

Signal Nonlinearity in fMRI: A Comparison Between BOLD and MION

Temujin Gautama, Danilo P. Mandic, *Member, IEEE*, and Marc M. Van Hulle, *Senior Member, IEEE*

Abstract—In this paper, we introduce a methodology for comparing the nonlinearities present in sets of time series using four different nonlinearity measures, one of which, the “delay vector variance” method, is a novel approach to the characterization of a time series. It is then applied to examine the difference in nonlinearity between functional magnetic resonance imaging (fMRI) signals that have been recorded using different contrast agents. Recently, an exogenous contrast agent, monocrySTALLINE iron oxide particle (MION), has been introduced for fMRI, which has been shown to increase the functional sensitivity compared with the traditional blood oxygen level dependent (BOLD) technique. The resulting fMRI signals are influenced by cerebral blood volume, whereas the more traditionally recorded BOLD signals are influenced not only by cerebral blood volume, but also by the cerebral blood flow and the metabolic rate of oxygen. The proposed methodology is applied to address the question whether this difference in the number of physiological variables is reflected in a difference in the degree of nonlinearity. We therefore analyze two sets of fMRI signals, one from a BOLD and the other from a MION monkey study with similar experimental designs. In the neuroimaging context, the proposed nonlinearity analyses are different from those described in the literature, since no *a priori* model is assumed: rather than pinpointing the source(s) of nonlinearity, nonparametric analyses are performed on BOLD and MION fMRI signals. Furthermore, we introduce a strategy for analyzing a population of fMRI signals, rather than focusing the analysis on one signal, as is traditionally done in the domain of nonlinear signal processing. Our results show that, overall, the BOLD signals are more nonlinear in nature than the MION ones, which is in agreement with current hypotheses.

Index Terms—BOLD, fMRI, MION, signal nonlinearity, surrogate data.

I. INTRODUCTION

FUNCTIONAL magnetic resonance imaging (fMRI) examines the concentration of a contrast agent as a measure for cerebral activity. Traditionally, the endogenous contrast agent deoxyhemoglobin is used for measuring the blood oxygen level dependent (BOLD) signal, which is a function of the cerebral blood volume, flow and metabolic rate of oxygen (for an overview, see [1]). The BOLD signal is a nonlinear observation of the cerebral activity, a property which has been addressed as such both experimentally and from a modeling perspective [2]–[6].

Recently, the use of monocrySTALLINE iron oxide nanoparticle (MION) as an exogenous contrast agent for fMRI, has been introduced [7], [8], allowing for the measurement of cerebral blood volume, which itself is a nonlinear function of the blood flow (see [9]). This approach has been shown to yield a better spatial localization of the active brain regions [8], [10], and has been demonstrated to yield a higher statistical power compared with experimental BOLD studies [8], [11].

Since the production model of a BOLD signal depends on more physiological systems that are coupled in a nonlinear manner than that of MION,¹ we hypothesize that this should be reflected in a lower degree of nonlinearity in the MION fMRI signals than in the BOLD ones. Indeed, both are “driven” by the same source, namely the cerebral activity and, thus, regardless of the fact whether or not the cerebral activity underlying the fMRI signals is linear, the BOLD signals should show a higher degree of nonlinearity than the MION ones, unless several nonlinear effects would cancel each other out. Linearity in the fMRI context is important, since it allows for simpler analysis methods, and is often assumed in conventional analyses [12]. The presence of nonlinearities has been addressed for BOLD [4], [5], [13] and MION [9], but it has been postulated that a linear model is sufficient in most cases [13], namely in those where there is no parametric variation in the form of the underlying evoked neuronal activity. The presence of nonlinearities in the fMRI signals has so far been investigated in a parametric manner (for an overview, see [5]), and, furthermore, no attempt has been made to quantify the difference in the degree of nonlinearity between BOLD and MION signals. Therefore, the approach adopted in this paper

Manuscript received August 29, 2002; revised November 29, 2002. The work of T. Gautama was supported by the Flemish Regional Ministry of Education under Scholarship GOA 2000/11, and by the Fund for Scientific Research under Research Grant G.0248.03. The work of D. P. Mandic was supported by the K.U. Leuven in part under Grant QLG3-CT-2000-30161 and in part under Visiting Fellowship F/01/079 while at the Laboratorium voor Neuro- en Psychofysiologie, K.U. Leuven, Belgium. The work of M. M. Van Hulle was supported in part by the Fund for Scientific Research under Research Grant G.0185.96N, in part by the National Lottery, Belgium, under Research Grant 9.0185.96 and Research Grant G.0248.03, in part the Flemish Regional Ministry of Education, Belgium, under Research Grant GOA 95/99-06 (2000/11), in part by the Flemish Ministry for Science and Technology under Research Grant VIS/98/012, and in part by the European Commission, 5th Framework Programme, under Research Grant QLG3-CT-2000-30161 and Research Grant IST-2001-32114. The authors are solely responsible for the contents of this paper. It does not represent the opinion of the Community, and the Community is not responsible for any use that might be made of data appearing therein. The Associate Editor responsible for coordinating the review of this paper and recommending its publication was A. Manduca. *Asterisk indicates corresponding author.*

*T. Gautama is with the Laboratorium voor Neuro- en Psychofysiologie, K.U. Leuven, Campus Gasthuisberg, Herestraat 49, B-3000 Leuven, Belgium.

D. P. Mandic is with the Department of Electrical and Electronic Engineering, Imperial College of Science, Technology and Medicine, SW7 2BT London, U.K.

M. M. Van Hulle is with the Laboratorium voor Neuro- en Psychofysiologie, K.U. Leuven, B-3000 Leuven, Belgium.

Digital Object Identifier 10.1109/TMI.2003.812248

¹Blood volume, which lies at the basis of the MION signal, is one of the factors contributing to the BOLD signal in a nonlinear fashion [5].

is novel: rather than pinpointing the source(s) of nonlinearity, nonparametric analyses are performed on BOLD and MION fMRI signals, and a comparative study is performed to test the hypothesis that BOLD signals convey a higher degree of nonlinearity than MION signals, irrespective of their source(s) of nonlinearity. Instead of analyzing the relationship between stimulus and fMRI signal variation [3], [4], [6], [14], [15], we analyze the fMRI signals *per se*. Notice that in this way, we analyze *signal*, rather than *system* nonlinearities.

Over the last decade, many nonparametric analysis techniques have been developed for the detection of nonlinearity in a signal (for an overview, see [16]). In this context, many biomedical signals, including heart rate variabilities (HRVs), electrocardiogram (ECG), hand tremor, and electroencephalogram (EEG) have been analyzed (see, e.g., [17]–[20]). In this paper, we introduce a novel “delay vector variance” (DVV) approach for the characterization of a time series and apply it to fMRI. Several other nonparametric nonlinearity analysis techniques are also considered in this context. Extensive experimentation and rigorous analysis show that the DVV results are more consistent than the those obtained using the other methods. We further propose a new methodology for a population analysis, i.e., for comparing sets of time series, rather than limiting the analysis to one signal per set. The proposed methodology is readily applicable to other type of biomedical signals.

