Journal of Neuroscience Methods 172 (2008) 27–33

Contents lists available at ScienceDirect

Journal of Neuroscience Methods

journal homepage: www.elsevier.com/locate/jneumeth

# Properties of multivariate data investigated by fractal dimensionality

Danko Nikolić[a,b,*], Vasile V. Moca[a,b,c], Wolf Singer[a,b], Raul C. Mureşan[a,c]

[a] Department of Neurophysiology Max-Planck-Institute for Brain Research, Deutschordenstr. 46, 60528 Frankfurt am Main, Germany
[b] Frankfurt Institute for Advanced Studies, Ruth-Moufang-Str. 1, 60438 Frankfurt am Main, Germany
[c] Center for Cognitive and Neural Studies (Coneural), Str. Cireşilor 29, 400487 Cluj-Napoca, Romania

ARTICLE INFO

*Article history:*
Received 11 December 2007
Received in revised form 31 March 2008
Accepted 7 April 2008

*Keywords:*
Fractal
Data model
Multivariate analysis
Dimensionality

ABSTRACT

Elaborated data-mining techniques are widely available today. Nevertheless, many non-linear relations among variables remain undiscovered in multi-dimensional datasets. To address this issue we propose a method based on the concept of fractal dimension that explores the structure of multivariate data and apply the method to simulated data, as well as to local field potentials recorded from cat visual cortex. We find that with changes in the analysis scale, the dimensionality of the data often changes, indicating first that the data are not simple fractals with one unique dimension and second, that, at a certain scale, important changes in the geometric structure of the data may occur. The method can be used as a data-mining tool but also as a method for testing a model's fit to the data. We achieve the latter by comparing the dimensionality of the original data to the dimensionality of the data reconstructed from a model's description of the data (here using the general linear model). The method provides indispensable help in estimating the complexity of non-linear relationships within multivariate datasets.



## 1. Introduction

Data analysis requires investigation of relations between data points, and by any type of such analysis, irrespectively of whether explorative or hypothesis-driven, only a limited subset of all possible relations can be addressed. Despite the elaborated data-mining procedures, many such relations in many datasets remain hidden. We can only guess how many (important) scientific insights have been missed just because patterns could not be easily detected in otherwise, perfectly reliable and legitimate sets of data. Therefore, we should welcome every new analysis method that is able to probe new relationships and present the results in an elegant and easily interpretable way. For such methods, reduction of dimensionality plays an important role (Brand, 2003; Levina and Bickel, 2004; Tenenbaum et al., 2000). In the present study we propose a method that is designed to explore relations across multivariate data points and that is based on the concept of fractal dimension.

### 1.1. The concept of fractal dimension

A fractal is an object with a high degree of self-similarity, whereby globally the object looks very similar to its details (Falconer, 2003). We show three example fractals in Fig. 1a–c and for one of them we illustrate how it is created by a simple iterative procedure (Fig. 1a) (Falconer, 2003, pp. xviii–xx). Perhaps the most common quantitative description of a fractal is the measure of its dimension. As the dimensionality of standard geometric objects, e.g., triangles and cubes, can be grasped easily by intuition, it is also easy to acquire intuitive understanding of fractal dimensions. Depending on the space that they occupy, fractals have different dimensionality and they can be given by real numbers. For example, the Koch fractal in Fig. 1a occupies a $D = 1.26$-dimensional space. The Sierpinsky fractal in Fig. 1b and another fractal in Fig. 1c (Landau and Paez, 1997) appear visually to occupy gradually more space, and this is consistent with their calculated fractal dimensions (indicated in Fig. 1).

In the present study we use the concept of fractal dimension to address the common scientific issue of the dimensionality of data—even if the data are, strictly speaking, not fractals. Data dimensionality is usually investigated by principal component analysis (PCA) or factor analysis (Gorsuch, 1983), but not with fractal dimension. The latter is normally used only if the analyzed objects are already known to have (or are expected to have) fractal properties (e.g., a chaotic attractor) (Strogatz, 1994). However, this need not be the case. Much insight about datasets commonly used in scientific research (e.g., those that are described typically by the general linear model—GLM) can be gained by investigating the dimensionality of non-fractal data with fractal dimension (Lutzenberger et al., 1992; Pereda et al., 1998; Woyshville and Calabrese, 1994).

* Corresponding author at: Department of Neurophysiology Max-Planck-Institute for Brain Research, Deutschordenstr. 46, 60528 Frankfurt am Main, Germany. Tel.: +49 69 96769 736; fax: +49 69 96769 327.
*E-mail address:* danko@mpih-frankfurt.mpg.de (D. Nikolić).

