NeuroImage 117 (2015) 449–455

Contents lists available at ScienceDirect

NeuroImage

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

Comments and Controversies

# Long-range dependencies make the difference—Comment on "A stochastic model for EEG microstate sequence analysis"☆

Markus Gschwind[a,b,*], Christoph M. Michel[b], Dimitri Van De Ville[c,d]

[a] Department of Neurology, University Hospital, Geneva, Geneva, Switzerland
[b] Functional Brain Mapping Laboratory, Department of Neuroscience, Biotech Campus, University of Geneva, Geneva, Switzerland
[c] Department of Radiology and Medical Informatics, University of Geneva, Geneva, Switzerland
[d] Institute of Bioengineering, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland

ARTICLE INFO

*Article history:*
Accepted 19 May 2015
Available online 30 May 2015

*Keywords:*
EEG microstates
Topographical fluctuations
Lognormal distribution
Power law
Long-range dependency
Markov chain
Fractality
Self-similarity
Autocorrelation
Detrended fluctuation analysis
DFA
Wavelet leaders

*Dear Editor,*

We have read with great interest the recent paper by Gärtner et al. (2015) on EEG microstate sequences. These authors propose a stochastic model termed sampled marked intervals (SMI), relating the observed microstate sequence to an assumed underlying stochastic process, similar to earlier work that proposed Hidden Markov Models (HMM)(De Lucia et al., 2011; Obermaier et al., 1999). We very much appreciate such a theoretical approach to enrich the classical EEG microstate analysis (Lehmann and Skrandies, 1984). In particular, we agree with the authors about the decisive procedure of restricting the analysis to local maxima of global field power (so called GFP peaks). However, in our opinion, one of the key features of microstate sequences is not considered in this paper: their long-range dependencies (Van De Ville et al., 2010). Conventional HMMs are well-known for their efficiency in modeling short-term dependencies between adjacent elements, but they cannot grasp long-range interactions between distant elements (Yoon and Vaidyanathan, 2006). Here we give an overview on evidence of long-range dependencies in microstate sequences of the resting brain.

## 1. Background—long-range dependency in complex systems

Long-range dependency (LRD) or "long memory effect" describes an exceptional behavior of a stochastic process, which appears in only a minority of stationary stochastic processes. The importance of LRD has become clear with the original work of *H. E. Hurst*, who developed the basic methods to assess the dynamics of floods in River Nile (Mandelbrot, 1965). LRD has become closely associated with self-similarity. In geometry, a self-similarity depicts a shape being composed of a basic pattern, which is repeated at multiple (or infinite) scale. In a self-similar process a scaling in time equivalents scaling in space (fractal structure) (Mandelbrot, 1983), and the connection between the two types of scaling is determined by a constant, often called the "Hurst exponent", which determines whether or not the self-similar process shows LDR (Samorodnitsky, 2006). Self-similar processes are based on power-law degree distributions (heavy-tailed), especially in failure-robust and self-regulating organizations of biological networks (Barabasi and Albert, 1999).

There is currently growing awareness about the importance of heavy-tailed distributions in a remarkable variety of biological and life-science related fields (Arita, 2005; He et al., 2010; Limpert and Stahel, 2011; Limpert et al., 2001). The lognormal distribution is increasingly recognized not only as being the underlying principle of psychophysics (i.e. the Weber–Fechner-Law of perception), but equally being present at multiple levels in neuronal structural–functional activity, starting from the axon caliber (Wang et al., 2008), over synaptic strength (Klinshov et al., 2014; Loewenstein et al., 2011; Yasumatsu et al., 2008), the neuronal firing pattern (Mizuseki and Buzsaki, 2013; Yasumatsu et al., 2008), up to the density of the brain's large-scale connections (Markov et al., 2014; Oh et al., 2014; Wang et al., 2012). The similarity of power-law distributions and other heavy-tailed distributions (e.g. lognormal) has been explained by our imperfect observation of

☆ Comment on: M Gärtner, V Brodbeck, H Laufs, G Schneider: A stochastic model for EEG microstate sequence analysis. NeuroImage 104 (2015) 199–208.

* Corresponding author at: Unité d'EEG et d'exploration de l'épilepsie & Laboratoire de cartographie cérébrale, Service de Neurologie, Hôpitaux Universitaires de Genève (HUG), Rue Gabrielle-Perret-Gentil 4, 1211 Genève, Switzerland. Fax: +41 22 3728340.
*E-mail address*: markus.gschwind@gmail.com (M. Gschwind).