II. SIGNALS AND METHODS

In this section, the data under analysis are briefly described. The concept of *signal* nonlinearity is compared with that of *system* nonlinearity, after which different nonlinearity measures are explained. These measures are computed for the original time series and compared with those obtained for linearized versions of the signal, called “surrogate time series,” or “surrogates” for short. The procedure for the statistical testing and the generation of the surrogates is discussed and, finally, we propose a novel methodology for the nonlinearity analysis of sets of signals (population analysis), and illustrate it by means of a synthetic example. We then apply this method to fMRI data in Section III.

A. Data Description

We analyze time series from two monkey fMRI motion studies, which are similar in experimental design, one from BOLD and the other from MION scanning sessions (for details, see the description of macaque monkey M1 in [8]). We analyze 14 fMRI signals of every study, corresponding to two of the foci of the study, namely motion areas left and right MT/V5 and their neighboring voxels, yielding a total of 28 signals. The voxel matching is based on anatomical coordinates, and we have taken the BOLD foci for reference. The BOLD signals consist of $N = 1920$ samples, whereas the MION signals comprise $N = 1200$ samples, only segments of which are shown in Fig. 1. All signals are mean subtracted and linearly detrended using linear regression. We shall adopt the following naming convention: BOLD left MT ($s_{b,l}$), BOLD right MT ($s_{b,r}$), MION left MT ($s_{m,l}$), and MION right MT ($s_{m,r}$).

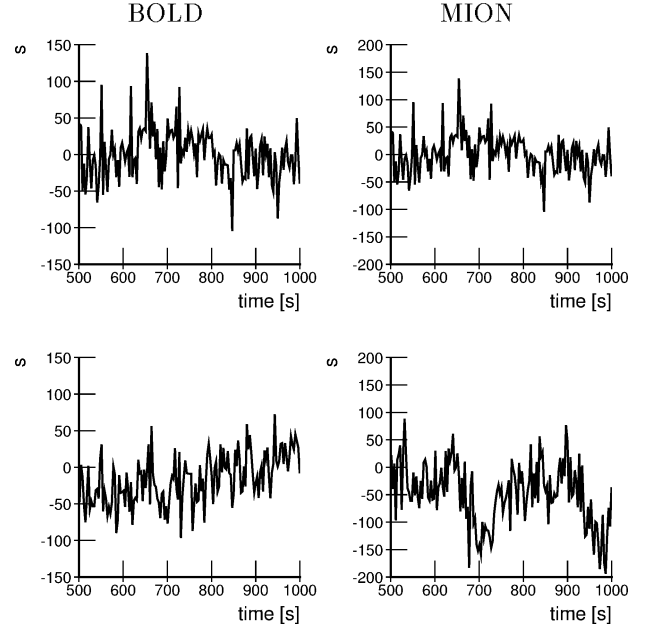


Fig. 1. Time slices of the four fMRI signals (after mean subtraction and detrending). Starting from the upper left and moving clockwise: BOLD left MT ($s_{b,l}$), MION left MT ($s_{m,l}$), MION right MT ($s_{m,r}$), and BOLD right MT ($s_{b,r}$).

B. Signal Versus System Nonlinearity

A linear system, $f(\cdot)$, is defined as one that obeys the superposition and scaling property, namely: $f(ax + by) = af(x) + bf(y)$. A system which violates these properties is nonlinear. By itself, this allows for a very strong tool for assessing *system* nonlinearity, referred to as “temporal summation,” as has been applied in the field of neuroimaging [3], [4], [6], [14], [15]: a short and a long pulse are presented to the system, and a system is linear if the response to the long pulse can be predicted from a summation of temporally shifted versions of the response to the short stimulus.

However, the principle of temporal summation for analysing the nonlinearity of a system implies that input and output time series can be measured simultaneously, while in many real-world settings, this is not possible. In applications of neuroimaging, for instance, the output can be recorded, but the input, i.e., the cerebral activity itself, cannot be measured directly. It is, therefore, approximated by an abstract stimulus representation (a square wave representing the alternation between conditions of stimulation and no stimulation).

The analysis of the nonlinearity of a *signal* can often provide insights into the nature of the underlying production system. However, care should be taken in the interpretation of the results, since the signal and system nonlinearities are confounded: the assessment of nonlinearity in a signal does not necessarily imply that the underlying production system is nonlinear. Indeed, if the input to the system were nonlinear and the system itself linear, the measured signal at the output would be nonlinear. Therefore, no unambiguous conclusions regarding an underlying system can be drawn from the nonlinearity analysis of one signal. However, this approach *does* allow for a comparative analysis between different systems, *driven by the same input*. Therefore, we examine the difference in signal nonlin-

earity between BOLD and MION fMRI signals, which have been recorded under comparable experimental designs.

C. Nonlinearity Measures

Many of the techniques described in this section rest upon the method of time delay embedding for representing a time series in phase space, i.e., by a set of delay vectors (DVs) of a given embedding dimension m , $\mathbf{x}(k) = [x_{k-m\tau} \dots x_{k-\tau}]$, where τ is a time lag, which for simplicity is set to unity in all simulations. In other words, $\mathbf{x}(k)$ is a vector containing m consecutive time samples. Every DV $\mathbf{x}(k)$ has a corresponding *target*, namely the next sample x_k .

1) *Traditional Nonlinearity Measures*: Two well-established measures for nonlinearity are considered, namely a third-order autocovariance (C3)

$$t^{C3}(\tau) = \langle x_k x_{k-\tau} x_{k-2\tau} \rangle \quad (1)$$

and the deviation due to *time reversibility* (REV)

$$t^{\text{REV}}(\tau) = \langle (x_k - x_{k-\tau})^3 \rangle \quad (2)$$

where τ is a time lag, which for simplicity and fair comparison, is set to unity in all analyses. For more details, we refer to [21] and [22].

2) *Correlation Exponent*: The approach is described in [23] and computes the correlation exponent, which yields an indication of the local structure of a strange attractor. For this purpose, the correlation integral is computed as

$$C(l) = \lim_{N \rightarrow \infty} \frac{1}{N^2} \{ \text{number of pairs } (i, j) \text{ whose distance } \|\mathbf{x}(i) - \mathbf{x}(j)\| \text{ is less than } l \}$$

where l is a length measure which is varied, and N is the number of DVs available for the analysis. In [23], it is established that the correlation exponent, i.e., the slope of $(\log(C(l)), \log(l))$, can be taken as a measure for the local structure of a strange attractor. Several methods exist for determining the range over which the slope is to be computed (“scaling region,” see, e.g., [16] and [24]). We adopt a pragmatic approach. The slope is computed over the l -interval $[\mu_d \pm \sigma_d]$, where μ_d and σ_d are the mean and standard deviation of all possible pairwise distances between different DVs. The resulting slope [referred to as correlation exponent approach (COR)] should not be interpreted as the actual correlation exponent, but it is sufficient in the context of surrogate data testing (see Section II-E), since it examines the correlation integral in a standardized scaling region. Indeed, since the surrogate time series have signal distributions identical to that of the original (see Section II-E), the distribution of pairwise distances and, thus, the mean and standard deviation, will be similar. Note that this distribution is approximately Gaussian for high embedding dimensions. Therefore, the correlation integral curve is examined in similar regions for both original and surrogate data, and a difference in the slope indicates a difference in local structure. Fig. 2(a) shows an example of the scaling region and the corresponding slope.