0165-0270/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.jneumeth.2008.04.007

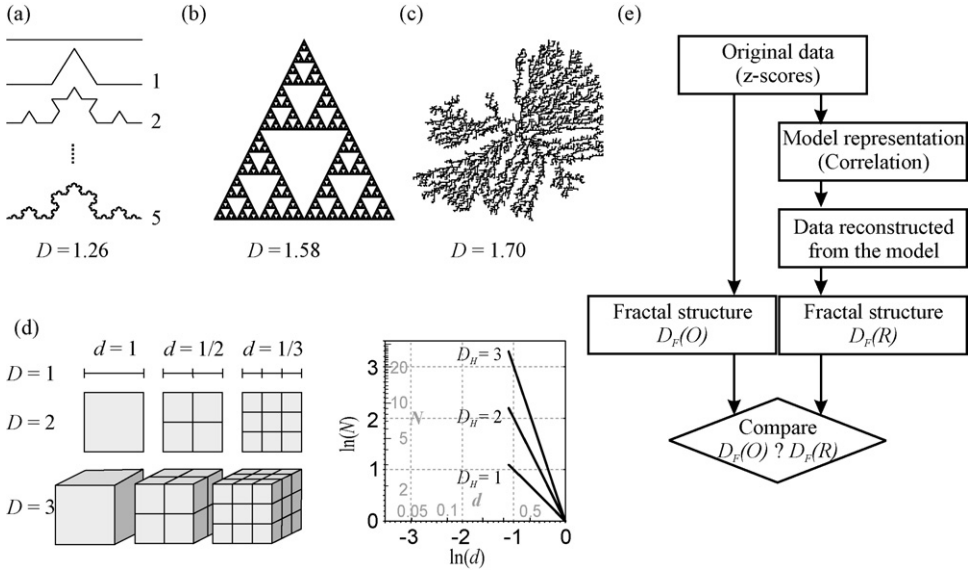


**Fig. 1.** The concept of fractal dimension, its measurement, and the proposed analysis steps necessary to investigate how well a model of a dataset accounts for the data's fractal structure. (a) Example Koch fractal and the process of its creation through iterative steps. (b) Sierpinsky (triangle) fractal with dimension larger than that in (a). (c) A fractal obtained through diffusion-limited aggregation of particles undergoing a random walk (Landau and Paez, 1997). The corresponding dimensions of fractals, $D$, are indicated. (d) The procedure for computing the fractal dimension is illustrated for three objects with true $D$-values = 1–3. The space occupied by the object is partitioned into 'boxes' of size $d$. The number of resulting boxes that intersect (or cover) the object, $N$, is counted. Finally, $N$ is plotted against the size $d$ in a log–log plot (right panel). If the object is a fractal, the plot results in a straight line and the slope of the line indicates the dimension of the object. (e) The dimensionality analysis of the data consists of two tracks. In one, the fractal structure of the original data is calculated and in the other, the fractal structure of data reconstructed from the model is calculated (in all examples we use GLM). Finally, the two structures are compared (e.g., by comparing their log–log plots).

### 1.2. Measuring fractal dimension

Fractal dimension is formally computed by one variant of the Hausdorf dimension, $D_H$ (Falconer, 2003). The principles of this calculation are shown in Fig. 1d by a variant of Hausdorf dimension known as box-counting method. Here, with a change in the analysis scale, $d$, a different number of boxes, $N$, is needed to cover the object. For example, for one-, two-, and three-dimensional objects in Fig. 1d (left panel) the counts are 2, 4, and 8 for $d = 1/2$ and 3, 9, and 27 for $d = 1/3$, respectively. The dimensions of the objects are then calculated by plotting $\ln(d)$ versus $\ln(N)$ and calculating the absolute values of the slopes of the fitted straight lines (right panel). Therefore, the dimension $D_H$ is the absolute value of the exponent in $N \approx d^{-D}$, describing how quickly the count $N$ grows with the decrease in $d$.

In the present study we estimate fractal dimensions by a numerical procedure that is more computation-effective than box-counting methods and that is known as the correlation dimension, $D_F$ (Camastra and Vinciarelli, 2002; Grassberger and Procaccia, 1983). In most cases, $D_F$ produces identical results as $D_H$ (within numerical limits) while in other cases $D_F < D_H$, the differences being very small. Thus, $D_F$ can be considered a lower estimate of $D$. Numerical details for the calculation of $D_F$ are provided in Section 2.

Central to our analyses are the log–log plots such as the one shown in Fig. 1d. For successful application of the method, it is not necessary that the data exhibit the actual properties of fractals. Fractals produce a straight line in the log–log plot (self-similarity) while plots for the data might have curvatures (changes in the slope). Curvatures provide important information about alterations in data dimensionality across different scales, indicating that the data are not simple fractals but could be instead described as multifractals, which can in turn lead to the discovery of interesting data properties (e.g., Feder, 1988, pp. 185–186). One important application is the comparison between the log–log plots for the original data and those for samples recreated by a model of the data. This allows one to test, in a novel way, how well the model accounts for the original data (for the present analyses only GLM models are tested, Fig. 1e). An example application to real data is made for local field potentials (LFP), simultaneously recorded with 16 electrodes from cat visual cortex.

