



**HAL**  
open science

## **Fractional Gaussian noise, functional MRI and Alzheimer's disease**

Voichita Maxim, Levent Sendur, Jalal M. Fadili, John Suckling, Rebecca Gould, Rob Howard, Ed Bullmore

► **To cite this version:**

Voichita Maxim, Levent Sendur, Jalal M. Fadili, John Suckling, Rebecca Gould, et al. Fractional Gaussian noise, functional MRI and Alzheimer's disease. *NeuroImage*, 2005, 25 (1), pp.141-158. <10.1016/j.neuroimage.2004.10.044>. <hal-00260264>

**HAL Id: hal-00260264**

**<https://hal.science/hal-00260264v1>**

Submitted on 10 Mar 2015

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

# Fractional Gaussian noise, functional MRI and Alzheimer’s disease

Voichița Maxim,<sup>a</sup> Levent Şendur,<sup>a</sup> Jalal Fadili,<sup>b</sup> John Suckling,<sup>a</sup> Rebecca Gould,<sup>c</sup>  
Rob Howard,<sup>c</sup> and Ed Bullmore<sup>a,\*</sup>

<sup>a</sup>Department of Psychiatry, Brain Mapping Unit, University of Cambridge, Department of Psychiatry, Addenbrooke’s Hospital, Cambridge CB2 2QQ, UK

<sup>b</sup>GREYC CNRS UMR 6072, Caen, France

<sup>c</sup>Institute of Psychiatry, King’s College London, UK

Fractional Gaussian noise (fGn) provides a parsimonious model for stationary increments of a self-similar process parameterised by the Hurst exponent,  $H$ , and variance,  $\sigma^2$ . Fractional Gaussian noise with  $H < 0.5$  demonstrates negatively autocorrelated or antipersistent behaviour; fGn with  $H > 0.5$  demonstrates  $1/f$ , long memory or persistent behaviour; and the special case of fGn with  $H = 0.5$  corresponds to classical Gaussian white noise. We comparatively evaluate four possible estimators of fGn parameters, one method implemented in the time domain and three in the wavelet domain. We show that a wavelet-based maximum likelihood (ML) estimator yields the most efficient estimates of  $H$  and  $\sigma^2$  in simulated fGn with  $0 < H < 1$ . Applying this estimator to fMRI data acquired in the “resting” state from healthy young and older volunteers, we show empirically that fGn provides an accommodating model for diverse species of fMRI noise, assuming adequate preprocessing to correct effects of head movement, and that voxels with  $H > 0.5$  tend to be concentrated in cortex whereas voxels with  $H < 0.5$  are more frequently located in ventricles and sulcal CSF. The wavelet-ML estimator can be generalised to estimate the parameter vector  $\beta$  for general linear modelling (GLM) of a physiological response to experimental stimulation and we demonstrate nominal type I error control in multiple testing of  $\beta$ , divided by its standard error, in simulated and biological data under the null hypothesis  $\beta = 0$ . We illustrate these methods principally by showing that there are significant differences between patients with early Alzheimer’s disease (AD) and age-matched comparison subjects in the persistence of fGn in the medial and lateral temporal lobes, insula, dorsal cingulate/medial premotor cortex, and left pre- and postcentral gyrus: patients with AD had greater persistence of resting fMRI noise (larger  $H$ ) in these regions. Comparable abnormalities in the AD patients were also identified by a permutation test of local differences in the first-order autoregression AR(1) coefficient, which was significantly more positive in patients. However, we found that the Hurst exponent provided a more sensitive metric than the AR(1) coefficient to detect these differences, perhaps because neurophysiological changes in early AD are naturally better described in terms of abnormal salience of long memory dynamics than a change in the strength of association between

immediately consecutive time points. We conclude that parsimonious mapping of fMRI noise properties in terms of fGn parameters efficiently estimated in the wavelet domain is feasible and can enhance insight into the pathophysiology of Alzheimer’s disease.

*Keywords:* Neuroimaging; Brain; Fractional Brownian motion; Spectral exponent; Neurodegeneration; Fractal

## Introduction

It is well-known that functional magnetic resonance imaging (fMRI) time series typically demonstrate complex and locally variable autocorrelation structure, even when the data have been acquired with the subject “at rest”. There is preliminary evidence that fMRI noise often has long memory in time, or  $1/f$  spectral properties, meaning it is positively autocorrelated and there is disproportionate power in the spectrum at low frequencies (Zarahn et al., 1997); for review, see Bullmore et al. (2004) and references therein. However, there may also be high-frequency events like spikes or transients, or more sustained bursts of negative autocorrelation, in fMRI noise.

Not only is fMRI noise diverse, its presumably multiple sources are incompletely known, and are likely to vary in impact from one data set to the next. Head movement, for example, is an individually variable but common source of long memory noise caused by slow rotation or translation of the subject’s head through an imperfectly homogeneous magnetic or RF field during scanning (Bullmore et al., 1999a). Cardiac and/or respiratory cycle-related pulsations may also contribute noise with properties depending in part on the sampling rate of data acquisition (repetition time, TR) and the proportion of cerebrospinal fluid represented in a voxel (Cordes et al., 2001; Purdon and Weisskoff, 1999). There are inevitably also instrumental and thermal sources of noise.

To date, one of the most successful modelling strategies for fMRI noise  $G = (G_1, G_2, \dots, G_n)$ ,  $n$  being the number of time

points, has been the adoption of autoregressive, linear time invariant models (Bullmore et al., 1996, 2001; Dale, 1999; Locascio et al., 1997; Purdon and Weisskoff, 1999; Worsley et al., 2002) of the form

$$G_t = \sum_{i=1}^p \eta_i G_{t-i} + \varepsilon_t, \quad \varepsilon_t \stackrel{iid}{\sim} N(0, \sigma^2), \quad (1)$$

where  $p$  is the order of the autoregressive AR( $p$ ) process and  $t = 1, 2, \dots, n$ . However, AR models will require many parameters to account for long-range autocorrelated processes. The variability of autocorrelation between voxels (Bullmore et al., 1996; Worsley et al., 2002) suggests that it might be appropriate to adapt the order of AR process to each individual time series, which can be automated using model selection criteria such as the Bayesian information criterion (Fadili and Bullmore, 2002), but this is not always done in practice.

In this paper, we consider an alternative class of models, called fractional Gaussian noise or fGn, as a new approach to statistical modelling of fMRI noise. More formally, we are interested in the model that fMRI noise is distributed as a fractional Gaussian noise specified completely by only two parameters, its Hurst exponent  $H$  and its variance  $\sigma^2$ , i.e.,  $G \sim N(0, \Sigma)$ , with covariance matrix  $\Sigma := \Sigma(H, \sigma^2)$ . The main potential advantage of fGn compared to AR( $p$ ) models is its simplicity and parsimony: Only two parameters need to be estimated, with no complications concerning optimal model order specification.

In the rest of this paper, we rehearse some key mathematical properties of fractional Gaussian noise; we comparatively evaluate four possible estimators of fGn parameters; and we apply the optimal, wavelet-based maximum likelihood estimator to analysis of fMRI noise properties in resting data acquired from healthy volunteers and patients with early Alzheimer's disease, motivated by the hypothesis that there might be disease-related changes in the Hurst exponent of fMRI noise (Jeong, 2004).

## Methods

### Terms and notation

A fractional Gaussian noise  $G = (G_t; t = 1, 2, \dots, n)$  is a zero mean stationary process characterised by two parameters: the Hurst exponent  $H \in (0, 1)$  and the variance  $\sigma^2 := \text{Var}(G_t)$ ,  $t \in \mathbb{N}$ . The distribution of fractional Gaussian noise is fully specified by its auto-covariances at lags  $\tau \in \mathbb{Z}$  (Beran, 1994),

$$c(\tau) := \frac{\sigma^2}{2} \left( |\tau + 1|^{2H} - 2|\tau|^{2H} + |\tau - 1|^{2H} \right). \quad (2)$$

As can be seen from Eq. (2), the Hurst exponent is a measure of the long-term correlation between the discrete time points  $G_t$ , whereas the variance is only a scale parameter.

Alternatively, a fractional Gaussian noise can be seen as the unique Gaussian process that is the stationary increment of a self-similar process, called fractional Brownian motion (Mandelbrot and van Ness, 1968). When  $H = 0.5$ , this generalisation is of course consistent with the familiar observation that a time series of white Gaussian noise cumulatively constitutes a sample of classical Brownian motion. It is easy to check from Eq. (2) that the samples are now independent:  $c(\tau) = 0$ , for all  $\tau \neq 0$ . The middle

value  $H = 0.5$  is also a special case because it defines two distinct regions in the interval  $(0, 1)$ . A Hurst exponent  $H \in (0, 0.5)$  means that the process is negatively correlated, or anti-persistent, whereas  $H \in (0.5, 1)$  means that it is positively correlated or has long memory. Some simulated fractional Gaussian noises, exemplifying these three kinds of behaviour, are shown in Fig. 1, along with their autocorrelations until lag 20.

The auto-covariances characterise fGn in the time domain. The equivalent information in the frequency domain is carried by the spectral density function (SDF), also called the power spectrum, and denoted here by  $S(f)$ ,  $f \in [-1/2, 1/2]$ .

The spectral density of an fGn is more precisely defined as the Fourier transform of its auto-covariance sequence,

$$S(f) := \sum_{\tau=-\infty}^{\infty} c(\tau) e^{-i2\pi f\tau} = 4\sigma^2 C_H \sin^2(\pi f) \times \sum_{j=-\infty}^{\infty} \frac{1}{|f+j|^{2H+1}}, \quad -\frac{1}{2} \leq f \leq \frac{1}{2}, \quad (3)$$

with  $C_H := \Gamma(2H+1)\sin(\pi H) / (2\pi)^{2H+1}$  (Beran, 1994; Sinai, 1976). By Taylor expansion, we can see that

$$S(f) \approx \sigma^2 C_H (2\pi)^2 |f|^{1-2H}, \quad (4)$$

in the neighbourhood of the origin. Thus,  $S(f)$  is integrable, and we have  $\int_{-1/2}^{1/2} S(f) df = \sigma^2$ . We note that the SDF is related to frequency by a power law with spectral exponent

$$\gamma = 1 - 2H, \quad -1 < \gamma < 1, \quad (5)$$

i.e.,

$$S(f) \propto |f|^\gamma \quad (6)$$

or, equivalently,

$$\log S(f) \propto \gamma \log |f|. \quad (7)$$

This approximately linear relation between  $\log S(f)$  and  $\log |f|$ , the log periodogram, provides a simple empirical estimator of the spectral exponent  $\gamma$ , useful especially in the case of general 1/f processes without entailing parametric assumptions.

Thus, when  $H > 0.5$ , the spectral exponent will be less than zero and power will attenuate as frequency increases (or the spectrum will be approximately  $1/f$ ); when  $H < 0.5$ , the spectral exponent will be greater than zero and power will increase with frequency; and when  $H = 0.5$ , all frequencies will be present with equal power in the spectrum (see Fig. 2).

Spectral densities defined by Eq. (3) for fractional Gaussian noises with various values of  $H$  can be compared directly to the SDFs approximated by truncated Taylor expansion, Eq.(4). It is clear that, especially when  $H < 0.5$ , the approximate SDF is only acceptable for small values of  $f$ ; see Fig. 2.

### Wavelets and fractional Gaussian noise

Wavelets provide flexible and increasingly widely used mathematical tools for time scale analysis of images and signals; see Mallat (1989), Percival and Walden (2000) for general background texts and Bullmore et al. (2004) for a recent review of wavelets applied to analysis of functional MRI data.