http://dx.doi.org/10.1016/j.neuroimage.2015.05.062
1053-8119/

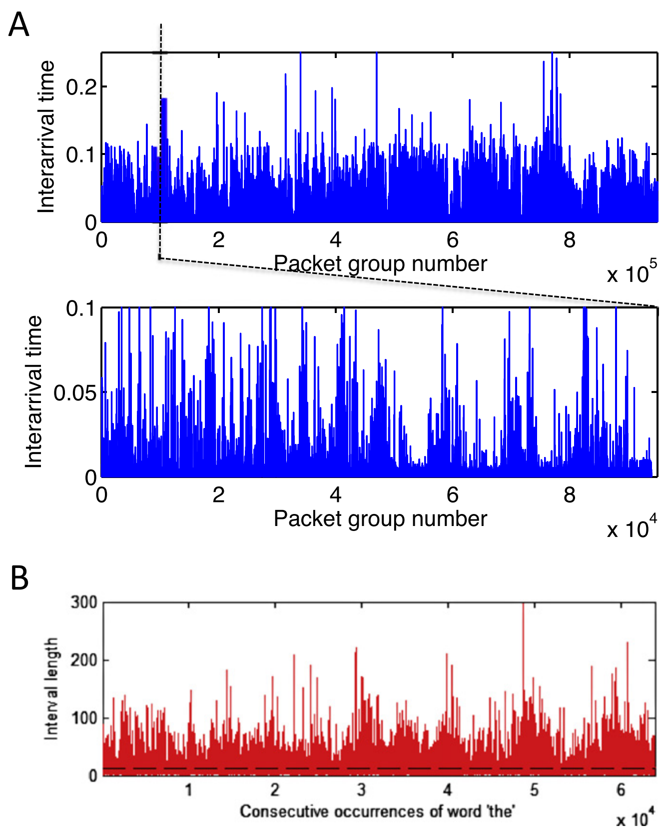


**Fig. 1.** Examples of scale-free behavior in sequences of other domains. A. Scale-free behavior has been repeatedly demonstrated in ethernet bytes-per-time arrival statistic, shown here at two different aggregation levels (graph recreated following Riedi et al. 1999) from the publicly available Bellcore internet traces (BC-pAug89)). B. Interval lengths between consecutive occurrences of the word "the" in the English language (Banchs, 2013) resembles strongly a scale-free behavior (original graph obtained by the author).

natural phenomena, which let power-law distributions arise from lognormal distributions upon small tweaks, e.g. by the setting of a lower boundary, or a non-uniform data sampling (Arita, 2005). In general, the lognormal distribution emerges as the collective fingerprint of a multitude of interactive processes, via a multiplication of a large number of variables (Buzsaki and Mizuseki, 2014).

Complex systems with inherent LRD have been described in many other fields, and its mathematical concepts are fruitfully employed. Numerous studies demonstrate LRD for example in geological and climate research (Scheffer et al., 2009; Varotsos and Kirk-Davidoff, 2006), finance market fluctuations (Matteo et al., 2005; Robinson, 2003), in Internet modeling and network traffic analysis (Abry et al., 2002; Karagiannis et al., 2004; Riedi et al., 1999), or statistic analyses of human language (Alvarez-Lacalle et al., 2006; Petersen et al., 2012). Fig. 1A shows an example of an ethernet bytes-per-time arrival, at two different aggregation levels. Short packages alternate with long packages in a typical manner, according to a heavy-tailed distribution (Riedi et al., 1999). Another example with very similar, "bursty" intervals originate from linguistic analysis: the frequency of word occurrence was at the origin of Zipf's Law in 1949 and shows similar LRD dynamics (Alvarez-Lacalle et al., 2006; Petersen et al., 2012). Fig. 1B plots the interval lengths between consecutive occurrences of the short word "the" in the English language (Banchs, 2013).

## 2. The heavy-tailed distributions of EEG microstate durations and intervals suggest long-range dependency

As Gärtner et al. (2015) mention, the duration of the microstates appears to be one of the most informative properties (Dierks et al., 1997; Lehmann et al., 2005; Wiedemann et al., 1998). However, it has rarely been acknowledged that the consecutive durations of a selected microstate occur in a very irregular "bursty" order (Fig. 2A). Moreover, the selected microstate also repeats at very irregular intervals, reflecting the same irregular and "bursty" structure as the durations (Fig. 2B).

The histograms of both microstate durations and between-microstate-intervals are reminiscent of a heavy-tailed distribution (not to be confounded with the histogram of intervals between local maxima of global field power, as shown in Gärtner et al. (2015). The underlying distribution was compared against plausible heavy-tailed distributions (lognormal, generalized pareto, gamma and exponential) using quantile–quantile plots (Q–Q plots; Fig. 1C). Lognormal and pareto distribution seem closest to the empirical measures, but notice that any of the four could be chosen based on different theoretical a priori assumptions.

## 3. Evidence of self-similarity and quantification of long-range dependency

In this section we demonstrate self-similar properties and LDR in the microstate sequence using popular methods. For comparison, we generate a pseudo sequence, by estimating the Markovian transition probabilities from our original microstate sequence using a simple Markov Model,[1] as used by Gärtner and colleagues (Gärtner et al., 2015). In the following we refer to this pseudo sequence as "HMM sequence". Both microstate and HMM sequences are then transformed into random walks as described in Van De Ville et al. (2010), in order to apply the below described methods.