## 2. Materials and methods

### 2.1. Experimental procedures

Intracranial LFP recordings were performed on an adult cat under anesthesia induced with ketamine and maintained with halothane and a mixture of $N_2O$ (70%) and $O_2$ (30%). The cats were paralyzed with intravenously applied pancuronium bromide (Pancuronium, Organon, 0.15 mg kg$^{-1}$ h$^{-1}$). LFP activity was recorded from area 17 with 16-channel silicon probes (organized in a $4 \times 4$ spatial matrix) which were supplied by the Center for Neural Communication Technology at the University of Michigan (Michigan probes). The inter-contact distances were 200 μm (0.3–0.5 MΩ impedance at 1000 Hz). Signals were amplified 1000× and filtered 1–100 Hz to extract local field potentials (LFP) (1 kHz sampling rate). To evoke visual responses drifting sinusoidal gratings were presented on a 21 in. computer screen (100 Hz refresh rate) using ActiveSTIM software for visual stimulation (ActiveSTIM, high precision stimulation tool, http://www.ActiveSTIM.com). One stimulus condition is presented in total 20 times. More details on methods for data acquisition can be found in (Biederlack et al., 2006).

### 2.2. Artificially generated data

The artificial datasets shown in Fig. 2a–c (2000 points each) were generated by a help of a Mersenne Twister pseudo random-number generator (Matsumoto and Nishimura, 1998). In Fig. 2a

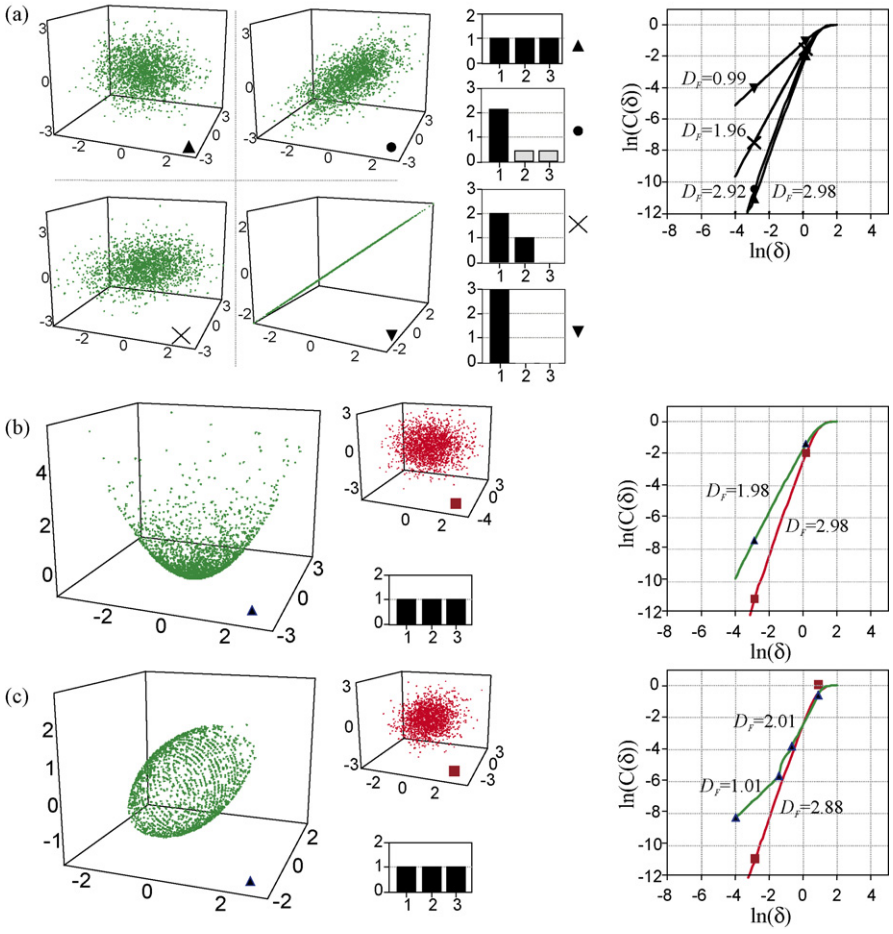


**Fig. 2.** The dimensionality of artificial datasets, each consisting of three-variables, are investigated with PCA and with fractal dimensionality. (a) Four different correlation patterns that satisfy the assumptions of GLM. All variables were either fully independent ('ball'), or only two were fully dependent (Pearson's correlation $r = 1.00$; 'disk'), or all were correlated with $r \sim 0.60$ ('egg'), or all were correlated with $r = 1.00$ ('line'). The histograms in the center indicate the eigenvalues of the PCA-components after varimax rotation. The log–log plot on the right-hand side indicates the dimensionality for the four respective datasets. (b) Three variables with a strong non-linear relationship forming a 'bowl'-like shape (Original) and the data scatter reconstructed after a GLM description of the original data (Reconstructed). (c) Data that have both one- and two-dimensional structure obtained by positioning the points on a spiral running along the surface of a spindle. The beginnings and the ends of the regions in the log–log plots that were used to compute the values of $D_F$ are given by the positions of the symbols indicating the corresponding datasets. Black bars: PCA-components above a significance threshold and subjected to varimax rotation. Grey bars: the loadings of PCA-components below the significance threshold (non-rotated).

all scatter-clouds were generated with three normally distributed (Gaussian) variables under the assumption of homoscedasticity, and were correlated to a different degree. In Fig. 2b, the quadratic, non-linear 'bowl' distribution was obtained by the formula $z = 4 + (x^2 + y^2)/3.5$ where $z$, $x$ and $y$ are the three variables. In Fig. 2c the data that form a shape of a spiral with expanding and contracting radius (a 'mandrel' shape) were created according to the formulas: $y = \cos(2\pi \times 0.065x) \times \sin(2\pi \times 5x)$, and $z = \cos(2\pi \times 0.065x) \cos(2\pi \times 5x)$. The pseudo-data for the questionnaire about beer quality were taken, with permission, from an SPSS tutorial on factor analysis (principal components analysis—SPSS; http://core.ecu.edu/psyc/wuenschk/MV/FA/PCA-SPSS.doc).

