectar um **verdadeiro negativo**, *true negative* (TN). Entretanto é possível que se detecte um sinal em uma região onde unicamente ruído está presente, caso este denominado **falso positivo**, *false positive* (FP). É possível, ainda, deixar de detectar um sinal na região de sinal presente, classificando-o como ruído, o que representaria um **falso negativo**, *false negative* (FN).

Seja  $f(x|R)$  a distribuição normalizada de probabilidade que descreve os resultados de alguma variável aleatória  $x$ , quando apenas ruído está presente,  $x_R$ , e seja  $f(x|S)$  a função de probabilidade correspondente ao caso em que sinal e ruído estão presentes,  $x_S$ . Cada resultado  $x$  se torna, então, a entrada de um sistema detector, que deve decidir se dado  $x$  deve ser classificado como ruído ou sinal. O problema de toda a detecção está no fato das distribuições  $f(x|R)$  e  $f(x|S)$  não serem disjuntas, como apresentado na figura 1.1.

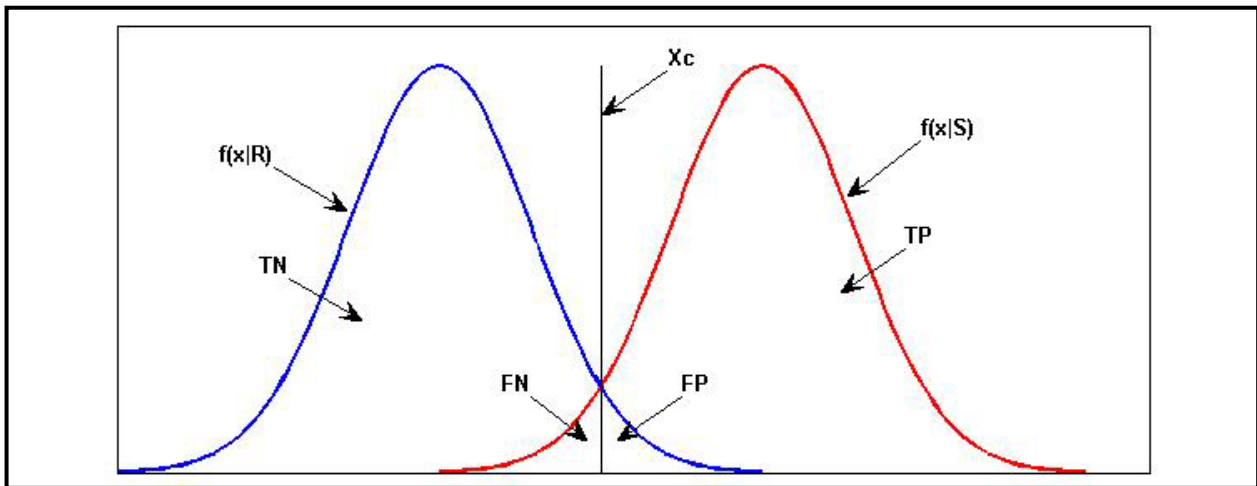


Figura 1.1.: Diagrama Esquemático representando as relação entre as distribuições  $f(x|S)$  e  $f(x|R)$  quando sobrepostas. ( $f(x|S)$  – Distribuição de Sinal,  $f(x|R)$  – Distribuição de Ruído, TP – Verdadeiro Positivo, TN – Verdadeiro Negativo, FN – Falso Negativo, FP – Falso Positivo,  $X_c$  – Valor de corte).

Vê-se, a partir da figura 1.1, que as distribuições de  $x_R$  e  $x_S$ , sobrepõe-se, levando a erros de detecção em que, casos inicialmente identificados como sinais poderão ser identificados como ruído, e vice-versa.

Desse modo, em qualquer teste de diagnóstico é fixado um valor de corte ( $x_c$ ) é fixado, para a variável em estudo, valor este que determina a classificação como sinal ou ruído. Assim o sistema de detecção deve responder como S, sinal presente, para cada valor de  $x \geq x_c$  e como R, ruído, cada vez que  $x < x_c$ .