### *3.a. Autocorrelation shows a slow decay*

Compared to the HMM sequence, the microstate sequence displays a slower decaying autocorrelation function (ACF; Fig. 3A). Such a high and slowly decaying ACF points to two properties of a stationary process, its self-similarity and its long-memory properties. The self-similar pattern also evokes scale invariance (scale-freeness), a synonym for "fractal structure" in a statistical sense.

### *3.b. Power spectral density shows a negative slope (power-law behavior)*

Time series of the type of Poisson or Markov rely on the assumption of *independence* between the elements. Classical limit theorems, such as the Law of Large Numbers, state that, at large scales, a Poisson process can be approximated by its mean arrival rate. In the real world, however, various real world phenomena carry essential information at different scales of observation; their traces are "spiky", throughout all scales. Such behavior is a sign of strong dependencies in the data: large values come in clusters and clusters of clusters and so on. Their power spectral density estimate, in double-logarithmic presentation, is not horizontally balanced across the different frequencies like in the HMM sequence, but it displays a negative slope, indicating with power-law like behavior (He et al., 2010), as shown for a selected microstate sequence in Fig. 3B.

### *3.c. Rescaled range analysis (R/S) shows a Hurst exponent >0.5*

Such a slowly decaying autocorrelation function and a negative slope of the periodogram was equally true for the classical hydrological data of the floods of River Nile from which H. E. Hurst, in the early 1950s, developed the estimation of LDR using the "pox plot" of the "rescaled

[1] The sequence was created using MATLAB statistical toolbox with the following commands:
[TransitionM, EmissionM ] = hmmestimate(orig_microstate_labelseq, orig_microstate_labelseq);
HMMSeq = hmmgenerate(length(orig_microstate_labelseq), TransitionM, EmissionM);