The embedding dimension for which the COR analysis is performed, is determined using Cao’s method [25], which is related to the false nearest neighbor method [26]. The measure, $E_1(d)$, stops changing when d exceeds the optimal embedding dimension, and quantifies the degree in which nearby DVs have similar targets. In our simulations, it is evaluated for embedding

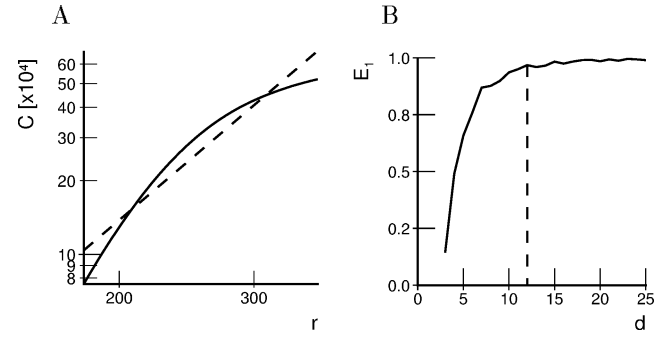


Fig. 2. Example analysis for fMRI signal $s_{b,1}$. (A) Grassberger-Proccaccia curve in the standardized scaling region. The dashed curve represents the output of the regression from which the slope is computed (in this case $\text{COR} = 3.45$). (B) Illustration of Cao’s method for determining the optimal embedding dimension. The determined point of convergence lies at 12.

dimensions $2 \leq m \leq 25$. We adopt the following criterion for determining the point of convergence: the difference between consecutive measurements should be lower than 0.01 and the measurement should exceed 0.95 times the mean of the last five measurements (to compensate for false plateaus). This method is exemplified for $s_{b,1}$ in Fig. 2(b).

D. DVV Method

Here, we introduce a novel analysis of a time series which examines a signal’s unpredictability by observing the variability of the targets belonging to sets of similar DVs. The approach is somewhat related to the false nearest neighbors [26], the $\delta - \epsilon$ methods [27], and the generalized synchrony approach proposed by [28]. The proposed DVV method can be summarized as follows for a given embedding dimension m .

- The mean μ_d and standard deviation σ_d are computed over all pairwise distances between DVs.
- The sets Ω_k are generated, which consist of all DVs that lie closer to $\mathbf{x}(k)$ than a certain distance. The distances are taken from the interval $[\mu_d - n_d \sigma_d; \mu_d + n_d \sigma_d]$, e.g., uniformly spaced, where n_d is a parameter controlling the span over which to perform the DVV analysis.
- For every set Ω_k , the variance of the corresponding targets, σ_k^2 , is computed. The average over all sets, divided by the variance of the time series σ_x^2 yields the measure of unpredictability σ^{*2}

$$\sigma^{*2} = \frac{\frac{1}{N} \sum_{k=1}^N \sigma_k^2}{\sigma_x^2}. \quad (3)$$

We refer to this measure as the “target variance,” and only compute the variance if Ω_k contains at least 30 DVs.

It is intuitively clear and mathematically sound that, when the embedding dimension and time lag are correctly determined and the signal exhibits some kind of structure, similar delay vectors (in terms of their Euclidean distance) have similar targets.² Therefore, for a correct choice of embedding parameters (which might not be unique), the target variance conveys information regarding one of the fundamental properties of a signal, namely its predictability. The two extreme cases are white noise (completely unpredictable) and a deterministic signal (completely

²This is related to the continuity of the mapping from DV to target, as described in [27].

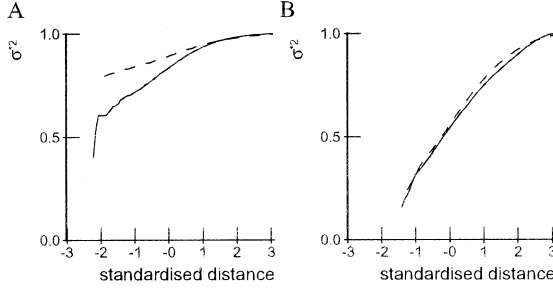


Fig. 3. (A) DVV-plots for BOLD left MT and (B) MION left MT shown as solid curves. The average DVV-plots for linearized versions (computed over 99 surrogates) are shown as dashed curves.

predictable). The resulting “DVV-plots” are obtained by plotting this measure of unpredictability (σ^{*2}) as a function of the standardized distances. Examples for BOLD left MT ($s_{b,l}$) and MION left MT ($s_{m,l}$) are shown in Fig. 3(a) and (b), respectively. The DVV-plots for linearized versions of these time series (surrogates, see further) are shown as dashed curves.

The optimal embedding dimension can be determined by running a number of DVV analyses for different values of m , and choosing that for which the minimal target variance is lowest, i.e., that which yields the best predictability. We have performed this analysis for embedding dimensions ranging from 2 to 25. Examples are shown in Fig. 7.

In the following step, the linear or nonlinear nature of the time series is examined by performing DVV analyses on both the original and a number of surrogate time series (see below), using the optimal embedding dimension of the original time series. Due to the standardization of the distance axis, these plots can be conveniently combined in a scatter diagram, where the horizontal axis corresponds to the DVV-plot of the original time series (Fig. 3, solid curves), and the vertical to that of the surrogate time series (Fig. 3, dashed curves). If the surrogate time series yield DVV-plots similar to the original, the “DVV scatter diagram” coincides with the bisector line, and the original time series is probably linear. The deviation from the bisector line is, thus, a measure of nonlinearity (for examples, see Fig. 8).

For statistical testing, it is convenient to quantify the DVV scatter diagram using a single measure. The deviation from the bisector line can be used for this purpose and can be computed as the root-mean-square error (RMSE) between the σ^{*2} 's of the time series (original or surrogate) and the σ^{*2} 's averaged over the DVV-plots of the surrogate time series.

E. Statistical Testing and Surrogate Data

The described techniques compare their results on the original signals to those obtained for linearized versions of these signals, the so-called *surrogate* time series (or “surrogates” for short). These surrogates are realizations of a composite null hypothesis, which in our case is that the original time series is linear (see further). If the analysis result for the original time series is significantly different from those of the surrogates, the null hypothesis is rejected and the original time series is hypothesized to be nonlinear. Since the analytical form of the probability distribution of the test statistics, i.e., the nonlinearity measures, is not known, a nonparametric rank-based test is used, as

suggested by [29]. For every original time series, we generate $N_s = 99$ surrogates for the nonlinearity tests. A right-tailed test (DVV) is rejected if rank r of the original time series exceeds 90, and a two-tailed test (C3, REV, and COR) is rejected if rank r is greater than 95 or less than or equal to 5. For the subsequent analyses, it is convenient to define the symmetrical rank r_{symm} as follows:

$$r_{\text{symm}}[\%] = \frac{r}{N_s + 1}, \text{ for right - tailed tests} \\ = \frac{\left| \frac{N_s + 1}{2} - r \right|}{\frac{N_s + 1}{2}}, \text{ for two - tailed tests.} \quad (4)$$

In this way, any of the described tests (right- or two-tailed) is rejected if $r_{\text{symm}} > 90\%$.