### 2.3. Data analysis

The analysis steps made for all artificial and real data in the present study are shown in Fig. 1e. Prior to the analysis, all data are normalized to $z$-scores (mean = 0; standard deviation, S.D. = 1). The data are either directly analyzed for fractal structure or the data are reconstructed from a GLM. In the latter case, Pearson correlation coefficients are first computed. Then, by making a Cholesky decomposition of the correlation matrix (Bock and Krischer, 1998) with routines from The GNU Scientific Library (http://www.gnu.org/software/gsl/), we reconstruct the dataset with the properties assumed by the model (e.g., normal distribution, homoscedasticity). The resulting scatter-clouds are shown in Figs. 2b, c and 3 in red color (For interpretation of the references to colour in this text, the reader is referred to the web version of the article.). The fractal structures (i.e., log–log plots) of reconstructed data are then compared to those of the original data.

The correlation dimension, $D_F$, uses a somewhat different approach for finding scale-versus-count relationships than the box-counting method. Instead of counting the number of boxes that cover a certain object, $D_F$ is based on a count of pairs $(x_i, x_j)$ of points, that belong to the object and can be bound within spheres of a certain size, $\delta$ (Grassberger and Procaccia, 1983). This count is known as the *correlation integral* and is theoretically given by:

$$C(\delta) = \lim_{n \to \infty} \frac{1}{n^2} |\{(x_i, x_j), \quad \text{where} |x_i - x_j| < \delta \text{ and } i \neq j\}|$$

For a set of $n$ empirical measurements, $n^2$ pairs are being considered for the count. If the analyzed object has a fractal structure,

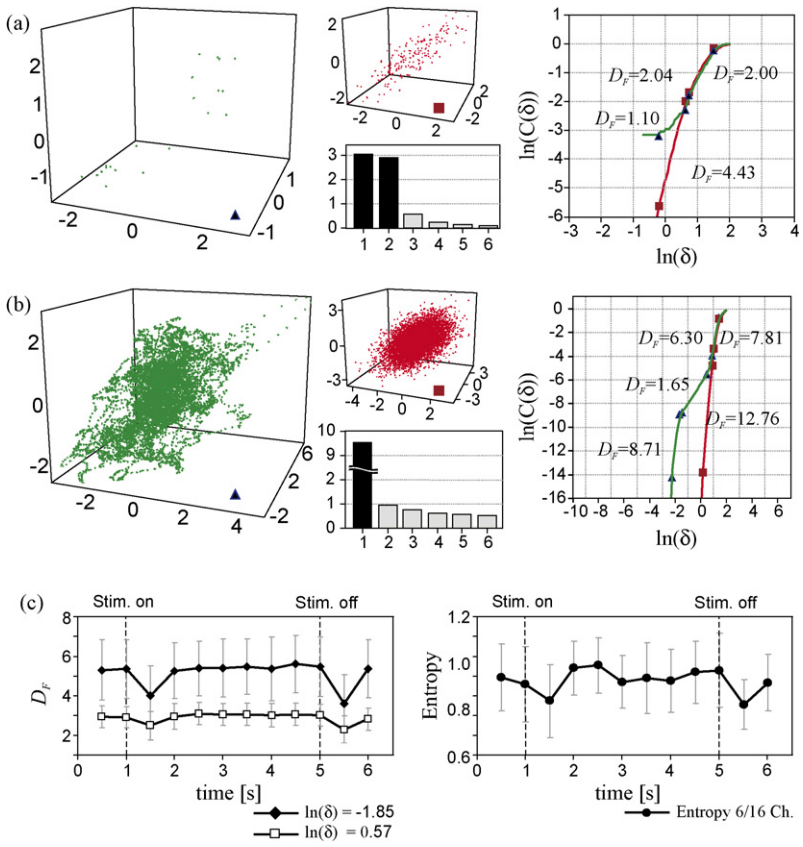


**Fig. 3.** Dimensionality of datasets consisting of larger number of variables. (a) Fictitious reports to a questionnaire on beer quality (7 variables, 200 data points). (b) Eight seconds of multi-channel recordings of LFP signals from cat visual cortex (7977 data points). In (a) and (b), all scatters are plotted for only three variables, while the dimensionality analyses were made for all 16 variables. The results for PCA analysis are shown only for the six factors with largest loadings. The organization of the plots and the notation are the same as in Fig. 2. (c) Temporal evolution of correlation dimension and entropy along an experimental trial during which a grating stimulus is presented. (Left panel) Correlation dimension computed for all 16 channels at two different scales and for non-overlapping time windows 500 ms in duration. (Right panel) Entropy of the system computed for all possible groups of six channels and averaged subsequently (the corresponding *y*-scale is shown on the right). Vertical lines: standard deviation estimated from 20 stimulus repetitions (trials). Thick vertical dotted lines indicate stimulus onset (at 1 s) and offset (at 5 s).

