Temporal and Spatial Complexity measures for EEG-based Brain-Computer Interfacing

Stephen J. Roberts, William Penny & Iead Rezek

Neural Systems Research Group
Department of Electrical & Electronic Engineering
Imperial College of Science, Technology & Medicine
London, UK
{s.j.roberts, w.penny, i.rezek}@ic.ac.uk

January 8, 1998

Abstract

There has been much interest recently in the concept of using information from the motor cortex region of the brain, recorded using non-invasive scalp electrodes, to construct a crude interface with a computer. It is known that movements of the limbs, for example, are accompanied by desynchronisations and synchronisations within the scalp-recorded electroencephalogram (EEG). These Event-Related Desynchronisations and Synchronisations (ERD & ERS), however, appear to be present when volition to move a limb occurs, even when actual movement of the limb does not in fact take place. The determination and classification of the ERD/S offers many exciting possibilities for the control of peripheral devices via computer analysis. To date most effort has concentrated on the analysis of the changes in absolute frequency content of signals recorded from the motor cortex. We present results in this paper which tackle the issues of both the interpretation of changes in signals with time and across channels with simple methods which monitor the temporal and spatial 'complexity' of the data. Results are shown on synthetic and real data sets.

Keywords: EEG analysis, signal processing, complexity analysis.

1 Introduction

There are still clear limitations in the way that human beings, able-bodied or not, interface with computers and associated peripheral equipment and the need exists for advances in novel computer interfacing techniques. One of the most challenging and exciting is in the appropriate use of biosignals directly recorded from the body of the user. We, in common with the terminology of other authors (most notably Pfurtscheller *et al.*, 1993), refer to an interface which uses the electrical activity of the brain (the electroencephalogram or EEG) as its information source as Brain-Computer Interfacing (BCI).

2 Signal recording

Non-invasive electrodes are used to record EEG potentials, in particular the Event-Related Desynchronisations and Synchronisations (ERD & ERS) which appear to be associated with the planning of muscle activity, especially of the fingers, limbs and tongue (Pfurtscheller *et al.*, 1993; Pfurtscheller *et al.*, 1994). The most-noted report of the ERD is that of a blocking of fundamental EEG rhythms, most notably the μ-rhythm (8-12Hz). It has also been shown recently (Pfurtscheller *et al.*, 1994) that as well as the μ-blocking phenomenon, there is a significant ERS within the γ-band (30-40Hz) in some subjects, although this is not generally present. The important point about such changes is that even if the intended movement does *not*, in fact, take place, the changes still appear to occur.

Previous publications in this field have concentrated on optimising the positioning of electrode sites on the head and some researchers have stated that 'optimal' electrode positioning is indeed over the motor cortex (Pfurtscheller *et al.*, 1993; Pfurtscheller *et al.*, 1994) and that only a small number of electrodes (2-8) appear to be required to extract some event related information.

Training of a subject to use the BCI, involving the correlation of these motor-related changes in EEG to desired movements of a cursor on a computer screen and can require extensive training periods of up to a day or more (Pfurtscheller *et al.*, 1993). Data are collated and a variety of filtering and preprocessing techniques employed to extract, primarily, frequency information from the EEG. Some form of subsequent statistical pattern classification is then utilised in an attempt to achieve a correlation with the desired cursor movements.

We record a total of 25 channels of information, of which 22 are EEG channels, positioned on the head over the primary motor cortex as shown in Figure 1. The other three channels consist of a single EOG channel (taken as a bipolar signal across the eyes from approximately 1cm to either side of the outer canthi) and two EMGs taken from the the belly of the forearm to check for any muscle activity during the experiments. All electrodes are standard Ag - AgCl and electrode impedances are checked throughout the experiment and kept below 5kΩ.

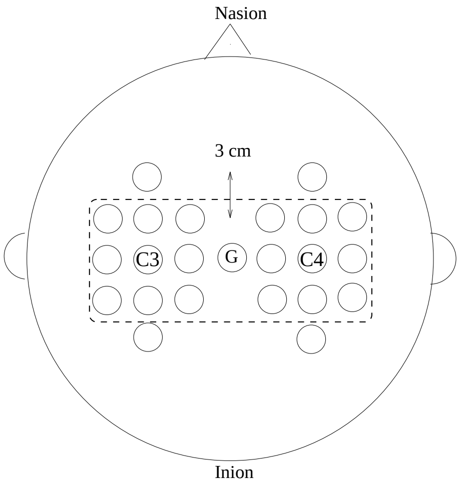


Figure 1: **Electrode placement:** Electrodes are placed in arrays over the primary motor cortex. 'G' denotes the ground electrode and the dashed box contains the 18 central electrodes used for the analysis in this paper. Electrodes at C3 and C4 are also indicated. The latter correspond to electrode sites in the standard 10-20 system.

Recordings from a total of eight subjects were made in this initial study (seven able-bodied volunteers and one consenting patient. The latter (subject GW) has motor impairment as a result of moderate brain injury). All recordings had formal ethical consent. Our recording protocol owes much to those published in (Pfurtscheller *et al.*, 1994). Subjects are sat before a screen on which they are presented with movement cues (Figure 2). Subjects perform a series of runs each with six cues after which they imagine opening and closing the fingers of either left or right hand (we refer to these runs as *externally paced*). Subjects are also asked to perform a series of *self- paced* imagined movements in which they imagine opening and closing the fingers of either right or left hand at intervals not less than 10 seconds (self-assessed). There is no external cueing present during these runs.