Seja X uma variável aleatória. Se X for discreta, então  $P(S|s, x_c)$  é a probabilidade condicional de X dado o acontecimento S;  $P(R|r, x_c)$  é a probabilidade condicional para o mesmo valor X, dado o acontecimento R. O desempenho de um observador em uma experiência num único intervalo é usualmente medido em termos de probabilidades condicionadas, por exemplo,  $P(S|s, x_c)$  e  $P(R|r, x_c)$ .

Se a variável for contínua, então os correspondentes elementos de probabilidades são  $f(x|s)dx$  e  $f(x|r)dx$ , em que  $f(x|s)dx$  e  $f(x|r)dx$  são as funções densidade de probabilidade associ-



ados, respectivamente, às duas distribuições da variável X. A probabilidade que a resposta seja S, denotada por  $P(S | s, x_c)$ , é dada por:

$$P(S | s, x_c) = \int_{x_c}^{\infty} f(x | s) dx . \quad \dots\dots\dots (1.1)$$

Analogamente, a probabilidade que um dado resultado de X, proveniente da região de ruído apenas resulte em R, denotada por  $f(R | r, x_c) dx$ , é dada por

$$P(R | r, x_c) = \int_{-\infty}^{x_c} f(x | r) dx . \quad \dots\dots\dots (1.2)$$

Uma vez que  $f(x | s)$  e  $f(x | r)$  são normalizadas, a probabilidade de que um resultado de X, proveniente da distribuição do sinal acrescido de ruído, resulte em uma resposta R é  $P(R | s, x_c) = 1 - P(S | s, x_c)$ , enquanto que a probabilidade de um resultado de x proveniente de uma distribuição de ruído resulte na detecção de um ruído R é dada por  $P(R | r, x_c) = 1 - P(S | r, x_c)$ . Logo, basta determinar uma das probabilidades da distribuição do sinal mais ruído ou de ruído para determinar as outras probabilidades associadas.

Em geral, um teste de diagnóstico tende a ser avaliado por duas medidas, fração de verdadeiros positivos, *True Positive Fraction* (TPF), também conhecido como sensibilidade, e fração de verdadeiros negativos, *True Negative Fraction* (TNF), designada por especificidade. Os valores TPF e TNF podem ser calculados através das seguintes equações:

$$\text{Sensibilidade} = \text{TPF} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad \dots\dots\dots (1.3)$$

e

$$\text{Especificidade} = \text{TNF} = \frac{\text{TN}}{\text{TN} + \text{FP}} \quad \dots\dots\dots (1.4)$$

Como consequência da definição de TPF e TNF, outras duas quantidades podem ser definidas: *False Positive Fraction* (FPF) e *False Negative Fraction* (FNF), dadas por:

$$\text{FPF} = \frac{\text{FP}}{\text{TN} + \text{FP}} = 1 - \text{Especificidade} \quad \dots\dots\dots (1.5)$$

e

$$FNF = \frac{FN}{TP + FN} \dots\dots\dots (1.6)$$

Em termos de diagnóstico, pode-se definir a sensibilidade como a capacidade que um teste tem para detectar o sinal e a especificidade como a capacidade que o teste tem para detectar o ruído. Contudo, essas medidas dependem do critério de diagnóstico ou do valor de corte, o qual é por vezes selecionado arbitrariamente. Assim, a mudança do critério pode aumentar a sensibilidade com a consequente perda de especificidade, e vice-versa. Considerando-se que as distribuições são normalizadas, e que todos os casos podem ser diagnosticados como positivos ou negativos, o número de decisões corretas mais o número de decisões incorretas é representados por:

$$TPF + FNF = 1 \dots\dots\dots (1.7)$$

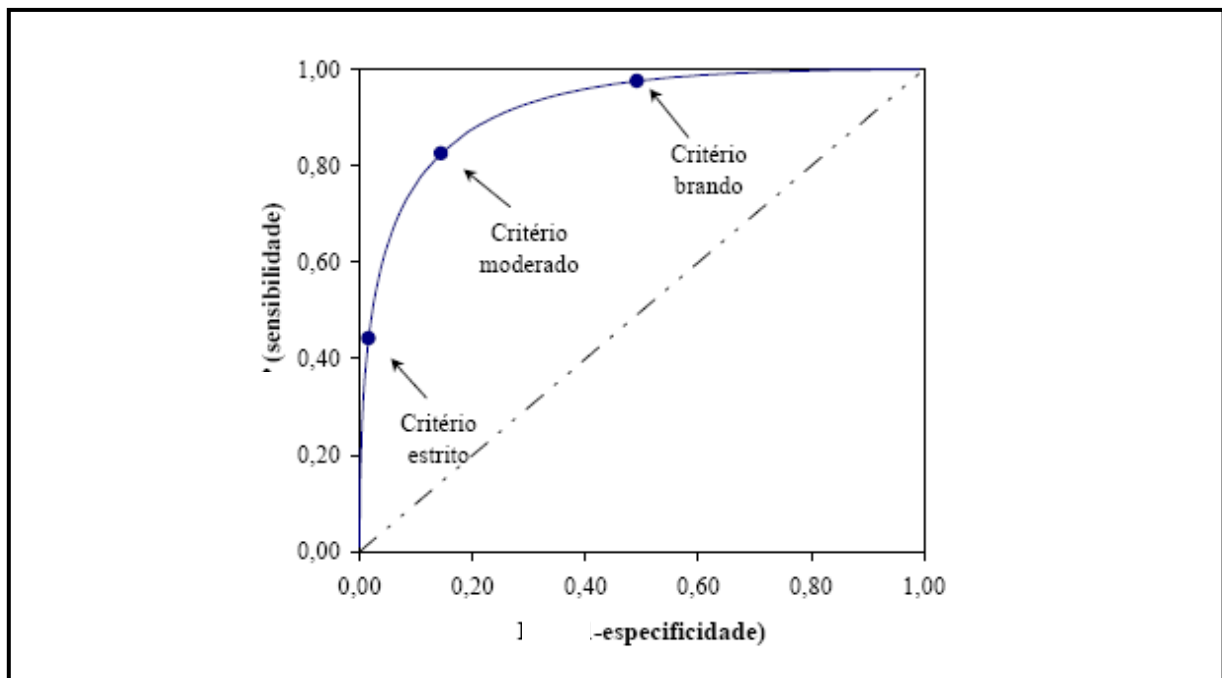
e

$$TNF + FPF = 1 \dots\dots\dots (1.8)$$

Um critério de decisão particular depende dos benefícios associados aos resultados corretos e dos custos associados aos incorretos. Por exemplo, se a principal preocupação é evitar o resultado falso-positivo (o resultado do teste pode indicar uma cirurgia arriscada para o paciente), então o ponto de corte deve objetivar o máximo de especificidade. Por outro lado, se a preocupação maior é evitar o resultado falso-negativo (o resultado do teste em suspeito de AIDS), então o ponto de corte deve objetivar o máximo de sensibilidade. Para contornar este tipo de situação, foi necessário desenvolver medidas alternativas de diagnóstico com propriedades mais robustas que a sensibilidade e a especificidade, a partir da análise ROC.

Por definição, uma curva ROC é a representação gráfica dos pares sensibilidades (ordenada) e 1 menos especificidades (abscissa), resultantes da variação do valor de corte ao longo de um eixo de decisão, X (Ver Figura 1.2). Cada ponto da curva representa um compromisso diferente entre a TPF e FPF, pela adoção de diferentes valores de corte ou nível crítico de confiança. Com efeito, uma curva ROC é uma descrição da capacidade de um sistema de diagnóstico discriminar dois estados.

Consoante aos critérios adotados para o valor crítico, um critério “estrito” é aquele que conduz a uma pequena fração de FP e uma pequena fração de VP, representado por um ponto situado no canto inferior esquerdo do espaço ROC. Progressivamente critérios mais lenientes conduzem a maiores frações de FP e VP, isto é, produzindo pontos situados no canto superior direito da curva no espaço ROC (Figura 1.2).