$C(\delta)$ can be approximated by a power law: $C(\delta) \approx \delta^D$, $D_F$ being given by the slope of a line in an $\ln(C(\delta))$ versus $\ln(\delta)$ plot. Note that due to the definition of $D_F$, the lines in the log–log plots of Figs. 2 and 3 have positive slopes rather than negative as in the theoretical plots of Fig. 1d. Also note that both $D_F$ and $D_H$ are positive quantities (dimensionality measures). Fractal dimensions were estimated separately for segments of the log–log plots that showed approximately constant slopes. The present methods for the estimation of fractal dimension are related to previous ones used for measuring fractal dimensions of electroencephalographic (EEG) signals. Lutzenberger et al. (1992) used hypercubes for counting the number of embedded points at different scales, Woyshville and Calabrese (1994) counted circles with increasing radii that had to cover the EEG traces, while Pereda et al. (1998) used a modified version of the correlation integral.

To compare fractal dimensions to the classical measures of dimensionality, we also compute PCA for each dataset. All PCA analyses were made with *varimax* rotation applied after the selection of factors with significant loadings. For this selection we used the criterion that the eigenvalue (total explained variance) should exceed the threshold of 0.5 for data with three variables or the threshold of 1.0 for data with more than three variables. PCA was computed using the Open Computer Vision Library (http://www.intel.com/technology/computing/opencv/).

Our multichannel recordings provided multi-dimensional datasets directly. Thus, in the present analysis we did not have to use Taken's embedding theorem to reconstruct the phase space from a single variable, as it has been done in some other analyses (e.g., Andrzejak et al., 2001; Liebovitch, 1988, p. 212). Our approach was expected to give more reliable results because it did not suffer (or suffered to a much lesser degree) from various issues such as an inappropriate choice of the embedding parameter 'lag', lack of stationarity or smoothness in the data, and the need for an excessively large number of data points to estimate the embedding dimension reliably (Liebovitch, 1988; Osborne and Provenzale, 1989). Another advantage of avoiding attractor reconstruction is that there is no need to generate surrogate data and test statistically whether the time-series contains potentially multi-dimensional structure that cannot be explained by a linear stochastic stationary process, usually Gaussian (e.g., Andrzejak et al., 2001; Schreiber and Schmitz, 2000). All presently discussed analyses assume that the multiple variables are measured simultaneously and hence, that the phase space is accessible directly.

We also compared the measures of fractal dimensionality with those of entropy. To compute entropy of the 16-channel signals, we first turned each continuous LFP into discrete values of 1 s and 0 s, depending on whether the measured voltage for a given sample had a value above or below the zero-line, respectively. The zero-line was estimated as the gross average of the signal over the entire recording period. Across all 16 signals recorded simultaneously this procedure would produce a 16-bit pattern for each sampling point and hence, a maximum of $I = 2^{16} = 65{,}536$ different patterns could be generated. With sufficiently long recordings, for each pattern $i$ (where $\{i \in 1, \ldots, I\}$) the probability of occurrence, $p_i$, could be

estimated. The entropy, $S$, of the system is then computed simply as:

$$S = -\sum_i p_i \log(p_i).$$

Given the 1 kHz sampling rate, 1000 sample patterns would be obtained per one second. In practice, this was not sufficient to estimate the probabilities for all the 16-bit patterns because, to detect the changes that might have occurred as a function of the various stimulus events along the trial, we needed to calculate entropy in a time-resolved manner, using short windows of 500 ms length. As a consequence, the large dimensionality of the 16-channel system became a particularly strong limitation. Therefore, we needed to limit our analysis to the number of channels that allowed accurate estimates of the probabilities with a number of samples that totaled only 500 samples/trial × 20 trials = 10,000 samples. Hence, the probabilities, $p_i$, could be calculated reliably only with about 6 channels at a time ($2^6 = 64$ patterns). To calculate the entropy for a 16-channel dataset, we averaged entropies obtained for all possible 6-channel subsets (in total 8008 combinations).

## 3. Results

### 3.1. Fractal structure of data satisfying GLM assumptions

We first investigated the fractal structure of three-dimensional data that were generated under the assumptions of GLM. To this end, we created four different datasets and applied PCA and fractal analysis (Fig. 2a). There are two important findings. First, all datasets showed straight lines in the log–log plots (the slopes of the lines stayed constant across different $\delta$s), indicating that the data satisfying GLM assumptions can sometimes have dimensionality-properties similar to those of fractals. As we will show later, this is not always the case.

