

# A Wavelet-Based Statistical Analysis of fMRI data: I. Motivation and Data Distribution Modeling

Ivo D. Dinov<sup>1,2</sup>, John W. Boscardin<sup>3</sup>, Michael S. Mega<sup>4</sup>,  
Elizabeth L. Sowell<sup>1</sup> and Arthur W. Toga<sup>1</sup>

<sup>1</sup>Laboratory of Neuro Imaging, Department of Neurology,

<sup>2</sup>Department of Statistics, <sup>3</sup>Department of Biostatistics,  
UCLA, Los Angeles, CA 90095.

<sup>4</sup>Pacific Health Research Institute, Honolulu, HI 96813.

## Correspondence:

Ivo D. Dinov

Department of Statistics

8130 Mathematical Sciences Bldg.

UCLA, Box 951554

Los Angeles, CA 90095-1554

E-mail: [dinov@stat.ucla.edu](mailto:dinov@stat.ucla.edu)

Tel. 310-825-8430

Fax: 310-206-5658

## Abstract

We propose a new method for statistical analysis of functional magnetic resonance imaging (fMRI) data. The discrete wavelet transformation is employed as a tool for efficient and robust signal representation. We use structural MRI and functional fMRI to empirically estimate the distribution of the wavelet coefficients of the data both across individuals and across spatial locations. *Heavy-tail* distributions are then proposed to model these data because these signals exhibit slower tail decay than the Gaussian distribution. There are two basic directions we investigate in the first part of this study: 1. Bayesian wavelet-based thresholding scheme, which allows better signal representation, and; 2. A family of heavy-tail distributions, which are used as models for the real MRI and fMRI timeseries data. We discovered that Cauchy, Bessel K-Forms and Pareto distributions provide the most accurate asymptotic models for the distribution of the wavelet coefficients of the data. In the second part of our investigation we will apply this technique to analyze a large fMRI data set involving repeated presentation of sensory-motor response stimuli in young, elderly and demented subjects.

**Key Words:** fMRI, wavelets, statistical analysis, brain mapping, Bayesian techniques, brain atlas

## Introduction

This is part one of a two-part report. Here we justify the utilization of wavelets for representation of fMRI functional brain data. In addition, we study empirically the distributions of the wavelet coefficients, the corresponding leptokurtic distribution models and the atlas-based technique for statistical analysis of the fMRI timeseries. In the second part of this manuscript we will present an extended analysis of a complex fMRI dataset [Buckner *et al.*, 2000] representing the BOLD response to paired and isolated sensory-motor trials in Alzheimer's disease subjects, age-matched normal controls and a similar number of young adults.

Our basic idea is to perform the statistical analysis of fMRI data in the wavelet domain. The reason for that is a two-fold: First, there are only a few wavelet coefficients that survive the wavelet shrinkage preprocessing, which reduces the rate of *type I* errors, and; Second, we effectively assess the variations in the *patterns* of activation in different groups, as opposed to studying the differences in the fMRI BOLD signal intensities, across subjects, anchored at specific voxel locations.

fMRI Imaging: Functional Magnetic Resonance Imaging (fMRI) uses blood oxygen level dependent (BOLD) signal to measure hemodynamic alterations of the volume and flow of the blood intravascular susceptibility [Ogawa *et al.*, 1990]. There are still issues about the relation of the observed BOLD signal, glucose metabolism and neuronal activation. Various neuro-physiological studies have been undertaken to validate the direct association between the detected fMRI signal and cellular patterns of activation. Some of these use optical intrinsic signal imaging [Hess *et al.*, 2000], electroencephalography [Krakow *et al.*, 1999, 2000] or microelectrode recordings [Logothetis *et al.*, 2001].

It is commonly accepted that there are 6-8 seconds of delay between stimulus onset and the peak of the observed BOLD fMRI signal [Bandettini *et al.*, 1993]. In addition, the percent modulation change in the data between rest and activation states is only in the order of 1-4%. Presence of noise, caused by normal physiological processes, the imaging equipment, between- and within-subject motion and habituation [Hajnal *et al.*, 1994], requires the use of robust statistical techniques for analyzing fMRI data. Poor signal-to-noise ratio can be addressed by designing better data acquisition protocols, including event related designs, pulse-sequences, novel radio-frequency coil models, different time, space and resolution

strategies, etc. [Logothetis, *et al.*, 2001]. One successful approach in the process of enhancing the signal at the time of data acquisition is to choose a certain on/off block design with specific time-length and stimulus conditions [Bandettini, *et al.*, 1993, Bullmore, *et al.*, 1996].

Analyses of fMRI data: Once the data is acquired there is only so much one can do to find significant signal temporal changes and their spatial locations. Various techniques have been proposed for post acquisition data analysis based on the general linear model [Friston *et al.*, 1995], mutual information [Tsai *et al.*, 1999], fuzzy logic [Golay *et al.*, 1998], neural networks [Chuang *et al.*, 1999] and Bayesian statistics [Kershaw *et al.*, 1999].

A number of other statistical approaches for analyzing fMRI data have been proposed in the literature, as well. These involve Student's  $T$  test [Worsley *et al.*, 1995],  $\chi^2$  and  $F$  tests [Cao and Worsley, 1999], Kolmogorov-Smirnov test [Aguirre *et al.*, 1998], correlation analysis [Kuppusamy *et al.*, 1997] and non-parametric tests [Bullmore *et al.*, 1996]. Most techniques, however, are applied in the spatial domain and either treat each voxel intensity as independent (an assumption oftentimes too artificial) or introduce a spatial smoothing by convolving the native data with a large (Gaussian) kernel. The latter scheme increases the statistical power, since it takes advantage of the fact that near-by voxels are highly correlated and decreases the noise by smoothing; however, it also lowers the already low fMRI signal gradient [Desco *et al.*, 2001].

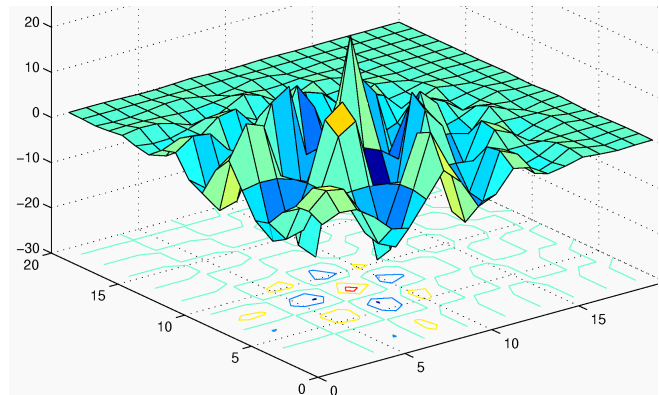


Figure 1. A 2D Daubechies wavelet illustrating the compact support, fast decaying and oscillatory properties of wavelets.

What are Wavelets? A wavelet basis is constructed using an oscillatory, compactly supported, (mother) wave function, usually  $C^k$  differentiable, which has rapidly decaying tails and satisfies an admissibility condition  $\int \psi(s) ds = 0$ . Regularity of the wave  $\psi$  is determined by the number of the continuously

vanishing moments. Hence,  $Reg(\psi) = M$ , if and only if  $\int s^r \psi(s) ds = 0, 0 \leq r \leq M-1$  and  $\int s^M \psi(s) ds \neq 0$ .

Daubechies [Daubechies, 1988] defined a class of mother wavelet functions, Figure 1, starting with a (father) wavelet,  $\phi$ , such that integral translations of  $\phi$  form an orthonormal basis for the space of the

piecewise step-functions (characteristic functions on intervals of length one). Then, we obtain the following wavelet function  $\psi(s) = \sqrt{2} \sum_{k=0}^{2M-1} g_k \phi(2s - k)$ . The basic problem is to find a non-trivial sequence of parameters  $\{g_k\}$ , such that  $\psi$  has a compact support, proper regularity and vanishing moments of order  $M$  ( $k \leq 2M-2$ ).

For orthonormal wavelet bases, if the number of wavelet vanishing moments is  $M$ , the Daubechies wavelet basis [Daubechies, 1993] has the fewest number of any class of wavelets for a fixed  $M$ . Higher order of non-trivial wavelet coefficients generally indicates a larger support. Daubechies wavelets,  $\psi$ , for example, are compactly supported on  $[-(M-1); M]$ . Therefore the dyadically-scaled and integrally-translated wavelets,  $\psi_{j,k} = 2^{j/2} \psi(2^{-j} - k)$ , are supported on  $[(k+1-M) 2^{-j}; (k+M) 2^{-j}]$ . Mallat [Mallat, 1989] proved that the set of the scaled and translated wavelets, induced by the mother wave, form an orthonormal basis for  $L^2(\mathbb{R})$ , which means that they are a complete set that any other square integrable function can be expressed in terms of. What is also very important is that the wavelet-based representation of signals is frequency and space localized. So, a domain-restricted perturbation of the original signal only affects the wavelets whose supports intersect the region of change of the function. Then, for  $f(x)$  in  $L^2(\mathbb{R})$ , the wavelet transform (WT) is defined by the collection of inner-products:  $\{w_{j,k} = \langle f, \psi_{j,k} \rangle = \int f(s) \psi_{j,k}(s) ds : j, k \in \mathbb{Z}\}$ . The discrete version of the wavelet transform is implemented [Dinov *et al.*, 2001] and freely available for download from <http://www.loni.ucla.edu/~dinov/WAIR.html>.