All signals are digitised to 16 bit accuracy at a sampling rate of 125 Hz. A digital 63 tap FIR (Hamming-window) anti-aliasing filter is used with a 3dB cut-off at 50Hz. During the validation phase of our techniques the data are stored and analysed off-line.

3 Methods

3.1 Why complexity measures?

Traditionally, the EEG signal has been divided, either visually or automatically, by means of a crude amplitude assessment (requiring, therefore, some calibration measures to be recorded intermittently)

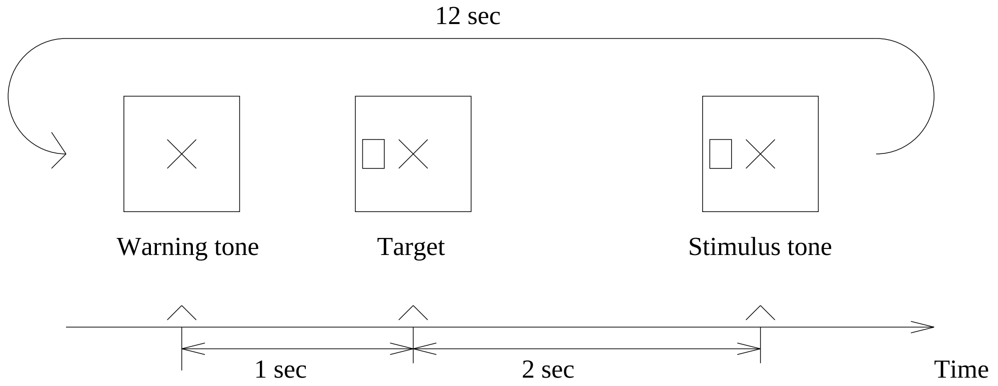


Figure 2: **BCI data recording protocol:** Subjects are given a 'get ready' warning, shown which hand to imagine movement of, and then given a further auditory prompt when movement is to be imagined.

or (more successfully) by means of its dominant frequency content. The latter gives rise to the set of frequency bands, *alpha, beta, delta* etc. commonly referred to in EEG publications. In Rezek & Roberts (1996) we have argued that changes in the dominant rhythms in the EEG are manifestations of more profound changes in the interactions between ensembles of neurones and as such are best analysed using methods which assess the 'complexity' of the data sequence. Such measures are, by and large, independent of the precise frequency content of the signal and are hence ideally suited in cases where measures which are robust across individuals are required (the 8-12Hz rhythm (mu), for example, can vary considerably from person to person).

3.2 There is no absolute 'complexity'

Many different methods and measures of signal, or system, 'complexity' have been proposed. Some are not well-suited to analysis of bio-signals, as they are notoriously variant in the presence of noise and non-stationarities (the classic examples being the correlation dimension (Grassberger & Procaccia, 1983) and the Lyapunov exponent (Wolf *et al.*, 1985)). Several methods of assessing the 'complexity' of data sequences were developed, presented and compared in Rezek & Roberts (1996). The methods investigated were:

1. Optimal model order (OMO) under autoregressive modelling.
2. Spectral entropy (SE).
3. Approximate entropy (ApEn) (see Pincus (1991) also).
4. Embedding-space Decomposition (ESD)

On a variety of data sets, it was observed empirically that OMO failed to give a reliable estimate of system complexity, ApEn and SE gave similar results (ApEn is computationally very expensive - so SE was preferred for this reason) and ESD methods gave good results over a variety of data sequences. We concentrate in this paper therefore on the ESD method, which is detailed next.

3.3 Embedding-space decomposition

We consider an arbitrary (real) matrix $\mathbf{X}$ (not normally square). We may decompose any such $\mathbf{X}$ as

$$\mathbf{X} = \mathbf{U}\mathbf{S}\mathbf{V}^T \tag{1}$$

where $\mathbf{U}$ and $\mathbf{V}$ are orthonormal, i.e. $\mathbf{V}^T = \mathbf{V}^{-1}$. The matrix $\mathbf{U}$ is the matrix of projections of $\mathbf{X}$ onto the eigenvectors of $\mathbf{X}\mathbf{X}^T$ and $\mathbf{S}$ is diagonal with elements $S_{ii} = \sigma_i$ where σ_i^2 is the i–th eigenvalue of $\mathbf{X}\mathbf{X}^T$ and $\sigma_i >= 0$. The resultant decomposition is usually referred to as singular value decomposition (SVD).

3.3.1 Temporal information

We define the construction of the data matrix, $\mathbf{X}$, using the method of delays (Takens, 1981) and consider some window as containing m samples taken at intervals of J samples from the observed time series x_t. The elements in the window represent components of an *embedding space* in $\Re^m$. As the data sequence (time series) is repeatedly windowed, the series of vectors obtained constitutes the trajectory, or embedding, matrix. Defining

$$\mathbf{x}_i = (x_i, x_{i+J}, \cdots, x_{i+(m-1)J})^T \tag{2}$$

we obtain an embedding matrix which may be written as

$$\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_{N-(i-1)})^T \tag{3}$$

The value m is referred to as the embedding dimension and must satisfy :

$$m \geq 2M + 1 \tag{4}$$

which sets the lower bound for m given an M-dimensional manifold in the phase space. Since M is not known *a priori* means that in practice the embedding dimension m is chosen large enough such that redundancy in the embedding results. This redundancy manifests itself as a rank deficiency in the embedding matrix $\mathbf{X}$. We may investigate this redundancy by means of a singular-value decomposition, whereby we decompose $\mathbf{X}$ via Equation (1).