Statistical Testing: A key issue in surrogate data testing is the definition of an appropriate null hypothesis. In this study, we have adopted a composite null hypothesis [30]. In contrast to a *simple* null hypothesis, which asserts that the analyzed time series is a realization of a specified and unique process, a *composite* null hypothesis specifies a family of processes and asserts that the process underlying the given time series is a member of that family [29]. In the case of the composite null hypothesis that the time series is generated by a Gaussian linear stochastic process, surrogates are *constrained* to produce autocorrelation functions identical to that of the original time series, e.g., by phase randomizing the frequency spectrum of the original time series [Fourier transform (FT)-based method]. However, the FT-based method could lead to a false rejection of the null hypothesis if, e.g., the signal distribution of the original time series would not match that of the surrogates, i.e., Gaussian, even if the process underlying the original time series were Gaussian and linear. This is the case for a time series generated by a Gaussian and linear process and subsequently passed through a zero-memory observation function, $h(\cdot)$ [30]. Such an effect can be incorporated into the composite null hypothesis by constraining the surrogates to have both identical autocorrelation functions *and* identical signal distributions, assuming that the zero-memory observation function $h(\cdot)$ is invertible.³ Thus, the composite null hypothesis H_0 is that of an underlying Gaussian, linear and stochastic process, followed by a zero-memory and invertible observation function.

Surrogate Data Generation: Various approaches have been proposed in the open literature for generating surrogate time series consistent with the aforementioned composite null hypothesis (for an overview, see [31]). For computational efficiency, we have opted for the iterative amplitude adjusted FT (iAAFT) method, introduced in [32]. It retains the linear properties of the data, namely the autocorrelation function (estimated by means of the amplitude spectrum) and the empirical signal distribution. A surrogate time series is generated using a fixed point iteration algorithm, which guarantees an exact replica of the signal distribution, and an approximation of the spectrum (the converse would also be possible, but its impact is beyond the scope of this paper). We do not include an endpoint matching procedure

³It is important to note that $h(\cdot)$ is not considered as an integral part of the generating process, since a memory-less and nonlinear function would violate the superposition principle. It should be regarded as part of the analysis, namely as a correction for deviations from a Gaussian signal distribution.

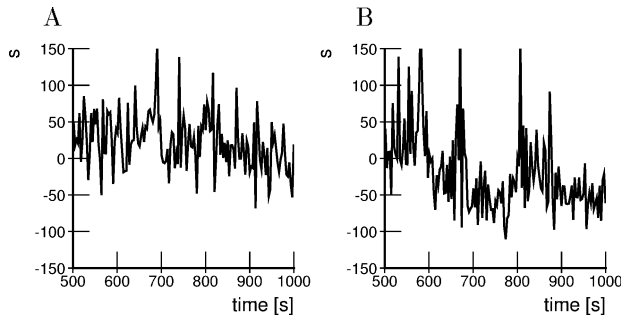


Fig. 4. Time slices of two surrogates, respectively, generated for (A) $s_{b,l}$ and (B) $s_{m,l}$.

for the fMRI time series,⁴ since they are periodic by design (the same stimulus sequence is repeated several times). Fig. 4(a) and (b) shows time slices of two surrogates, respectively, generated for $s_{b,l}$ and $s_{m,l}$. Where necessary, the optimal embedding dimension is only determined for the original time series, after which this value is also used for the analysis of the surrogates.

The COR analysis is more specific than the other nonlinearity analyses, since it examines the local structure of a strange attractor in a certain scaling region. On the other hand, there exist time series that do not follow a strange attractor, but *do* lead to a rejection of a null hypothesis of linearity. Still, if one time series follows a strange attractor, and another does not, there should be a difference in COR results. However, a difference in COR results as such does not provide sufficient evidence for the presence of a strange attractor.

F. Population Analysis

We propose a methodology for comparing different sets of signals on the basis of their degrees of nonlinearity. In this way, conclusions can be drawn with respect to a population of signals, rather than of single time series. Each set contains N_p time series, for each of which $N_s = 99$ surrogates are generated for performing the rank-test. For each of the four nonlinearity tests (C3, REV, COR, DVV), the test statistic is computed for the original time series and its $N_s = 99$ corresponding surrogates, after which it is assigned a symmetrical rank r_{symm} [(4), where C3, REV, and COR are two-tailed, and DVV right-tailed tests]. Thus, every set of time series yields N_p symmetrical ranks, which are visualized in a “rank histogram,” the rightmost bin of which corresponds to a significant rejection of the null hypothesis at the level of $\alpha = 0.10$ (i.e., $r_{\text{symm}} > 90\%$). We denote the rejection rate as the percentage of ranks that fall in the rightmost bin. Ideally, the rank histograms should display a 100% rejection rate for a set of nonlinear signals, and for a set containing linear signals, every bin should contain αN_p ranks (thus, a rejection rate of 10%).

To illustrate the proposed method, consider the following example of three synthetically generated sets of time series, one linear, the other nonlinear and the third inherently chaotic. The linear time series were generated from an autoregressive model of order four [AR(4)] which was driven using Gaussian white noise. The nonlinear time series were generated from a bilinear

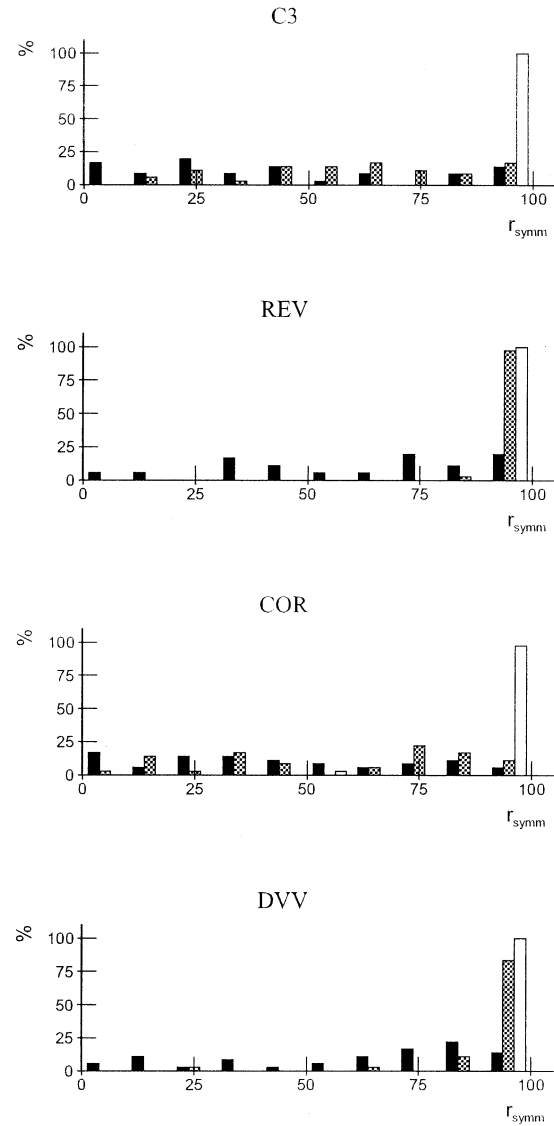


Fig. 5. Histogram of the rank results of the considered nonlinearity analyses for the sets of linear (black bars), nonlinear (gray bars) and chaotic time series (white bars), using bins of 10%. The horizontal axis denotes the symmetrical rank defined in (4).