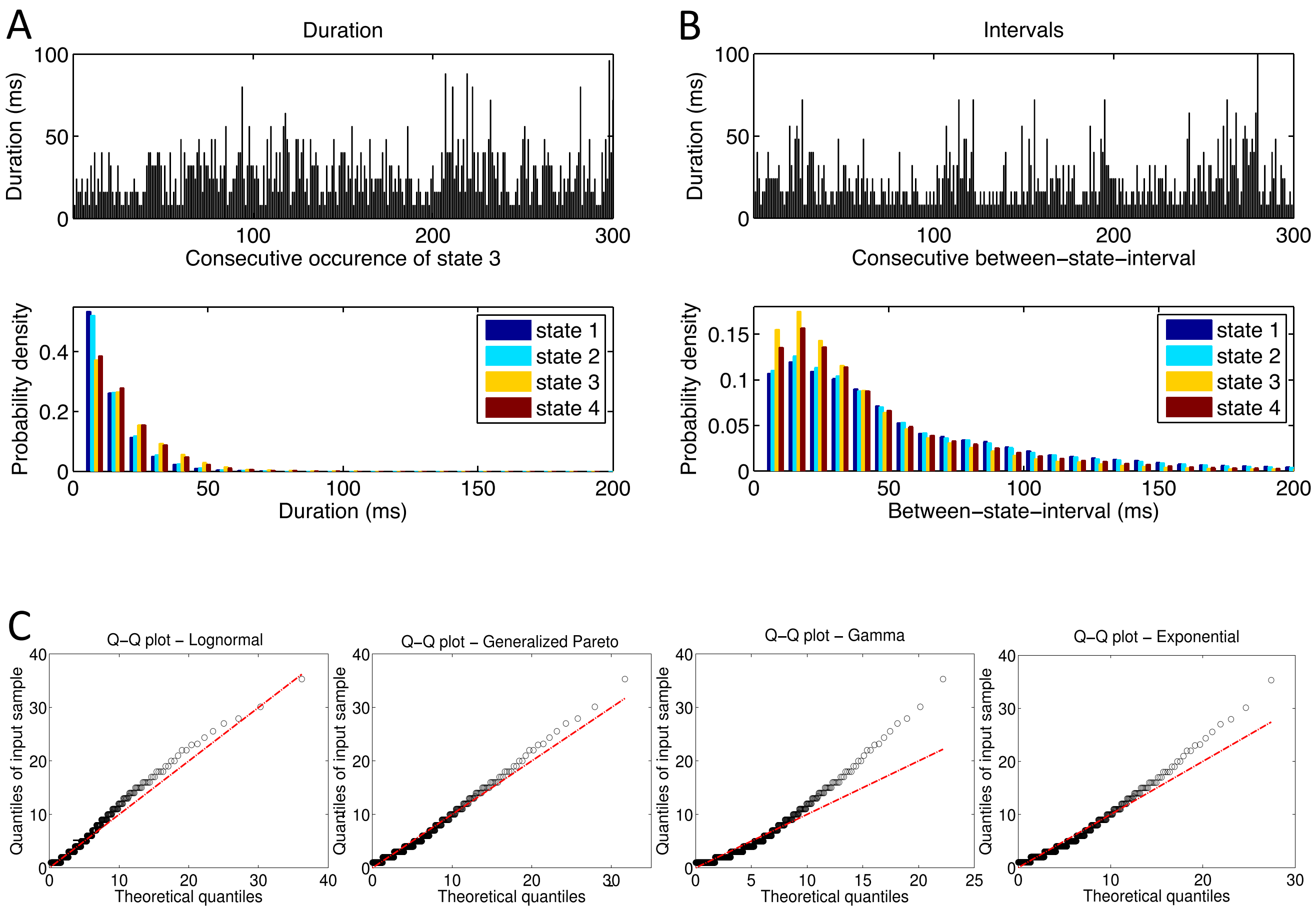


**Fig. 2.** Distribution of durations of microstates and between-state intervals. A – Top: The durations of consecutive occurrences of one selected microstate. They are very irregular, with high variance. Bottom: The histograms of durations for a set of 4 microstates. All 4 histograms show a heavy-tailed shape. B – Top: The intervals between the same first consecutive occurrences of the same microstate as in A. These between-state-intervals resemble the durations in irregularity and variance. Bottom: The corresponding histograms: all between-state-intervals show similar histograms as the durations. C. Q–Q plots illustrating the comparison of the histogram of durations of a selected microstate with heavy-tailed distributions such as lognormal, generalized pareto (power-law), gamma and exponential.

range". In this method the range of the data is rescaled at increasing time blocks, using a division by the standard deviation (range/SD or R/S). When plotting R/S over the increasing time blocks in logarithmic axes, a least-square line can be fitted. The slope of this fitting line indicates the degree of LRD and is called the "Hurst exponent". A slope of 0.5 is indicating random behavior equivalent to white noise, while a slope of >0.5 indicates the presence of LRD. Fig. 4A shows such a pox plot for a selected microstate sequence (the slope of 0.65 indicates LRD), compared to the HMM sequence (the slope is not different from 0.5).

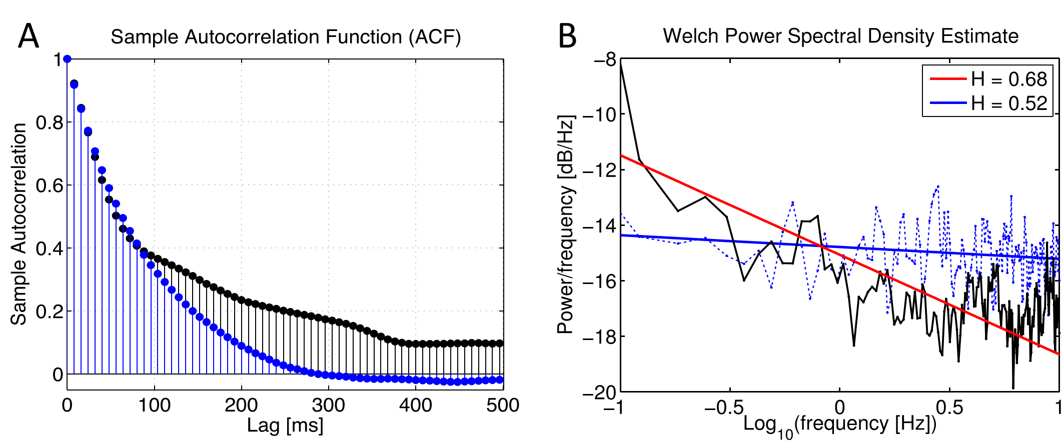


**Fig. 3.** Basic methods of determining long-range dependency in microstate sequences. A. The sample autocorrelation function shown for a selected microstate sequence for a time lag 1 to 500 ms (black), and the sample autocorrelation function of the HMM sequence (blue). Both ACF exhibit a slow decay; however, the HMM sequence arrives at 0 at the lag of 260 ms, while the microstate sequence decays much slower at lags >100 ms, and multiple Ljung-Box Q-tests demonstrated a significant autocorrelation for all possible lags (mean $p = 0$). B. The power spectral density estimation between 0.1 Hz and 10 Hz of the whole microstate sequence (Welch's method, $L = 2048$, overlap 87.5%, black line), and the power spectral density estimation of the HMM sequence (blue line). The microstate sequence shows a strong negative slope (red line) consistent with a power-law like pattern, while the HMM sequence is almost flat.