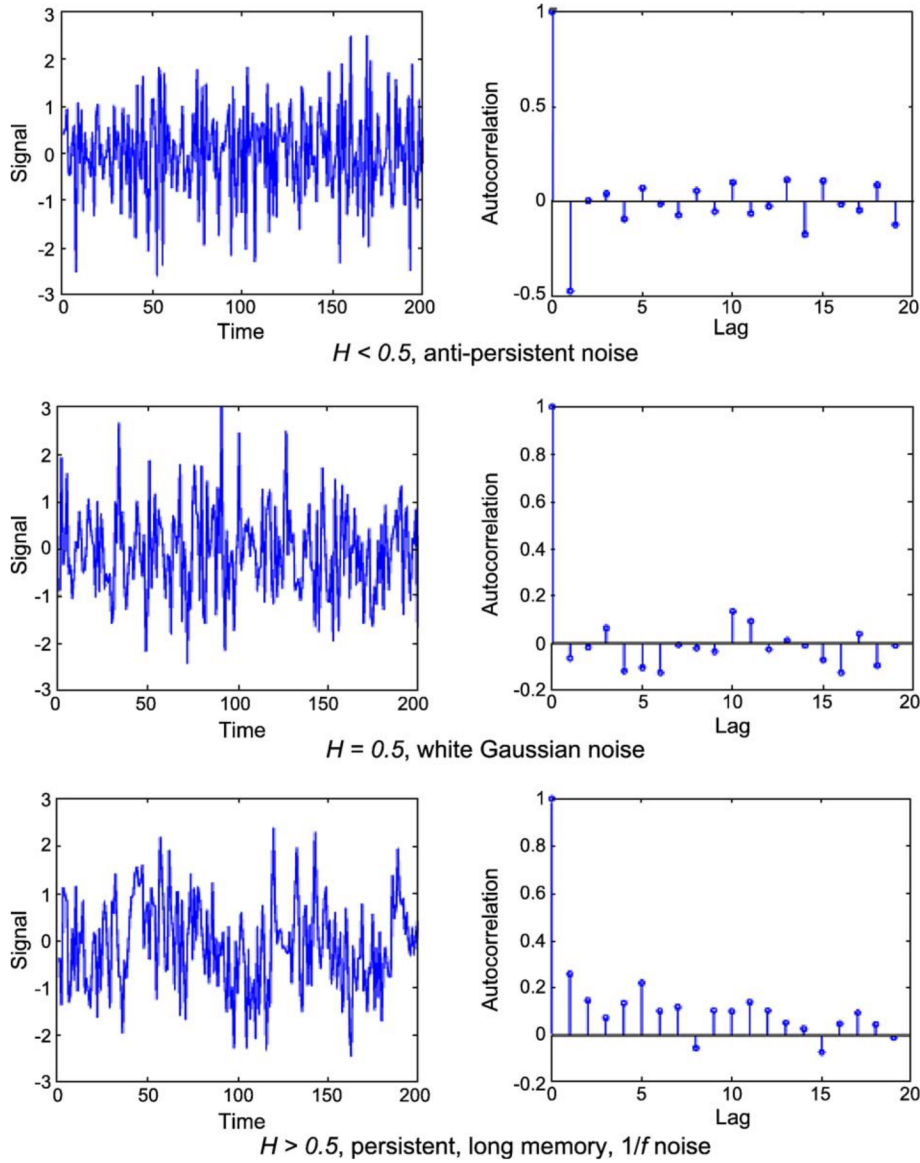


Fig. 1. Simulated fractional Gaussian noises with various values of the Hurst exponent (left column). From top to bottom, the values of  $H$  are 0.1 (antipersistent), 0.5 (white Gaussian noise) and 0.7 (persistent). The right column shows their estimated autocorrelation sequences. Slow autocorrelation decay with increasing lag can be observed for the first and the last time series, with negative autocorrelation at first lag for the Hurst exponent  $H = 0.1 < 0.5$  and positive dominance for the Hurst exponent  $H = 0.7 > 0.5$ . The corresponding spectral density functions are plotted in Fig. 2.

### The discrete wavelet transform

Like the Fourier transform, which decomposes a signal into a spectrum of frequencies, wavelet analysis decomposes a signal into a hierarchy of scales, ranging from the coarsest scale (defined by the approximation coefficients) to a series of finer scales (defined by the detail or wavelet coefficients). However, the important distinction between these two transforms is that wavelets are little waves, localised in space or time, unlike the sine and cosine waves of the Fourier basis, which extend infinitely. Thus, a wavelet transform is inherently more sensitive to nonstationary aspects of data and may provide a particularly natural basis for analysis of scale-invariant or fractal processes.

More precisely, the wavelet transform can be seen from two perspectives, which are unified by the formalism of multiresolution analysis developed by Mallat (1989). On the one hand, if the signal of interest is continuous and square integrable, the wavelet

transform can be understood to project it onto a basis, the elements of which are shifted and dilated versions of a scaling function and a wavelet function  $\psi$ ,

$$\varphi_{j,k}(x) := 2^{j/2} \varphi(2^j x - k), \quad \psi_{j,k}(x) := 2^{j/2} \psi(2^j x - k), \quad j, k \in \mathbb{Z}. \quad (8)$$

Wavelets can be distinguished by their smoothness (also called regularity) and by their number of vanishing moments. The number of vanishing moments of a wavelet  $\psi$  is the largest integer  $R$  which satisfies  $\int t^r \psi(t) dt = 0$ , for all  $r = 0, \dots, R-1$ .

On the other hand, if the signal is discrete, the wavelet transform is implemented by a two-channel filtering algorithm (Mallat, 1989), summarised in Fig. 3. At each step, the low-scale component from the preceding step is split in two, by passing it through a low-scale filter  $h$  and a high-scale filter  $g$ . Let us suppose that the length of the signal  $n$  is some power of

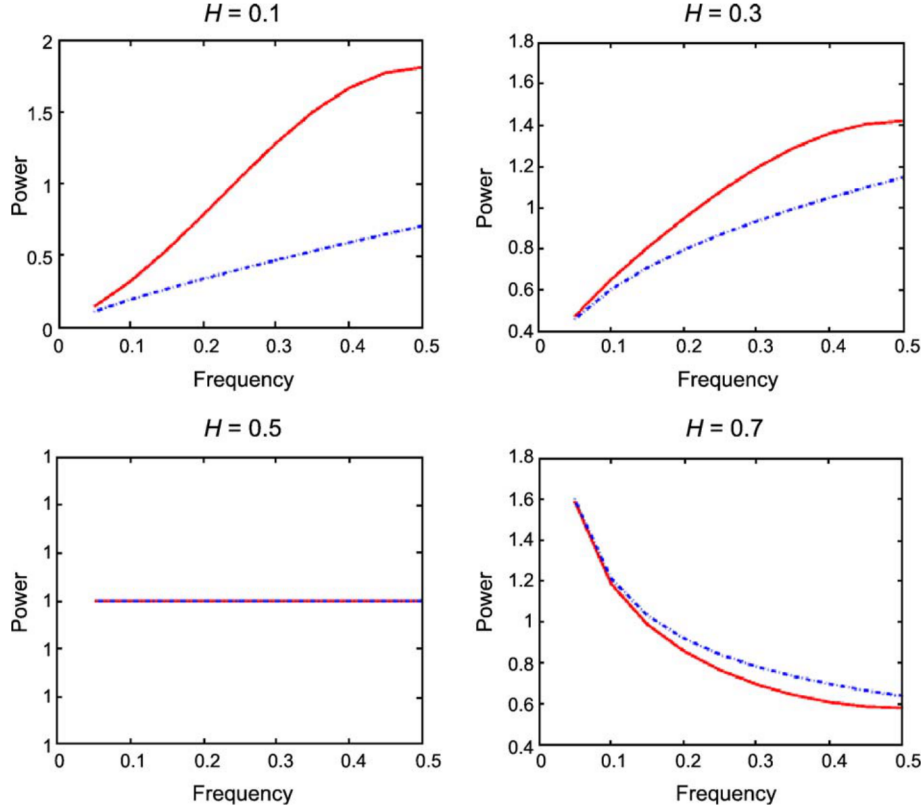


Fig. 2. Spectral density functions (SDFs) of fractional Gaussian noise with various values of the Hurst exponent  $H$ . The exact SDF (Eq. (3)) is shown as a red continuous line and can be compared to the approximation by truncated Taylor series expansion (Eq. (4)), which is shown as a blue dotted line. Fractional Gaussian noise with  $H > 0.5$  has greater power at low frequencies, whereas fGn with  $H < 0.5$  has greater power at high frequencies; fGn with  $H = 0.5$  has equal power at all frequencies. Note that the approximation to the exact SDF is only tenable in the neighbourhood of the origin and for  $H \geq 0.5$ . (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

2 and we repeat the preceding operation  $j_0$  times. The result is a set of approximation (or scaling) coefficients  $a_{-j_0,k}$ ,  $k = 1, \dots, n/2^{j_0}$  and a sequence of detail (or wavelet) coefficients  $d_{j,k}$ ,  $j = -j_0, \dots, -1$ ,  $k = 1, \dots, n/2^j$ . If the scaling function has finite support, the required number of operations for DWT of a discrete signal is  $O(n)$ , compared to  $O(n \log n)$  for the Fast Fourier Transform.

#### Wavelets as decorrelating filters

It was shown by several authors (Dijkerman and Mazumdar, 1994; Flandrin, 1992; Tewfik and Kim, 1992; Wornell, 1996) that the covariance of any two wavelet coefficients  $d_{j,k}$  and  $d_{j',k'}$  of a continuous  $1/f$  process decays rapidly to zero when the distance

$|k2^{-j} - k'2^{-j'}|$  increases to infinity. Moreover, assuming the wavelet filter is bandpass, i.e.,

$$F\psi(f) = \begin{cases} 1, & \frac{1}{2} \leq |f| \leq 1 \\ 0, & \text{otherwise,} \end{cases} \quad (9)$$

where  $F\psi$  denotes the Fourier transform of the wavelet  $\psi$ , it can be shown that wavelet coefficients from different scales are uncorrelated (Wornell, 1996). Due to this result and to empirical observation, it is widely accepted that the wavelet coefficients of a correlated  $1/f$  time series are, to a good approximation, uncorrelated. Wornell's estimator (Wornell and Oppenheim, 1992) and the wavelet-ML estimator (Fadili and Bullmore, 2002) of fGn parameters we describe hereafter are based on this assumption.

#### Estimators of fractional Gaussian noise parameters

There are many possible estimators of fractional Gaussian noise parameters  $H$  and  $\sigma^2$ ; see Beran (1994) and Taqqu et al. (1995) for a detailed treatment of the issues. Here we are concerned primarily with comparative evaluation of four estimators, one in the time domain and three in the wavelet domain; additionally, we will briefly discuss the maximum likelihood (ML) estimator in the time domain:

- Time
  1. Maximum likelihood estimator
  2. Whittle's estimator (Beran, 1994; Whittle, 1953)

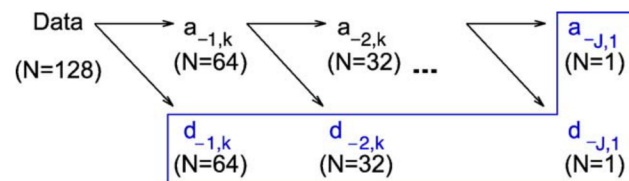


Fig. 3. Mallat's pyramid algorithm for the discrete wavelet transform. The data are iteratively decomposed into high- and low-scale components by a two-channel filter bank. The wavelet or detail coefficients are retained at each scale, together with the approximation or scaling coefficients at the coarsest scale.

- Wavelet
  1. Estimation by discrete variations of filtered fractional Brownian motion (fBm)
  2. Wornell’s estimator (Wornell, 1996; Wornell and Oppenheim, 1992)
  3. Wavelet-based maximum likelihood estimator (wavelet-ML)
    - with approximate SDF (Fadili and Bullmore, 2002)
    - with exact SDF

Each of these estimators is technically described in greater detail in Appendix A.

#### *Simulation of fractional Gaussian noise*

Several algorithms exist to simulate fGn (Beran, 1994) and these have been comparatively evaluated (Coeurjolly, 2000a). On this basis, we adopted the Davies–Harte algorithm (Davies and Harte, 1987), which is both exact and fast, to generate the fGn simulations used here. For each value of  $H = 0.1, 0.2, \dots, 0.9$ , we simulated 1000 realisations of fGn with 512 time points in each series; we set  $\sigma^2 = 1$  for all simulations.

#### *Functional MRI data acquisition*

Gradient-echo echo planar imaging (EPI) data depicting T2\*-weighted BOLD contrast were acquired at 1.5 and 3.0 T from participants lying quietly “at rest” in the scanners with their eyes closed. All participants gave informed consent in writing. The studies were approved by the Local Research Ethics Committee, Addenbrooke’s NHS Trust (3.0 T data set) and the Research (Ethics) Committee, South London and Maudsley NHS Trust (1.5 T data sets).

#### *3.0 T data set: younger volunteer*

A single healthy male volunteer (age = 45 years) was scanned using a Bruker Medspec 3.0 T system at the Wolfson Brain Imaging Centre, Cambridge, with the following parameters: TR = 1100 ms, time to echo (TE) = 27.5 ms, slice thickness = 4 mm, interslice gap = 1 mm, in-plane resolution = 3.1 mm. Five hundred eighteen image volumes, each comprising 21 slices collectively providing whole brain coverage, were acquired over the course of 9 min and 30 s; the first six images were discarded because of T1 equilibration effects.

#### *1.5 T data sets: case-control study of Alzheimer’s disease (AD)*

Nine patients who met National Institute of Neurological and Communicative Disorders criteria and Stroke–Alzheimer’s Disease and Related Disorders Association (McKhann et al., 1984) criteria for mild probable Alzheimer’s Disease (6 female, 3 male; mean age = 75.9 years, range = 69–85 years) and 12 healthy, elderly controls (7 female, 5 male; mean age = 77.3 years, range = 69–82 years) were recruited into the study. The patients were diagnosed at an early stage and were not severely impaired cognitively: mean mini-mental state examination (MMSE) score was 26.1 (range 24–29) in the patients compared to 29.1 (range 27–30) in the healthy volunteer group. Subjects were right-handed apart from two patients and three controls. All subjects were screened for concomitant serious medical diagnoses (such as Parkinson’s disease, diabetes, stroke, and epilepsy) and previous psychiatric

history. Seven patients were receiving treatment for cognitive impairment (6 were being treated with donepezil and 1 with rivastigmine). Twelve patients and 12 controls were included into the study but data were missing or incomplete for some of them.

Each participant was scanned using a General Electric (GE) Signa LX Horizon 1.5-T system at the Maudsley Hospital, London, with the following parameters: TR = 2000 ms, TE = 40 ms, slice thickness = 7 mm, interslice gap = 0.7 mm, in-plane resolution = 3.75 mm. One hundred fifty-four image volumes, each comprising 16 oblique axial slices collectively providing whole brain coverage, were acquired over the course of 5 min and 8 s; the first 4 volumes were discarded because of T1 equilibration effects, leaving 150 images available for analysis.

All images were mapped into standard space (Talairach and Tournoux, 1988) using affine transforms to minimise sum of square differences between the mean EPI volume for each individual and a contrast-matched template image. The resulting transformation matrices were identically applied to  $H$  and  $\sigma$  parameter maps individually estimated in native space to co-register them in standard space.