For a noiseless system with redundancy (i.e. m is larger than the intrinsic dimensionality of the embedded data sequence) some σ_i will zero. In real world situations, however, the observed time series will be corrupted by noise (including quantisation noise). This results in a shifting of the singular values such that

$$\sigma_i^2 \leftarrow \sigma_i^2 + \langle \xi^2 \rangle \tag{5}$$

where $\langle \xi^2 \rangle$ is the expected signal noise variance. Hence, no singular value will be zero.

As the expected noise is randomly distributed, however, it will take no preferred direction in the embedding space, thus giving rise to a noise-floor in the singular value spectrum. This is in contrast to those singular values associated with the deterministic system which will be significantly larger (see Broomhead & King, 1986). Unless stated otherwise all the examples shown in this paper used an embedding dimension of $m = 20$ and lag $J = 1$. For a more detailed discussion regarding these parameters see Kember & Fowler (1993).

As a simple example we apply this methodology to a one-dimensional sequence of 1000 points generated from the following equation

$$x_t = \sin(t/11) + \sin(t/5) + \eta_t \tag{6}$$

where η_t is a Gaussian noise source of unity variance and $t = 1, 2...1000$. The first 200 samples of the noiseless sequence is shown in plot (a) of Figure (3), plot (b) shows the noisy data sequence x_t over the same range, plot(c) the set of singular values (on a logarithmic scale) and plot (d) the first two components of the projection matrix, $\mathbf{U}$.

To define a measure of dynamic complexity, we note that, as the number of apparent dynamic components[1] increases, so the number of singular values above the noise floor increases. In the limit, when the system becomes complex enough to be indistinguishable from any additive noise, all singular values lie on the noise floor and have similar magnitude. We may define, therefore, a pragmatic measure of this trend using the entropy of the singular value spectrum. Normalising the singular values such that $\overline{\sigma}_i = \sigma_i / \sum_{i'} \sigma_{i'}$ we define

$$H = -\sum_{i=1}^{N} \overline{\sigma}_i \log \overline{\sigma}_i \tag{7}$$

[1]This may not, of course, be the true number of components - the apparent number is conditional on the sample size, the embedding dimension etc.

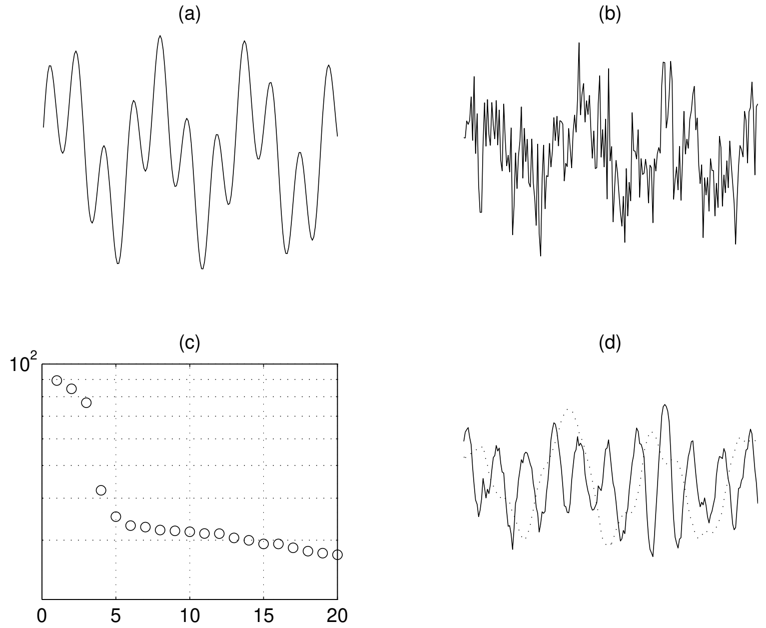


Figure 3: **Simple sine wave example:** (a) Noiseless data sequence, (b) Noisy data sequence, (c) Spectrum of singular values, σ_i, and (d) first two components of the projection matrix, $\mathbf{U}$.

We choose this form so that systems of decreasing complexity will give rise to decreasing entropy measures. Entropy may be defined equivalently in terms of a distribution or as a function of available system states. For a system with Ω available states the entropy is given as $H = \log \Omega$. If we choose 2 to be the logarithmic base then we may write

$$\Omega = 2^H \tag{8}$$

We will use this entropy-based measure as an indicator of system complexity.

Figures 4(a) and (b) respectively show the singular-value spectra for increasing numbers of sinusoidal components and for chaotic data sequences of increasing 'complexity' (degrees of freedom). Note that we did not add noise to the signals for this demonstration. Note also that the number of 'significant' singular values increases with increasing signal 'complexity' and that this trend is seen in the case of periodic and non-periodic systems. An increasing number of significant singular values is equivalent to an increase in the entropy defined via Equation (7).