model, with linear coefficients that were identical to the linear AR(4) model, also driven by Gaussian white noise. The chaotic time series were different realizations of the Hénon map. In total, we generated three sets of $N_p = 50$ time series of $N = 1000$ samples, for each of which $N_s = 99$ surrogate time series were generated.⁵

Fig. 5 shows the results for the linear (black bars), nonlinear (gray bars) and chaotic (white bars) sets. The time reversal strategy (REV) shows excellent performance: the rejection rate is 19% for the linear set, 98% for the nonlinear set, and 100% for the chaotic set. The performance of the proposed DVV is lower: 14% for the linear and 78% for the nonlinear sets, but also 100% for the chaotic set. The third-order cumulant (C3) and the COR are unable to detect nonlinearity in the nonlinear

⁴The endpoint matching procedure is suggested in [31] as a compensation for the bias toward a flatter spectrum due to the periodicity assumption of the fast Fourier transform (FFT).

⁵In contrast to those generated for the fMRI signals, we *have* used the end-point matching for reasons described in [31], since the synthetic signals are not periodic by design. The first and last 40 samples are scanned and the best matching signal values are used as the signal's start and end points.

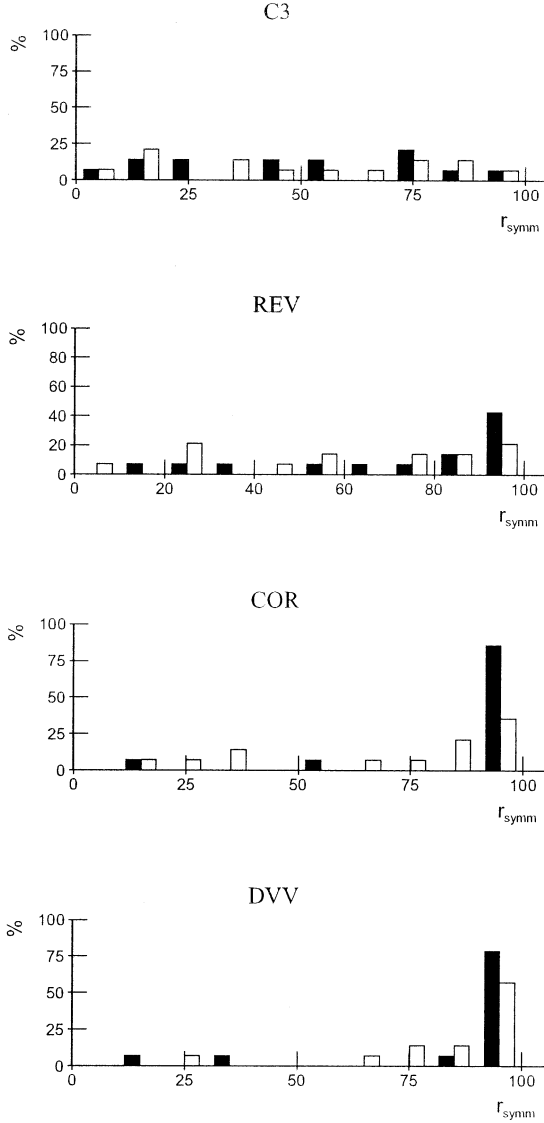


Fig. 6. Histogram of the rank results of the considered nonlinearity analyses for the two sets of fMRI signals, using bins of 10%. Results for the BOLD signals are shown as black bars and those for the MION signals as white bars. As is the case in Fig. 5, the horizontal axis denotes the symmetrical rank defined in (4).

set: the rank histograms for the linear and nonlinear sets are similar, i.e., flat with rejection rates of 14% and 17% for C3, respectively, 8% and 6% for COR. However, both detect the nonlinearity in the chaotic set (100% for C3 and 97% for COR). This is not surprising for COR, since, as explained in Section II-E, it is based on the presence of a strange attractor, a property which is absent in the bilinear model that has been used for generating the nonlinear time series.

III. RESULTS

Population Analysis: The population analysis for nonlinearity is now performed on the two sets of $N_p = 14$ fMRI time series, one from the BOLD, \mathcal{S}_b , and the other from the MION study, \mathcal{S}_m . The rank histograms are shown in Fig. 6 (BOLD results in black, MION results in white) and the rejection rates in Table I (columns labeled “iAAFT”), showing that with the

TABLE I
REJECTION RATES (REJ), IN PERCENT AND RESULTS OF THE TWO-TAILED WILCOXON RANK SUM TEST (P_W) FOR A DIFFERENCE IN DISTRIBUTIONS. THE COLUMNS ARE LABELLED “iAAFT” AND “FT” FOR THE CORRESPONDING COMPOSITE NULL HYPOTHESES (SEE SECTION II-E). THE LABELS A AND R IN THE LAST TWO COLUMNS DENOTE ACCEPTANCE OR REJECTION OF THE NULL HYPOTHESIS AT THE LEVEL OF 0.05

NL measure	Rej \mathcal{S}_b		Rej \mathcal{S}_m		P_W	
	iAAFT	FT	iAAFT	FT	iAAFT	FT
C3	7	7	7	7	0.8743 (A)	
REV	43	57	21	28	0.0008 (R)	
COR	86	93	36	43	0.0241 (R)	
DVV	79	79	57	57	0.0350 (R)	0.0042 (R)

TABLE II
RESULTS OF THE NONLINEARITY ANALYSES. THE RANK r FOR EVERY SIGNAL CORRESPONDS TO THAT IN THE RANK TEST (USING $N_s = 99$ SURROGATES). THE BOXES INDICATE THE TESTS THAT REJECT THE NULL HYPOTHESIS AT THE 10% LEVEL OF SIGNIFICANCE

signal	C3	REV	COR	DVV
left MT BOLD ($s_{b,l}$)	65	87	1	100
right MT BOLD ($s_{b,r}$)	58	100	1	100
left MT MION ($s_{m,l}$)	99	79	83	83
right MT MION ($s_{m,r}$)	27	52	93	100

exception of C3, all analyses have a higher rejection rate for \mathcal{S}_b than for \mathcal{S}_m . The latter rejection rates are still higher than the 10% which is expected for linear signals, therefore indicating that, on average, the MION signals are nonlinear, but less so compared with the BOLD signals. The results for C3 are *not* consistent with our hypothesis.