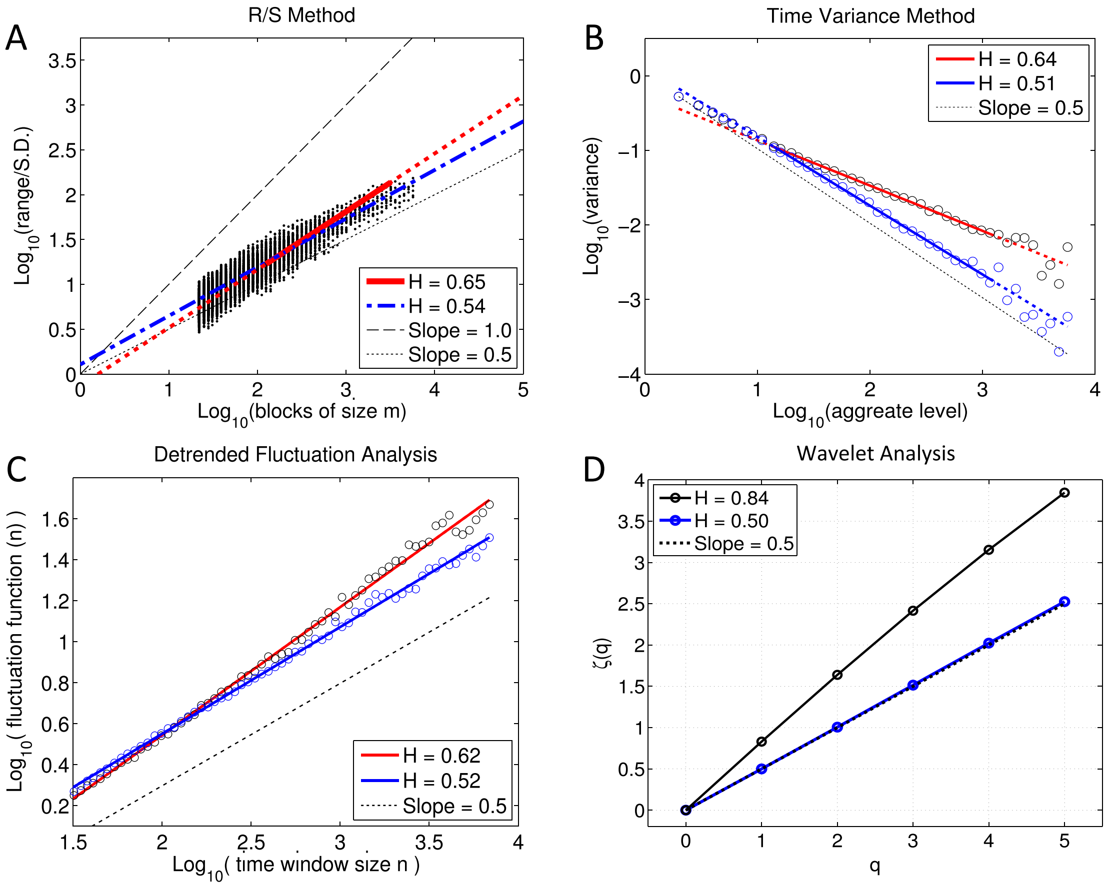


**Fig. 4.** Different quantification of long-range dependency (Hurst exponent estimators). A. The pox plot of "rescaled range" over increasing blocks of time shows a linear fit line with a slope of 0.65 indicating LRD (red line), the fitting of the "rescaled range" of the HMM sequence has a slope of 0.54, which is close to random white noise (blue dotted line). B. Time-variance analysis relates variance and aggregate levels of time in a double logarithmic plot. The values of an HMM sequence (blue) follow the diagonal dotted line ($\beta = 1, H = 0.5$), while the values of the microstate sequence follow a line (red), which lies more horizontal ($1 > \beta > 0$, $0.5 < H < 1$), indicating present long-range dependency. C. Detrended fluctuation analysis (DFA) of a microstate sequence. The fluctuation function $F(n)$ plotted against the time window size $n$ in log–log presentation. The values display a slightly curved line with a slope of 0.65 (red line), while the HMM sequence follows the slope of 0.51 (blue). D. The wavelet leader framework based method used in (Van De Ville et al., 2010) estimates the Hurst exponent across 6 different orders of power exponents $q$ (0 to 5), resulting in a more stable and complete estimation of the Hurst exponent. The resulting scaling spectra build a straight line (black) with a slope 0.84 showing strong LRD. The HMM sequence (blue) displays a slope of exactly 0.50, consistent with the expected behavior of random white noise.

### *3.d. Time-variance analysis shows long-range dependency*

There is no unique way to calculate the Hurst exponent, but several techniques can be used for its estimation (Sarker, 2007). Another variant of LRD sensitive analysis is the time-variance analysis. When generating an temporal aggregation at different levels $m$ and calculating the variance at each level of aggregation, the logarithmic dependence diagram will represent a descending line with a slope equal to $-\beta$, and the Hurst exponent is $1 - \beta / 2$. As shown in Fig. 4B, the HMM sequence follows directly the descending diagonal ($\beta = 1$ and $H = 0.5$), whereas the selected microstate sequence stays clearly more horizontal, with $1 > \beta > 0$ and $0.5 < H < 1$, confirming LRD in the data.