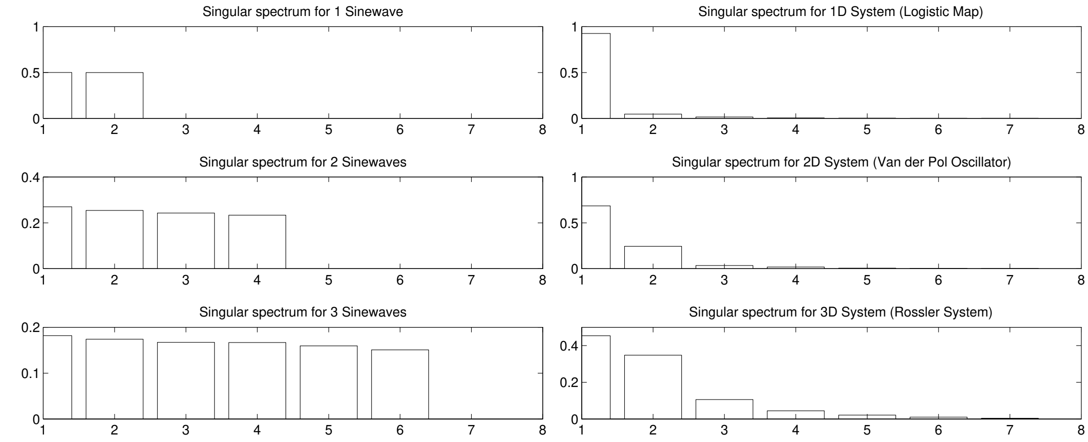
As a final example we show the entropy measure taken over a sliding window on a simple 1-D signal generated from $x_t = \sin(\omega t)$ for $t < 3000$ and

$$x_t = (1 - t/T)\sin(\omega t) + (t/T)\eta_t$$

for $t > 3000$ in which $T = 5000$ and η_t is, as before, a unit variance, zero mean noise source. Figure (5) shows x_t and the resultant time course of the Ω measures, each taken over a sliding window of length 500 samples with 25 samples overlap. We see this entropy-based measure clearly indicates the system transition and subsequent increase in the signal 'complexity'. Note that the estimated number of system states saturates in noise to the upper limit given by m, the embedding dimension.

3.3.2 Spatial information

We introduce this concept by re-defining the data matrix $\mathbf{X}$. In the previous section, this matrix was formed by embedding a *single* time series using the method of delays. We now consider the case where, instead of a single channel of information, we have a set of N such channels.



(a) (b)

Figure 4: **(a) Sinusoidal components:** one, two and three sinusoids – note that each sinusoid has two degrees of freedom. **(b) Chaotic time series:** One, two and three degree of freedom systems. Note the fact that, unlike the case of sinusoids, we could not *per se* determine the 'dimensionality' of the system. A clear trend is however obvious.

Let $\mathbf{x}_n(t)$ be n–th such signal sequence (channel), we now define

$$\mathbf{X} = (\mathbf{x}_1(t), \mathbf{x}_2(t), ..., \mathbf{x}_N(t))^T \tag{9}$$

and, as before, decompose this matrix using singular value decomposition, i.e.

$$\mathbf{X} = \mathbf{U}\mathbf{S}\mathbf{V}^T \tag{10}$$

We now consider the matrix $\mathbf{U}$ to be composed of a series of data vectors, such that:

$$\mathbf{U} = (\mathbf{u}_1(t), \mathbf{u}_2(t), ..., \mathbf{u}_N(t))^T \tag{11}$$

Consider now the one of these component vectors, $\mathbf{u}_i$ say. From Equation (10) we have:

$$\mathbf{u}_i = \mathbf{X}(\mathbf{S}\mathbf{V}^T)_i^{-1} \tag{12}$$

and, as $\mathbf{S}$ is diagonal and $\mathbf{V}^T = \mathbf{V}^{-1}$ so this may be written as:

$$\mathbf{u}_i(t) = \frac{1}{\sigma_i} \sum_{n=1}^{N} \mathbf{x}_n(t) v_{i,n} \tag{13}$$

i.e. a linear combination of the N channels $\mathbf{x}_n(t)$. The weighting fractions for each $\mathbf{u}_i(t)$ being given by the appropriate column of the matrix $\mathbf{V}$.

If the set of signals are associated with particular spatial locations (on the head, for example) then the matrix $\mathbf{V}$ provides information regarding the spatial location of information in each $\mathbf{u}_i(t)$.

We may, furthermore, describe the complexity of the spatial distribution using the entropy function of Equation (7) and estimate the number of system states via Equation (8). We give a simple example of this approach in Figure (6). Plot (a) shows the 3-dimensional data sequence $\mathbf{x}(t)$ which is synthesised via the Equations:

$$\begin{aligned} x_1(t) &= (1 - t/T)\sin\omega_c t + (t/T)\sin(\omega t/2)\cos(\omega t/3) \\ x_2(t) &= (1 - t/T)\sin\omega_c t + (t/T)\sin(\omega t/3) \\ x_3(t) &= (1 - t/T)\sin\omega_c t + (t/T)\sin(\omega t/5)\cos^2(\omega t/4) \end{aligned} \tag{14}$$