The second finding was that the slopes in log–log plots did not always match the dimensionality estimated by PCA. In the extreme examples, having all variables correlated with $r = 0$ or $r = 1$ (see Section 2.2 and Fig. 2a, 'ball' and 'line'), the match was good (i.e., a three-factor PCA was associated with $D_F = 2.98$ and a one-factor PCA with $D_F = 0.99$). Also, for the case with two perfectly correlated variables and one uncorrelated (Fig. 2a, 'disk'), PCA revealed two factors, in agreement with $D_F = 1.96$. However, when the variables had more realistic correlations of $r = 0.6$ (Fig. 2a, 'egg'), only one significant factor was found by PCA while fractal analysis indicated larger dimensionality ($D_F = 2.92$). Therefore, the results of fractal analysis are not straightforward given the known GLM properties of the data, indicating that the former provides different and thus, complementary information to that of the latter.

### 3.2. Non-linear data

An interesting case for illustrating data-to-model comparisons are the deviations from model assumptions. In the case of GLM, this can be made by introducing non-linear relationships between variables. In Fig. 2b a two-dimensional dataset is curved in a three-dimensional space forming a shape of a bowl. PCA indicates incorrectly that the data have three dimensions and so is the case for the fractal analysis of a GLM-reconstructed dataset ($D_F = 2.98$). Only the fractal dimension of the original data reveals the true two-dimensionality of the data ($D_F = 1.98$).

In Fig. 2c an even more interesting case of a non-linear relation in the data is introduced. Here, the data points are located on a spiral (a one-dimensional object) that runs around the surface of an imaginary spindle (a two-dimensional shape), resulting in a 'mandrel'-like shape. In this case, a scatter-plot for any pair of variables is hardly distinguishable from scatters that satisfy GLM assumptions (not shown). Consequently, both the factor loadings of PCA and the fractal dimension of GLM-reconstructed data ($D_F = 2.88$) indicate three dimensions (incorrectly). However, the analysis of the fractal structure of the original dataset reveals the geometric properties of the data much more accurately. This analysis shows both underlying dimensions of the data: at smaller scales, the analysis suggests that the data are one-dimensional ($D_F = 1.01$) and at larger scales, the slope of the log–log plot changes into a two-dimensional structure ($D_F = 2.01$), correctly describing the real geometric structure of the data.

### 3.3. Application to realistic datasets

The examples in Fig. 2 used data that either fitted ideally GLM assumptions or had perfect non-linear properties (i.e., no measurement errors). In practice however, data will not have such properties exclusively but will instead combine different features. The two datasets in Fig. 3 are examples of such, more realistic cases, in which a satisfactory match to GLM is achieved only at large scales while, at small ones, the data depart from GLM assumptions.

In Fig. 3a, 200 hypothetical subjects have been asked to grade seven qualities of beer (cost, size, alcohol, reputation, color, aroma and taste) on a scale 0–100 in steps of five (Wuensch, 2005; principal components analysis—SPSS; http://core.ecu.edu/psyc/wuenschk/MV/FA/PCA-SPSS.doc). The cloud-scatter (plotted only for three out of seven variables) shows a much more sparse structure than that of the GLM-reconstructed data. This difference could be tracked to the poor measurement resolution of the questionnaires: due to the discrete nature of the variables (taking values with increments of 5), many measurement points overlapped—resulting in a sparse scatter. The analysis of fractal structure has captured this difference in a form of strong reduction in the data dimensionality at small scales: while at large $\delta$s, the original and reconstructed datasets had very similar dimensionalities ($D_F = 2.00$ vs. 2.04), indicating high suitability for GLM-representation, at small $\delta$s, the two measures strongly disagreed. At small scale, the reconstructed dataset increased its dimensionality to about double the value at the large scale (i.e., $D_F = 4.43$ compared to 2.00), indicating that GLM does not necessarily have the fractal property of self-similarity across different scales (as was the case in Fig. 2a). In contrast to this increase in the dimensionality, the decrease in $\delta$ resulted in a reduction of the dimension of the original dataset, first to ∼1 and then to 0 (indicating complete overlap in data points). PCA analysis suggested the dimensionality of two, which was only consistent with the fractal estimates of data-dimensionality at large scales.

Our most important example application of fractal analysis is to LFP data recorded from cat visual cortex in response to visual stimulation. In Fig. 3b we show the analysis of a data-segment with a length of about 6 s and recorded across 16 electrodes positioned along a regular grid 600 μm × 600 μm in size. The dimensionality analyses (both $D_F$ and PCA) were made with all 16 electrodes while only the first three channels are shown in the scatter plot. The scatter plot reveals a rich structure that has many details and varies in the properties across scales but also across positions in space. In some cases the plot suggests line-like trajectories (with a dimensionality close to 1), while in others, it suggests highly dense knots of what appears to be a complex structure. The scatter plot of the reconstructed data indicates that this rich structure is not captured well by GLM already at the three-variable level and this is confirmed by the analysis of fractal structure at the level of all 16 variables. The reconstructed data show the typical monotonic change in dimensionality (slope) from larger ($D_F = \sim 13$) to smaller ($D_F = \sim 9$) with an

increase in $\delta$. At large scales the two types of datasets had relatively similar dimensions ($D_F = 6.3$ vs. 7.81), indicating that correlations based on GLM described well the properties of the data at the global level (e.g., the amount of variance shared by the variables).