We further test for a difference in distribution between the nonlinearity measures for \mathcal{S}_b and \mathcal{S}_m , using the two-tailed Wilcoxon rank sum test. The results are shown in Table I (columns labeled “iAAFT”). The fourth column shows P_W , the probability of observing a result equally or more extreme if the null hypothesis were true (thus, $P_W \approx 0$ indicates a difference in distributions). Note that a one-tailed version of the test, for $H_0 : t_{\text{BOLD}} < t_{\text{MION}}$, is rejected for the same three tests with P_W values that are half of those in Table I. The results for REV, COR, and DVV support the hypothesis that there is, indeed, a difference in nonlinearity between BOLD and MION signals. The C3 test does not detect this difference as it did not distinguish between the linear and the nonlinear sets in our synthetic example (Fig. 5).

The COR rejection rates for \mathcal{S}_b (86%) and \mathcal{S}_m (36%) suggest that the difference in nonlinearity could be attributed to the presence of a strange attractor, indicating possible chaotic behavior, albeit the evidence is not conclusive (see Section II-E2). Cao’s method yields an optimal embedding dimensions between 9 and 13, with no significant difference in distributions between \mathcal{S}_b and \mathcal{S}_m .

Detailed Analysis for Example Cases: The foci (left and right MT) of the two studies are now examined in more detail to illustrate the various analysis methods. The rank r of the measurement for the original time series is shown in Table II. The table illustrates that it is not evident to detect nonlinearity, since the various tests (at the level of 0.10) yield different results.

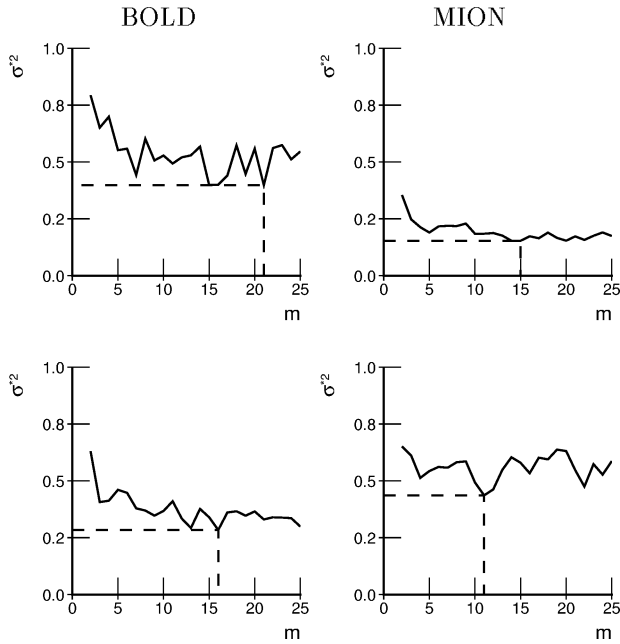


Fig. 7. Target variance, σ^2 for the four fMRI signals as a function of the embedding dimension. Starting from the upper left and moving clockwise: $s_{b,l}$, $s_{m,l}$, $s_{m,r}$, and $s_{b,r}$. The dashed line indicates the minimal target variance and, thus, the optimal embedding dimension.

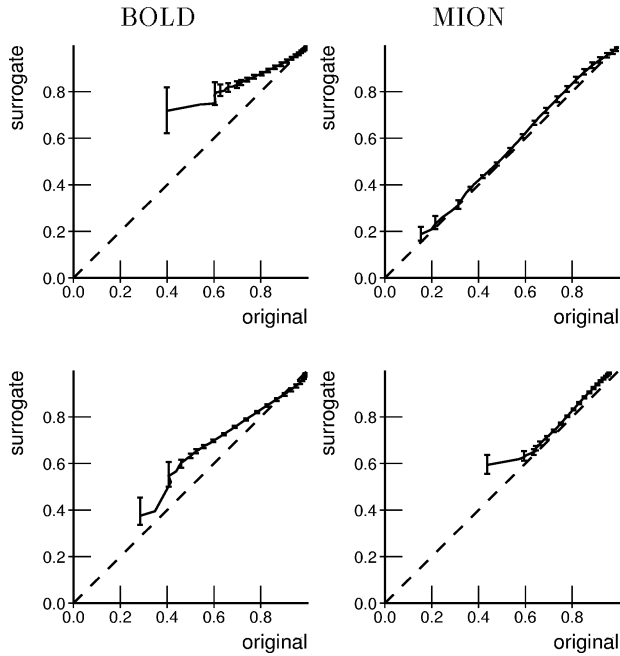


Fig. 8. DVV scatter diagrams of the four fMRI signals for the optimal embedding dimension (median and quartiles). Starting from upper left and moving clockwise: $s_{b,l}$, $s_{m,l}$, $s_{m,r}$, and $s_{b,r}$, with corresponding RMSEs of 0.0804, 0.0214, 0.0275, and 0.0639.

Both the COR and DVV analyses indicate the BOLD signals to be nonlinear. The REV analysis only shows nonlinearity for $s_{b,r}$, and the C3 analysis indicates the presence of nonlinearity in $s_{m,l}$. The DVV analysis also detects nonlinearities in $s_{m,r}$.

The DVV analyses yield optimal embedding dimensions of 21, 16, 15, and 11, respectively, $s_{b,l}$, $s_{b,r}$, $s_{m,l}$, and $s_{m,r}$ (see Fig. 7). The DVV scatter diagrams are shown in Fig. 8 (so as not to overload the figures, only one in three error bars are

shown). Qualitatively, the DVV scatter diagrams in Fig. 8 indicate that the nonlinearities present in the MION signals are less pronounced than those in the BOLD signals, which is also evident in the RMSE values, which are 0.0804 and 0.0639 for BOLD, and 0.0214 and 0.0275 for MION, respectively, for left and right MT.

IV. DISCUSSION

It has been shown both experimentally and from a modeling perspective that the BOLD fMRI signal is a complex function of the underlying cerebral activity, and depends on the cerebral blood flow, blood volume, and metabolic rate of oxygen [2]–[6]. Recently, the use of MION as a contrast agent has been shown to yield a better spatial localization of the active brain regions and a higher statistical power [8], [11]. The MION fMRI signal is dependent on fewer (nonlinearly) interrelated physiological systems than the BOLD one, namely predominantly on blood volume, which, in turn, is a function of blood flow. This has led to the working hypothesis of the paper, namely that BOLD signals show a higher degree of nonlinearity than the MION signals.

Rather than analyzing and identifying the nonlinearities present in the production systems of BOLD and MION, we quantify the signal nonlinearities in the fMRI signals, and perform a comparative study between BOLD and MION. Since both are “driven” by the underlying cerebral activity, a difference in signal nonlinearity can be attributed to a difference in system nonlinearity. For the assessment of nonlinearity in a single time series, we have introduced a novel way for characterizing a time series, the DVV method, and have compared the results with three other, well-established nonlinearity measures, namely those based on C3, REV, and COR, the latter of which stems from chaos theory. Furthermore, we have introduced a methodology for comparing and testing the degree of nonlinearity between populations of signals, rather than limiting the analysis to one time series per set, as is traditionally done in the nonlinear signal processing literature.