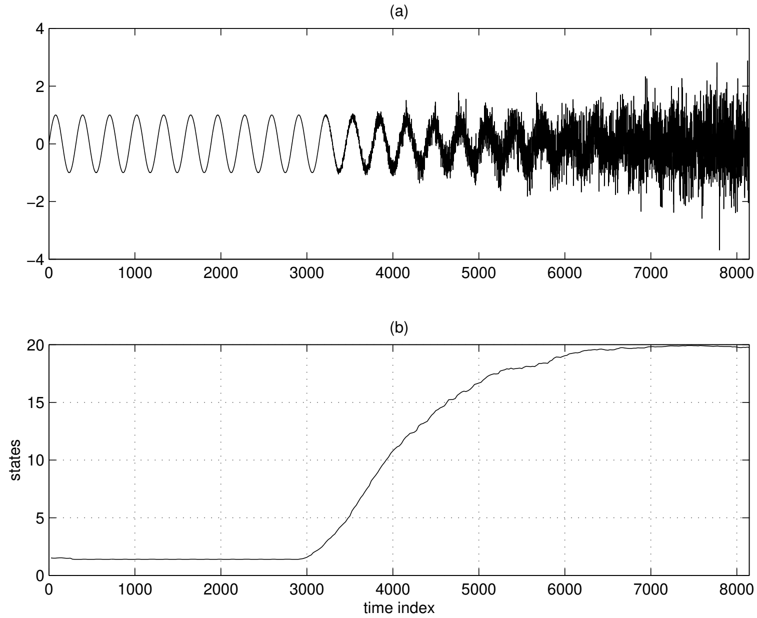


Figure 5: **Deteriorating sinusoid:** (a) x_t and (b) measures of number of system states (Ω) taken over a sliding window on x_t.

where ω_c is common to all three signals. Plot (b) shows the resultant Ω measure taken over a series of sliding windows of length 500 samples and overlap 25 samples. Note that as the three components decouple Ω increases and that the state estimates are correct. [2]

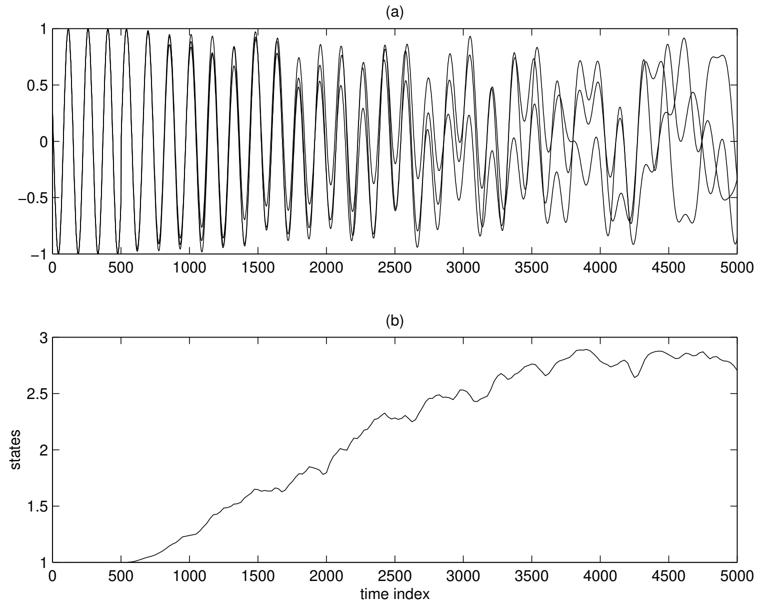


Figure 6: **De-synchronising 3-D system:** (a) $\mathbf{x}(t)$ & (b) resultant number of states measure (Ω).

4 Results on BCI data

We present results of temporal, spatial and spatio-temporal complexity in this section. We apply all methods to blocks of data in which *imagined* finger movements are present. An example of spatial

[2]This is, in practice, a nicety and we merely require that the Ω measure shows a correct and consistent trend.

localisation using the $\mathbf{V}$ matrix is also presented. Finally, we observe the utility of the complexity features as BCI features and compare the results with those obtained using a frequency-based measure of the signals.

When we perform a decomposition of the 18-D data space, we obtain a set of 18 weighting vectors (the columns of $\mathbf{V}$) via Equation (13). In this study we are interested in investigating events which are expected to show some spatial localisation. We will concentrate, therefore, on the projection of the data onto the most localised spatial distribution defined over the entire data set. This chooses some column of $\mathbf{V}$, $\mathbf{v}^*$ say, and a corresponding column of $\mathbf{U}$. The vector $\mathbf{v}^*$ may be reformatted and represented as spatial array of size 6×3 corresponding to the spatial locations of the 18 central electrodes of Figure (1). Figure (7) shows the spatial distribution of $\mathbf{v}^*$ over four one-second windows at from 2 seconds prior to a right imagined movement to 1 second after the event. We see a clear indication that localisation increases at the event on the contralateral side.