**Figura 1.2.** Curva ROC com a indicação do critério de decisão e respectiva capacidade de discriminação.

No que diz respeito ao desempenho de diferentes sistemas de diagnóstico, o sistema com a curva ROC mais próxima do canto superior esquerdo, fornece um maior poder discriminante. Entretanto, podem-se classificar os sistemas para um conjunto de frações de falsos positivos ou verdadeiros positivos de interesse, adotando-se um valor de corte que leva em consideração os custos e benefícios de um diagnóstico.

### 3. Índices de precisão das curvas ROC

Existe uma variedade de índices que foram propostos para especificar e/ou resumir as curvas ROC empíricas [METZ, 1986]. As frações de verdadeiros positivos e falsos positivos representam dois

tipos de precisão de diagnóstico, e podem ser utilizados para representação da curva ROC. Existem outros índices que podem ser utilizados como, por exemplo, o valor da fração de verdadeiros positivos num ponto ótimo de operação,  $FVP_0$ , e o índice área abaixo da curva ROC,  $A_z$ . O ponto ótimo de operação é o ponto sobre a curva ROC que está mais distante da reta  $x=y$  do plano ROC unitário. A área abaixo da curva ROC está associada ao poder discriminante de um teste de diagnóstico. Quanto maior a área, melhor o teste. Um  $A_z$  igual a 0,5 corresponde a uma performance sem importância, e um  $A_z$  igual a 1 representa um teste perfeito. Dentre os métodos para o cálculo de área abaixo de uma curva ROC, os seguintes são os mais utilizados:

- Métodos de resolução numérica, tipo regra do trapézio;
- Aproximação a estatística U de Wilcoxon-Mann-Whitney.

# Livros Grátis

( <http://www.livrosgratis.com.br> )

Milhares de Livros para Download:

[Baixar livros de Administração](#)

[Baixar livros de Agronomia](#)

[Baixar livros de Arquitetura](#)

[Baixar livros de Artes](#)

[Baixar livros de Astronomia](#)

[Baixar livros de Biologia Geral](#)

[Baixar livros de Ciência da Computação](#)

[Baixar livros de Ciência da Informação](#)

[Baixar livros de Ciência Política](#)

[Baixar livros de Ciências da Saúde](#)

[Baixar livros de Comunicação](#)

[Baixar livros do Conselho Nacional de Educação - CNE](#)

[Baixar livros de Defesa civil](#)

[Baixar livros de Direito](#)

[Baixar livros de Direitos humanos](#)

[Baixar livros de Economia](#)

[Baixar livros de Economia Doméstica](#)

[Baixar livros de Educação](#)

[Baixar livros de Educação - Trânsito](#)

[Baixar livros de Educação Física](#)

[Baixar livros de Engenharia Aeroespacial](#)

[Baixar livros de Farmácia](#)

[Baixar livros de Filosofia](#)

[Baixar livros de Física](#)

[Baixar livros de Geociências](#)

[Baixar livros de Geografia](#)

[Baixar livros de História](#)

[Baixar livros de Línguas](#)

[Baixar livros de Literatura](#)  
[Baixar livros de Literatura de Cordel](#)  
[Baixar livros de Literatura Infantil](#)  
[Baixar livros de Matemática](#)  
[Baixar livros de Medicina](#)  
[Baixar livros de Medicina Veterinária](#)  
[Baixar livros de Meio Ambiente](#)  
[Baixar livros de Meteorologia](#)  
[Baixar Monografias e TCC](#)  
[Baixar livros Multidisciplinar](#)  
[Baixar livros de Música](#)  
[Baixar livros de Psicologia](#)  
[Baixar livros de Química](#)  
[Baixar livros de Saúde Coletiva](#)  
[Baixar livros de Serviço Social](#)  
[Baixar livros de Sociologia](#)  
[Baixar livros de Teologia](#)  
[Baixar livros de Trabalho](#)  
[Baixar livros de Turismo](#)