Why Use Wavelets? The advantages of employing wavelet analysis over other conventional spatial-domain approaches include efficient signal representation [Dinov *et al.*, 2002], noise reduction and decorrelation [Donoho *et al.*, 1998] and time-frequency localization [Daubechies, 1991]. From an empirical point of view, thresholding the wavelet coefficients provides a way to extract

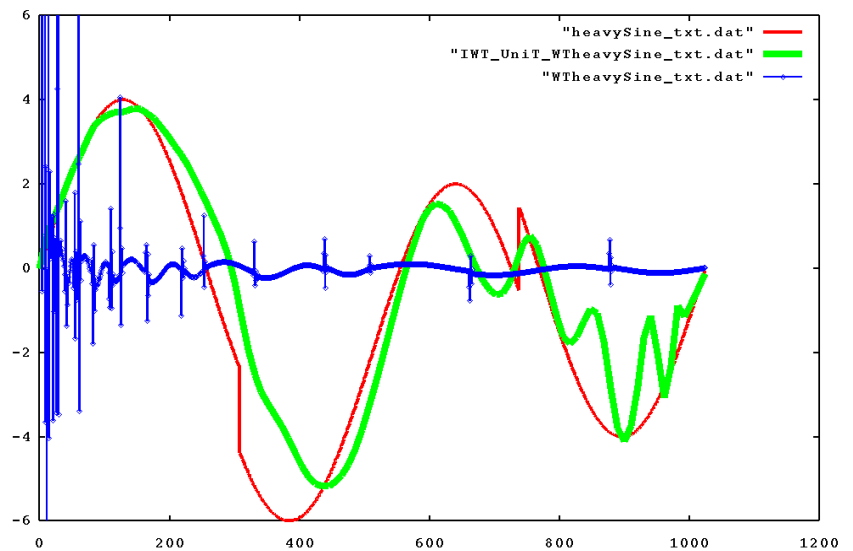


Figure 2. The three curves on this graph represent the original signal (*HeavySine* function, thin red curve), its wavelet transform (thin blue curve) and the reconstructed function estimator using only the largest 2% of the wavelet parameters. Note the space-frequency decorrelation of the original data in the wavelet-space (blue curve).

the essence of the data content of a signal. In Figure 2, we have an original image (red smooth thin curve, *HeavySine* function), its wavelet transform (WT) (thin blue curve) and a wavelet estimate of the original signal (thick green smooth curve). A couple of interesting phenomena quickly emerge from this 1D plot. First, it is clear that very few of the wavelet coefficients have magnitudes larger than 0.3, across the range of frequencies and locations. A natural question then is: *If we set to zero all wavelet coefficients smaller than 0.3 and then synthesize the image back (by inverting the wavelet transform), will we get a reasonable representation of the original function?* One heuristic example illustrating the answer to that question is shown by the thick green smooth curve superimposed on Figure 2. This leads to the second important observation; using only the largest 2% of the wavelet coefficients to recover the *HeavySine* function we obtain a reasonable estimate of the true signal. Even though the new estimate is not a perfect approximation of the original data it does capture the main trend of the signal, at 50:1 compression ratio. The next question then is: *How do we select a few of the wavelet parameters to compactly represent the data (in our case 3D/4D brain volumes)?* We will provide partial answers to this question in the methods section. Another empirical observation worth pointing out now is the fact that the original signal is often times spatially correlated, whereas its WT is uncorrelated in the wavelet-space, see the blue curve [WT(*HeavySine*)] in Figure 2.

Prior Studies of fMRI Data Using Wavelets: Bullmore and coworkers presented a wavelet-based methodology for characterizing the noise structure in short to medium length fMRI timeseries [Fadili and Bullmore, 2002]. The same group also introduced wavelet-generalized least squares (WLS), which are used to obtain the best linear unbiased estimator of regression model parameters in the context of long-term memory errors. These theoretical developments were compared with results obtained by ordinary least squares and estimators based on auto regressive moving averages. The authors show superiority of the generalized WLS in terms of *type I* error control.

Wavelet-based analysis of fMRI data has been proposed as early as 1996, by Hilton and coworkers [Hilton *et al.*, 1996] who were the first to suggest signal denoising of fMRI data using a non-linear soft-thresholding. The group of Ruttimann [Ruttimann *et al.*, 1998] also employed a multiresolution approach for increasing the sensitivity without sacrificing the probability for committing a *type I* error. They utilized a uniform thresholding in the wavelet domain followed by a Z-test, Bonferroni corrected for multiple comparisons. Yet another study, by Melke and coworkers, [Mekle *et al.*, 2000] used multi-scale singularity detection in the temporal domain. Their data analysis is based on the constructed maxima

lines, using wavelet expansions at integral scales to trace the wavelet modulus maxima. From the maxima lines the signal singularities are then localized by the beginning and ending of the stimulus.

Another method for detection of activation in fMRI data using wavelet analysis was introduced by Feilner's group [Feilner *et al.*, 2000]. The extension proposed in this work was to incorporate non-stationary Gaussian noise and the subsequent T-testing, required by the assumption of unstable variances at different locales. No wavelet-space shrinkage is applied in their analysis, however, they use of a post-inverse wavelet transform signal thresholding, based on the residual noise-level. The same study investigated the effects of the choice of a WT relying on various information measures (e.g., Kullback-Leibler information number, which measures the similarity between a statistical model and the true distribution).

Desco and coworkers investigated the performance of the wavelet analysis of fMRI data on a *gold* standard (a phantom functional volume) by making use of different wavelet decomposition schemes, location, size and level of activation and the presence of smoothness [Desco *et al.*, 2001]. When measuring *sensitivity* (percentage of true activation areas detected) and *specificity* (percentage of locations correctly detected as non-activated), the authors clearly identified dependences in the results of the fMRI wavelet analysis, based on the statistical tests, the shape of the model-activation and the wavelet-family used.

## Methods

In this section we describe two different wavelet shrinkage strategies, which may be employed prior to statistical analysis: frequency adaptive and Bayesian thresholding. Both schemes induce theoretically almost optimal signal estimators, but are quite distinct in their nature. Then we present the basics of stable and heavy-tail distributions, which we choose to use as models for the distributions of the wavelet coefficients of the functional MRI data in 3D and 4D. Finally, we introduce the framework for statistical analysis of fMRI timeseries utilizing a functional and anatomical sub-volume probabilistic atlas (F&A SVPA atlas) of the brain.

Frequency-Adaptive Wavelet Shrinkage: The process of wavelet shrinkage begins with a collection of observations  $\{y_k\}$ ,  $1 \leq k \leq N$ , obtained by adding noise to some unknown signal  $f(t_k)$ , for each time point  $t_k$ . That is,  $y_k = f(t_k) + \varepsilon_k$ , where  $\varepsilon_k \sim \mathcal{N}(0, \sigma^2)$  are independent (and identically distributed) Gaussian random variables, most of the time. To recover the unknown function  $f(t)$ , having the data  $y_k$  alone, we construct function estimators  $\hat{f}$ . The *Risk* function  $R(f, \hat{f}) = \frac{1}{N} E(\|f - \hat{f}\|^2) = \frac{1}{N} E\left(\sum_{k=0}^{N-1} \|f(t_k) - \hat{f}(t_k)\|^2\right)$  measures the performance of the estimator  $\hat{f}$  by the average quadratic loss at the sample points where  $E$  is the expectation of the misfit between the function and its estimator. Notice the similarity between this *Risk* function,  $R(f, \hat{f})$ , and the variance of unbiased estimators. Preference will be given to estimators having small *Risk* measures (i.e., small “variances”).

If  $W$  represents the matrix of the discrete wavelet transform [Donoho and Johnstone, 1995; Dinov and Summers, 2001], the time-domain model  $y=f+\varepsilon$  becomes  $w=\theta+z$  in wavelet space, where  $w = W y$ ,  $\theta = W f$  and  $z = W \varepsilon$ . Suppose  $U$  is a subset of the set of all wavelet coefficients  $\{w\}$  of the data, one defines *selective-wavelet (SW) reconstruction* estimators by  $\hat{f} = \sum_{j,k \in U} w_{j,k} W_{j,k}$ , where  $w = \{w_n\} = \{w_{j,k} \mid 0 \leq j \leq J-1; 0 \leq k \leq 2^j-1\}$ ,  $U$  is an index subspace,  $j$  codes for the frequency and  $k$  indexes the location of the corresponding wavelet coefficients,  $W_{j,k}$  represents the  $(j, k)^{\text{th}}$  row of the orthogonal wavelet transform matrix  $W$  and  $N=2^J-1$  is the sample size of the data [Dinov *et al.*, 2002].

Suppose now  $f(t)$  is the true 4D BOLD signal and  $y(t) = f(t) + \varepsilon$  be the observed fMRI timeseries. Wavelet shrinkage is applied by filtering the wavelet transform  $w = W(y)$ , through the soft thresholding filter  $\eta_{\lambda_j}(w_{j,k}) = \text{sign}(w_{j,k}) \times \max\{|w_{j,k}| - \lambda_j, 0\}$ . The major challenge is to determine a meaningful threshold level  $\lambda_j$ , which will induce function estimators that are optimal and efficient across the space of functions of interest (these spaces include  $L^2$ , Besov spaces, Sobolev spaces, Hilbert spaces, etc.) [Donoho and Johnstone, 1996]. One very promising example is  $\lambda_j = \sigma \sqrt{2 \ln(2n_j + 4)}$ , where  $n_j$  is the number of wavelet coefficients at the  $j$ -frequency band. *This particular thresholding scheme is appealing because noise is most likely to be present in the high-frequency bands of the signal; hence the steadily increasing (with the frequency index  $j$ ) shrinkage imposed in this case.* From the theoretical point of view this scheme also induces an almost-optimal estimator ( $\hat{f} = W^{-1} \eta_j W(y)$ ) in  $L^2$  of the unknown signal  $f(t)$ . First, we have

the following upper bound on the risk measure  $R(f, \hat{f}) \leq (1 + 2 \ln(N + 4)) \times \left( \sigma^2 \ln N / N + R_{N, \sigma}(SW, f) \right)$ , indicating that this estimator is within  $\ln^2(N)/N$  of the *ideal risk* ( $R_{N, \sigma}(SW, f)$ ), where  $N$  is the size of the signal,  $\sigma$  is the signal noise level and  $SW$  is selective wavelet reconstruction scheme indicating a few of the wavelet parameters are used to synthesize the estimator. Secondly, we have a lower bound of any function estimator, again in terms of the ideal risk:  $\inf_{\hat{\theta}} \sup_{\theta} \left\{ \left( \frac{1}{N} E(\|\theta - \hat{\theta}\|^2) \right) \times \left( \sigma^2 \ln N / 2N + \frac{1}{N} R_{N, \sigma}(SW, f) \right)^{-1} \right\} \sim 2 \ln N$ . In other words, no estimator can be closer to the ideal risk than  $\ln^2(N)/N$  across all possible fMRI signals. Therefore,  $\lambda_j = \sigma \sqrt{2 \ln(2n_j + 4)}$  is a robust and optimal wavelet shrinkage scheme, that naturally increases with the increase of the frequency index of the wavelets representing the signal. Our goals here are to improve on this thresholding approach so that we obtain a cut off scheme that is not necessarily monotonically increasing.

Bayesian Wavelet Shrinkage: The second wavelet shrinkage scheme we utilize is Bayesian wavelet thresholding, which allows for empirically obtained signal representation. We use two component mixture prior distributions on the wavelet coefficients  $\theta_{j,k}$  with  $\theta_{j,k} | \pi_j \tau_j \sim \pi_j \mathcal{N}(\mathbf{0}, \tau_j^2) + (1 - \pi_j) \delta(\mathbf{0})$ , where  $\pi_j$  is a proportion between 0 and 1,  $\delta(0)$  is the Dirac point mass at zero and  $\tau_j > 0$ . In other words, there is a level-dependent positive probability  $\pi_j$  *a priori* that each wavelet coefficient will be exactly zero. If not, the coefficient will be normally distributed with mean zero and a level-specific standard deviation  $\tau_j$ . Abromovitch and coworkers [Abromovitch *et al.*, 1998] details the following interesting relationship: if  $\tau_j^2 = C_1 \times 2^{-\alpha j}$  and  $\pi_j = \min(1, C_2 \times 2^{-\beta j})$ , then the choices for  $\alpha$  and  $\beta$  correspond to requiring  $f$  to belong to certain Besov spaces. The parameter  $\alpha$  governs the typical size of the non-zero wavelet coefficients (large values mean that the non-zero values at higher frequencies tend to be extremely small in absolute value) and  $\beta$  controls the proportion of non-zero coefficients (large values mean that nearly all of the high frequency coefficients are exactly zero).  $C_1$  and  $C_2$  can be estimated empirically from the data and represent the typical size and proportion of non-zero wavelet coefficients at low frequencies.