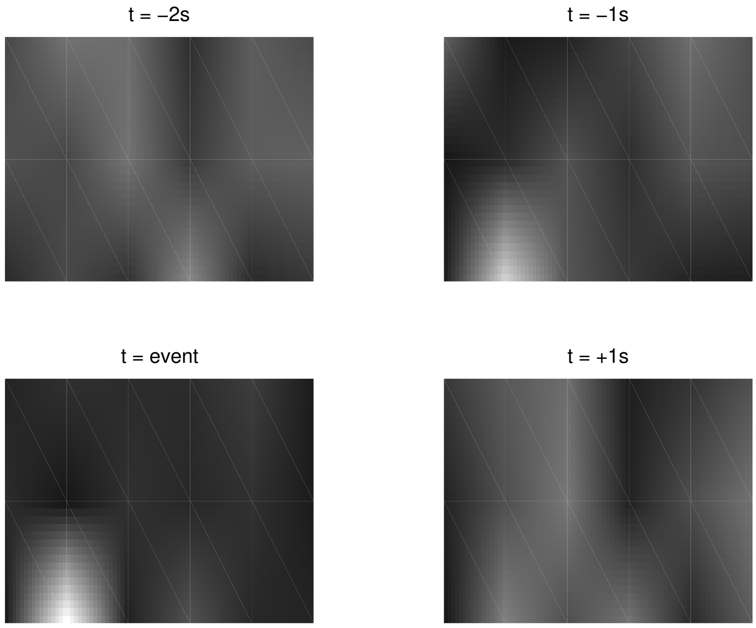


Figure 7: **Spatial localisation - right imagined movement:** The localisation increases at the event. Note also that the distribution has been interpolated and smoothed for presentation purposes. Results are obtained over successive one-second windows with no overlap.

The next example we show (Figure 8) is of spatial and temporal complexity measures (the entropy-based state estimate, Ω) applied to the 18-electrode signal array (for the spatial measure) and to the column of $\mathbf{U}$, $\mathbf{u}^*$, defined in the previous section (in the case of the temporal measure). Results are obtained using a four-second running window with overlap between successive windows of 3s. A clear periodicity is present in the complexity traces, although it is significantly phase-shifted from the external movement cues. Of note is the anti-correlation between the temporal and spatial measures. Although such an observation may be insignificant on a single subject, it is interesting to observe that as activity becomes more spatially focused (reduced spatial complexity) the activity over the 'focal site' becomes more complex (raised temporal complexity).

As a final example of the method we evaluate the complexity of data for which we build up an embedding matrix both across channels and through time (spatio-temporal complexity). The embedding matrix which we analyse is hence built up through time over N channels and is:

$$\mathbf{X}_{tot} = (\mathbf{X}_1, ..., \mathbf{X}_N)$$

where each $\mathbf{X}_i$ is defined on each channel via Equation (3). We thence evaluate the entropy and estimated number of states of the resultant decomposition of $\mathbf{X}_{tot}$ using Equations (7, 8). We evaluate the embedding matrix using only channels C3 and C4 each of which is embedded in a five-dimensional

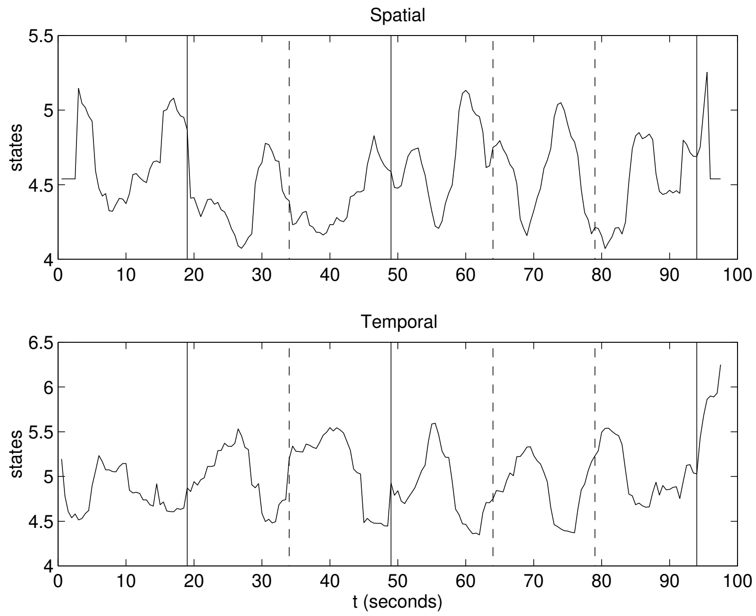


Figure 8: **Imagined finger movements from BCI recording, subject GW:** Both spatial and temporal system-state (Ω) traces show clear event-related changes. Cues for movements occur at 15s intervals shown by the vertical lines (dashed lines correspond to left and solid lines to right imagined movements). The flat sections at beginning and end of the upper trace are run-in / run-out effects. Note also that a cued imagined movement occurs at $t = 4$s but that we exclude it due to possible transients at the start of the recording.

space using a lag of $J = 5$. These electrodes are chosen as the resultant complexity measure is a combined descriptor of inter-hemispheric coupling as well as the temporal complexity of each channel. Results are shown in Figure (9). The upper trace shows the state measure (Ω) evaluated over four-second windows with an overlap of 3s and the lower trace the probability values obtained using a one-sided t-test which assessed the significance of dips in complexity between successive one-second sections of the Ω trace. Correlation with the external movement cues is high. For comparison we present results over the same data block (subject HH) using power in the 8-12 Hz band which has been extensively used for BCI classification (Pfurtscheller *et al.*, 1993). Figure (10) shows the fraction of signal power in the 8-12 Hz band as estimated via the power spectral density over four-second sliding windows with overlap of 3s. As can be seen, there is little indication of movement planning.