The proposed methodology has been applied to time series obtained from two monkey motion fMRI studies, one BOLD and one MION, which are similar in experimental design. The results for the two sets of fMRI signals show that the ratio of BOLD signals in which nonlinearity is detected, is higher than that of MION signals, but that in both cases, there is an indication of nonlinearity. This is not surprising, since the neuronal activity underlying the BOLD and MION signals has been conjectured to be nonlinearly related to the stimulus (see, e.g., [1], [6]). Furthermore, we have tested for a difference in distributions between the nonlinearity measures for BOLD and MION using the Wilcoxon rank sum test, and found that for all measures except the C3, the difference is significant at the level of 0.05. Finally, the COR analysis, which is related to the presence of a strange attractor, rejects the composite null hypothesis, namely that the time series is generated from an amplitude transformed Gaussian linear stochastic process. This indicates that the nonlinearities present in fMRI signals could be attributed to the presence of a strange attractor, indicating possible chaotic behavior, as described in Section II-E.

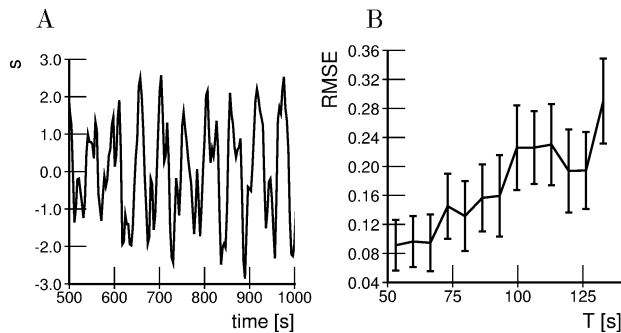


Fig. 9. A) Time segment of a hypothetical fMRI response to an on/off stimulus, for a period $T = 53$ s; B) Average degree of nonlinearity (RMSE), computed over 50 time series, as a function of the period T . The vertical bars indicate the standard deviations.

The validation of our working hypothesis supports and, to a certain extent, explains the findings described in [8] and [11], where it has been demonstrated that fMRI experiments using MION yield higher statistical power than those using the BOLD effect. Indeed, since conventional fMRI analyses use a general linear model, the lower degree of nonlinearity of MION studies will yield a lower regression error, and subsequently, higher statistical power compared with BOLD studies. Furthermore, the validity of the correction for the haemodynamic impulse-response function is violated in a lesser degree by the MION than by the BOLD signals.

One could argue that the difference in nonlinearity between BOLD and MION could be attributed to different stimulus durations, which in our case are 24 s for BOLD and 66 s for MION. To shed further light on this, hypothetical responses to on/off stimuli are generated, the period T of which is varied. They consist of square waves, sampled every 3.321 s, with an amplitude of 0.5, contaminated by Gaussian, unit-variance white noise, and convolved with the haemodynamical impulse-response function traditionally used for analyzing BOLD studies. A time slice of an example signal with $T = 53$ s is shown in Fig. 9(a). The period T is varied from 53 (BOLD) to 133 s (MION), and the DVV method is used for measuring the degree of nonlinearity (RMSE; the average DVV-plot is computed over 19 surrogates). For every period, 50 time series are generated, each of which with a length of 27.678 min, using different random seeds for the noise. The results are summarized in Fig. 9(b). The degree of nonlinearity *increases* with the period T . Since in the comparative study on the real recordings, the period T for BOLD is *lower* than for MION, the degree of nonlinearity of the BOLD set would be expected to be lower than that of the MION set. However, the analysis results indicate the converse. Therefore, it can be concluded that the difference in stimulus timing is not responsible for the higher degree of nonlinearity in BOLD signals.

Finally, there remains one issue regarding our methodology. The definition of nonlinearity in the strict sense, e.g., related to the validity of the superposition principle or to the modeling of, and compensation for the haemodynamic impulse-response function, is generally considered too stringent in the context of surrogate data testing, as explained in Section II-E. Therefore, we have performed an amplitude transformation, such that rejections of the null hypothesis due to a mismatch in signal

distribution are avoided (see, e.g., [31]). The results of this analysis confirmed our working hypothesis. However, the question remains whether the difference in degree of nonlinearity between BOLD and MION can be attributed to the amplitude transformation, or to the fMRI signals themselves. In order to address this question, we have also performed the population analysis *without* including the amplitude transformation, namely by using FT-based surrogates (see Section II-E). As a result, the rejection rates should not be smaller than those in the case of the iAAFT-based surrogates, since the null hypothesis is less stringent. This is confirmed by the results shown in Table I (columns labeled “FT”). When the rejection rates change, they are higher for the BOLD signals than for the MION ones. Moreover, the difference between nonlinearity measures for BOLD and MION remain statistically significant. This is trivial for C3, REV, and COR, since only the measures for the original time series are used and are, thus, not influenced by the surrogate data. The result for DVV changes, since this method uses the Euclidean distance between the DVV-plot of the original and the average DVV-plot of the surrogates, and the difference in distribution becomes even more significant. Thus, even when performing the nonlinearity analyses with the strict definition of nonlinearity, BOLD signals are more nonlinear than MION signals.

The proposed methodology has been applied to the case of fMRI data recorded using different contrast agents. The results support the current hypotheses that, indeed, BOLD fMRI signals are more nonlinear and, thus, more complex, than the MION signals. The same methodology can be readily applied to other sets of biomedical signals where a difference in nonlinearity can, e.g., convey information about the health status of the subject, as has been suggested in the cases of EEG (see, e.g., [33]) and HRV (see, e.g., [34]) signals. The DVV method is expected to provide useful information in these settings, since it yields clear information regarding the degree to which the signals in question differ from linear ones, namely the distance to the bisector line in the scatter diagrams, and since the sensitivity of the method is overall higher than that of other methods.

ACKNOWLEDGMENT

The authors wish to thank W. Vanduffel and G. A. Orban for providing the fMRI data.

REFERENCES

- [1] D. J. Heeger and D. Ress, “What does fMRI tell us about neuronal activity?,” *Nat. Rev. Neurosci.*, vol. 3, pp. 142–151, 2002.
- [2] R. D. Hoge, J. Atkinson, B. Gill, G. R. Crelier, S. Marrett, and G. B. Pike, “Stimulus-dependent BOLD and perfusion dynamics in human VI,” *NeuroImage*, vol. 9, pp. 573–585, 1999.
- [3] A. L. Vasquez and D. C. Noll, “Nonlinear aspects of the BOLD response in functional MRI,” *NeuroImage*, vol. 7, pp. 108–118, 1998.
- [4] G. Rees, A. Howseman, O. Josephs, C. D. Frith, K. J. Friston, R. S. J. Frackowiak, and R. Turner, “Characterizing the relationship between BOLD contrast and regional cerebral blood flow measurements by varying the stimulus representation rate,” *NeuroImage*, vol. 6, pp. 270–278, 1997.
- [5] K. J. Friston, A. Mechelli, R. Turner, and C. J. Price, “Nonlinear responses in fMRI: The balloon model, Volterra kernels, and other hemodynamics,” *NeuroImage*, vol. 12, pp. 466–477, 2000.