The prior family for the wavelet coefficients is conjugate. Given the observed data  $y = f + \varepsilon$ , the corresponding wavelet representation  $w = \theta + z$ , where  $W$  is the discrete WT,  $w = Wy$ ,  $\theta = Wf$  and  $z = W\varepsilon$ , and the above prior distribution for the true wavelet coefficients, the posterior distributions of  $\theta_{j,k}$  are again independent two-component mixtures

$$p(\theta_{j,k} | w_{j,k}) \sim \lambda_{j,k} \mathcal{N} \left( \frac{w_{j,k} \tau_j^2}{\sigma^2 + \tau_j^2}, \frac{\sigma^2 \tau_j^2}{\sigma^2 + \tau_j^2} \right) + (1 - \lambda_{j,k}) \delta(\mathbf{0}),$$

where  $\lambda_{j,k} = 1/(1 + \rho_{j,k})$ , and the posterior odds that  $\theta_{j,k}$  is exactly zero are

$$\rho_{j,k} = \frac{1 - \pi_j}{\pi_j} \frac{\sqrt{\tau_j^2 + \sigma^2}}{\sigma} \exp\left(\frac{-\tau_j^2 w_{j,k}}{2\sigma^2(\tau_j^2 + \sigma^2)}\right).$$

This procedure (called *BayesThresh*) replaces the  $w_{j,k}$  by their posterior medians and then transforms back the data into the time domain to get the signal estimate  $\hat{f}$ . It turns out that *BayesThresh* is a compromise between the hard and soft thresholding procedures most commonly used in the field [Donoho, 1996], essentially borrowing the advantageous features of both approaches in that the thresholding function is continuous (like soft thresholding), but greatly reduces the amount of shrinkage needed to achieve this. Figure 3 shows a comparison of these three techniques.

Heavy-Tail Distributions: There is a volume of work on heavy-tail distributions in the field of signal processing. We consider some of the most popular and practically useful *leptokurtic* distributions (distributions with tails heavier as compared to a Normal (Gaussian) curve with similar variance) [Gourieroux and Monfort, 1992].

As described earlier, we carry our statistical analyses in the *compressed wavelet space* (on wavelet coefficients that survive the frequency-adaptive or Bayesian thresholding scheme). Because these wavelet coefficients are large in magnitude it is extremely important to have a distribution model that is asymptotically accurate. In the results section we provide evidence that the empirical distributions of the wavelet coefficients of the MRI and fMRI data appear to be leptokurtic<sup>1</sup>. This motivates our interests in utilizing various heavy-tail distributions as proper models for the distribution of these wavelet parameters.

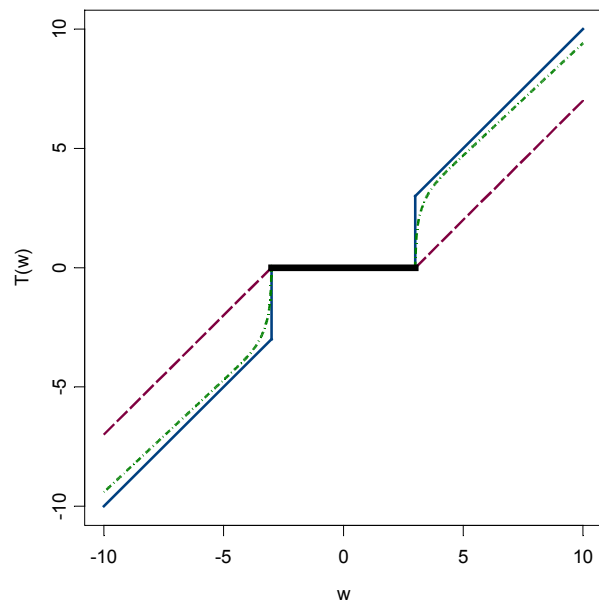


Figure 3. This figure shows the soft, hard and Bayesian ( $\sigma=1$ ,  $\tau=4$ ,  $\pi=0.055$ ) thresholding rules for a threshold-level of 3.0. Notice that the Bayes thresholding function (smooth dashed curve) is continuous (like the soft rule), but uses much less shrinkage (represented by vertical distance from the dotted line).

The double-exponential distribution is defined by the following probability density function

$$f(x) = \frac{1}{2\beta} \times \exp\left(-\left|\frac{x - \mu}{\beta}\right|\right).$$

<sup>1</sup> In fact, the distribution of the wavelet coefficients of the 3D structural, 3D and 4D functional MRI data appears to be leptokurtic both across subjects at fixed location-scale indices, and within subjects across location-scale indices.

at the rate of  $\exp(-|x|)$ . The density function for the double-Pareto distribution with a tail-size parameter  $\alpha$  and a scale parameter  $c$  is defined by  $f(x) = \frac{\alpha c^\alpha}{|x - \mu|^{1+\alpha}}$ , for  $|x - \mu| > c$ . Another theoretically important heavy

tail distribution is the Cauchy distribution with density function  $f(x) = \frac{\gamma}{\pi} \times \frac{1}{\gamma^2 + (x - \mu)^2}$ . Cauchy

distribution is frequently used to provide counterexamples to various phenomena, which appear quite natural for light-tail distributions. One transition between light- and heavy-tail distributions is provided by the T distribution which serves as a one-parameter homotopy family connecting the heavy-tailed Cauchy and the standard normal distribution. The density function for a T distribution with  $df=n$  is given by

$$f(x) = \frac{\Gamma\left(\frac{1}{2}(n+1)\right)}{\sqrt{n\pi} \times \Gamma\left(\frac{n}{2}\right)} \times \left(1 + \frac{x^2}{n}\right)^{-\frac{1}{2}(n+1)}, \text{ where the Gamma function is defined by } \Gamma(\alpha) = \int_0^\infty \exp(-y) \times y^{\alpha-1} dy.$$

T ( $df=1$ ) coincides with the Cauchy distribution and, as  $n \rightarrow \infty$ , T ( $df=n$ )  $\rightarrow$  Normal(0,1). The last large class of leptokurtic distributions we consider as possible models for the wavelet coefficient distribution of the fMRI functional data is referred to as Bessel K forms [Granender and Srivastava, 2001]. There exists an asymptotic analytic approximation of the density of Bessel K forms, which can be written by

$$\hat{f}(x; p, c) = \frac{1}{\Gamma(p)} \left(\frac{2}{c}\right)^{\frac{p}{2}} |x - \mu|^{p-1} \exp\left(-\sqrt{\frac{2}{c}} |x - \mu|\right). \text{ The parameters } p \text{ and } c \text{ can be estimated from the data}$$

by  $\hat{p} = \frac{3}{\text{skewness}(WT(f)) - 3}$  and  $\hat{c} = \frac{\hat{\sigma}_{WT(f)}^2}{\hat{p}}$ , where the skewness and the sample variance of the wavelet

transformation of the data ( $WT(f)$ ) can be empirically obtained. It is known that this approximation is sub-optimal near the mean,  $\mu$ . However, the asymptotic behavior of  $\hat{f}(x; p, c)$  is very similar to that of the corresponding Bessel K form.

Atlas-based Statistical Analysis of fMRI data: Developing an atlas of functional brain data has been challenging due to image registration, intensity normalization, physiological brain variation and other unresolved signal localization issues. We have developed a functional brain for Alzheimer's Disease (AD) in the frequency space using the wavelet signal representation. This atlas, AD functional and structural sub-volume probabilistic atlas (F&A SVPA) [Dinov *et al.*, 2001], was based on rest-state PET data and employed in conjunction with our probabilistic anatomical atlas for the AD population [Mega *et al.*, 2000; Crabtree *et al.*, 2000; Mega *et al.*, 2001]. Analogously, one can design a framework for construction of paradigm- and population-specific stochastic atlases of brain activation using fMRI data.

Figure 4 illustrates schematically the protocol for the construction of a new fMRI functional atlas, in general. The structural MRI and functional fMRI data are first anatomically tessellated to obtain the individuals' regions of interest. This involves delineating a number of regions of interest, which can be done manually, semi-manually or completely automatically [Mega *et al.*, 2000]. The next step involves applying the 4D WT to the fMRI data over each anatomical region separately. In the wavelet domain, we then threshold the wavelet coefficients to get the denoised and decorrelated representation of the timeseries in compressed wavelet space. Next, we calculate the joint cross-subject-and-time distribution of the 4D functional data. Finally, the induced functional and anatomical atlas (which is population and paradigm specific) is employed to statistically analyze new fMRI data.

The development of the F&A SVPA is directly coupled to the construction of an anatomic ASVPA atlas; however, it is accomplished in the wavelet-space. We are interested in quantifying the paradigm-specific variability in functional data (fMRI) for diverse populations. More precisely, quantify functional variability, as measured by differences in the observed BOLD signal between subjects, at different spatial locations. Suppose we have  $N$  subjects scanned in an on/off block-design paradigm using fMRI. To simplify the presentation assume we are using our anatomical AD ASVPA atlas, which contains 60 regions of interest. Every individual's fMRI timeseries over an ROI is analyzed separately by transforming it into frequency-

space using the discrete WT. For each ROI, the *mean* and *variance* of the transformed wavelet coefficients, across subjects, will be computed at each location-frequency index,  $w=(j,k)$ .

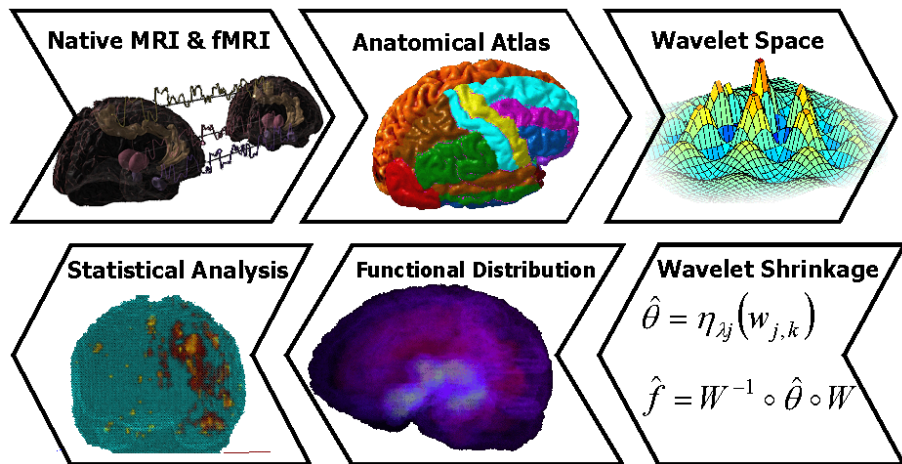


Figure 4. Design protocol for the wavelet-based construction and utilization of the fMRI functional atlas (F&A SVPA). The construction of the anatomical probabilistic atlas is augmented by including the regional distributions of the wavelet coefficients of the functional signals for one population. Statistical analysis of functional variability between a new fMRI volume and the F&A SVPA atlas is assessed following wavelet-space shrinkage.