## **Results**

### *Comparative evaluation of fGn estimators*

Using simulated fGn with known parameters, we comparatively evaluated the performance of Whittle’s estimator in the time domain and three estimators in the wavelet domain: the wavelet-ML estimator incorporating the exact expression for the SDF; the estimator based on discrete variations of a filtered fractional Brownian motion; and Wornell’s algorithm. The wavelet-ML algorithm was also used to estimate the parameter vector  $\beta$  of an arbitrary design matrix  $X$  fitted to simulated fGn and real fMRI data acquired at rest, i.e., under the null hypothesis that  $\beta = 0$ .

### *Bias and efficiency*

The results are summarised in Figs. 4 and 5, which show boxplots for estimated values of  $H$  and  $\sigma$  for 1000 simulated fGn series at each value of  $H$ .

All methods are reasonably good estimators of  $H$ . Whittle’s estimator and the fBm-based estimator are the least biased; Wornell’s estimator and the wavelet-ML algorithm are both slightly biased in estimating extreme values of  $H$ . The most efficient estimators are the wavelet-ML algorithm and Whittle’s estimator; both Wornell’s estimator and the fBm-based algorithm are marginally less efficient.

However, Whittle’s algorithm incorporates the usual variance estimator in the time domain, which is known to be biased towards underestimation of  $\sigma$  in the context of long memory autocorrelation (Eq. (12)), and our results confirm this prediction empirically. Wavelet-ML and fBm-based algorithms are less biased estimators of  $\sigma$  and, of the two, wavelet-ML is more efficient. Wornell’s algorithm performs poorly as an estimator of  $\sigma$ .

### *Type one error calibration curves*

Using the wavelet-ML algorithm, we fitted an arbitrary experimental design (a periodic contrast between epochs each comprising 10 consecutive images) to simulated fGn and the 3.0 T resting fMRI data set. On this basis, we estimated  $\hat{\beta}$  and its standard error, referred the ratio  $t = \hat{\beta}/SE(\hat{\beta})$  to critical values of

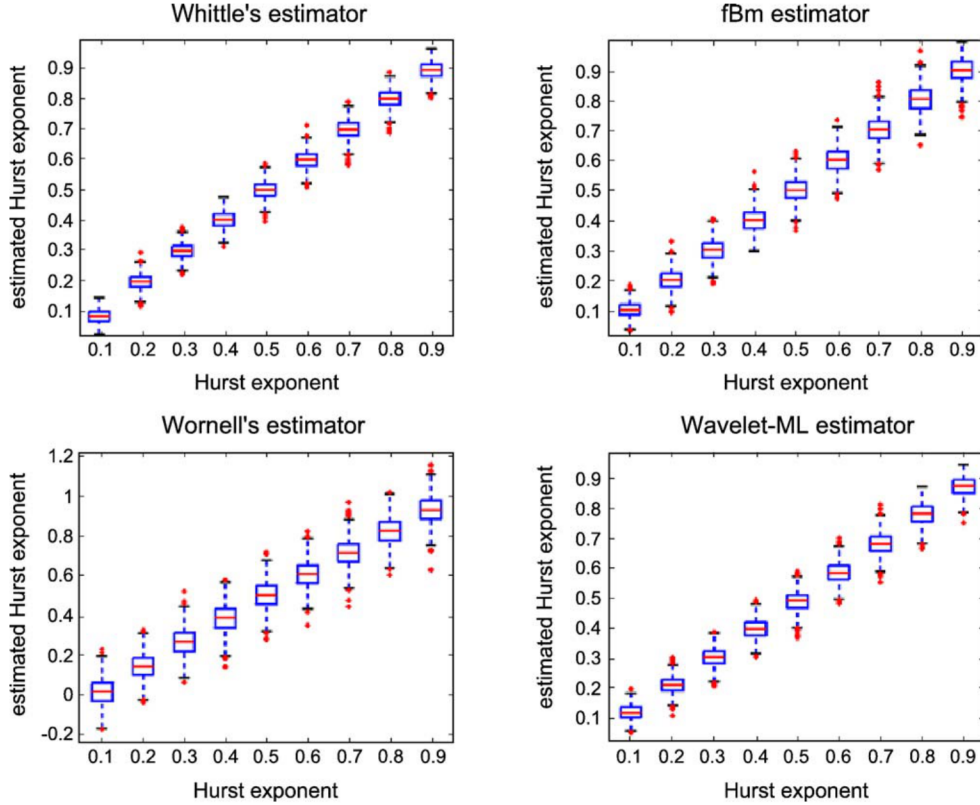


Fig. 4. Comparative evaluation of bias and efficiency of four possible estimators of the Hurst exponent  $H$  of fractional Gaussian noise. Simulated realisations of fGn ( $n = 512$ ) were generated using the Davies–Harte algorithm with  $H$  ranging in increments of 0.1 from  $H = 0.1$  to  $H = 0.9$ ; 1000 realisations were simulated for each value of  $H$ . The boxplots in each panel summarise the median and interquartile range of the estimated values of  $H$  for the classical time domain estimator and for three alternative estimators in the wavelet domain. Variance of all simulations was set  $\sigma^2 = 1$ .

the  $t$  distribution, and could directly compare the nominal threshold for statistical significance, or size of test  $0 < \alpha < 1$ , to the empirical size of test estimated by the proportion of hypothesis tests that were (false) positives.

As shown in Fig. 6, wavelet-ML estimation of  $t$  provided exact or slightly conservative type 1 error control for tests on simulated fGn over all values of  $H$ , and exact type 1 error control for tests on the resting fMRI data set. Note that these results provide an alternative validation of the estimation of  $\sigma^2$  by the wavelet-ML algorithm because if  $\sigma^2$  was, for example, substantially underestimated this would lead to an underestimate of  $SE(\beta)$  with consequent inflation of the  $t$  statistic and an invalid hypothesis test with an empirical significance threshold more lenient (or less conservative) than the nominal size of test.

#### Mapping fGn parameters in “resting” state fMRI data

As a result of these evaluation exercises, we considered the wavelet-ML algorithm to represent the best available solution for efficient estimation of both fGn parameters  $H$  and  $\sigma^2$  (as well as  $\beta$ ); though see Discussion for a fuller treatment of the relative merits of all four estimators. To explore the noise properties of the resting fMRI data sets in greater detail, we will therefore report primarily results estimated by the wavelet-ML algorithm. Additionally, we will report some results obtained using the log periodogram (Eq. (6)) to estimate the spectral exponent  $\gamma$  in each fMRI time series, and some comparative results of using a first-order autoregression AR(1) model to characterize serial dependencies in the data.

#### Effects of head movement and its correction

Let us first consider the effects of head movement on fMRI noise. It is well known that head movement is almost inevitable in fMRI studies and even translations  $< 1$  mm or rotations of a few degrees can have profound effects on fMRI time series; this issue may be particularly problematic in studies of patients rather than healthy young volunteers (Bullmore et al., 1999a). Movement correction preprocessing algorithms generically consist of two steps: first, the series of image volumes is realigned, e.g., by rigid body or affine transformation, so that each volume is geometrically approximated to the first image volume in the series; second, the realigned time series may be regressed on the time series of three-dimensional rotations and translations of the head’s center of gravity estimated in the course of realignment (Friston et al., 1996), or the time series of local voxel displacements trigonometrically derived from the estimated movements of the head’s center of gravity (Bullmore et al., 1999a).

If we estimate  $\gamma$  at each voxel of one of the 1.5-T data sets, acquired from a healthy elderly volunteer, after movement correction by realignment (but not regression), we can see that there are many large negative values of the spectral exponent, i.e.,  $\gamma < -1$ , on the edge of the cortex, where we might expect the residual effects of head movement to be most salient; see Fig. 7. Similarly, the frequency distribution of  $\gamma$ , estimated over all voxels in this slice of the image, confirms that values of  $\gamma < -1$  are quite common; see Fig. 7. This is technically important because a fractional Gaussian noise model assumes  $-1 < \gamma < 1$ ; values of  $\gamma < -1$  are consistent with fractional Brownian motion, rather than fGn, as an appropriate noise

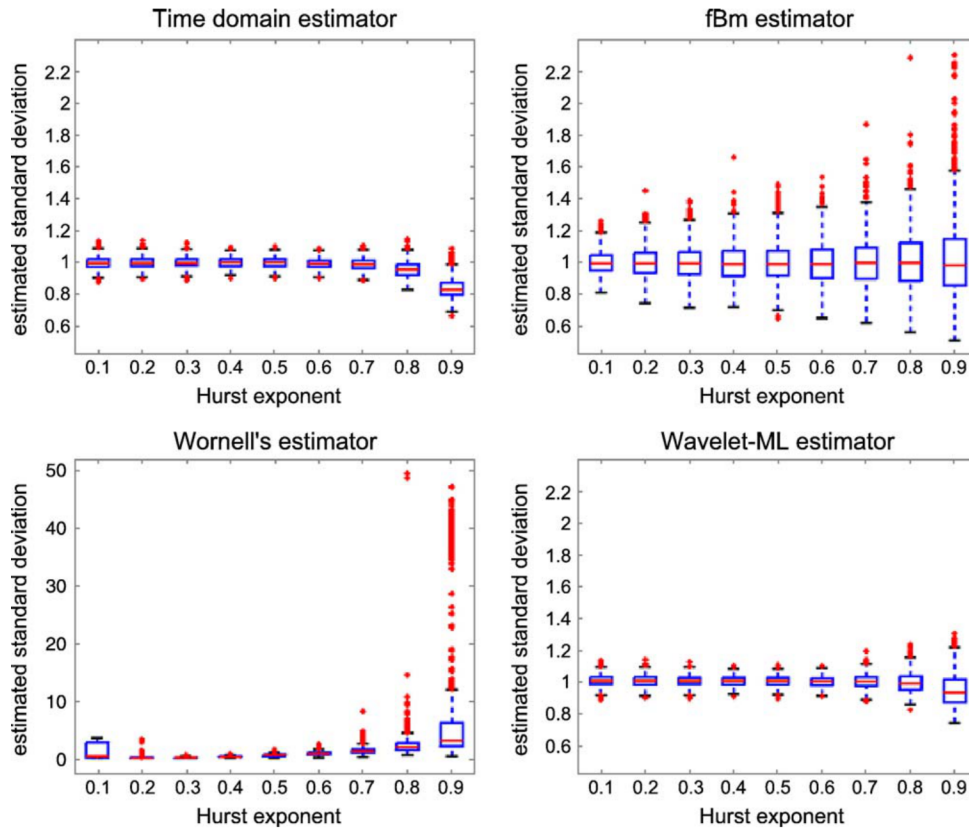


Fig. 5. Comparative evaluation of bias and efficiency of four possible estimators of the variance  $\sigma^2$  of fGn. Simulated realisations of fractional Gaussian noise ( $n = 512$ ) were generated using the Davies–Harte algorithm with  $H$  ranging in increments of 0.1 from  $H = 0.1$  to  $H = 0.9$ ; 1000 realisations were simulated for each value of  $H$ . The boxplots in each panel summarise the median and interquartile range of the estimated values of  $\sigma^2$  for Whittle’s estimator in the time domain and for three alternative estimators in the wavelet domain. Variance of all simulations was set  $\sigma^2 = 1$ .

model. However, the second stage of movement correction applied to these data by regressing each realigned time series on the estimated voxel displacements (Bullmore et al., 1999a) substantially changes the anatomical and frequency distribution of  $\gamma$ ; see Fig. 7. Many of the large negative spectral exponents on the cortical rim are substantially reduced in absolute value and the frequency distribu-

tion of  $\gamma$  is now centered on zero and falls almost entirely within the range associated with fGn.

In short, head movement and its correction can impact strongly on fMRI noise parameters, and a two-stage correction algorithm, involving regression as well as realignment, seems advisable if we wish to model the noise as fGn. We therefore adopted this

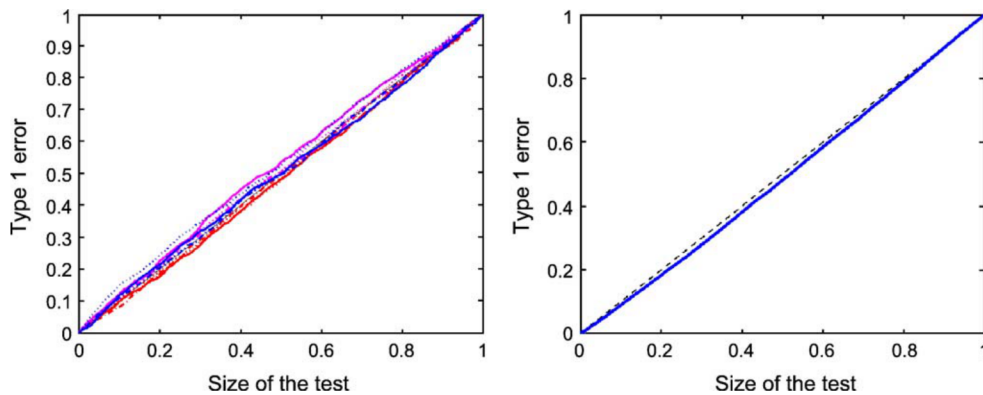


Fig. 6. Type I error calibration curves for the wavelet-maximum likelihood estimator of the GLM model parameter vector  $\beta$  in the context of simulated fGn (left panel) and biological fMRI noise (right panel). For the simulated data,  $V = 1000$  realisations of fGn with  $n = 512$ ,  $\sigma^2 = 1$  and  $0.1 \leq H \leq 0.9$  were generated using the Davies–Harte algorithm. A blocked periodic design matrix was fitted to each realisation and the test statistic  $\beta/\text{SE}(\beta)$  was tested against the  $t$  distribution for various sizes of test  $0 < \alpha < 1$ . The observed frequency of positive tests ( $y$ -axis) was almost exactly identical to the nominal size of test ( $\alpha$ ,  $x$ -axis), indicating that the test based on wavelet-ML estimation is exact and has nominal type I error control. Each curve corresponds to a different value of  $H$ . For the biological data, the same design matrix was fit to all voxels of a null fMRI data set, acquired at 3.0 T with the subject “at rest”; the observed number of positive tests was slightly below the number expected under the null hypothesis, indicating that the test is valid and slightly conservative.