### *3.e. Detrended fluctuation analysis shows self-similarity and power-law behavior*

Detrended fluctuation analysis (DFA), is another, very popular, approach of capturing self-similarity and the presence of LRD, which has proven very useful in the analysis of a variety of complex physiological signals of any kind (Goldberger et al., 2000; Peng et al., 1994). In this method, the integrated sequence is divided into time windows of equal length $n$. Each window is detrended and the fluctuation function $F(n)$ is measured as the average of all standard deviations in all $n$ detrended time windows. This measure is repeated for a whole scale of different $n$ and presented in a log–log plot. The fluctuation function $F(n)$ displays a straight line, with, in case of a LRD sequence, a slope between 0.5 and 1, similar to the Hurst exponent (Fig. 4C). The linear relationship on the log–log plot indicates the presence of power law (fractal) scaling.

### *3.f. Wavelet framework shows fractal behavior*

Wavelet frameworks (Arneodo et al., 1988; Jaffard, 2004; Wendt et al., 2007) combine analyses at multiple scales in a robust way and provide therefore a more natural and elegant analysis of scale invariant properties of a complex system. In Van De Ville et al. (2010), the Hurst exponent was estimated using scaling spectra ($\zeta$) of 11 different power exponents $q$ ($-5$ to 5), allowing to go beyond second-order statistics as the time variance analysis. Fig. 4D shows the result of this method, the plot of the estimates of the scaling exponents $\zeta$ as a function of $q$. The slope of the microstate sequence is significantly $>0.5$, the slope of the HMM sequence is exactly 0.5 (dotted line), which is like a white noise sequence.

Van De Ville et al. (2010) have ruled out that the observed long memory effect appeared as an artifact of applied filters. The comparison of the original sequence with a shuffled-labels sequence (keeping the temporal relations unchanged), and with an equalized-duration sequence, demonstrated the importance of the temporal properties of the microstate sequence, with a large variety of durations, in order to maintain fractal organization and LRD. The duration-equalized sequence has similar dynamics as white noise, although the exact sequence of transitions from one to the next state was maintained. The transition probabilities of a simple Markov chain only reflect the proportion of the overall presence for each of the four maps. They are not influenced by short-range interactions, but they cannot capture the important long-range interaction.

### *3.g. Multiscale permutation entropy shows the importance of local variation of higher order*

Entropy is an information-theoretic measure of higher-order system complexity that attracted much attention from various fields. Different methods have been proposed to calculate the entropy of a complex system (Bandt and Pompe, 2002; Costa et al., 2002; Ouyang et al., 2013; Pincus, 1991). Here we use a recent approach, termed “multiscale permutation entropy” (MPE) (Ouyang et al., 2013; Wu et al., 2014), combining different scales (Costa et al., 2002) and permutation orders (Bandt and Pompe, 2002), in order to demonstrate the relationship between local and global information content in the microstates sequence. Entropy of a single discrete random variable is a measure of its average uncertainty (Costa et al., 2002):

$$H(X) = -\sum_{x_i \in \Theta} p(x_i) \log p(x_i).$$

MPE uses a scaling factor $s$ defining the granularity, and an embedding factor $m$ defining the order of permutation. For example, to calculate MPE with $s = 10$ and $m = 5$: first, every 10 non-overlapping time points $(x(t) \ldots x(t+10))$ are averaged into a coarse grained surrogate sequence; second, for every 5 of the resulting surrogate elements the entropy is calculated for all possible permutations and averaged. Fig. 5 shows the MPE of a selected microstate sequence using permutations from 2nd to 7th order, and for a scaling from 1 to 125 (corresponding to 1 s; red circles), compared to the entropy of the HMM sequence (blue dots). Globally, entropy almost linearly decreases with growing permutation order. For each $m$, entropy is lowest in the fine scales, increasing rapidly with coarser granularity.

This provides evidence that local variation (low $s$) of high order (high $m$) contains most of the information in a microstate sequence; the coarser the granulation is (high scale), the more flattened and the more similar to noise it becomes, and the lower the permutation order, the more noise like the sequence is. This converges with the high importance of local variety for long-range dependency in microstate sequence.

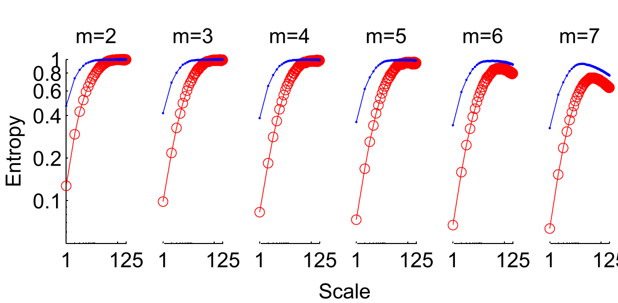


**Fig. 5.** Multiscale permutation entropy (MPE) using permutations from 2nd to 7th order ($m$), shown for different scales from 1 to 125 (sampling rate 125 Hz) for a microstate sequence (red) and HMM sequence (blue) in log–log presentation. For each $m$, entropy is lowest in the fine scales, increasing rapidly with coarser granularity. With growing permutation order, entropy decreases and is clearly lower than for the HMM sequence.