Together with the probability maps encoded in the anatomical ASVPA atlas, these distribution maps of functional variability in

frequency space represent the desired functional and anatomical F&A SVPA atlas. If  $P_n(v,t)$  is the

intensity of subject  $n$ , at voxel location  $\mathbf{v}$  and time  $t$ ,  $1 \leq n \leq N$ , and  $A_k$  is the indicator function  $A_k(\mathbf{v}) = \begin{cases} 1, \mathbf{v} \in ROI_k \\ 0, \mathbf{v} \notin ROI_k \end{cases}$  of the  $k^{\text{th}}$  ROI of the ASVPA, then  $P_{n,k}(\mathbf{v}, t) = P_n(\mathbf{v}, t) \times A_k(\mathbf{v})$  represents the value of the BOLD fMRI signal at location  $\mathbf{v}$  and time  $t$ , for subject  $n$  over the  $ROI_k$ ,  $1 \leq k \leq 60$ ,  $1 \leq n \leq N$ . Denoting by  $\hat{P}_{n,k} = WT(P_{n,k})$  the WT of the regional 4D fMRI signal we obtain an estimate of the sample mean and standard deviation of the intensities in frequency-space, over  $ROI_k$ :  $\mu_k(\mathbf{w}) = \frac{1}{N} \sum_{n=1}^N \hat{P}_{n,k}(\mathbf{w})$ ;  $\sigma_k(\mathbf{w}) = \sqrt{\frac{1}{(N-1)} \sum_{n=1}^N (\hat{P}_{n,k}(\mathbf{w}) - \mu_k(\mathbf{w}))^2}$ . Finally, the collection of volumes  $\{\mu_k\}$  and  $\{\sigma_k\}$ ,  $1 \leq k \leq 60$ , represents the *functional component* of the F&A SVPA atlas, for each scale-shift wavelet-space index  $w$ .

Suppose  $P$  is a new fMRI timeseries, with  $P_k(\mathbf{v}, t) = P(\mathbf{v}, t) \times A_k(\mathbf{v})$  being the restriction of  $P$  over  $ROI_k$ , and  $\hat{P}_{n,k} = WT(P_{n,k})$  is the WT of  $P_k$ . Then we employ the following wavelet *shrinkage* strategy,  $\hat{T}$ .

Where  $\mu_k$  and  $\sigma_k$  are the mean and standard deviation of the distribution encoded in the F&A SVPA at location-frequency index  $w$ , and  $z_\alpha$  is the Z-score at significance level  $\alpha$ .

$$\hat{T}(\hat{P}_k(\mathbf{w})) = \begin{cases} \hat{P}_k(\mathbf{w}), & |\hat{P}_k(\mathbf{w}) - \mu_k(\mathbf{w})| \geq z_\alpha \times \sigma_k(\mathbf{w}), \\ 0, & |\hat{P}_k(\mathbf{w}) - \mu_k(\mathbf{w})| < z_\alpha \times \sigma_k(\mathbf{w}) \end{cases}$$

In our experiments we will investigate thresholds corresponding to  $\alpha = \{0.05, 0.001, 0.0001\}$ . In general, one needs to design a model for the distribution of the frequency-space data, having mean and variance equal to the sample mean and variance,  $\bar{\mu}_k$  and  $\hat{\sigma}_k^2$ , and then use this model to estimate  $z_\alpha$ .

The operator  $\hat{T}$  either *shrinks* a wavelet coefficient or leaves it unchanged, depending on whether it is significantly different from the corresponding F&A SVPA atlas mean at  $w$ . This will be the first step of the statistical assessment. The second step will involve inverting the WT and *constraining* the resulting wave patterns to the corresponding ROI's, in the time-domain. This procedure is followed by applying a time-domain thresholding, which selects only the highest intensities (0.01-0.0001%, precise assignment will have to be closely studied) within each separate ROI. Finally, we reconstitute the statistical image in anatomical 3D space from the *collage* of regional statistical maps representing the spatial-domain

thresholded differences, which were significant in the wavelet-space. This volume can be expressed

symbolically as  $\hat{P}(v) = \bigcup_{k=1}^{60} \hat{P}_k(v)$ , where  $\hat{P}_k = T(W^{-1}(\hat{T}(W(P_k))))$  and

$$T(P_k(v,t)) = \begin{cases} P_k(v,t) & , |P_k(v,t)| \geq \beta_\alpha^{(k)} \\ 0 & , |P_k(v,t)| < \beta_\alpha^{(k)} \end{cases}, \text{ where } \alpha \text{ is a significance level, as before, } \beta_\alpha^{(k)} \text{ is the threshold}$$

separating the highest  $\alpha \times 100\%$  from the lowest  $(1 - \alpha) \times 100\%$  image intensities (uniform threshold over  $ROI_k$ ) and  $W^{-1}$  is the inverse wavelet transform operator. This methodology, is clearly directly extendible to, and applicable for, various anatomical tessellation schemes, disease populations and fMRI block designs. One needs to exercise caution, however, when employing an F&A SVPA atlas constructed for one population, or one study paradigm, to study another unrelated fMRI timeseries.

## Results

We now present the empirical evidence that the distributions of the fMRI and MRI data in the wavelet domain have heavy tails. This fact is used in determining a robust method for statistical analysis of the wavelet coefficients trying to identify significantly different patterns of activation across the 4D fMRI volume, instead of mapping statistical differences at individual *voxels* (volume-time-elements).

Data: The large structural MRI database [Mazziotta *et al.*, 1995], part of the International Consortium for

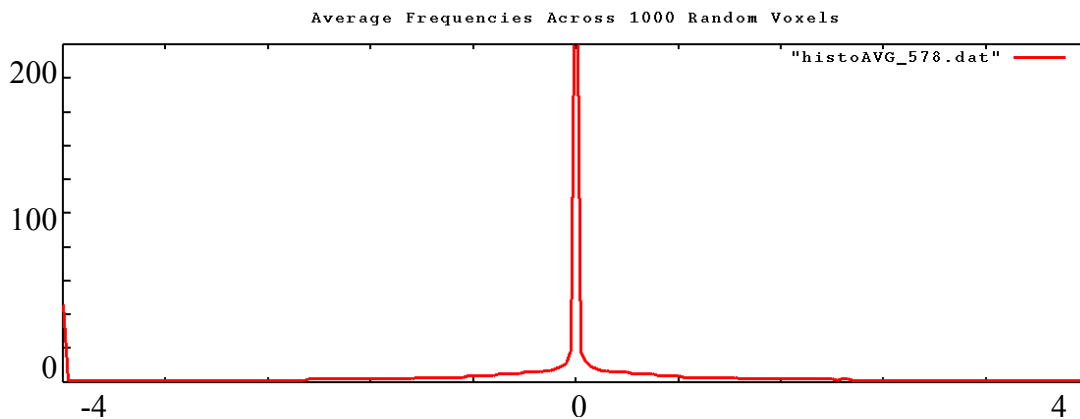


Figure 5. Displayed is the frequency histogram of the wavelet coefficients at 1,000 randomly selected locations (random indices of  $w_{j,k}$ ), averaged across all 578 MRI volumes part of the ICBM database [Mazziotta *et al.*, 1995]. Notably, most of the wavelet coefficients are near the origin, with some having sporadic, but large magnitudes. Heavy-tail distributions models will be appropriate for these data.

Brain Mapping (ICBM), was utilized to obtain 578 anatomical volumes for testing the characteristics of the distribution of the wavelet coefficients within and between subjects. These data represents young healthy adult MRI brains drawn from diverse ethnic, gender and socio-economic populations. The functional fMRI data used in our experiments consisted of randomly selected subjects from the pool of 41 participants (14 young adults, 14 non-demented older adults and 13 demented older adults), Buckner *et al.*, 2000. In that study the task paradigm consisted of presentation of a 1.5 second visual stimulus. Subjects were required to press a key with their right index fingers upon stimulus onset. The visual stimulus was an 8-Hz counterphase flickering checkerboard subtending approximately  $12^\circ$  of the visual angle. Runs were structured so that for every eight-image acquisition, one of two kinds of trial condition was presented (15 trials per run for a total of 60 trials per subject). The trials either involved stimuli presented in isolation or in pairs with and inter-trial interval of 5.36 seconds. One trial and two-trial conditions were pseudo-randomly intermixed so that eight trials of one type and seven trials of the other appeared in each run. All subjects were English speaking, normal visual acuity right-handers.

Wavelet coefficient distribution of the MRI data: Figure 5 indicates the frequency histogram of the average magnitude, across all 578 MRI volumes part of the ICBM database [Mazziotta *et al.*, 1995], of the wavelet coefficients at 1,000 randomly selected locations in the wavelet domain (random indices of  $w_{j,k}$ ). On the horizontal axis is the magnitude of the average wavelet coefficient between  $[-4; 4]$  and on the

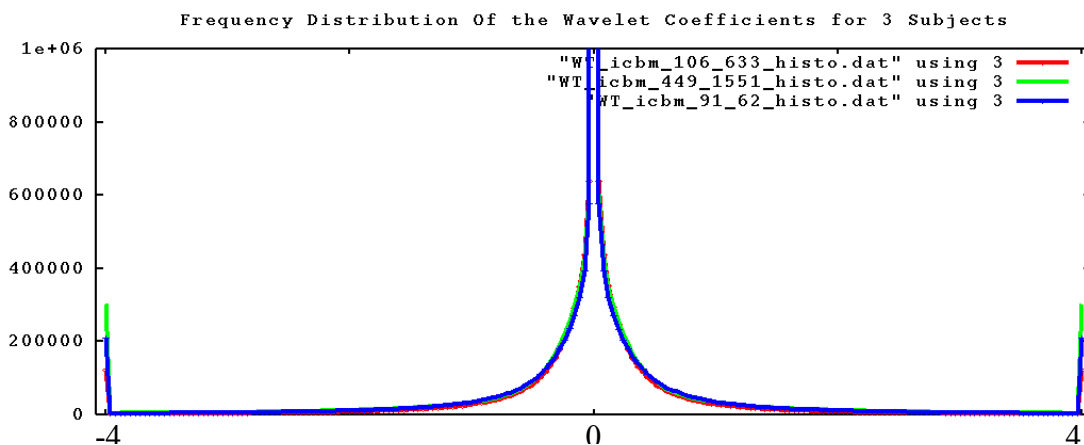


Figure 6. Shown here are the frequency distributions of the wavelet coefficients for three separate MRI volumes (randomly selected from the pool of 578). There is little variation between these three individuals, however the overall shape of the distribution of the magnitudes of the wavelet coefficients of each individual across the 1,000 random locations is more regular, still heavy-tailed, than the averaged distribution across subjects depicted in the previous Figure 5.

vertical axis (range  $[0; 578]$ ) are the frequencies of these average magnitudes<sup>2</sup>. Clearly one observes an empirical distribution with heavy tails.