When looking into more detail (smaller scales) large differences in dimensionality could be found. For small scales ($\ln(\delta) < 0$), GLM-reconstructed data could not be measured because of insufficient number of data points. Below a certain scale ($\ln(\delta) < 1$), GLM-reconstructed data had a constant, very high dimensionality ($D_F = 12.76$; still less than the maximum of 16 dimensions), while the dimensionality of the original data first dropped to a very small value ($D_F = 1.65$). This low dimensionality most likely reflects the temporal dependencies in the LFP signals that produce a trajectory-like behavior in scatter plots—a feature of the signal that is not assumed by GLM but that is detectable by our fractal analysis. Finally, with further decrease in $\delta$, LFP signals show again an increase in dimensionality ($D_F = 8.71$). This high dimensionality for very small changes in the voltage of LFPs reflects the complex 'knots' of the activity seen in the scatter plot. PCA analysis is totally uninformative regarding this rich structure, as it suggests that a single factor explains most of the variability in the data. Moreover, the dimensionality of LFP signals is not necessarily stationary over time. In Fig. 3c we show that the correlation dimension can change along experimental trials as a function of the stimulus dynamics. Here, a transient but strong drop in $D_F$ is observed following onset and offset of a sinusoidal grating stimulus. Therefore, fractal analysis can extract a lot of information that is not accessible to GLM. The results indicate that, in the case of LFPs, GLM produces a rather incomplete picture of the true property of signals. In other words, we can say that, by the criteria of fractal dimension, GLM is not an appropriate model for LFP data.

### 3.4. Comparison to a measure of entropy

It is to be expected that a decrease in fractal dimensionality is concomitant with a decrease in the entropy of the data. We calculated entropy across the simultaneously recorded LFP-signals by estimating the frequencies of the binary patterns generated by the zero-crossings of the electrical signals (see Section 2 for details). The comparison of the results to those of fractal dimensionality indicated that the two measures provided similar results (Fig. 3c, right panel). The stimulation periods that had smallest fractal dimensionality had also smallest entropy, suggesting relatedness in the type of information probed by these two measures. This was supported further by the inability of the data reconstructed by a GLM to reproduce, at least in some cases, the entropy of the original data (results not shown).

## 4. Discussion

We have shown that measures of fractal dimension can be applied to multivariate data that are not fractals and that, by doing so, one can extract important information about the properties of the data. In this analysis, the dimensionality of the data is investigated across different scales. A change in dimensionality is indicative of a change in the data's geometric properties, the detailed identification of which might require further analysis steps (e.g., inspecting scatter-plots or phase diagrams). As a result, the present method can help to discover relationships between variables that could not be discovered by other, standard methods (e.g., GLM). Here, small fractal dimensions suggest correlations between variables and large dimensions can, but do not have to, indicate noise.

The present method is also suitable for testing whether models of the data account for (fit) the geometric properties of the data. We illustrate this point by using GLM only, but an application to any other type of model would be identical: the fractal structure of the original data needs to be compared to the fractal structure of the data reconstructed from the model of the data. A mismatch allows us to detect the scale at which the model and the data disagree. In the case of GLM, this allows us to detect, for example, non-linear relationships in the data.

It is important to note that the present analysis method is not free of assumptions, as it assumes that data have spatio-temporal stationarity (Woyshville and Calabrese, 1994), i.e., the geometric properties of the data are similar across different locations in the state-space. If data clearly violate this assumption, one might consider applying the analysis only over the parts of the state-space that exhibit stationarity or, in time-series, applying a time-resolved analysis (as we have done).

Importantly, information obtained from the numeric analysis of dimensionality is highly consistent with the relationships observable in scatter-plots (phase diagrams). The advantage of the numeric methods is that the dimensionality can be quantified and can be probed for spaces that exceed grossly the maximum of three dimensions that can be visualized in graphical representations (in our case, up to 16 dimensions). We have also shown that fractal dimensionality provides information related to the complexity of a system assessed by entropy. Higher dimensionality is associated with higher entropy. However, this does not mean that the two provide identical information or that entropy can be considered a replacement for fractal dimensionality. Accurate assessment of entropy in many cases requires prohibitively large number of samples. In addition, entropy does not normally provide information associated with the analyses at different scales in the way provided by fractal dimensionality. Advanced methods for estimation of entropy can also incorporate some scale-based information, which is usually defined in the temporal dimension. For example, multiscale entropy (Costa et al., 2005) applies first a 'coarse graining' of the time-series using time-windows of a given size (scale) and then computes entropy using techniques less sensitive to the size of the dataset, such as sample entropy (Richman and Moorman, 2000). Thus, entropy measures can complement the analysis of fractal dimensionality in cases in which the temporal scale is an important factor. For applications in which the amplitude of the signal is most relevant, fractal dimension should be the method of choice for determining the complexity of a dataset.

The highly dimensional knots of small changes in voltages (small scales) in our analyses are not likely to reflect noise. Instead, their sporadic nature and intermittent exchange with the periods of smooth, large changes in voltages may reflect computationally relevant events that establish new phase relations between the signals and/or mark new stages in the processing of the stimulus.