## 4. Need of enhanced Markov models to adequately describe microstate sequences

Markov chains have been used before to describe microstate sequences (Lehmann et al., 2005; Nishida et al., 2013; Wackermann et al., 1993). These models seduce on the first glance by their simplicity and direct application to the sequence of microstate. However, in a 1st-order Markov chain, as Gärtner et al. (2015) mention themselves, the probability that the next state is s, given a current state $r$, only depends on the current $r$, and is given by the transition probabilities $P(s|r)$. Time is considered homogenous, meaning that the transition probability matrix stays constant during the whole sequence. The probability of the $n$ following steps will spread out away from the initial random point, and, with increasing $n$, approach the same limiting distribution $\pi$ when $n \to \infty$ (i.e. the probability of 0.25 for each in a set of 4 microstates), given it is an aperiodic, irreducible, transient and stationary Markov chain (central limit theorem). If $n$ is limited, the transition probability matrix will simply reflect the occurrences per microstates (Chang, 2007). Long-range dependency is not captured with these Markov models (Callut and Dupont, 2005; Park et al., 2012; Yoon and Vaidyanathan, 2006).

One of the classic models based on local interactions, but capable of long-range dependency and generation of scale-free signals is the Ising model. It is a ferromagnetic model consisting of discrete integer variables (spins, taking values $-1$ or $+1$), organized in a lattice. Normally this system behaves only according to short-term interactions, but there is a critical parameter adjustment, by which long-range dependence does emerge. Similar to the Ising model, where a spatial arrangement (lattice) and a contextual behavior of the ferromagnetic spins is a precondition in order to produce, under certain circumstances, long-range dependencies, there have been several propositions to improve the long-range capability of Markov models. For example, a context-sensitive HMM was introduced (Yoon and Vaidyanathan, 2006), associated at a dynamic programming in order to find the optimal state sequence, as well as a parameter re-estimation algorithm for optimizing the model parameters, given training sequences. Another example is the introduction of structure, associated at non-linear optimization and iterative state splitting (Callut and Dupont, 2005). A third example is the proposition to include a variational Gaussian process dynamical system in the HMM used in phoneme classification (one of the most successful practical fields of HMM application), enabling the complex dynamic structure and long-range dependency of speech to be better represented than that by an HMM (Park et al., 2012).

## 5. Conclusion

In sum, we retain that scale-free dynamics in EEG microstate sequences are reminiscent of the complex system of the brain, likely as a result of self-organizing mechanisms (Bak et al., 1987; Barabasi and Albert, 1999) with uncountable activators (Buzsaki and Mizuseki, 2014), all operating far from homeostasis (Breakspear and Stam, 2005). Scale-freeness is a sign of efficient and flexible information flow between multiple sources (Kello et al., 2010), and from an evolutionary point of view might enable the system to adapt and reconfigure rapidly, so that learning can occur on this basis (de Arcangelis and Herrmann, 2010; Lewis et al., 2009). Fractal organization has been demonstrated in an increasing number of physiological measurements, such as heartbeat rhythm (Goldberger et al., 2000) and gait stride (Peng et al., 2000).

The microstate analysis of the resting brain provides therefore an ideal macroscopic observational window of these global temporal aspects of spontaneous brain activity. In the light of the above-presented evidence, further adequate modeling of microstate sequences needs to go beyond the step-by-step short-term interactions of the states. As demonstrated here and initially by Van De Ville et al. (2010), the long-range dependency of the microstate sequence is not based on the sequence of the labels alone, but much more on the temporal relations within and between the states. As a consequence, an empirical microstate sequence, which was arbitrarily "re-labeled" by an applied HMM, still would display LRD in case the durations and intervals are preserved, i.e. the "shuffled label" condition in Van De Ville et al. (2010). It is the microstate durations and between-state intervals that involve higher-order relationships with quantifiable long-range dependency.

We would therefore like to encourage the authors of Gärtner et al. (2015) to solicit their HMM model like approach of "sampled marked intervals" (SMI) also for its capability to capture long-range dependency. In case the presented SMI model indeed lets emerge LRD, this property should also be described. If, however, the SMI model does not allow to capture LRD, we would suggest the authors to consider an extension of their model that would allow for LRD, a key feature of the EEG microstate sequence.

## Acknowledgment

This work was supported by the Swiss National Science Foundation, grants number 33CM30_124115/2SPUM and 326030_128775.

The authors do not report any conflict of interest.