Would the average histogram on Figure 5 significantly change if we focus on a single subject and plot the histogram of the wavelet coefficients across 1,000 randomly selected locations in the wavelet domain (random indices of  $w_{j,k}$ )? Figure 6, shows these frequency distributions for three separate MRI scans (randomly selected from the pool of 578). There

is little variation between these three individuals, and the overall shape of the distribution of the magnitudes of the wavelet coefficients for one individual across 1,000 random locations is more regular, still heavy-tailed, than the averaged distribution in the previous Figure 5.

The 2D image in Figure 7 illustrates a small portion of all of the individual distributions. The horizontal axis again represents the magnitude of the wavelet coefficients in the range  $[-4; 4]$ , vertical axis labels the subject index and the color map on each row indicates the frequencies of occurrence of a wavelet coefficient of certain magnitude across all 1,000 voxels. Bright colors indicate high, and dark colors represent low, frequencies. Note that Figure 6 effectively represents the across-subjects (row) average of these individual frequency histograms shown in Figure 7.

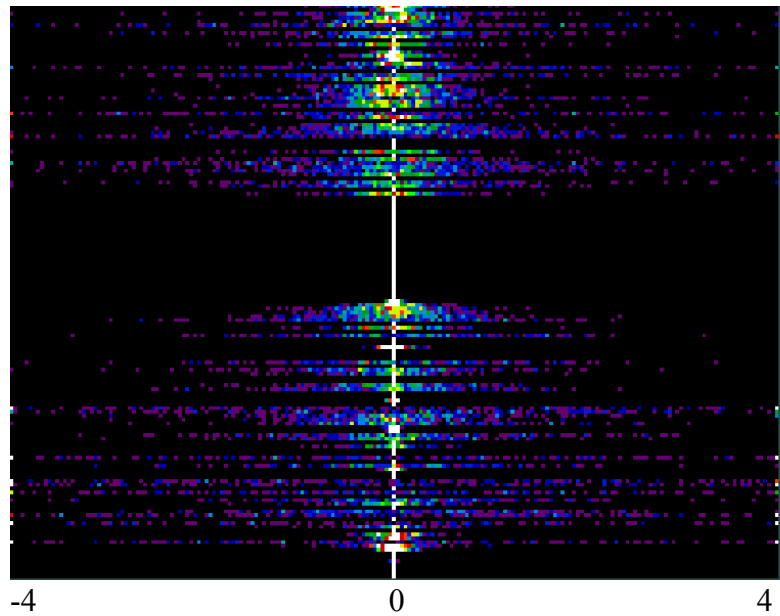


Figure 7. This image illustrates a portion of all of the individual distributions of the wavelet parameters for the 578 ICBM MRI volumes. The horizontal axis represents the magnitude of the wavelet coefficients in the range  $[-4; 4]$ , vertical axis labels the subject index and the row color map indicates the frequencies of occurrence of a wavelet coefficient of certain magnitude across all 1,000 voxels. Bright colors indicate high, and dark colors represent low, frequencies.

<sup>2</sup> There is nothing special with the range  $[-4; 4]$  which we have chosen to work with. The same phenomena were observed as we extended the range on which we calculate the distribution of the wavelet coefficients – the distribution remains *heavy-tailed*.

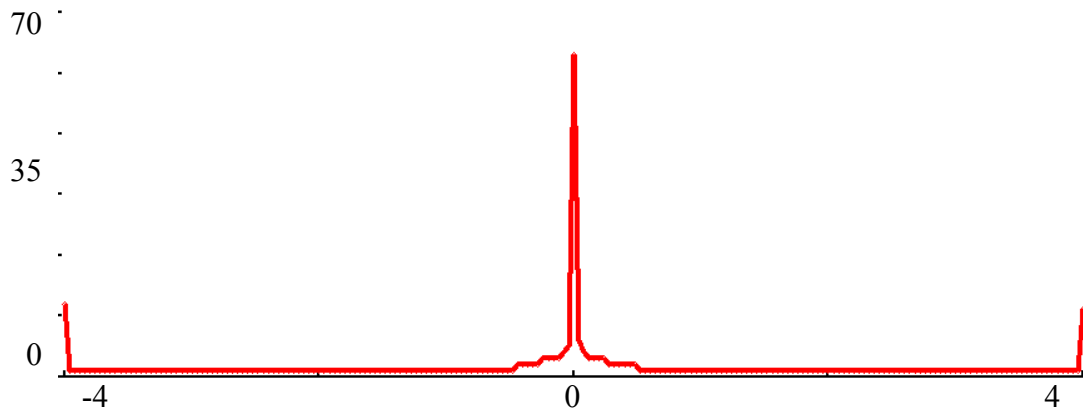


Figure 8. This diagram depicts the frequency histogram of the average magnitude of a wavelet coefficient, across all 128 fMRI 3D time-volumes (part of the fMRI study of Buckner and colleagues [Buckner *et al.*, 2000]) at 1,000 randomly selected locations (random indices of  $w_{j,k}$ ). Again, we observe the heavy-tailness of the data.

Wavelet coefficient distribution of the functional MRI data: Similar results are obtained if one studies functional fMRI instead of the structural MRI volumes. Figure 8 indicates the frequency histogram of the average wavelet coefficient magnitude, across all 128 fMRI 3D time-volumes, part of the fMRI study of Buckner and colleagues [Buckner *et al.*, 2000], at 1,000 randomly selected locations in the wavelet domain (random indices of  $w_{j,k}$ ).

And the 2D image in Figure 9 shows all of the individual distributions for all 128 time points (one run, six trials). Horizontal axis represents the magnitude of the wavelet coefficients in the range  $[-4; 4]$ , the vertical axis labels the time point of the fMRI epoch and the color map on each row indicates the frequencies of occurrence

128

64

1

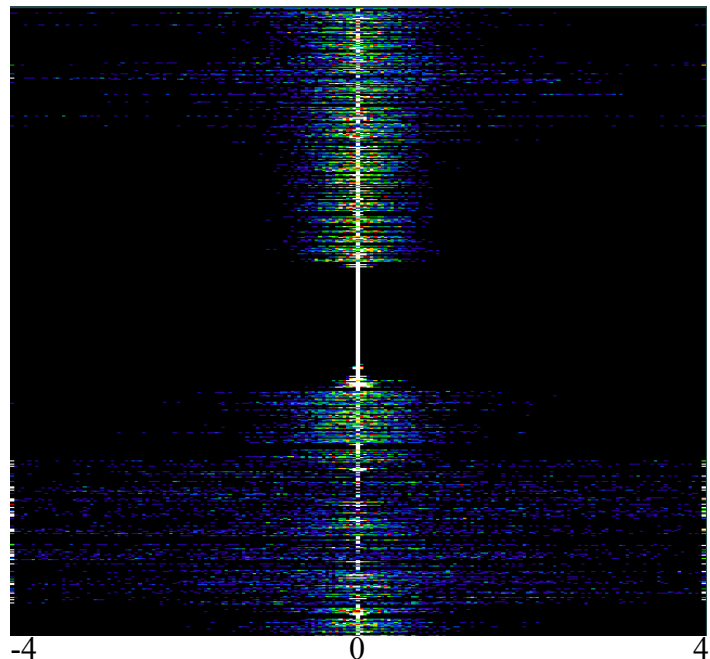


Figure 9. A 2D image displaying all of the individual distributions for each 128 3D time points (one run, six trials) of an fMRI timeseries. On the horizontal axis is the magnitude of the wavelet and the vertical axis labels the time point of the fMRI epoch. Colors indicate the frequencies of occurrence of a wavelet coefficient of certain magnitude across all 1,000 voxel locations. The across row average is effectively shown on Figure 8.



Bessel K forms and double-Pareto distribution models, in that order, provide increasingly heavier tails with Cauchy being the most likely candidate for the best fit to the observed wavelet coefficients across the entire range. The double-Pareto and the Bessel K forms densities provide the heaviest tails, however, they

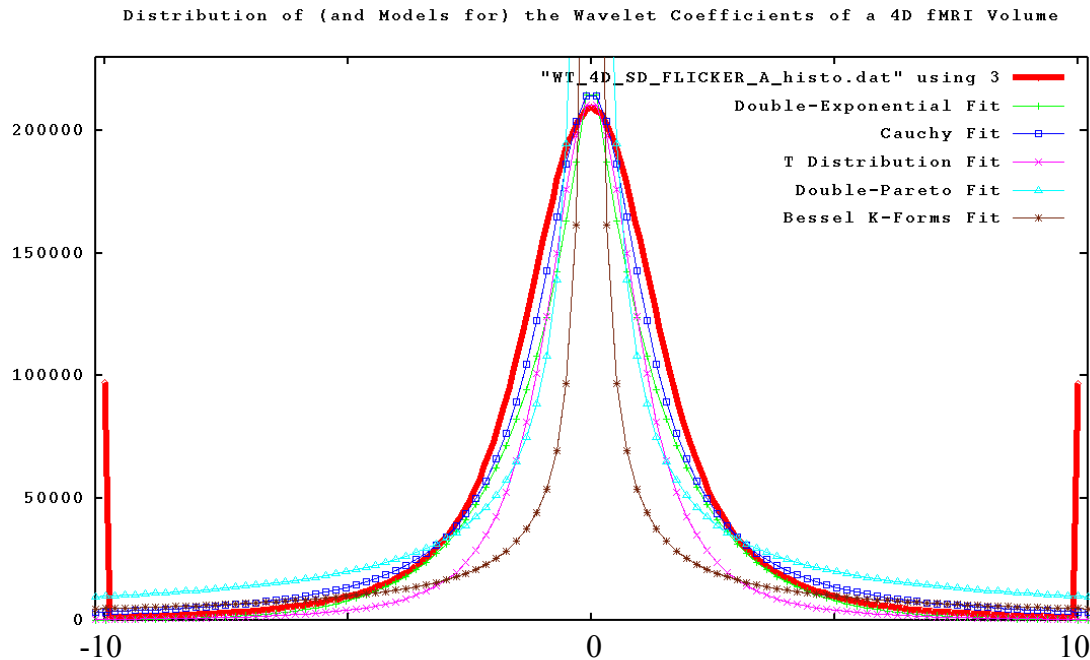


Figure 11. Several heavy-tail distribution models are fitted to the frequency histogram of 4D wavelet coefficients. These include *double-exponential*, *Cauchy*, *T*, *double-Pareto* and *Bessel K form* models. Because our statistical tests will be applied on the coefficients that survive wavelet shrinkage it is important to utilize a distribution model that provides accurate asymptotic approximation to the real data in the tail regions.

are inadequate in the central range  $[-3; 3]$  and undefined near the mean.