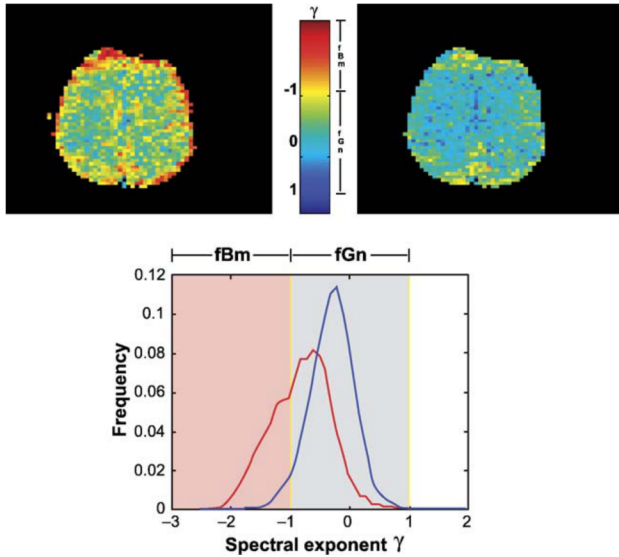


Fig. 7. Effects of head movement on spectral exponent of fMRI noise. Top panel: Maps of spectral exponent  $\gamma$  estimated by the log periodogram in “resting” state fMRI data after one-stage (left) and two-stage (right) correction for head movement. One-stage correction involved geometric realignment by sinc interpolation of all images volumes relative to the first volume in the series; two-stage correction involved realignment followed by regression of the realigned time series on the estimated time series of voxel displacements. As indicated by the color bar, for fractional Gaussian noise (fGn)  $-1 < \gamma = 1 - 2H < 1$ ; whereas for fractional Brownian motion (fBm)  $-1 > \gamma$ . These data demonstrate that uncorrected head movement can exaggerate the persistence of fMRI noise, causing the distribution of  $\gamma$  to include frequent values  $\gamma < -1$ , which is indicative of fractional Brownian motion. Bottom panel: The frequency distributions of spectral exponent  $\gamma$  in the “resting” state fMRI data. The red line indicates the distribution of  $\gamma$  following one-stage movement correction by realignment alone; the blue line indicates the distribution of  $\gamma$  following two-stage correction by realignment and regression. The more rigorous movement correction results in a distribution of  $\gamma$  that is centered approximately on zero and largely falls within the expected range for fractional Gaussian noise.

movement correction strategy consistently in all subsequent studies of “resting” fMRI data sets.

#### Anatomical distribution of fGn and AR(1) parameters

Fractional Gaussian noise parameters were estimated by wavelet-ML at each voxel of the “resting” 3.0-T data set; a single slice sampled from this single subject’s pair of fGn parameter maps is shown in Fig. 8. It is clear in this example that the anatomical distribution of  $H$  is not random. Larger values of the Hurst exponent  $H > 0.5$  are symmetrically concentrated in lateral and medial cortical regions whereas values of  $H \leq 0.5$  predominate in the central white matter and in the vicinity of the lateral ventricles. It is also notable that the anatomical distribution of  $\sigma$  demonstrates particularly high variance in the ventricles and locally on the cortical rim.

This map of the Hurst exponent is clearly comparable to a map of the first-order autoregression AR(1) coefficient estimated at each voxel of the same data; see Fig. 8. As expected theoretically, voxels with  $H > 0.5$  tend to have large positive AR(1) coefficients, whereas voxels with  $H < 0.5$  tend to have negative AR(1) coefficients. A similar, cortically concentrated anatomical distribution of the AR(1) coefficient (Worsley et al., 2002), the Hurst exponent and the Hölder exponent (Shimizu et al., 2004) has

previously been reported in single slices of resting fMRI data. However, although  $H$  is correlated with AR(1), it is evident from the scatterplot in Fig. 8 that the relationship between these two metrics is not simply linear. In particular, it seems that the AR(1) coefficient provides a greater dynamic range to characterize moderately autocorrelated voxels, whereas  $H$  provides a greater dynamic range to characterize the most strongly autocorrelated voxels. Both metrics provide equivalent characterisations of white noise sampled from voxels in the image outside the brain, which tended to have  $AR(1) \sim 0$  and  $H \sim 0.5$ . The variance map estimated in the context of a first-order autoregressive model was very similar to the map of variance estimated as an fGn parameter (data not shown). However, we again noted that the residual fMRI time series defined by subtraction of an estimated AR(1) process often demonstrated significant temporal autocorrelation. For example, approximately 95% of residual time series demonstrated significant autocorrelation by the Box–Pierce test for whiteness after AR(1) modeling (Fig. 8), and 85% of residual series were significantly autocorrelated after AR(3) modeling.

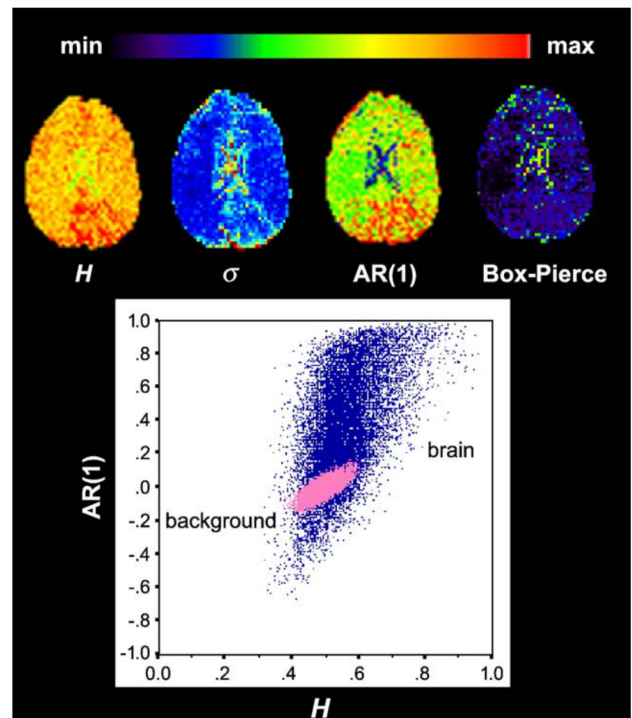


Fig. 8. Fractional Gaussian noise parameters, Hurst exponent  $H$  and  $\sigma$ , are mapped for a single section of “resting” FMRI data (top row, left). Also shown is a comparable map of the first order autoregression coefficient AR(1) and a map of the Box–Pierce statistic (a test for white noise) estimated after subtracting an AR(1) process from each time series (top row, right). The key points to note are that voxels with large (positive) Hurst exponents (or AR(1) coefficients) tend to be concentrated symmetrically in cortical regions, whereas voxels with small values of  $H$  or negative values of AR(1) tend to be concentrated around ventricles and are associated with high variance. The Box–Pierce statistic map indicates that a simple AR(1) model is generally inadequate to account for all correlation structure in these data. Bottom row: Scatterplot of  $H$  versus AR(1) for all voxels shows that the two metrics are correlated ( $r = 0.54$ ) but  $H$  provides greater dynamic range for characterization of strongly autocorrelated time series. Also shown in this plot is the swarm of (pink) points corresponding to estimates of  $H$  and AR(1) in voxels located outside the brain, which show that the background instrumental noise in fMRI is approximately white by both metrics.

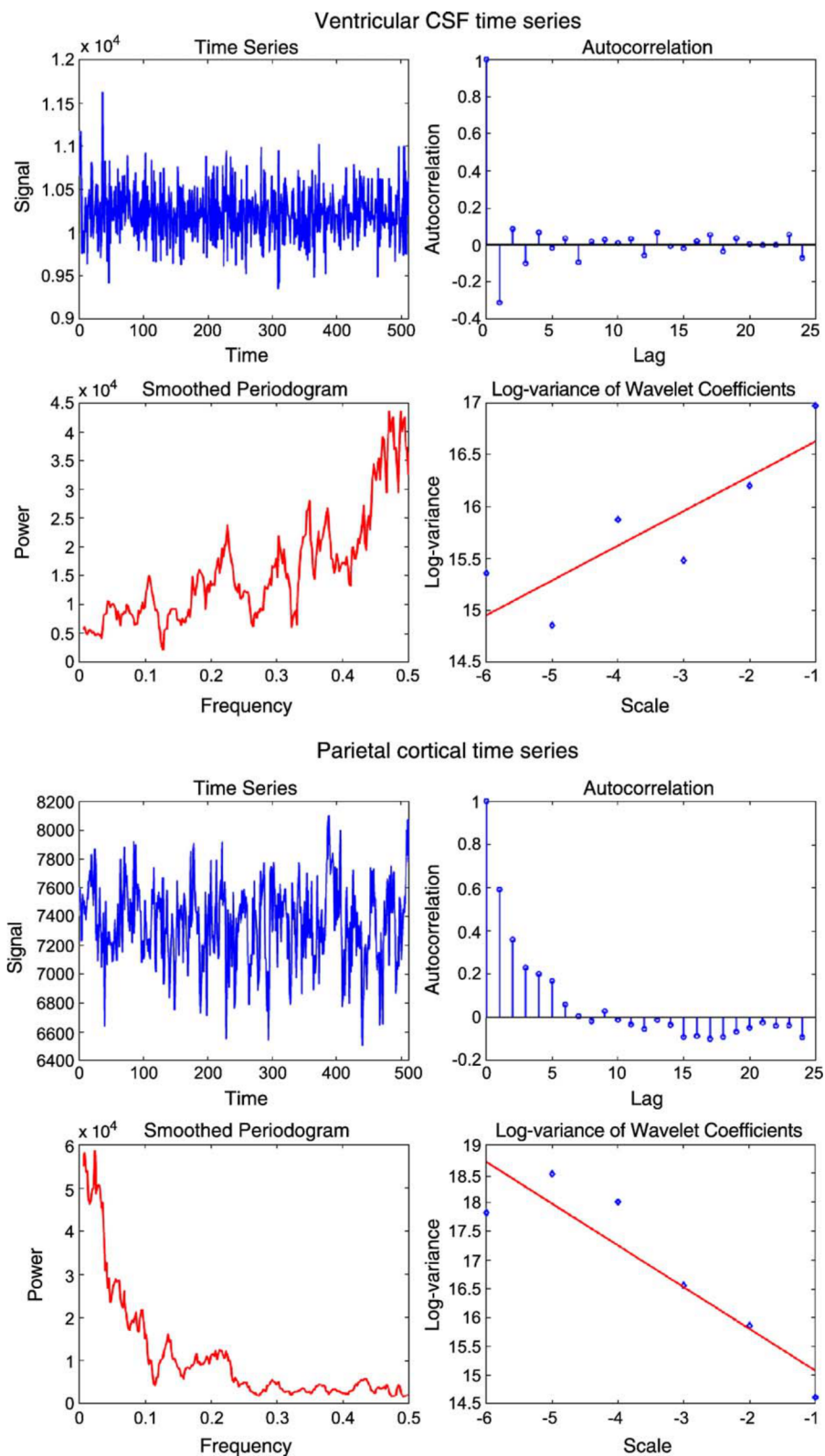


Fig. 9. Examples of single-voxel fMRI time series exhibiting antipersistence (top panel) and persistence (bottom panel). Each panel comprises four diagnostic plots, from left to right, top to bottom: the time series, its autocovariance sequence, smoothed periodogram, and log-variance of the wavelet coefficients at each scale. The slope of the straight line fitted to the log wavelet variance plots is an indirect estimator of  $H$ , leading to the values  $\hat{H} = 0.33$  and  $\hat{H} = 0.86$ , compared to  $\hat{H} = 0.31$  and  $\hat{H} = 0.92$  by wavelet-ML, in the case of the antipersistent and persistent time series, respectively. The persistent series was sampled from a region of medial posterior parietal cortex and the antipersistent series was sampled from the left lateral ventricle.

To illustrate in greater detail the wide range of fMRI noises that are subsumed by the fGn model, we extracted time series from two voxels with different fGn parameters: one time series was extracted from a voxel in medial posterior parietal cortex that had  $H = 0.92$  and high variance; another time series was extracted from a voxel in the lateral ventricles that had  $H = 0.31$  and high variance. As shown in Fig. 9, and as theoretically expected, the cortical time series was persistent in time and had positive autocorrelation coefficients for lags  $< 5$ . The spectral density function of this series demonstrated peak power at low frequencies and a plot of the logarithm of the wavelet coefficient variances at each scale versus scale showed an approximately linear decay in log variance as a function of increasingly detailed scale of the DWT. Collectively, these results can be summarised by saying that this cortical time series with  $H > 0.5$  demonstrated long memory or  $1/f$  properties. The ventricular time series (Fig. 9) demonstrated a very different behavior characterised by anti-persistence in time, a negative autocorrelation coefficient at lag = 1, a spectral density function

dominated by high-frequency power, and an approximately linear increase in the log of the wavelet coefficient variances as a function of increasingly detailed scale of the DWT.