- [6] K. L. Miller, W. M. Luh, T. T. Liu, A. Martinez, T. Obata, E. C. Wong, L. R. Frank, and R. B. Buxton, "Nonlinear temporal dynamics of the cerebral blood flow response," *Human Brain Mapping*, vol. 13, pp. 1–12, 2001.
- [7] J. B. Mandeville, J. J. Marota, B. E. Kosofsky, J. R. Keltner, R. Weissleder, B. R. Rosen, and R. M. Weisskoff, "Dynamic functional imaging of relative cerebral blood volume during rat forepaw stimulation," *MRM*, vol. 39, pp. 615–624, 1998.
- [8] W. Vanduffel, D. Fize, J. B. Mandeville, K. Nelissen, P. Van Hecke, B. R. Rosen, R. B. H. Tootell, and G. A. Orban, "Visual motion processing investigated using contrast-agent enhanced fMRI in awake behaving monkeys," *Neuron*, vol. 32, no. 4, pp. 565–577, 2001.
- [9] J. B. Mandeville, J. J. Marota, C. Ayata, G. Zararchuk, M. A. Moskowitz, B. Rosen, and R. M. Weisskoff, "Evidence of a cerebrovascular postarteriole windkessel with delayed compliance," *J. Cereb Blood Flow Metab.*, vol. 19, pp. 679–689, 1999.
- [10] J. B. Mandeville and J. J. A. Marota, "Vascular filters of functional MRI: Spatial localization using BOLD and CBV contrast," *MRM*, vol. 42, pp. 591–598, 1999.
- [11] F. P. Leite, D. Tsao, W. Vanduffel, D. Fize, L. Wald, A. Dale, K. Kwong, G. A. Orban, B. R. Rosen, R. B. H. Tootell, and J. B. Mandeville, "Repeated fMRI using iron oxide contrast agent in awake, behaving macaques at 3 tesla," *NeuroImage*, vol. 16, pp. 283–294, 2002.
- [12] K. J. Friston, P. Jezzard, and R. Turner, "Analysis of functional MRI time series," *Human Brain Mapping*, vol. 1, pp. 153–171, 1994.
- [13] K. J. Friston, O. Josephs, G. Rees, and R. Turner, "Nonlinear event-related responses in fMRI," *MRM*, vol. 39, pp. 41–52, 1998.
- [14] M. G. Boynton, S. A. Engel, G. H. Glover, and D. J. Heeger, "Linear systems analysis of functional magnetic resonance imaging in V1," *J. Neurosci.*, vol. 16, no. 13, pp. 4207–4221, 1996.
- [15] A. Mechelli, C. J. Price, and K. J. Friston, "Nonlinear coupling between evoked rCBF and BOLD signals: A simulation study of hemodynamic responses," *NeuroImage*, vol. 14, pp. 862–872, 2001.
- [16] R. Hegger, H. Kantz, and T. Schreiber, "Practical implementation of nonlinear time series methods: The TISEAN package," *Chaos*, vol. 9, pp. 413–435, 1999.
- [17] C.-S. Poon and C. K. Merrill, "Decrease of cardiac chaos in congestive heart failure," *Nature*, vol. 389, pp. 492–495, 1997.
- [18] J. Timmer, S. Häubler, M. Lauk, and C. H. Lücking, "Pathological tremors: Deterministic chaos or nonlinear stochastic oscillators?," *Chaos*, vol. 10, pp. 278–288, 2000.
- [19] J. Theiler, "On the evidence for low-dimensional chaos in an epileptic electroencephalogram," *Phys. Lett. A*, vol. 196, pp. 335–341, 1995.
- [20] T. Schreiber, "Interdisciplinary application of nonlinear time series methods," *Phys. Rep.*, vol. 308, no. 1, pp. 1–64, 1999.
- [21] T. Schreiber and A. Schmitz, "On the discrimination power of measures for nonlinearity in a time series," *Phys. Rev. E*, vol. 55, pp. 5443–5447, 1997.
- [22] C. Diks, J. C. van Houwelingen, F. Takens, and J. DeGoede, "Reversibility as a criterion for discriminating time series," *Phys. Lett. A*, vol. 201, pp. 221–228, 1995.
- [23] P. Grassberger and I. Procaccia, "Measuring the strangeness of strange attractors," *Physica D*, vol. 9, pp. 189–208, 1983.
- [24] J. Theiler and T. Lookman, "Statistical error in a chord estimator of correlation dimension: The 'rule of five'," *Int. J. Bif. Chaos*, vol. 3, pp. 765–771, 1993.
- [25] L. Cao, "Practical method for determining the minimum embedding dimension of a scalar time series," *Physica D*, vol. 110, pp. 43–50, 1997.
- [26] M. B. Kennel, R. Brown, and H. D. I. Abarbanel, "Determining embedding dimension for phase-space reconstruction using a geometrical construction," *Phys. Rev. A*, vol. 45, pp. 3403–3411, 1992.
- [27] D. T. Kaplan, "Exceptional events as evidence for determinism," *Physica D*, vol. 73, pp. 38–48, 1994.
- [28] S. J. Schiff, P. So, T. Chang, R. E. Burke, and T. Sauer, "Detecting dynamical interdependence and generalized synchrony through mutual prediction in a neural ensemble," *Phys. Rev. E*, vol. 54, no. 6, pp. 6708–6724, 1996.
- [29] J. Theiler and D. Prichard, "Constrained-realization Monte-Carlo method for hypothesis testing," *Physica D*, vol. 94, pp. 221–235, 1996.
- [30] J. T. Theiler, S. Eubank, A. Longtin, B. Galdrikian, and J. D. Farmer, "Testing for nonlinearity in time series: The method of surrogate data," *Physica D*, vol. 58, pp. 77–94, 1992.
- [31] T. Schreiber and A. Schmitz, "Surrogate time series," *Physica D*, vol. 142, no. 3–4, pp. 346–382, 2000.
- [32] —, "Improved surrogate data for nonlinearity tests," *Phys. Rev. Lett.*, vol. 77, pp. 635–638, 1996.
- [33] R. G. Andrzejak, K. Lehnertz, F. Mormann, C. Rieke, P. David, and C. E. Elger, "Indications of nonlinear deterministic and finite-dimensional structure in time series of brain electrical activity: Dependence on recording region and brain state," *Phys. Rev. E*, vol. 64, p. 061 907, 2001.
- [34] D. J. Christini, F. M. Bennett, K. R. Lutchén, H. M. Ahmed, J. M. Hausdorff, and N. Oriol, "Applications of linear and nonlinear time series modeling to heart rate dynamics analysis," *IEEE Trans. Biomed. Eng.*, vol. 42, pp. 411–415, Apr. 1995.