## Discussion and Conclusions

Atlas-based studies: The main advantage in using a frequency-spatial F&A SVPA atlas is that small across-subjects spatial intensity variations, within each ROI, will have less of an effect on the statistical maps, as compared to other time-domain approaches [Friston *et al.*, 1995; Dinov *et al.*, 2000]. Note that for any two subjects the bumps and valleys of the signal may have similar frequency-distributional characteristics but be slightly offset in the spatial-domain. This approach also avoids the low-pass filtering of the data required by other such techniques [Friston *et al.*, 1995]. The process of our frequency-spatial analysis begins by employing the N-dimensional discrete wavelet transform provided in the wavelet analysis of image registration package [Dinov *et al.*, 2001]. Then an appropriate distribution model is selected in a data-driven manner. And finally, the regional distribution of the wavelet parameters of the

data is compared against the distribution of the across subject wavelet coefficients saved in the F&A SVPA atlas. One can perform both individual vs. atlas and group/population comparisons in this atlas-based framework.

In the second part of this investigation we will apply this methodology to analyze the fMRI data acquired and processed by [Buckner et al., 2000]. This dataset consisted of 41 right-handed, English-speaking gender-matched volunteers separated in 3 groups: 14 young participants (5 males mean age  $21.1 \pm 3$

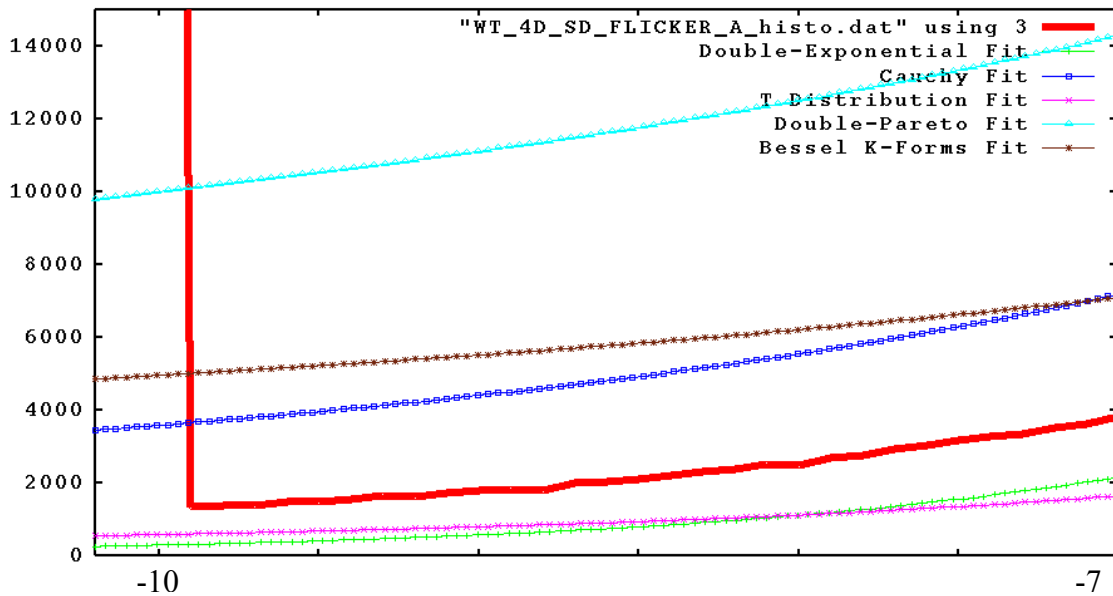


Figure 12. Shows the extreme left trail of the data distribution (data is symmetric). The *double-exponential* and the *T* distribution models underestimate the tails of the data asymptotically, but provide good fits around the mean. *Cauchy*, *Bessel K forms* and *double-Pareto* distribution models, in that order, provide increasingly heavier tails with *Cauchy* being the most likely candidate for the best fit to the observed wavelet coefficients across the entire range. The *double-Pareto* and the *Bessel K form* densities provide the heaviest tails, however, they are inadequate in the central range  $[-3 : 3]$  and undefined near the mean of zero.

years); 14 old subjects (five males) were non-demented with (mean age of  $74.9 \pm 9$  years) and; 13 old demented patients (six males) were demented (mean age of  $77.2 \pm 8$  years). Older adults were excluded if they had neurologic, psychiatric, or medical illnesses that could manifest as dementia. All older adults had normal (corrected) visual acuity. The Buckner's group employed the *T* and *F* statistics to assess for group effects (young adults, normal elderly and normal demented). These investigators found a significant difference between the groups of old demented and young adult subjects using Mann–Whitney nonparametric test. The individual subject and group analyses were performed in an event-related fashion using selective-averaging methods. To determine the added contribution of the second event in the two-trial conditions, the one-trial conditions were simply subtracted from the two-trial conditions. Secondly, Buckner and coworkers extracted the evoked hemodynamic responses for each subject from *a priori* defined regions and compared these using a random-effects statistical model. Peak activations in visual

and motor cortex were then derived from this composite activation map. The hemodynamic response and variance were obtained for each region and each subject and entered into subsequent analyses based on analysis of variance.

In the atlas-based method we are proposing, an F&A SVPA atlas of one population will be generated and the individual fMRI data from the remaining two groups will be tested against the distributions of the wavelet coefficients of the atlas. The statistical test we employ will be quite distinct from the general linear model used by others, because first our statistics will be computed on the data in compressed wavelet space. And secondly, the statistical tests we'll carry will involve one (or several) heavy-tail distribution models as discussed in the Methods and Results sections.

Heavy-tailness: In general, heavy tail signals are more likely to exhibit large observations and oftentimes may have an impulsive nature. Our empirical results indicate that 3D MRI/fMRI and 4D fMRI data have leptokurtic sporadic nature. The term *heavy tail* refers to the fact that the probability density functions (histograms) of the signals have relative large mass in the extreme *tails*. Because, the tails of leptokurtic distributions decay much slower than the Gaussian density many of the commonly used statistical tests are not appropriate for analyzing such data. Notably, in Figure 12 we see that the  $P$ -values corresponding to a statistical score of say  $-10.0$  for the  $T$  and  $Cauchy$  distributions will be different by several orders of magnitude. This means that one is bound to commit *type I* error by not using the correct heavy-tail distribution model. So, studying a small number of coefficients employing T-testing may be inefficient.

## Acknowledgements

This research is supported by grants from NIA P50 AG16570, K08 AG100784; NLM R01 2R01 LM05639-06; NIH/NCRR 2 P41 RR13642 and NIH/NIMH 5 P01 MN52176. We are also indebted to De Witt L. Sumners, Dimitre Stefanov and Anuj Srivastava at Florida State University for helpful hints and discussions.

## References

Abramovich F., Sapatinas T. and Silverman B. *Wavelet thresholding via a Bayesian approach*. Journal of the Royal Statistical Society, Series B, 60, 725-749 (1998).

- Aguirre G., Zarahn E., D'Esposito, M. *A critique of the use of the Kolmogorov-Smirnov (KS) statistic for the analysis of BOLD fMRI data*. Magn Reson Med., 39, 500–505 (1998).
- Agoncillo, A., Mejino, J. and Rosse, C. *Influence of the Digital Anatomist Foundational model on traditional representations of anatomical concepts*. In Proceedings, American Medical Informatics Association Fall Symposium Washington, D.C., 2-6, (1999).
- Alsberg, B., Woodward, A., Winson, M., Rowland, J. and Kell, D. *Variable selection in wavelet regression models*. Analytica Chimica Acta, 368, 29-44, (1998).
- Amit, Y. *Graphical shape templates for automatic anatomy detection with applications to MRI brain scans*. IEEE Transactions on Medical Imaging, 16(1), 28-40 (1997).
- Baleydier C. and Mauguière F. *The duality of the cingulate gyrus in monkey. Neuroanatomical study and functional hypothesis*. Brain. 103, 525-554 (1980).
- Bandettini, P., Wong, E. *A hypercapnia-based normalization method for improved spatial localization of human brain activation with fMRI*. NMR in Biomedicine, 10(4-5), 197-203 (1997).
- Bandettini, P., Jesmanowicz, A., Wong, E. and Hyde, J. *Processing strategies for time-course data sets in functional MRI of the human brain*. Magnetic Resonance in Medicine, vol. 30, 161-173, (1993).
- Barber, S., Nason, G. and Silverman, B. *Posterior probability intervals for wavelet thresholding*. To appear in the Journal of the Royal Statistical Society, Series B, 64 (2002).
- Bradley, S., Rosse, C. and Brinkley, J. *Web-based access to an online atlas of anatomy: the Digital Anatomist Common Gateway Interface*. In 19th Symposium on Computer Applications in Medical Care, New Orleans, 512-516, (1995).
- Brown, P., Fearn, T. and Vannucci, M. *Bayesian wavelet regression on curves with application to a spectroscopic calibration*. Journal of the American Statistical Association, 96, 398-408 (2001).
- Buckner R., Snyder A., Sanders A., Raichle M. and Morris J. *Functional Brain Imaging of Young, Nondemented, and Demented Older Adults*. Journal of Cognitive Neuroscience, 12(suppl. 2), 24-34 (2000).
- Bullmore, E., Brammer, M., Williams, S., Rabe-Hesketh, S., Janot, N., David, A., Mellers, J., Howard, R. and Sham, P. *Statistical methods of estimation and inference for functional MR image analysis*. Magnetic Resonance in Medicine, vol. 35, 261-277, (1996).
- Bullmore, E., Long, C., Suckling, J., Fadili, J., Calvert, G., Zelaya, F., Carpenter, T., Brammer, M. *Colored noise and computational inference in neurophysiological (fMRI) time series analysis: Resampling methods in time and wavelet domains*. Human Brain Mapping, 12(2): 61-78 (2001).
- Buzás, P., Eysel, U. and Kisvárdy, Z. *Functional topography of single cortical cells: an intracellular approach combined with optical imaging*. Brain Res. Brain Res. Protoc. 3 (2), 199-208 (1998).
- Cannestra, A., Pouratian, N., Bookheimer, S., Martin, N., Becker, D., Toga, A. *Temporal spatial differences observed by functional MRI and human intraoperative optical imaging*. Cerebral Cortex, 11 (8), 773-782 (2001).
- Cao, J. and Worsley, K. *The detection of local shape changes via the geometry of Hotelling's  $T^2$  fields*. Annals of Statistics, 27, 925-942 (1999).
- Cavada C. and Glodman-Rakic P. *Posterior parietal cortex in rhesus monkey: I. Parcellation of areas based on distinctive limbic and sensory corticocortical connections*. J. Comp. Neurol., 287, 393-421 (1989).
- Chuang, K.-H., Chiu, M.-J. and Lin, C.-C. *Model-Free Functional MRI Analysis Using Kohonen Clustering Neural Network and Fuzzy C-Means*. IEEE Transactions on Medical Imaging, vol. 18, 1117-1128, (1999).
- Cocosco, C., Kollokian, V., Kwan, R. and Evans, A. *BrainWeb: Online Interface to a 3D MRI Simulated Brain Database*. NeuroImage, vol.5, no.4, part 2/4, S425 (1997).
- Crabtree, E., Mega, M., Linshield, C., Dinov, I., Thompson, P., Felix, J., Cummings, J., Toga, A. *Alzheimer grey matter loss across time: unbiased assessment using a probabilistic Alzheimer brain atlas*. Soc. for Neurosci. Abs. 26:294 (2000).
- Davis, G. *A wavelet-based analysis of fractal image compression*. IEEE Transactions on Image Processing. 7(2), 141-154 (1998).
- Daubechies, I. *Orthonormal bases of compactly supported wavelets*. Comm. Pure & Appl. Math. 41, 909-996 (1988).
- Daubechies, I. *The wavelet transform: a method for time-frequency localization*. In the book Advances in Spectrum Analysis and Array Processing, Vol. 1, ed. S. Haykin, Prentice-Hall, 366-417, (1991).
- Daubechies, I. *Ten Lectures on Wavelets*. CBMS-NSF Lecture Notes nr. 61, SIAM, 1992.