To provide a more generally representative map of fGn parameters, the individual subject maps of  $H$  and  $\sigma$  estimated in the nine healthy elderly volunteers were co-registered in Talairach space (Brammer et al., 1997) and averaged at each voxel to produce a pair of group mean fGn parameter maps; see Fig. 10. These group maps confirm that the anatomical distribution of  $H$  is not random: Cortical regions are generally associated with larger  $H$  and there is some evidence for right-left symmetry in the cortical distribution of long memory noise.

*A case-control study of long memory noise: patients with Alzheimer's disease compared to healthy volunteers*

Group mean maps of  $H$  and  $\sigma^2$  were constructed in exactly the same way using the data acquired from patients with early AD; see

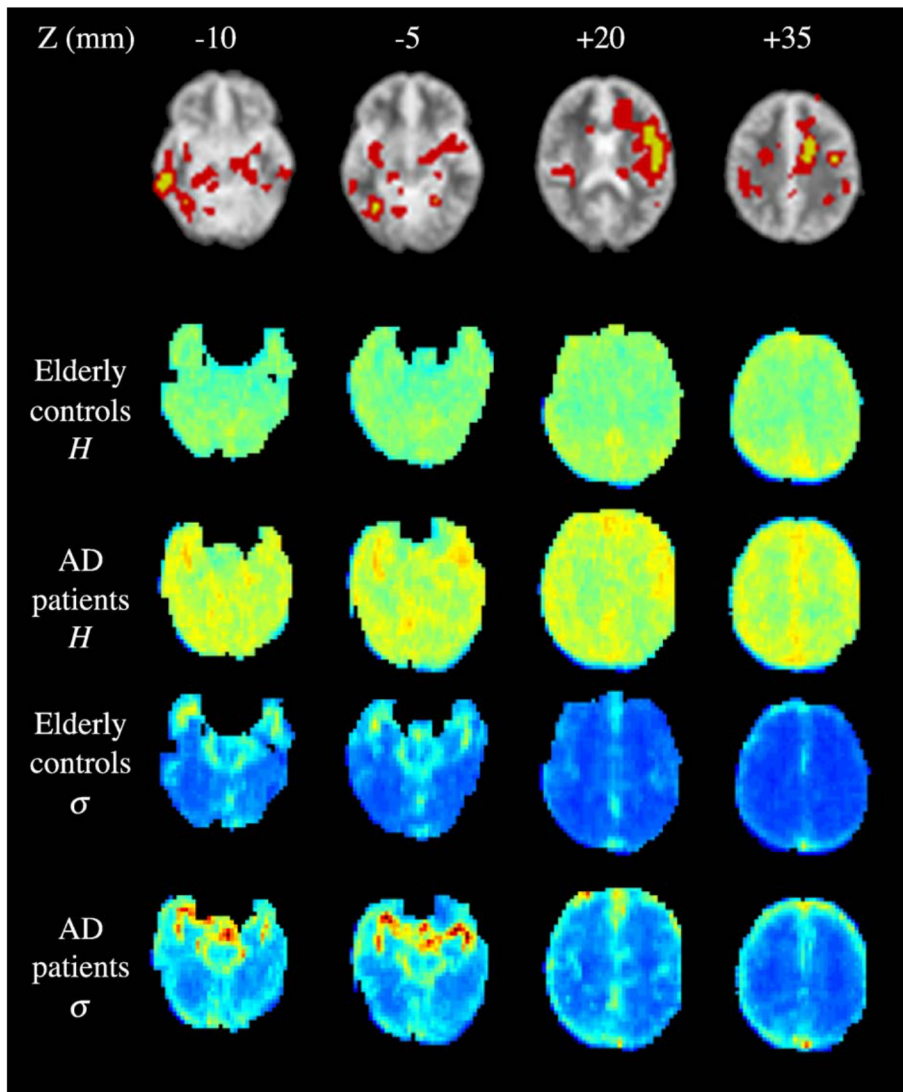


Fig. 10. Effects of Alzheimer's disease (AD) on Hurst exponent estimated by wavelet-ML in "resting" state fMRI data. Group average maps of  $H$  are shown for AD group ( $N = 9$ ; third row) and age-matched volunteers ( $N = 12$ ; second row). The top row localises significant between-group differences in  $H$  (greater in AD; cluster-wise  $\alpha < 0.007$  by permutation test) to medial and lateral temporal cortex, left somatosensorimotor cortex and dorsal cingulate. All maps are orientated with the right brain represented by the left side of each section; the  $z$ -coordinates (mm) of each section in relation to the intercommissural line are indicated.

Fig. 10. The anatomical distributions of  $H$  and  $\sigma^2$  can be simply compared between groups by inspection of the two sets of group mean parameter maps, which suggest that the patients with AD may have higher values of  $H$  in various cortical regions. We used a permutation test to make this comparison more formally (Bullmore et al., 1999b; Suckling and Bullmore, 2004). Briefly, the observed between-group difference in  $H$  at each voxel was compared to critical values of its permutation distribution sampled by repeatedly, randomly reassigning each subject to two groups (regardless of diagnostic status) and estimating the voxel-wise differences between groups after each permutation. Voxels in the observed group difference map, which exceeded the critical value for a two-tailed test of size  $\alpha = 0.05$ , comprised a set of suprathreshold voxel clusters. The “mass” of each observed voxel cluster, i.e., the sum of the suprathreshold voxel statistics it comprised, was then tested for significance by comparison to a critical value of the permutation distribution for cluster mass sampled by applying the same voxel threshold to the between-group difference maps derived from the permuted data. The significance threshold for this cluster-level test was set at  $\alpha < 0.005$ , meaning that we expect less than one positive cluster over the whole map under the null hypothesis. The same methods were used to test the significance of between-group differences in  $\sigma^2$ , estimated by the wavelet-ML algorithm at each voxel. To further explore the relationship between the Hurst

exponent and the AR(1) coefficient, which in principle provide equally parsimonious (single parameter) descriptions of fMRI noise, we also compared the AD patients to the healthy volunteer group in terms of the AR(1) coefficient, using an identical permutation test and cluster-wise probability threshold to define regions of significant between-group difference.

As shown in Fig. 11, we found significant between-group differences for the Hurst exponent in several cortical regions including bilateral inferior, middle and superior temporal gyri, left precentral and postcentral gyri, left dorsal cingulate and medial premotor cortex, insula and bilateral medial temporal structures including hippocampus, amygdala and parahippocampal gyrus. In all these regions  $H$  was significantly greater in the patients with AD than in the comparison subjects. These differences are also represented more extensively in Fig. 11.

Comparable maps of between-group difference in  $\sigma^2$  show a different anatomical distribution, with significantly greater variance in the patient group located in the vicinity of the ventricles and sulcal CSF spaces at the base of the brain, on the cortical rim and in proximity to the interhemispheric fissure.

As shown in Fig. 11, many of the regions identified as significantly different between groups by testing the Hurst exponent also showed a significant difference in terms of the AR(1) coefficient. However, the map of differences defined by the AR(1)

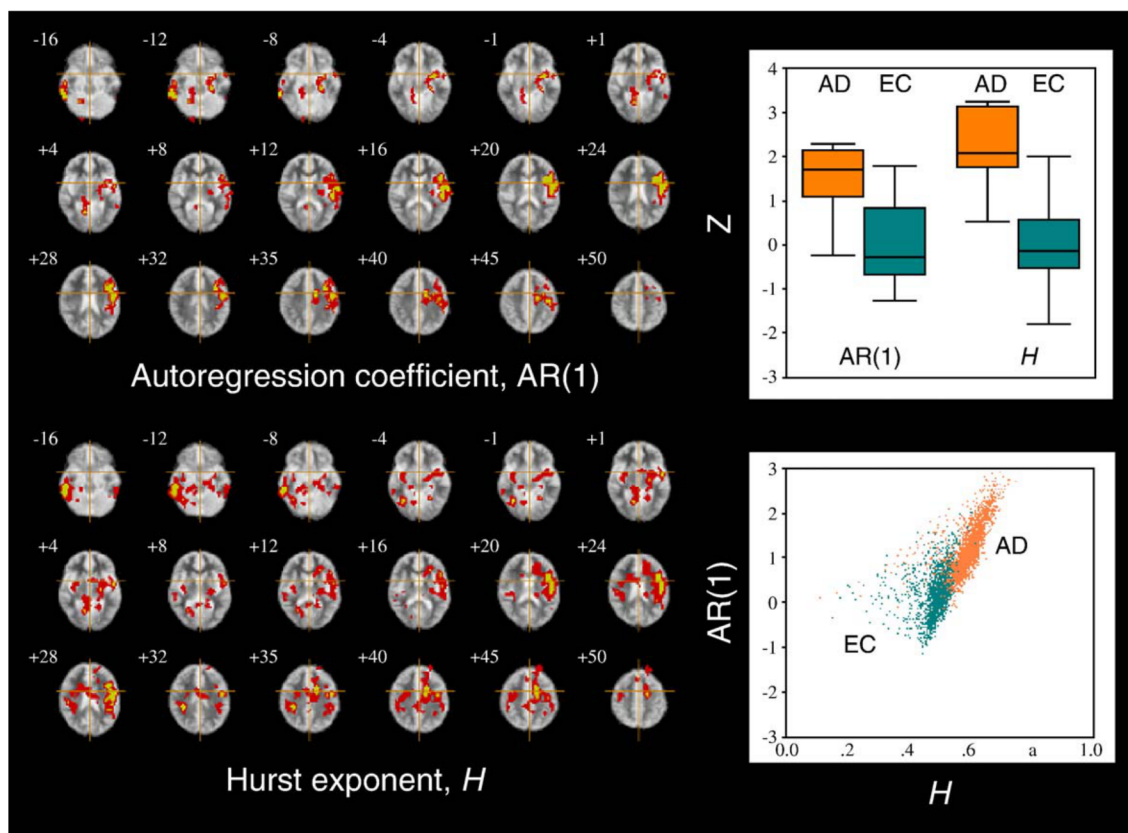


Fig. 11. Whole brain maps showing location of significant differences between patients with Alzheimer’s disease (AD) and an elderly control group of age-matched volunteers (EC) in terms of the Hurst exponent  $H$  and the first-order autoregression coefficient AR(1). Note that the anatomical pattern of case-control differences is comparable between metrics but the set of regions identified by the AR(1) coefficient is essentially a subset of the regions identified by the Hurst exponent; orientation of maps and thresholds for significance are as in Fig. 10. Greater sensitivity of the Hurst exponent to detect emergence of long memory dynamics in AD is corroborated by the box-plots (bottom left) of standardised AR(1) and  $H$  (averaged over all significantly different voxels within each subject), which show smaller within-group variability and greater between-group difference in terms of  $H$ . Superior sensitivity of the Hurst exponent is also demonstrated by the scatterplot (bottom right) of voxel values of AR(1) versus  $H$ , color-coded according to group, which shows clearer separation of the groups in terms of  $H$ .

coefficient was essentially a subset of the between-group differences mapped by testing the Hurst exponent, implying greater sensitivity of the Hurst exponent to detect pathological changes in these data. This impression was substantiated by box-plots of AR(1) and  $H$  (standardised for each subject over the set of voxels identified as abnormal by the Hurst exponent test), which show between-group differences in both parameters but relatively reduced normal variability and greater between-group difference in terms of  $H$ . This differential sensitivity between the two metrics can be quantified by comparing the results of two unpaired  $t$  tests for between-group difference in terms of the  $H$  and AR(1) data summarised in the box-plots: for  $H$ ,  $t_{19} = 4.66$ ;  $P < 0.001$ ; whereas for AR(1),  $t_{19} = 3.18$ ;  $P = 0.005$ . The clearer separation between groups in terms of the Hurst exponent is evident also in a simple scatterplot of  $H$  and AR(1) estimated in all voxels defined by the Hurst exponent test as significantly different between groups (Fig. 11).

## Discussion

The problem of fMRI noise is not new and there have been several previous efforts to address it statistically, many of them based on linear time-invariant, autoregressive and/or moving average models (Bullmore et al., 1996; Dale, 1999; Locascio et al., 1997; Purdon and Weisskoff, 1999; Worsley et al., 2002). Here we have introduced fractional Gaussian noise as an alternative modelling strategy. We have shown empirically that the spectral exponents of movement-corrected fMRI data acquired at rest are mainly distributed in the range  $-1 < \gamma < 1$ , which is compatible with a fractional Gaussian noise model that economically encompasses several different forms of noise naturally arising in fMRI data. The most attractive feature of fGn as a candidate model for fMRI noise, we suggest, is its parsimony. Only two parameters,  $H$  and  $\sigma^2$ , are required to describe completely both persistent, long memory,  $1/f$  processes with  $H > 0.5$  and antipersistent, negatively autocorrelated, predominantly high-frequency processes with  $H < 0.5$ , as well as classical Gaussian white noise with  $H = 0.5$ . An equivalently parsimonious AR model would necessarily be first order and it has previously been shown that AR(1) models may not always be sufficient to provide an adequate account of long memory and other complex forms of noise that occur particularly in high-field fMRI (Bullmore et al., 2001). We have replicated this observation in these data, showing that a large proportion of voxels remain autocorrelated after AR(1) or AR(3) prewhitening, which would invalidate a test of  $\beta/SE$  ( $\beta$ ) estimated by autoregressive prewhitening in these data. Nonetheless, we have also directly compared the AR(1) coefficient to the Hurst exponent in terms of sensitivity to detect pathological differences in resting state fMRI data.