To quantify the performance of the complexity-based measure for subsequent classification we use the p-value from the t-test and examine the 'centre of mass' of each peak in this trace. If this centre lies within three seconds of a movement cue (for externally-paced data) we consider this a 'correct' detection. The choice of three seconds is based on our own, empirical, observations regarding movement planning prior to a movement (real or imagined) and also those of other researchers (Pfurtscheller *et al.*, 1993). We note that this is a simple method of classification and that, in cases where multiple features are to be used, flexible discriminant methods (such as neural networks) are preferred. This paper does not cover this area, however, as its aim is to introduce and discuss a particular analysis method.

In all our experiments we record a control block during which external cues are provided but subjects are asked *not* to imagine any limb movement. In all eight subjects tested there were significant ($p < 0.05$) event-related changes associated with imagined movements. In only four of these cases, however, were the associated control blocks *insignificant* ($p \not< 0.05$). If we exclude these four subjects then we can detect imagined hand movements with an average accuracy of 80% (range = [70,100]%, $N = 120$ events). We should comment that asking subjects *not* the imagine hand movements when

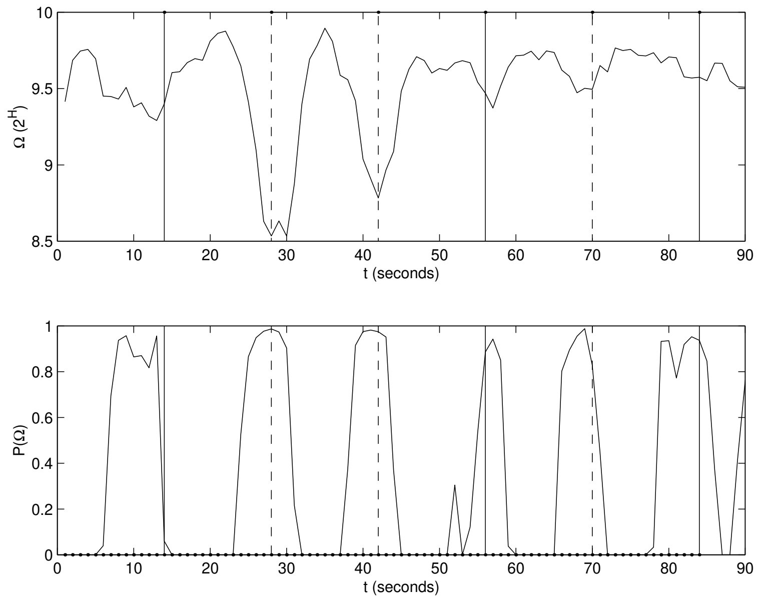


Figure 9: **Imagined finger movements from BCI recording, subject HH:** The top trace shows Ω estimated over a spatio-temporal embedding matrix. The bottom trace shows the resultant probability of change in Ω using a one-sided t-test. Dashed lines correspond to left and solid lines to right imagined movements.

cues are given may not lead to ideal control blocks as some psychologists argue that this action requires some primary imagination of the movement itself. What we may infer from the subset of subjects whose control blocks are insignificant, however, is that the detection of imagined movements in these subjects is not due to evoked responses to the visual or auditory cues.

Clearly, the goal of any BCI system is to operate without external cueing. We next present results obtained from recording blocks in which imagined finger movements are self-paced. The subject is asked merely to imagine movement at intervals of not less than ten seconds (self-judged). Results from two subjects are shown in Figure (11). Although we cannot objectively validate these results, as there are no reference cues, it is clear that significant peaks in the p-value obtained from the Ω measure (using the same spatio-temporal processing of C3 and C4 as in the previous section) occur at semi-regular intervals.

5 Conclusions / discussion

We have presented in this paper some preliminary results obtained from the use of measures of system 'complexity' on BCI data. The results are encouraging. As good results are obtained using only spatio-temporal information from electrodes C3 and C4 a small number of electrodes may be used. Furthermore, the methods discussed do not require a large computational overhead and may hence be used as part of a real-time system. Such a system would then enable issues such as the use of bio-feedback to be resolved in a consistent manner. The complexity-based methods, we would argue, perform well compared to more traditional frequency-based information over the data we have collected. In general, however, we see them as extra information sources and not as a replacement for generic frequency-based features, such as autoregressive model coefficients.

We point out, furthermore, that we have not reported in this paper on our classification of imagined movements on the basis of left or right hand closure. Whilst this is an important area, much of the current literature does not report on the accuracy of systems to detect the imagined movements themselves, but relies on external cues to indicate the location of events. We feel that any practical