- Daubechies, I. *Orthonormal bases of compactly supported wavelets, II. Variations on a theme*, SIAM J. Math. Anal. **24**, 499-519 (1993).
- Desco, M., Hernandez, J., Santos, A. and Brammer, M. *Multiresolution analysis in fMRI: Sensitivity and specificity in the detection of brain activation*. Human Brain Mapping, 14 (1): 16-27 (2001).
- Dinov, I., *Mathematical and statistical techniques for modeling and analysis of medical data*. xi, 159 leaves, Libraries: Florida State University, Florida USA (1998).
- Dinov, I., Mega, M., Thompson, P., Lee, L., Woods, R., Holmes, C., Sumners, D.-W., Toga, A. *Analyzing functional brain images in a probabilistic atlas: a validation of subvolume thresholding*. Journal of Computer Assisted Tomography, 24(1), 128-138 (2000).
- Dinov, I., Mega M. and Toga A. *Wavelet Analysis of Image Registration*. Abs. of Papers Presented to the American Mathematical Society, Vol. 21, Issue 3, No. 956-92-74 (2000a).
- Dinov, I. and Sumners, D-W. *Applications of Frequency Dependent Wavelet Shrinkage to Analyzing Quality of Image Registration*. SIAM J. Appl. Math. (SIAP), 62(2), 367-384 (2001).
- Dinov, I., Mega, M., Manese, M., Felix, J., Tran, N., Lindshield, C., Cummings, J. and Toga, A. *Construction of the First Rest-State Functional Sub-Volume Probabilistic Atlas of Normal Variability in the Elderly and Demented Brain*. Neurology, Abstracts of the 53-rd AAN meeting, Philadelphia, PA, May 06-11, (2001).
- Dinov, I., Mega, M., Thompson, P., Woods, R., Sumners, D.-W., Sowell, E., Toga, A. *Quantitative Comparison and Analysis of Image Registration Using Frequency-Adaptive Wavelet Shrinkage*. IEEE-TITB 6(1), 73-85, (2002).
- Donoho, David L. *De-noising by soft-thresholding*. IEEE Transactions on Information Theory, 41(3), 613 -628 (1995).
- Donoho, D. and Johnstone, I. *Adapting to unknown smoothness via wavelet shrinkage*. J. Am. Statist. Ass., 90, 1200-1224 (1995).
- Donoho, D., Vetterli, M., DeVore, R. and Daubechies, I. *Data compression and harmonic analysis*. IEEE Trans. Inf. Theory, 44 (6), 2435-2476, (1998).
- Espinal, F., Huntsberger, T., Jawerth, B., Kubota, T., *Wavelet-based fractal signature analysis for automatic target recognition*. SPIE, Opt. Eng. 37(1), 166-174 (1998).
- Evans, A. Collins, D. and Holmes, C. *Computational approaches to quantifying human neuroanatomical variability*. In *Brain Mapping: The Methods*, Mazziotta J. and Toga A. eds., Academic Press, 343-361 (1996).
- Fadili, M. and Bullmore, E. *Wavelet-Generalized Least Squares: A New BLU Estimator of Linear Regression Models with 1/f Errors*. NeuroImage, 15, 217-232 (2002).
- Falconer, K. *Fractal geometry: Mathematical foundations and applications*. John Wiley & Sons, 1990.
- Feilner, M.; Blu, T.; Unser, M. *Optimizing wavelets for the analysis of fMRI data*. Proceedings of the SPIE - The International Society for Optical Engineering, vol.4119, pt.1-2, SPIE-Int. Soc. Opt. Eng., 626-37 (2000).
- Fowler, J. and Hua, L. *Omnidirectionally Balanced Multiwavelets for vector wavelet transforms*. In Proceedings of IEEE Data Compression Conference, Storer, J. and Cohn, M., eds., April, Utah, USA, 422-431 (2002)
- Friston, K., Frith, C., Liddle, P. and Frackowiak, R. *Functional connectivity: The principal-component analysis of large (PET) data sets*. Journal of Cerebral Blood Flow and Metabolism, 13:5-14 (1993).
- Friston K., Worsley K., Frackowiak R., Mazziotta J., Evans A. *Assessing the Significance of Focal Activations Using their Spatial Extent*. Human Brain Mapping, 1, 214-220 (1994).
- Friston, K., Holmes, A., Worsley, K., Poline, J., Frith, C. and Frackowiak, R. *Statistical Parametric Maps in Functional Imaging: A General Linear Approach*. Human Brain Mapping, vol. 2, 189-210, (1995).
- Friston, K., Frith, C., Frackowiak R. and Turner, R. *Characterizing dynamic brain responses with fMRI: a multivariate approach*. NeuroImage, 2, 166-172 (1995a).
- Friston, K., Ashburner, J., Poline, J., Frith, C., Heather, J. and Frackowiak, R. *Spatial registration and normalization of images*. Hum. Brain Mapping. 2, 165-189 (1995b).
- Friston, K., Buchel, C., Fink, C., Morris, G., Rolls, J. and Dolan, R. *Psychophysiological and modulatory interactions in neuroimaging*. NeuroImage, 6, 218-229 (1997).
- Friston, K. and Buchel, C., *Attentional Modulation of Effective Connectivity from V2 to V5/MT in Humans*. PNAS, 97(13), 7591-7596 (2000).
- Golay, X., Kollias, S., Stoll, G., Meier, D., Valvanis, A. and Boesiger, P. *A new correlation-based fuzzy logic clustering algorithm for fMRI*. Magnetic Resonance in Medicine, vol. 40, pp. 249-60, (1998).

- Gourieroux, C. and Monfort, A. *Qualitative threshold ARCH models. (Autoregressive Conditional Heteroscedasticity) (ARCH Models in Finance)* Journal of Econometrics v52, n1-2 (April-May, 1992):159-200.
- Grenander, U and Srivastava, A. *Probability Models for Clutter in Natural Images*. IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 23, no. 4, 424-429 (2001).
- Hajnal, J., Myers, R., Oatridge, A., Schwieso, J., Young, I. and Byder, G. *Artifacts due to stimulus correlated motion in functional imaging of the brain*. Magnetic Resonance in Medicine, vol. 31, 283-291, (1994).
- Haller, J., Banerjee, A., Christensen, G., et al. *Three-dimensional hippocampal MR morphometry with high-dimensional transformation of a neuroanatomic atlas*. Radiology, 202, 504-510 (1997).
- Hess, A., Stiller, D., Kaulisch, T., Heil, P., Schliech, H. *New insights into the hemodynamic blood oxygenation level-dependent response through combination of functional magnetic resonance imaging and optical recordings in gerbil barrel cortex*. J. Neurosci., 20, 3328-3338 (2000).
- Hilton, M., Odgen, T., Hattery, D., Eden, G. and Jawerth, J. *Wavelet Denoising of Functional MRI Data*. In: Wavelets in Medicine and Biology, A. Aldroubi and M. Unser, Eds. Boca Raton: CRC Press, 93-114 (1996).
- Horwitz, B., Grady, C., Mentis, M., Pietrini, P., Ungerleider, L., Rapoport, S. and Haxby, J. *Brain functional connectivity changes as task difficulty is altered*. NeuroImage, 3, S248 (1996).
- Huang-Hellinger, F., Breiter, H., McCormack, G., Cohen, M., Kwong, K., Sutton, J., Savory, R., Weisskoff, R., Davis, T., Baker, J., Belliveau, J., Rosen, B. *Simultaneous functional magnetic resonance imaging and electrophysiological recording*. Human Brain Mapping, 3(1), 13-23 (1995).
- Huttenlocher, P. *Synaptic density in human frontal cortex - developmental changes and effects of aging*. Brain Res 163(2), 195-205 (1979).
- Huttenlocher, P. and de Courten, C. *The development of synapses in striate cortex of man*. Hum Neurobiol 6(1), 1-9 (1987).
- Jack, C., Petersen, R., Xu, Y., et al. *Prediction of AD with MRI-based hippocampal volume in Mild cognitive impairment*. Neurology, 2, 97-1403 (1999).
- Jack, C., Petersen, R., Xu, Y., et al. *Medial temporal atrophy on MRI in normal aging and very mild Alzheimer's disease*. Neurology, 49, 786-794 (1997).
- Johnstone, I. and Silverman, B. *Wavelet threshold estimators for data with correlated noise*. J. Roy. Statist. Soc. B., 59, 319-351 (1997).
- Juottonen, K., Laakso, M., Insausti, R., et al. *Volumes of the entorhinal and perirhinal cortices in Alzheimer's disease*. Neurobiol. Aging., 19, 15-22 (1998).
- Jensen, M. *An alternative maximum likelihood estimator of long-memory processes using compactly supported wavelets*. J. Econ. Dynamics and Control, 24, 361-387 (2000).
- Jernigan, T., Trauner, D., Hesselink, J. and Tallal, P. *Maturation of human cerebrum observed in vivo during adolescence*. Brain 114(5)(10), 2037-49 (1991).
- Johnson, K. and Becker, A. *The whole brain atlas*. Philadelphia : Lippincott Williams & Wilkins, 1 computer optical disc : col. (1999).
- Jones, K. and Smith, D. *Recognition of the fetal alcohol syndrome in early infancy*. Lancet 2(7836), 999-1001 (1973).
- Kaye, J., Swihart, T., Howieson, D., et al. *Volume loss of the hippocampus and temporal lobe in healthy elderly persons destined to develop dementia*. Neurology, 48, 1297-1304 (1997).
- Kershaw, J., Ardekani, B. and Kanno, I. *Application of Bayesian inference to fMRI data analysis*. IEEE Transactions on Medical Imaging, vol. 18, 1138-53 (1999).
- Kikinis et al. *A Digital Brain Atlas for Surgical Planning, Model Driven Segmentation and Teaching*. IEEE Transactions on Visualization and Computer Graphics, 2(3), 232-241, (1996).
- Kimeldorf, G. and Wahba, G. *A correspondence between Bayesian estimation of stochastic processes and smoothing by splines*. Ann. Math. Statist. 41(2), 495-502 (1970).
- Kjems, U., Strother, S., Anderson, J., Law, I., Hansen, L. *Enhancing the multivariate signal of [<sup>15</sup>O] water PET studies with a new nonlinear neuroanatomical registration algorithm [MRI application]*. IEEE Transactions on Medical Imaging, 18(4), 306-19 (1999).
- Konishi, S., Nakajima, K., Uchida, I., Kikyo, H., Kameyama, M. and Miyashita, Y. *Common inhibitory mechanism in human inferior prefrontal cortex revealed by event-related functional MRI*. Brain, 122(5), 981-91, (1999).