### Neurobiological interpretation of the Hurst exponent

One of the obvious benefits of parsimony in the context of functional MRI is that it simplifies the problems of understanding simultaneously the noise properties of tens of thousands of time series. For example, we can provide a complete account of the noise properties of the resting brain in the form of two parameter maps ( $H$  and  $\sigma$ ), which can then be spatially transformed, averaged, and compared between groups using techniques already familiar in the multiple univariate analysis of regression model parameters estimated at each voxel. In this way, we have shown that the anatomical distribution of the Hurst exponent is not

random: larger values of  $H$ , indicative of  $1/f$  processes, are concentrated symmetrically in lateral and medial cortex; whereas smaller values, indicative of antipersistence, are concentrated in ventricles and other CSF spaces. We propose that the predominance of long memory processes in cortical regions is compatible with recent human and primate electrophysiological studies demonstrating the existence of long memory oscillations in spontaneous neuronal firing rates and local field potentials (Leopold et al., 2003; Linkenkaer-Hansen et al., 2001) and suggesting that these long-term trends in neuronal activity should be indirectly measurable by associated hemodynamic changes on the time scale of fMRI. In other words, it seems likely that the phenomenon of low-frequency fluctuations in cortical BOLD signals represented, for example, by  $H$  maps may be indicative of synchronised, long memory neuronal oscillations rather than merely noise in a neurophysiological sense.

### Pathophysiological interpretation of the Hurst exponent

Some further evidence that the Hurst exponent of fMRI noise may be substantively informative is provided by our demonstration that patients with probable early Alzheimer's disease have significantly greater values of  $H$  in regions of medial and lateral temporal cortex, dorsal cingulate and premotor cortex, and left pre- and postcentral gyrus. These regions are among those that have previously been implicated in the early stages of AD (Thompson et al., 2003) and it seems plausible that changes in persistence of resting state fMRI time series reflect neurodegenerative changes in the long memory dynamics of neuronal systems in these structures. Electrophysiological methods have previously provided the only window on abnormal neurophysiological dynamics in Alzheimer's disease but there is considerable evidence from this literature in support of our observation of abnormal long memory or  $1/f$  dynamics in AD. The EEG in AD is characterised by a shift of the power spectrum to lower frequencies and a decrease in coherence of fast rhythms (Jeong, 2004). It is perhaps particularly relevant that nonlinear dynamical systems modeling of EEG and MEG data in AD has repeatedly shown that the disease is associated with reductions in the first Lyapunov exponent  $L_1$  and the fractal (correlation) dimension  $D_2$  (Stam et al., 1995; van Capellen van Walsum et al., 2003). These changes indicate that the electrical activity of the brain becomes less dynamically complex, or more predictable, as a consequence of AD, which is entirely consistent with the emergence of more persistent dynamics that we have described in terms of the Hurst exponent. More exactly, the fractal (Hausdorff) dimension  $D_0$  is simply related to the Hurst exponent by the relation  $D_0 = T + 1 - H$ , where  $T = 1$  is the topological dimension of a time series (Schroeder, 1991); so we can see that increasing  $H$  means decreasing  $D_0$ . Since  $D_0$  and  $D_2$  are monotonically consistent, it follows that the previously reported decreases in fractal dimension  $D_2$  associated with AD directly predict increases in the Hurst exponent of neurophysiological time series. It is perhaps also worth noting that increasingly predictable physiological dynamics, quantified by reductions in  $L_1$  or fractal dimensions, have been described as a hallmark of other pathophysiological processes including some epilepsies, cardiac arrhythmias and normal ageing; see Goldberger et al. (2002) for review.

Severity of Alzheimer's disease correlates with the extent of EEG abnormalities (Hughes et al., 1989; Kowalski et al., 2001) and in longitudinal studies of disease progression, delta and theta activity increase and alpha and beta decrease (Duffy, 1984).

Changes in EEG power are most prominent in left-sided temporal regions (Breslau et al., 1989; Rice et al., 1990). This is consistent with reports of reductions in left temporal and temporo-occipital alpha (Adler et al., 2003; Leuchter et al., 1992; Locatelli et al., 1998) and interhemispheric theta coherence (Adler et al., 2003). EEG changes in AD probably reflect functional disconnections among cortical areas resulting from death of cortical neurons, axonal pathology and cholinergic deficit. Since the basal forebrain's acetylcholinergic projections maintain desynchronised EEG activity (Metherate et al., 1992), loss of cholinergic innervation of the neocortex early in the course of AD (Bartus et al., 1992; Coyle et al., 1983) is likely to play a critical role in the emergence of low-frequency dynamics in the EEG. Damage to other neurotransmitter systems besides the cholinergic system is also likely to be important in AD (Dringenberg, 2000). However, both cholinergic and anticholinergic drugs have acute normalizing effects on the EEG in AD (Agnoli et al., 1983; Neufeld et al., 1994) and long-term treatment with cholinesterase inhibitors decreases the mean absolute power of theta activity (Kogan et al., 2001) and reduces EEG deterioration (Rodriguez et al., 2002).

Several experimentally refutable predictions are generated by the hypothesis that abnormally long-memory fMRI noise, indicated by large values of the Hurst exponent, represents the same pathophysiological process previously measurable only in terms of altered EEG spectral and nonlinear dynamical parameters in AD. For example, we might expect that EEG and fMRI measures of low-frequency dynamics should be correlated and both should show progressive deterioration with increasing disease severity. We might also anticipate that pharmacological manipulations emulating the cholinergic model of AD, such as administration of a cholinergic antagonist like scopolamine to healthy elderly volunteers, will cause transiently increased Hurst exponents in fMRI time series recorded from the temporal, cingulate and subcortical regions found to be dynamically abnormal in this study. Indeed, bearing in mind that we were able to demonstrate significant fMRI abnormalities in such a small group of patients with early AD and mild cognitive impairment, many of whom were already receiving treatment, it is conceivable that fMRI may prove to be a more sensitive marker than EEG of disease- and treatment-related changes in AD.

#### *Mapping between-group differences in low-frequency fMRI noise*

It is notable that these case-control differences in correlation structure of fMRI noise were evident in data that had been rigorously preprocessed to eliminate effects of head movement, making this an unlikely alternative interpretation. It is interesting also that pathological changes in correlation structure were not generally accompanied by local changes in variance, although there was evidence for pathologically increased noise variance in left prefrontal and parietal cortex, thalamus and basal forebrain; only in the dorsal cingulate cortex was increased variance associated with locally increased persistence.

We have carefully compared the sensitivity of the Hurst exponent in detecting these differences to the sensitivity of the equally parsimonious AR(1) model and shown, convincingly we hope, that the Hurst exponent is the more discriminating metric. Comparable group differences are evident in terms of the AR(1) coefficient, as expected theoretically, but these represent a subset of the differences mapped in terms of the Hurst exponent. The most likely explanation for this disparity, in our view, is that the brain changes in AD do not involve simply a greater strength of

association between immediately consecutive time points, which would be optimally described by a first order autoregression model. Rather, the neurophysiological signature of AD is more likely to involve greater salience of low-frequency or long-memory components in neurophysiological data, consistent with previously reported changes in nonlinear dynamical system parameters indicating pathologically increased predictability or reduced fractal dimensionality (Jeong, 2004), and dynamic changes of this sort are naturally better described by the Hurst exponent parameterising the covariance of a fractional Gaussian noise model.

#### *Relative merits of various fGn estimators*

Given that fractional Gaussian parameters of fMRI noise may be interesting and useful quantities, it is important to be sure that they are estimated with low bias and high efficiency. We have evaluated several possible estimators and the fairest overall conclusion is probably that no single estimator was unquestionably superior by all criteria. Whittle's approximation to the time domain ML estimator was the least biased and most efficient estimator of  $H$ ; its performance was particularly good compared to alternatives in the wavelet domain when  $H$  was close to its limiting values (0, 1). However, Whittle's algorithm incorporates the usual empirical estimator of  $\sigma^2$ , which is biased by long memory, and this is a serious problem potentially if we are interested in using estimates of  $\sigma^2$  to construct standardised statistics like  $\beta/SE(\beta)$  for tests of brain activation (such statistics will be overestimated, leading to loss of nominal type 1 error control, if  $\sigma^2$  is underestimated). Similar considerations counted against Wornell's algorithm, which was a slightly biased estimator of  $H$  but a severely biased estimator of  $\sigma^2$ . The fractional Brownian motion-based algorithm performed almost as well as Whittle's algorithm for estimation of  $H$  (it was the best wavelet-based method in this respect) and, in principle, it could be generalised to estimate the parameters of more complex models such as fGn plus additive white noise, or multifractional Gaussian noise with nonstationary  $H$  and  $\sigma^2$ . However, the wavelet maximum likelihood algorithm, refined by incorporation of the exact SDF for fGn, was only slightly more biased in estimation of  $H$ , and was marginally more efficient in estimation of  $\sigma^2$ , than the fBM-based algorithm. Additionally, the wavelet-ML algorithm was the most readily generalised to estimate simultaneously the regression model parameters  $\beta$  and was shown empirically to provide a valid basis for significance testing of possibly activated voxels. On balance, we have preferred to use the wavelet-ML algorithm because of our interest potentially in both signal and noise parameters of fMRI time series. If one was interested more purely in the Hurst exponent of the noise then Whittle's estimator in the time domain, or the fBM-based algorithm in the wavelet domain, would be good alternative choices. Moreover, if one was interested only in the spectral exponent of the noise  $\gamma$ , without wishing to make any prior assumptions about whether the noise was fGn or fBM, then the log periodogram is a simple, quick and model-free estimator of  $\gamma$ . Another class of estimators that we did not discuss here are the semiparametric estimators. They allow for weaker assumptions on the model and may be used when the parameter of interest is the spectral exponent (Bardet et al., 2003).

It almost goes without saying that the efficiency of all estimators improved as a function of increasing length of the time series, at least up to  $n = 512$  in simulated data. For estimation of long-memory parameters in resting fMRI data, it is therefore probably advisable to acquire time series comprising at least several hundred time points and preferably perhaps one thousand or more.

## Acknowledgments

This neuroinformatics research was supported by a Human Brain Project grant from the National Institute of Biomedical Imaging and Bioengineering and the National Institute of Mental Health. fMRI data acquisition at 1.5 T was supported by a Wellcome Trust grant to RH and at 3.0 T was supported by an MRC cooperative group grant to the Wolfson Brain Imaging Centre (WBIC). We thank colleagues in the MRI Unit at the Maudsley Hospital, London, and the WBIC, Cambridge, for technical assistance with fMRI; and we gratefully acknowledge the cooperation of the patients who participated in this study. Finally, we valued the constructive criticism of an earlier draft by two anonymous reviewers.

## Appendix A

### A.1. Wavelet analysis of the variance of a fractional Gaussian noise

The aim of this section is to emphasize how important it is to know the mean of a long-range dependent process when estimating its variance and, therefore, how wavelet-based estimators of the variance may be less biased than traditional estimators when applied to long memory processes such as fractional Gaussian noise.

Let us first assume that the process we observe is a pure fGn. In this case, its mean is known to be 0 and an unbiased estimator of the variance is

$$\hat{\sigma}^2 := \frac{1}{n} \sum_{t=1}^n G_t^2. \quad (10)$$

Unfortunately, the mean of a natural process is not usually known a priori so it must be estimated. This is the case, for instance, with fMRI noise, which has a spatially varying, unknown mean. One possible estimator of the variance is then

$$\hat{\sigma}^2 := \frac{1}{n} \sum_{t=1}^n (G_t - \bar{G})^2. \quad (11)$$

However, this estimator is biased. Indeed,

$$\mathbb{E}(\hat{\sigma}^2) = \sigma^2 - \text{Var}(\bar{G}) = \sigma^2(1 - n^{2H-2}), \quad (12)$$

and it can be seen that the bias can be arbitrarily close to  $\sigma^2$  when  $H$  gets close to 1. Let us have a closer look at this formula for the special case  $H = 0.5$ , corresponding to a Normal white noise. In this case, one obtains the well-known result that  $n/(n-1)\hat{\sigma}^2$  is an unbiased estimator of the variance. Following this example, an appealing estimator would be  $\hat{\sigma}^2/(1 - n^{2H-2})$ , but one has to bear in mind that the Hurst exponent is also unknown, and errors in its estimation will exponentially affect the estimation of the variance.

We will now show that the inconveniences caused by an unknown mean can be easily circumvented by filtering the process, in our case by a wavelet transform. The key benefit of this approach is that the wavelet coefficients of any stationary process have zero mean, leading to a better estimation of the variance (Percival and Walden, 2000, pp. 299–301). Moreover, even if some low-order polynomial trend is added to an fGn, the resulting signal will still have zero mean wavelet coefficients. This is because the filtering operations implicit in computation of the discrete wavelet transform (see Fig. 3) act like multiple backward differentiations (depending on the number of vanishing moments  $R$ ), which

destroy the positive correlation in the data due to both noise and trend. As a consequence, the estimated variance of the detail coefficients converges quickly to the theoretical value.