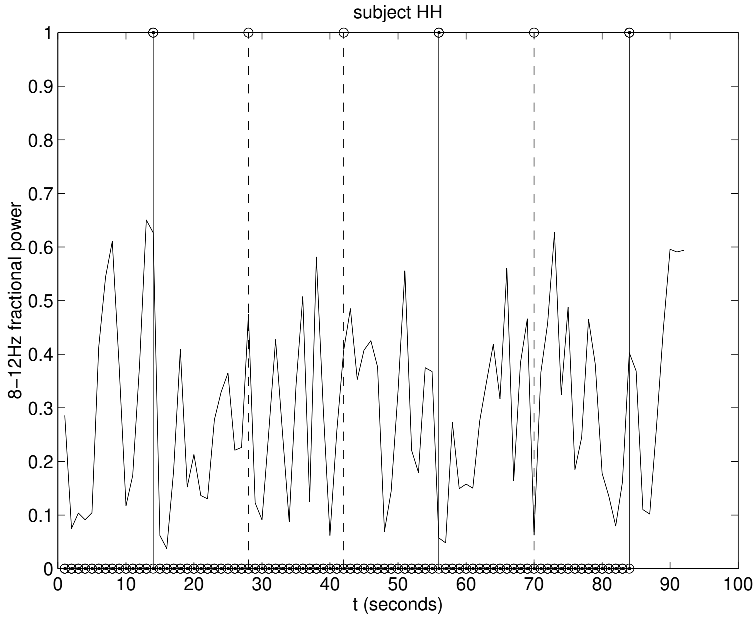


Figure 10: **Imagined finger movements from BCI recording, subject HH:** Fractional spectral power in the 8-12Hz band (alpha or mu rhythm). Right and left event cues are shown as solid and dashed lines respectively. Note that no clear event-related changes in spectral power are seen.

BCI system must not rely upon external cueing and so event detection in its own right is an important issue.

6 Acknowledgements

Iead Rezek and Will Penny are, respectively, funded via the Commission of the European Community (project SIESTA, grant BMH4-CT97-2040) and the UK EPSRC (grant GR/K79062) whose support we gratefully acknowledge. We thank Maria Stokes, Director of Research at the Royal Hospital for Neuro-disability, Putney, UK, for help with the protocol and for providing facilities at the hospital. We would finally like to thank the Living Again Trust for their support. This work was completed whilst SJR was on sabbatical leave from Imperial College at the Newton Institute for Mathematical Sciences in Cambridge.

References

[1] D. Broomhead and G. King. Extracting qualitative dynamics from experimental data. *Physica D*, 20:217–236, 1986.

[2] P. Grassberger and I. Procaccia. Measuring the Strangeness of Strange attractors. *Physica-D*, 9D:189–208, 1983.

[3] G. Kember and A.C Fowler. A correlation function for choosing time delays in phase portrait reconstructions. *Physics Letters A*, 179(2):72–80, 1993.

[4] G. Pfurtscheller, D. Flotzinger, and J. Kalcher. Brain-Computer Interface – a new communication device for handicapped people. *Journal of Microcomputer Applications*, 16:293–299, 1993.

[5] G. Pfurtscheller, D. Flotzinger, and C. Neuper. Differentiation between finger, toe and tongue movement in man based on 40 Hz EEG. *Elctroenceph. and Clin. Neur.*, 90:456–460, 1994.

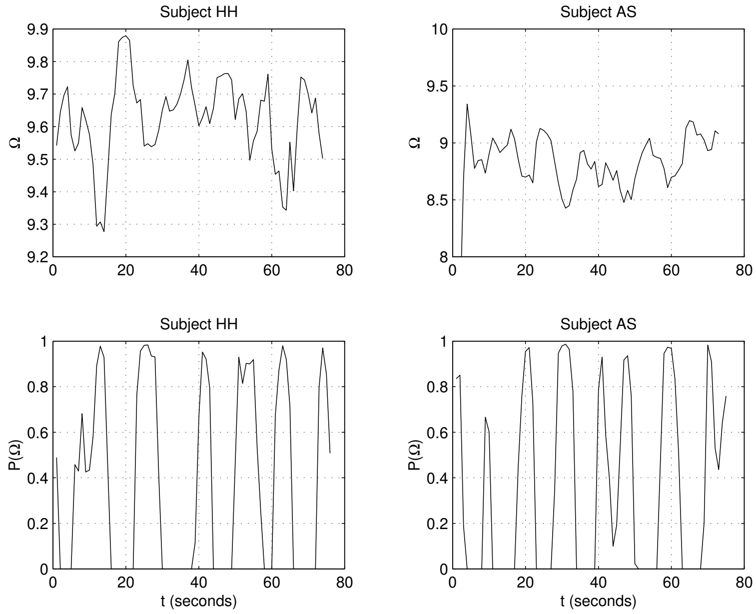


Figure 11: **Self-paced imagined movements, subjects HH & AS:** The upper traces show the state estimates (Ω) from spatio-temporal embedding of electrodes C3 and C4. The lower traces show probability measures taken via a one- sided t-test looking for significant dips in the traces. Semi-regular events appear to be present, although validation against cues is not possible.

[6] S.M. Pincus. Approximate entropy as a measure of system complexity. *Proceedings of the National Academy of Science, USA*, 88:2297–2301, March 1991.

[7] I.A. Rezek and S.J. Roberts. Stochastic complexity measures for physiological signal analysis. Technical report, Imperial College, University of London, November 1996. Available from `http://www.ee.ic.ac.uk/hp/staff/sroberts.html`.

[8] Takens, F. eds : Rand, D.A. and Young, L.S. *Dynamical Systems and Turbulence, Lecture Notes in Mathematics*, volume 898, pages 366–381. Springer, 1981.

[9] A Wolf, J.B. Swift, H.L. Swiney, and J.A. Vastano. Determining Lyapunov exponents from a time series. *Physica-D*, 16D(3):285–317, 1985.