- Kötter R., Staiger J.F., Zilles K., Luhmann H.J. *Analyzing functional connectivity in brain slices by a combination of infrared videomicroscopy, flash photolysis of caged compounds, and scanning methods*. Neuroscience 86, 265-277 (1998).
- Krakow, K. et al. *EEG-triggered functional MRI of interictal epileptiform activity in patients with partial seizures*. Brain 122, 1679-1688 (1999).
- Krakow, K. et al. *EEG recording during fMRI experiments: image quality*. Hum. Brain Mapp. 10, 10-15 (2000).
- Kundur D. and Hatzinakos, D. *Digital Watermarking for Telltale Tamper Proofing and Authentication*. Proc. of IEEE, 87(7), 1167-1180 (1999).
- Kuppusamy K., Lin W. and Haacke E. *Statistical assessment of cross-correlation and variance methods and the importance of electrocardiogram gating in functional magnetic resonance imaging*. Magn. Reson. Imaging 15:169-181 (1997).
- Langer, M. *Large scale failures of  $f^a$  scaling in natural image spectra*. Journal of the Optical Society of America A, 17(1), 28-33 (2000).
- Logothetis, N., Pauls, J., Augath, M., Trinath, T. and Oeltermann, A. *Neurophysiological investigation of the basis of the fMRI signal*. Nature, vol. 412, 150-157 (2001).
- Lega, H., Scholl, H., Alimi, J.M., Bijaoui, A., Bury. *A parallel algorithm for structure detection based on wavelet and segmentation analysis*. Parallel Computing, vol. 21, 265-285 (1995).
- Mai, J., Assheuer, J. and Paxinos, G. *Atlas of the Human Brain*. Academic Press, San Diego, (1997).
- Mandelbrot, B. and van Ness, J. *Fractional Brownian motion, fractional noises and applications*. SIAM Rev. 10(4), 422-437 (1968).
- Mallat, S. *A theory for multi-resolution signal decomposition: The wavelet representation*. IEEE-TPAMI, vol. 11, 674-693 (1989).
- Mattson, S. and Riley, E. *A review of the neurobehavioral deficits in children with fetal alcohol syndrome or prenatal exposure to alcohol*. Alcoholism, Clinical and Experimental Research 22(2), 279-94 (1998).
- Mazziotta, J., Toga, A., Evans, A., Fox, P. and Lancaster, J. *A probabilistic atlas of the human brain: theory and rationale for its development. the international consortium for brain mapping*. NeuroImage, 2(2), 89-101 (1995).
- McIntosh, A. and Gonzalez-Lima, F. *Structural equation modeling and its application to network analysis of functional brain imaging*. Human Brain Mapping, 2, 2-22 (1994).
- McIntosh, A., Bookstein, F., Haxby, J. and Grady, C. *Spatial pattern analysis of functional brain images using Partial Least Squares*. NeuroImage, 3, 143-157 (1996).
- Mega M., Fennema C., Dinov I., Thompson P., Archibald S., Lindshield C., Felix J., Toga A., Jernigan T. *Construction, testing, and validation of a sub-volume probabilistic human brain atlas for the elderly and demented populations*. NeuroImage, 11(5), S597 (2000a).
- Mega, M., Thompson, P., Toga, A. and Cummings, J. *Neuroimaging in Dementia*. In: Mazziotta, J., Toga, A., Frackowiak, R., ed. *Brain Mapping: The Disorders*. San Diego: Academic Press, 217-293 (2000).
- Mega, M., Cummings, J., Masterman, D., Dinov, I., Felix, J., O'Connor, S., Phelps, M., Small, G., Toga, A. *Cognitive and metabolic responses to metrifonate therapy in Alzheimer's disease*. J. Neuropsychiatry, Neuropsychology, and Behavioral Neurology, 14(1), 63-68, 2001.
- Mekle, R., Laine, A., Perera, G., DeLaPaz, R. *Activation detection in fMRI data via multi-scale singularity detection*. Proceedings of the SPIE - The International Society for Optical Engineering, vol.4119, pt.1-2, SPIE-Int. Soc. Opt. Eng., 615-25 (2000).
- Murtagh, F., Starck, J.L., Bijaoui, A. *Multiresolution in astronomical image processing: a general framework*. Intern. J. Image Systems and Technology, vol. 6, 332-338 (1995).
- Nyberg, L., McIntosh, A., Cabeza, R., Nilsson, L-G, Houle, S., Habib, R., Tulving, R. *Network analysis of PET rCBF data: Ensemble inhibition during episodic memory retrieval*. Journal of Neuroscience, 16(11), 3753-3759 (1996).
- Ogawa, S; Lee, T M; Kay, A R; Tank, D W. *Brain magnetic resonance imaging with contrast dependent on blood oxygenation*. PNAS, v.87, n.24, 9868-9872 (1990).
- O'Rahilly, R. and Müller, F. *The Embryonic Human Brain: An Atlas of Developmental Stages*, John Wiley & Sons, 464 pages, (1999).
- Pandya, D., Van Hoesen, G. and Mesulam, M-M. *Efferent connections of the cingulate gyrus in the rhesus monkey*. Exp. Brain Res., 42, 319-330 (1981).

- Rashkovskiy, O., Sadovnik, L. and Caviris, N. *Scale, rotation and shift invariant wavelet transform*. In Proceedings of SPIE, vol. 2237, 390-401 (1994).
- Ruttimann, U., Unser, M., Rawlings, R., Rio, D., Ramsey, N., Mattay, V., Hommer, D., Frank, J. and Weinberger, D. *Statistical analysis of functional MRI data in the wavelet domain*. IEEE Trans. Medical Imaging 17 (2), 142–154 (1998).
- Silverman, J. *Bayesian confidence intervals for the cross validated smoothing spline*. J. Roy. Stat. Soc. B., 45:133-150 (1983).
- Steinberg, D. *A Bayesian approach to flexible modeling of multivariate response for functions*. Journal of Multivariate Analysis, 34:157-172 (1990).
- Sweldens, W. *The lifting scheme: A construction of second generation wavelets*. SIAM J. Math. Anal. 29(2), 511–546 (1997).
- Talairach J. and Tournoux P., *Co-Planar Stereotactic Atlas Of The Human Brain*. Stuttgart: Georg. Thieme Verlag, 1988
- Toga, A. and Thompson, P. *The Role of Image Registration in Brain Mapping*. Image and Vision Computing Journal, 19, 3-24 (2001).
- Thompson P., Schwartz C., Lin R., Khan A. and Toga A. *Three-dimensional statistical analysis of sulcal variability in the human brain*. J. Neurosci. 16(13), 4261-4274 (1996a).
- Thompson P., MacDonald, D., Mega, M., Holmes, C., Evans, A. and Toga, A. *Detection and mapping of abnormal brain structure with a probabilistic atlas of cortical surfaces*. J. Comp. Asst. Tech., 21, 567-581 (1997).
- Thompson, P., Moussai, J., Zohoori, S., et al. *Cortical variability and asymmetry in normal aging and Alzheimer's disease*. Cerebral Cortex, 8, 492-509 (1998).
- Thompson, P., Woods, R., Mega, M., Toga, A. *Mathematical/Computational Challenges in Creating Deformable and Probabilistic Atlases of the Human Brain*. Human Brain Mapping 9(2), 81-92, (2000).
- Thompson, P., Mega, M., Woods, R., et al. *Cortical change in Alzheimer's disease detected with a disease-specific population-based brain atlas*. Cerebral Cortex, 11, 1-16 (2001).
- Tsai, A., Fisher, J., Wible, C., Wells, W., Kim, J. and Willsky, A. *Analysis of Functional MRI Data Using Mutual Information*. Presented at Medical Imaging and Computer-Assisted Intervention - MICCAI'99, C. Taylor, A. Colchester, Eds. Cambridge, England, 473-480 (1999).
- Van Hoesen, G., Morecraft, R., Vogt, B. *Connections of the monkey cingulate cortex*. In: Vogt BA, Gabriel M, ed. Neurobiology of cingulate cortex and limbic thalamus: a comprehensive handbook. Boston: Birkhäuser, 249-284 (1993).
- Van Horn, J. and Gazzaniga, M. *Databasing fMRI studies – towards a 'discovery science' of brain function*. Nature Reviews Neurosci., 3, 314-318, (2002).
- Vannier, M., Brunsdon, B., Hildebolt, C., et al. *Brain surface cortical sulcal lengths: Quantification with three-dimensional MR imaging*. Radiology, 180, 479-484 (1991).
- Vogt, B. and Pandya, D. *Cingulate cortex of the rhesus monkey: II. Cortical afferents*. J. Comp. Neurol., 262, 271-289 (1987).
- Woods R., Cherry S. and Mazziotta J. *Rapid automated algorithm for aligning and reslicing PET images*. J. Comput. Assist. Tomogr., 16(4), 620-633 (1992).
- Wornell, G. *Signal processing with fractals: a wavelet-based approach*. Prentice Hall (1996).
- Worsley, K. *Local maxima and the expected Euler characteristic of excursion sets of  $X^2$ ,  $F$  and  $T$  fields*. Advances in Applied Probability, 26, 13-42 (1994).
- Worsley, K., Poline, J-B., Friston, K. and Evans, A. *Characterizing the response of fMRI data using multivariate linear models*. NeuroImage, 6, 305-319 (1997).
- Worsley, K.J. *An overview and some new developments in the statistical analysis of PET and fMRI data*. Human Brain Mapping, 5, 254-258 (1997a).
- Worsley, K.J., Cao, J., Paus, T., Petrides, M. and Evans, A. *Applications of random field theory to functional connectivity*. Human Brain Mapping, 6, 364-367 (1998).
- Yu, T. *Approximation order/smoothness tradeoff in Hermite subdivision schemes*. Proc. SPIE, Wavelets: Applications in Signal and Image Processing IX, Andrew F. Laine; Michael A. Unser; Akram Aldroubi; Eds., Vol. 4478, p. 117-128 (2001).