Let us consider now a wavelet decomposition of a finite sample  $G$  of an fGn, from the finer level 0 to some coarser level  $-j_0, j_0 \in \mathbb{N}$ . At each scale, the wavelet coefficients and the approximation coefficients form a pair of zero mean stationary time series, with variances (Wornell, 1996)

$$\begin{aligned} \text{Var}(d_{j,\cdot}) &= \int S(f) |F\psi_{j,\cdot}(f)|^2 df, \\ \text{Var}(a_{j,\cdot}) &= \int S(f) |F\varphi_{j,\cdot}(f)|^2 df. \end{aligned} \quad (13)$$

Eq. (8) and a change of variables under the integral gives

$$F\psi_{j,\cdot}(f) = \int e^{-2\pi i f x} \psi_{j,\cdot}(x) dx = 2^{-j/2} F\psi(2^{-j}f). \quad (14)$$

Eqs. (13) and (14) suggest an approximation to the wavelet coefficients variance, exploiting the fact that the wavelet filters are almost bandpass [(Percival and Walden, 2000; Wornell, 1996); and see Eq. (9)]:

$$\text{Var}(d_{j,\cdot}) \approx v_d^2(j) := 2^{-j+1} \int_{2^{j-1}}^{2^j} S(f) df, \quad j = -j_0, \dots, -1. \quad (15)$$

The same argument holds for

$$\text{Var}(a_{-j_0,\cdot}) \approx v_a^2(-j_0) := 2^{j_0+1} \int_0^{2^{-j_0-1}} S(f) df. \quad (16)$$

Note that this is effectively a wavelet decomposition of the variance of the process, since it is easy to see that

$$\text{Var}(G) = \int_{-1/2}^{1/2} S(f) = 2^{-j_0} v_a^2(-j_0) + \sum_{j=-j_0}^{-1} 2^j v_d^2(j). \quad (17)$$

A simple approximation to the variances of the wavelet and scaling coefficients (Fadili and Bullmore, 2002; Wornell, 1996) can be calculated by replacing Eq. (4) in Eqs. (15) and (16), thus:

$$\text{Var}(d_{j,\cdot}) \approx \tilde{v}_d^2(j) := \sigma^2 K_d(H) 2^{-j(2H-1)}, \quad j = -j_0, \dots, -1 \quad (18)$$

$$\text{Var}(a_{-j_0,\cdot}) \approx \tilde{v}_a^2(-j_0) := \sigma^2 K_a(H) 2^{j_0(2H-1)}, \quad (19)$$

with

$$K_d(H) := \frac{\Gamma(2H+1) \sin(\pi H) (1 - 2^{2H-2})}{(2\pi)^{2H-1} (1-H)}, \quad (20)$$

$$K_a(H) := \frac{\Gamma(2H+1) \sin(\pi H) (2^{2H-2})}{(2\pi)^{2H-1} (1-H)}. \quad (21)$$

From Eq. (18), we can see that there is an almost linear relation between the log of the variances  $\text{Var}(d_{j,\cdot})$  and the scales  $j$

$$\begin{aligned} \log_2 [\text{Var}(d_{j,\cdot})] &\approx -(2H-1)j + \log_2(\sigma^2 K_d(H)), \\ j &= -j_0, \dots, -1 \end{aligned} \quad (22)$$

which can be exploited by empirical estimators of  $H$  (Percival and Walden, 2000), such as some of those described in the following section.

## A.2. Estimators of fractional Gaussian noise parameters

This section provides technical detail concerning the fGn estimators comparatively evaluated in this paper.

### A.2.1. Maximum likelihood estimator in the time domain

The likelihood function of the fGn  $G = (G_1, \dots, G_n)'$ , with covariance matrix  $\Sigma$  depending on an unknown parameter vector  $\theta := (H, \sigma^2) \in (0, 1) \times \mathbb{R}_+$ , is

$$L(G; \theta) := (2\pi)^{-\frac{n}{2}} |\Sigma(\theta)|^{-\frac{1}{2}} e^{-\frac{1}{2} G' \Sigma^{-1}(\theta) G}, \quad (23)$$

where  $|\Sigma(\theta)|$  denotes the determinant of the matrix  $\Sigma(\theta)$ . The maximum likelihood estimator (MLE) of  $\theta$  is obtained by maximising  $L(G; \theta)$  with respect to  $\theta$ , or, equivalently, by maximising the log-likelihood function

$$\begin{aligned} LL(G; \theta) := \log L(G; \theta) &= -\frac{n}{2} \log(2\pi) - \frac{1}{2} \log |\Sigma(\theta)| \\ &\quad - \frac{1}{2} G' \Sigma^{-1}(\theta) G. \end{aligned} \quad (24)$$

This estimator in the time domain converges almost surely and is asymptotically Normal (Dahlhaus, 1989). However, it entails inversion of the covariance matrix of the fractional Gaussian noise  $\Sigma(\theta)$ , which is computationally costly and numerically unstable. For this reason, the method has not been widely used and will not be further considered here.

### A.2.2. Whittle's estimator

An alternative to the ML estimator in the time domain was proposed by Whittle (1953). The key idea is to replace the two components of  $LL(G; \theta)$  that are dependent on  $\Sigma(\theta)$  with their asymptotic equivalents. More precisely,  $\log |\Sigma(\theta)|$  is approximated by  $n \int_{-1/2}^{1/2} \log S(f) df$  and  $\Sigma^{-1}(\theta)$  is approximated by the  $n \times n$  matrix  $\Omega(\theta) := (\omega(i-l))_{i,l=1,\dots,n}$  with elements

$$\omega(k) := \int_{-\frac{1}{2}}^{\frac{1}{2}} \frac{1}{S(f)} e^{i2\pi k f} df. \quad (25)$$

This substitution numerically simplifies maximisation of the log-likelihood function, yet Whittle's estimator enjoys the same asymptotic properties as the exact MLE.

### A.2.3. Estimation by discrete variations of filtered fractional Brownian motion

As previously noted, a fGn can be regarded as the increment process of a fractional Brownian motion (fBm). Equally, if  $G = (G_1, \dots, G_n)$  is a finite sample of a fGn with Hurst exponent  $H$ , the variable  $B = (B_1, \dots, B_n)$ ,

$$B_t = \sum_{\ell=1}^t G_\ell \quad (26)$$

is a finite sample from the path of a fBm, with the same Hurst exponent  $H$ . We can thus identify the fGn by estimating the parameters of the corresponding fBm.

One of the methods that may be used in this context exploits the simple dependence on  $H$  of the discrete variations of the filtered sample of fBm (Coeurjolly, 2000a; Istas and Lang,

1997; Kent and Wood, 1997; Poggi and Viano, 1998). In brief, we can derive an approximately linear relationship between the log of the variance of the path of fBm passed by a filter  $g$  and the number of filter dilatations. We can thus obtain a simple least-squares estimator of  $H$  and  $\sigma^2$ , by considering  $M$  dilatations  $g^1, \dots, g^M$  of the wavelet filter  $g$ .

This estimator is unbiased, converges almost surely, and is asymptotically normal. The algorithm is computationally fast, simple and it automatically eliminates polynomial trends. Extensions are available to the case of fBm contaminated with additive white noise, and to the estimation of the parameters of multifractional Brownian motion (Coeurjolly, 2000b). It is specified in more detail as follows.

Let  $B_H$  be a fractional Brownian motion (a continuous process) sampled at discrete time instants  $t/n$ ,  $t = 1, \dots, n$ . Let

$$V^g(t/n) := \sum_{q=0}^{\ell} g_q B_H \left( \frac{t-q}{n} \right), \quad t = \ell + 1, \dots, n \quad (27)$$

be the sequence obtained after filtering it with some filter  $g$  having length  $\ell + 1$  and satisfying the relations

$$\sum_{q=0}^{\ell} g_p q^i = 0, \quad \text{for } i = 0, \dots, R-1 \quad \text{and} \quad \sum_{q=0}^{\ell} g_p q^R \neq 0, \quad (28)$$

for some integer  $R \geq 1$ . If  $g$  is a wavelet filter,  $R$  is the number of vanishing moments of the wavelet. Let

$$s_n(g) := \frac{1}{n-\ell} \sum_{t=\ell+1}^n |V^g(t/n)|^2 \quad (29)$$

be the empirical variance (second order moment) of the filtered sequence. If  $g^j$  denotes the filter  $g$  dilated  $j$  times, i.e.,

$$g_i^j = \begin{cases} g_k & \text{if } i = kj \\ 0 & \text{otherwise,} \end{cases} \quad i = 0, \dots, j\ell + 1, \quad (30)$$

we have the relation

$$\mathbb{E}(s_n(g^j)) = j^{2H} \mathbb{E}(s_n(g)) = A_{H,\lambda^2} \left( \frac{j}{n} \right)^{2H}, \quad (31)$$

with  $A_{H,\lambda^2} = -\lambda^2/2 \sum_{q,r=0}^{\ell} g_q g_r |q-r|^{2H}$ . Here  $\lambda^2$  denotes the variance of the increments of  $B_H(t)$ ,  $t = 0, 1, 2, \dots$ ; in other words, the stationary time series of fGn ( $B_H(1) - B_H(0), B_H(2) - B_H(1), \dots$ ). Given  $M$  dilatations  $g^1, \dots, g^M$  of the filter  $g$ , the parameters  $H$  and  $\lambda^2$  can be easily estimated by ordinary least squares after taking the logarithm in Eq. (31).

We will not discuss in detail the fundamental Eq. (26); however, we note briefly that in applying this estimator one is implicitly assuming that

$$G_t =_d B_H(t/n) - B_H((t-1)/n), \quad 1, \dots, n, \quad (32)$$

where  $=_d$  denotes equality in distribution. In other words, this would mean an increase in the sampling rate as  $n$  increases but the total duration of the experiment is constant since

$$\mathbb{E} \left[ \left( B_H \left( \frac{t}{n} \right) - B_H \left( \frac{t-1}{n} \right) \right)^2 \right] = \lambda^2 \frac{1}{n^{2H}}, \quad (33)$$



estimation of the regression model parameter vector  $\beta$ , at least for the case when the experimental design matrix  $X$  includes low-frequency periodic contrasts between experimental conditions; second, including the approximation coefficients will slightly improve efficiency of estimation of  $H$  (Fadili and Bullmore, 2002). A permissible compromise is to exclude the approximation coefficients in estimation of the noise parameters  $H$  and  $\sigma^2$  (steps 2 and 3 of the wavelet-ML algorithm) but use all scales of the wavelet decomposition in estimating  $\beta$  (step 4; see below).

#### A.2.6. Wavelet maximum likelihood estimator: algorithmic details

Let us denote, respectively, by  $T_a(-j_0; H)$  and  $T_d(j; H)$  the factors not depending on  $\sigma^2$  in the definitions of  $\hat{v}_a^2$  and  $\hat{v}_d^2$  [see Eqs. (18) and (19)], i.e.,

$$T_a(-J; H) := K_a(H)2^{J(2H-1)}, \quad T_d(j; H) := K_d(H)2^{-j(2H-1)},$$

$$j = -J, \dots, -1. \quad (39)$$

To find the parameters maximising the likelihood function (38), we can maximise the log-likelihood function

$$LL(G; \theta)$$

$$:= -\frac{1}{2} \left( n \log(2\pi\sigma^2) + \log T_a(-J; H) + \sum_{j=-J}^{-1} 2^{J-j} \log T_d(j; H) \right)$$

$$- \frac{1}{2\sigma^2} \left( \frac{a_{-J,1}^2}{T_a(-J; H)} + \sum_{j=-J}^{-1} \sum_{k=1}^{2^{J-j}} \frac{d_{j,k}^2}{T_d(j; H)} \right), \quad (40)$$

with  $\theta = (\beta^t, H, \sigma^2)$  and  $a_{-j,1}, d_{j,k}$  the wavelet coefficients of the fractional Gaussian noise  $G$ .

The algorithm is:

1. Initialisation: Define the wavelet basis to be used. Estimate  $\hat{\beta}$  by ordinary least-squares fit of the model (36) and  $\hat{H}$  by an arbitrary value in  $(0,1)$ .
2. Estimate  $\sigma^2$ : Using  $\hat{H}$  and  $\hat{\beta}$ , define

$$\hat{\sigma}^2 := \frac{1}{n} \left( \frac{\hat{a}_{-J,1}^2}{T_a(-J; \hat{H})} + \sum_{j=-J}^{-1} \sum_{k=1}^{2^{J-j}} \frac{\hat{d}_{j,k}^2}{T_d(j; \hat{H})} \right), \quad (41)$$

where  $\hat{a}_{-j,1}, \hat{d}_{j,k}$  are the wavelet coefficients of the residual  $E := Y - X\hat{\beta}$ .

3. Estimate  $H$  by numerical maximisation of the log-likelihood function (40), after replacing  $\beta$  and  $\sigma^2$  with their estimated values. As we expect  $H$  to be in the interval  $(0,1)$ , the search may be restricted to this range of values.
4. Estimate  $\beta$  and its covariance matrix: With  $\hat{H}$  and  $\hat{\sigma}^2$  calculated at steps 2 and 3, let

$$\hat{\beta} := \left( X_w^t \hat{\Sigma}(\hat{H}, \hat{\sigma}^2)^{-1} X_w \right)^{-1} X_w^t \hat{\Sigma}(\hat{H}, \hat{\sigma}^2)^{-1} Y_w, \quad (42)$$

and

$$\text{var}(\hat{\beta}) := \text{diag} \left( X_w^t \hat{\Sigma}(\hat{H}, \hat{\sigma}^2)^{-1} X_w \right)^{-1}. \quad (43)$$

where  $X_w$  and  $Y_w$  are the wavelet decomposition coefficients of vectors  $X$  and  $Y$ .

5. Loop: Go to step 2 and iterate until the change in successive parameter estimates is arbitrarily small, e.g., smaller than  $10^{-2}$ .