The analysis of dimensionality is also important as an estimate of the brain's ability to represent information. A state of very low dimensionality (i.e., low entropy) would indicate that the system lacks the degrees of freedom necessary to represent information (e.g., Schneidman et al., 2003), which would then also prevent its processing. Thus, if the system is not able to assume a variety of different states, the computational flexibility is lost, which means loss of brain's function. In the case of epileptic attacks the low dimensionality (Babloyanz and Destexhe, 1986) may be even the cause of the loss of consciousness. It is therefore, mandatory to understand how the dimensionality of the brain activity changes as a function of oscillations that are considered healthy and beneficial to information processing (e.g., those in the beta/gamma range 20–80 Hz), as a function of stimulus properties (e.g., those inducing strong and

weak oscillations), and under different brain states (e.g., anesthesia, sleep, awake, focused attention).

In conclusion, the present method should be useful for analysis of multivariate data and in particular for those data that involve relationships between a large number of variables. The method can be used as a data-mining tool, as a tool for testing how well a model accounts for the data, or as an estimate of the system's ability to represent information.

## Acknowledgements

The authors would like to thank Julia Biederlack for help with the acquisition of LFP data and Robert Terry and Shan Yu for fruitful discussions. We are thankful to Karl L. Wuensch for the permission to use the 'beer' data. This study was partially supported by The Hertie Foundation, Alexander von Humboldt Stiftung, the computing facilities of the University of Oklahoma, a grant from the Deutsche Forschungsgemeinschaft number NI 708/2-1, a Max Planck—Coneural Partner Group, and by the following grants of the Romanian Government (RP-5/2007: contract 1/01.10.2007, ID_48/2007: contract 204/01.10.2007, NEUROBOT: contract 11039/18.09.2007).

## References

Andrzejak RG, Lehnertz K, Mormann F, Rieke C, Peter David P, Elger CE. Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: dependence on recording region and brain state. Phys Rev E 2001;64:061907.

Babloyanz A, Destexhe A. Low-dimensional chaos in an instance of epilepsy. Proc Natl Acad Sci USA 1986;83:3513–7.

Biederlack J, Castelo-Branco M, Neuenschwander S, Wheeler DW, Singer W, Nikolić D. Brightness induction: rate enhancement and neuronal synchronization as complementary codes. Neuron 2006;52:1073–83.

Bock RK, Krischer W. The data analysis briefbook. Berlin, Heidelberg: Springer-Verlag; 1998.

Brand M. Charting a manifold. Proceedings of neural information processing systems, vol. 15. MIT Press; 2003.

Camastra F, Vinciarelli A. Estimating the intrinsic dimension of data with a fractal-based method. IEEE Trans Pattern Anal Mach Intell 2002;24:1404–7.

Costa M, Goldberger AL, Peng C-K. Multiscale entropy analysis of biological signals. Phys Rev E 2005;71:021906.

Falconer K. Fractal geometry: mathematical foundations and applications. New York, NY: John Wiley & Sons; 2003.

Feder J. Fractals. New York: Plenum Press; 1988.

Gorsuch RL. Factor analysis. 2nd ed. Hillsdale, NJ: Erlbaum; 1983.

Grassberger P, Procaccia I. Measuring the strangeness of strange attractors. Physica D 1983;9:189–208.

Landau RH, Paez MJ. Projects in computational physics. New York, NY: John Wiley & Sons; 1997.

Levina E, Bickel PJ. Maximum likelihood estimation of intrinsic dimension. Adv Neural Inform Process Syst 2004:17.

Liebovitch LS. Fractal and chaos simplified for the life sciences. Oxford, New York: Oxford University Press; 1988.

Lutzenberger W, Elbert T, Birbaumer N, Ray WJ, Schupp H. The scalp distribution of the fractal dimension of the EEG and its variation with mental tasks. Brain Topogr 1992;5:27–34.

Matsumoto M, Nishimura T. Mersenne twister: a 623-dimensionally equidistributed uniform pseudo-random number generator.ACM transactions on modeling and computer simulation, vol. 1. New York, NY: ACM Press; 1998, 3–30.

Osborne AR, Provenzale A. Finite correlation dimension for stochastic systems with power-law spectra. Physica D 1989;35:357–81.

Pereda E, Gamundi A, Rial R, González J. Non-linear behaviour of human EEG: fractal exponent versus correlation dimension in awake and sleep stages. Neurosci Lett 1998;250:91–4.

Richman JS, Moorman JR. Physiological time-series analysis using approximate and sample entropy. Am J Physiol Heart Circ Physiol 2000;278:2039–49.

Schneidman E, Still S, Berry II MJ, Bialek W. Network information and connected correlations. Phys Rev Lett 2003;91:238701.

Schreiber T, Schmitz A. Surrogate time series. Physica D 2000;142:346–82.

Strogatz SH. Nonlinear dynamics and chaos: with applications to physics, biology, chemistry and engineering. New York, NY: Addison–Wesley Publishing Company; 1994.

Tenenbaum JB, Silva VD, Langford JC. A global geometric framework for nonlinear dimensionality reduction. Science 2000;290:2319–23.

Woyshville MJ, Calabrese JR. Quantification of occipital EEG changes in Alzheimer's disease utilizing a new metric: the fractal dimension. Biol Psychiatry 1994;35:381–7.