## References

- Adler, G., Brassens, S., Jajcevic, A., 2003. EEG coherence in Alzheimer's dementia. *J. Neural Transm.* 110, 1051–1058.
- Agnoli, A., Martucci, N., Mana, V., Conti, L., Fioravanti, M., 1983. Effect of cholinergic and anticholinergic drugs on short-term memory in Alzheimer's dementia: a neuropsychological and computerized electroencephalographic study. *Clin. Neuropharmacol.* 6, 311–323.
- Bardet, J.-M., Lang, G., Oppenheim, G., Philippe, A., Stoev, S., Taqqu, M.S., 2003. Semi-parametric estimation of the long-range dependence parameter: a survey. In: Doukhan, P., Oppenheim, G., Taqqu, M.S. (Eds.), *Theory and Applications of Long-Range Dependence*. Birkhäuser, Boston.
- Bartus, R.T., Dean, R.L., Beer, B., Lippa, A.S., 1992. The cholinergic hypothesis of geriatric memory dysfunction. *Science* 217, 408–414.
- Beran, J., 1994. *Statistics for Long-Memory Processes*. Chapman and Hall, London.
- Brammer, M.J., Bullmore, E.T., Simmons, A., Williams, S.C.R., Grasby, P.M., Howard, R.J., Rabe-Hesketh, S., 1997. Generic brain activation mapping in functional magnetic resonance imaging: a nonparametric approach. *Magn. Reson. Imaging* 15, 763–770.
- Breslau, J., Starr, A., Sicotte, N., Higa, J., Buchsbaum, M.S., 1989. Topographic EEG changes with normal ageing and SDAT. *Electroencephalogr. Clin. Neurophysiol.* 72, 281–289.
- Bullmore, E.T., Brammer, M., Williams, S.C.R., Rabe-Hesketh, S., Janot, N., David, A.S., Mellers, J.D.C., Howard, R., Sham, P., 1996. Statistical methods of estimation and inference for functional MR image analysis. *Magn. Reson. Med.* 35, 261–277.
- Bullmore, E.T., Brammer, M.J., Rabe-Hesketh, S., Curtis, V., Williams, S.C.R., Sharma, T., McGuire, P.K., 1999a. Methods for diagnosis and treatment of stimulus correlated motion in generic brain activation studies using fMRI. *Hum. Brain Mapp.* 7, 38–48.
- Bullmore, E.T., Suckling, J., Overmeyer, S., Rabe-Hesketh, S., Taylor, E., Brammer, M.J., 1999b. Global, voxel and cluster tests, by theory and permutation, for a difference between two groups of structural MR images of the brain. *IEEE Trans. Med. Imag.* 18, 32–42.
- Bullmore, E.T., Long, C., Suckling, J., Fadili, J., Calvert, G., Zelaya, F., Carpenter, T.A., Brammer, M., 2001. Colored noise and computational inference in neurophysiological (fMRI) time series analysis: resampling methods in time and wavelet domains. *Hum. Brain Mapp.* 12, 61–78.
- Bullmore, E.T., Fadili, J., Maxim, V., Şendur, L., Whitcher, B., Suckling, J., Brammer, M., Breakspear, M., 2004. Wavelets and functional magnetic resonance imaging of the human brain. *NeuroImage* 23, S234–S249.
- Coeurjolly, J.F., 2000a. Simulation and identification of the fractional Brownian motion: a bibliographical and comparative study. *J. Stat. Softw.* 50 (7).
- Coeurjolly, J.F., 2000b. *Inférence statistique pour les mouvements browniens fractionnaires et multifractionnaires*. PhD thesis, Université Joseph Fourier, Grenoble.
- Cordes, D., Haughton, V.M., Arfanakis, K., Carew, J.D., Truski, P.A., Moritz, C.H., Quigley, M.A., Meyerand, M.E., 2001. Frequencies contributing to functional connectivity in the cerebral cortex in “resting state” data. *Am. J. Neuroradiol.* 22, 1326–1333.
- Coyle, J.T., Price, D.L., DeLong, M.R., 1983. Alzheimer's disease: a disorder of cortical cholinergic innervation. *Science* 219, 1184–1190.
- Dahlhaus, R., 1989. Efficient parameter estimation for self-similar processes. *Ann. Stat.* 17, 1749–1766.
- Dale, A.M., 1999. Optimal experimental design for event-related fMRI. *Hum. Brain Mapp.* 8, 109–114.
- Davies, R.B., Harte, D.S., 1987. Tests for Hurst effect. *Biometrika* 74, 95–102.
- Dempster, A.P., Laird, N.M., Rubin, D.B., 1977. Maximum likelihood from incomplete data via the EM algorithm. *J. R. Stat. Soc.* 39, 1–38.
- Dijkerman, R.W., Mazumdar, R.R., 1994. On the correlation structure of

- the wavelet coefficients of fractional Brownian motion. *IEEE Trans. Inf. Theory* 40 (5), 1609–1612.
- Dringenberg, H.C., 2000. Alzheimer's disease: more than a cholinergic disorder—evidence that cholinergic–monoaminergic interactions contribute to EEG slowing and dementia. *Behav. Brain Res.* 115, 235–249.
- Duffy, F.H., 1984. Brain electrical activity in patients with presenile and senile dementia of the Alzheimer type. *Ann. Neurol.* 16, 439–448.
- Fadili, M.J., Bullmore, E.T., 2002. Wavelet-generalised least squares: a new BLU estimator of linear regression model with  $1/f$  errors. *NeuroImage* 15, 217–232.
- Flandrin, P., 1992. Wavelet analysis and synthesis of fractional Brownian motion. *IEEE Trans. Inf. Theory* 38, 910–917.
- Friston, K.J., Williams, S.C.R., Howard, R., Frackowiak, R.S.J., Turner, R., 1996. Movement-related effects in fMRI time series. *Magn. Reson. Med.* 35, 346–355.
- Goldberger, A.L., Amaral, L.A.N., Hausdorff, J.M., Ivanov, P.C., Peng, C.-K., Stanley, H.E., 2002. Fractal dynamics in physiology: alterations with disease and aging. *Proc. Natl. Acad. Sci. U. S. A.* 99, 2466–2472.
- Hughes, J.R., Shanmugham, S., Wetzel, L.C., Bellur, S., Hughes, C.A., 1989. The relationship between EEG changes and cognitive functions in dementia: a study in a VA population. *Clin. Electroencephalogr.* 20, 77–85.
- Istas, J., Lang, G., 1997. Quadratic variations and estimation of the Hölder index of a Gaussian process. *Ann. Inst. Henri Poincaré* 33 (4), 407–436.
- Jeong, J., 2004. EEG dynamics in patients with Alzheimer's disease. *Clin. Neurophysiol.* 115, 1490–1505.
- Kent, J.T., Wood, A.T.A., 1997. Estimating the fractal dimension of a locally self-similar Gaussian process using increments. *J. R. Stat. Soc.* 590 (3), 679–700.
- Kogan, E.A., Korczyn, A.D., Virchovsky, R.G., Klimovizky, S.S., Treves, T.A., Neufeld, M.Y., 2001. EEG changes during long-term treatment with donepezil in Alzheimer's disease patients. *J. Neural Transm.* 108, 1167–1173.
- Kowalski, J.W., Gawel, M., Pfeffer, A., Barcikowska, M., 2001. The diagnostic value of EEG in Alzheimer's disease: correlation with severity of mental impairment. *J. Clin. Neurophysiol.* 18, 570–575.
- Leopold, D.A., Murayama, Y., Logothetis, N.K., 2003. Very slow activity fluctuations in monkey visual cortex: implications for functional brain imaging. *Cereb. Cortex* 13, 422–433.
- Leuchter, A.F., Newton, T.F., Cook, I.A., Walter, D.O., Rosenberg-Thompson, S., Lachenbruch, P.A., 1992. Changes in brain functional connectivity in Alzheimer-type and multi-infarct dementia. *Brain* 115, 1543–1561.
- Linkenkaer-Hansen, K., Nikulin, V.V., Palva, J.M., Kaila, K., Ilmoniemi, R.J., 2001. Long-range temporal correlations and scaling behavior in human brain oscillations. *J. Neurosci.* 21, 1370–1377.
- Locascio, J.L., Jennings, P.J., Moore, C.I., Corkin, S., 1997. Time series analysis in the time domain and resampling methods for studies of functional magnetic resonance imaging. *Hum. Brain Mapp.* 5, 168–193.
- Locatelli, T., Cursi, M., Liberati, D., Franceschi, M., Comi, G., 1998. EEG coherence in Alzheimer's disease. *Electroencephalogr. Clin. Neurophysiol.* 106, 229–237.
- Mallat, S., 1989. Multiresolution approximation and wavelet orthonormal bases of  $L^2(\mathbb{R})$ . *Trans. Am. Math. Soc.* 315, 69–88.
- Mandelbrot, B.B., van Ness, J.W., 1968. Fractional Brownian motions, fractional noises and applications. *SIAM Rev.* 10, 422–437.
- McKhann, G., Drachman, D., Folstein, M., Katzman, R., Price, D., Stadlan, E.M., 1984. Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA work group under the auspices of Department of Health and Human Services task-force on Alzheimer's disease. *Neurology* 34, 939–944.
- Metherate, R., Cox, C.L., Ashe, J.H., 1992. Cellular bases of neocortical activation: modulation of neural oscillations by the nucleus basalis and endogenous acetylcholine. *J. Neurosci.* 12, 4701–4711.
- Neufeld, M.Y., Rabey, M.J., Parmet, Y., Sifris, P., Treves, T.A., Korczyn, A.D., 1994. Effects of a single intravenous dose of scopolamine on the quantitative EEG in Alzheimer's disease patients and age-matched controls. *Electroencephalogr. Clin. Neurophysiol.* 91, 407–412.
- Percival, D.B., Walden, A.T., 2000. *Wavelet Methods for Time Series Analysis*. Cambridge University Press, Cambridge.
- Poggi, J.M., Viano, M.C., 1998. An estimate of the fractal index using multiscale aggregates. *J. Time Ser. Anal.* 19 (2), 221–233.
- Purdon, P.L., Weisskoff, R.M., 1999. Effect of temporal autocorrelation due to physiological noise and stimulus paradigm on voxel-level false-positive rates in fMRI. *Hum. Brain Mapp.* 6, 239–249.
- Rice, D.M., Buchsbaum, M.S., Starr, A., Auslander, L., Hagman, J., Evans, W.J., 1990. Abnormal EEG slow activity in left temporal areas in senile dementia of the Alzheimer type. *J. Gerontol. Med. Sci.* 45, 145–151.
- Rodriguez, G., Vitali, O., De Leo, C., De Carli, F., Girtler, N., Nobili, F., 2002. Quantitative EEG changes in Alzheimer patients during long-term donepezil therapy. *Neuropsychobiology* 46, 49–56.
- Schroeder, M., 1991. *Fractals, Chaos, Power Laws: Minutes from an Infinite Paradise*. W.H. Freeman and Co., New York.
- Shimizu, Y., Barth, M., Windischberger, C., Moser, E., Thurner, S., 2004. Wavelet-based multifractal analysis of fMRI time series. *NeuroImage* 22, 1195–1202.
- Sinai, Y.G., 1976. Self-similar probability distributions. *Theory Probab. Appl.* 21, 64–80.
- Stam, C.J., Jelles, B., Achtereekte, H.A., Slaets, J.P., Keunen, R.W., 1995. Investigation of EEG non-linearity in dementia and Parkinson's disease. *Electroencephalogr. Clin. Neurophysiol.* 95, 309–317.
- Suckling, J., Bullmore, E.T., 2004. Permutation tests for functional neuroimaging experiments. *Hum. Brain Mapp.* 22, 193–205.
- Talairach, J., Tournoux, P., 1988. *A Coplanar Stereotaxic Atlas of the Human Brain*. Thieme, Stuttgart, Germany.
- Taqqu, M., Teverovsky, V., Willinger, W., 1995. Estimators for long-range dependence: an empirical study. *Fractals* 3 (4), 785–798.
- Tewfik, A.H., Kim, M., 1992. Correlation structure of the discrete wavelet coefficients of fractional Brownian motion. *IEEE Trans. Inf. Theory* 38, 904–909.
- Thompson, P.M., Hayashi, K.M., de Zubicaray, G., Janke, A.L., Rose, S.E., Herman, D., Hong, M.S., Dittmer, S.S., Doddrell, D.M., Toga, A.W., 2003. Dynamics of gray matter loss in Alzheimer's disease. *J. Neurosci.* 23, 994–1005.
- van Capellen van Walsum, A.-M., Pijnenburg, Y.A.L., Berendse, H.W., van Dijk, B.W., Knol, D.L., Scheltens, P., Stam, C.J., 2003. A neural complexity measure applied to MEG data in Alzheimer's disease. *Clin. Neurophysiol.* 114, 1034–1040.
- Whittle, P., 1953. Estimation and information in stationary time series. *Ark. Mat.* 2, 423–434.
- Wornell, G.W., 1996. *Signal Processing with Fractals: A wavelet based approach*. Signal Processing. Prentice Hall, Upper Saddle River, NJ.
- Wornell, G.W., Oppenheim, A.V., 1992. Estimation of fractal signals from noisy measurements using wavelets. *IEEE Trans. Signal Process.* 40, 611–623.
- Worsley, K.J., Liao, C., Aston, J., Petre, V., Duncan, G.H., Morales, F., Evans, A.C., 2002. A general statistical analysis for fMRI data. *NeuroImage* 15, 1–15.
- Zarahn, E., Aguirre, G.K., D'Esposito, M., 1997. Empirical analyses of BOLD fMRI statistics I. Spatially unsmoothed data collected under null hypothesis conditions. *NeuroImage* 5, 179–197.