## References

Abry, P., Baraniuk, R.G., Flandrin, P., Riedi, R.H., Veitch, D., 2002. Multiscale nature of network traffic. IEEE Signal Process. Mag. 19, 28–46.
Alvarez-Lacalle, E., Dorow, B., Eckmann, J.P., Moses, E., 2006. Hierarchical structures induce long-range dynamical correlations in written texts. Proc. Natl. Acad. Sci. U. S. A. 103, 7956–7961.
Arita, M., 2005. Scale-freeness and biological networks. J. Biochem. 138, 1–4.
Arneodo, A., Grasseau, G., Holschneider, M., 1988. Wavelet transform of multifractals. Phys. Rev. Lett. 61, 2281–2284.
Bak, P., Tang, C., Wiesenfeld, K., 1987. Self-organized criticality: an explanation of the 1/f noise. Phys. Rev. Lett. 59, 381–384.
Banchs, R.E., 2013. Text Mining with MATLAB®. Springer, New York.
Bandt, C., Pompe, B., 2002. Permutation entropy: a natural complexity measure for time series. Phys. Rev. Lett. 88.
Barabasi, A.L., Albert, R., 1999. Emergence of scaling in random networks. Science 286, 509–512.
Breakspear, M., Stam, C.J., 2005. Dynamics of a neural system with a multiscale architecture. Philos. Trans. R. Soc. Lond. B Biol. Sci. 360, 1051–1074.
Buzsaki, G., Mizuseki, K., 2014. The log-dynamic brain: how skewed distributions affect network operations. Nat. Rev. Neurosci. 15, 264–278.
Callut, J., Dupont, P., 2005. Inducing hidden Markov models to model long-term dependencies. 16th European Conference on Machine Learning (ECML), Lecture Notes in Artificial Intelligence, pp. 513–521.
Chang, J.T., 2007. Stochastic Processes. (http://www.scribd.com/doc/206500071/Stochastic-Processes-by-Joseph-T-Chang).
Costa, M., Goldberger, A., Peng, C.K., 2002. Multiscale entropy analysis of complex physiologic time series. Phys. Rev. Lett. 89.
de Arcangelis, L., Herrmann, H.J., 2010. Learning as a phenomenon occurring in a critical state. Proc. Natl. Acad. Sci. U. S. A. 107, 3977–3981.
De Lucia, M., Constantinescu, I., Sterpenich, V., Pourtois, G., Seeck, M., Schwartz, S., 2011. Decoding sequence learning from single-trial intracranial EEG in humans. PLoS One 6, e28630.
Dierks, T., Jelic, V., Julin, P., Maurer, K., Wahlund, L.O., Almkvist, O., Strik, W.K., Winblad, B., 1997. EEG-microstates in mild memory impairment and Alzheimer's disease: possible association with disturbed information processing. J. Neural Transm. 104, 483–495.
Gärtner, M., Brodbeck, V., Laufs, H., Schneider, G., 2015. A stochastic model for EEG microstate sequence analysis. Neuroimage 104, 199–208.
Goldberger, A.L., Amaral, L.A., Glass, L., Hausdorff, J.M., Ivanov, P.C., Mark, R.G., Mietus, J.E., Moody, G.B., Peng, C.K., Stanley, H.E., 2000. PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. Circulation 101, E215–E220.
He, B.J., Zempel, J.M., Snyder, A.Z., Raichle, M.E., 2010. The temporal structures and functional significance of scale-free brain activity. Neuron 66, 353–369.
Jaffard, S., 2004. Wavelet techniques in multifractal analysis, fractal geometry and applications: a jubilee of Benoît Mandelbrot. In: Lapidus, M., van Frankenhuijsen, M. (Eds.), Proceedings of Symposia in Pure Mathematics. AMS, Providence, RI, pp. 91–152.
Karagiannis, T., Molle, M., Faloutsos, M., 2004. Long-range dependence—ten years of Internet traffic modeling. IEEE Internet Computing 6, 57–64.
Kello, C.T., Brown, G.D., Ferrer, I.C.R., Holden, J.G., Linkenkaer-Hansen, K., Rhodes, T., Van Orden, G.C., 2010. Scaling laws in cognitive sciences. Trends Cogn. Sci. 14, 223–232.
Klinshov, V.V., Teramae, J.N., Nekorkin, V.I., Fukai, T., 2014. Dense neuron clustering explains connectivity statistics in cortical microcircuits. PLoS One 9, e94292.
Lehmann, D., Skrandies, W., 1984. Spatial analysis of evoked potentials in man—a review. Prog. Neurobiol. 23, 227–250.
Lehmann, D., Faber, P.L., Galderisi, S., Herrmann, W.M., Kinoshita, T., Koukkou, M., Mucci, A., Pascual-Marqui, R.D., Saito, N., Wackermann, J., Winterer, G., Koenig, T., 2005. EEG microstate duration and syntax in acute, medication-naive, first-episode schizophrenia: a multi-center study. Psychiatry Res. 138, 141–156.
Lewis, C.M., Baldassarre, A., Committeri, G., Romani, G.L., Corbetta, M., 2009. Learning sculpts the spontaneous activity of the resting human brain. Proc. Natl. Acad. Sci. U. S. A. 106, 17558–17563.
Limpert, E., Stahel, W.A., 2011. Problems with using the normal distribution—and ways to improve quality and efficiency of data analysis. PLoS One 6, e21403.
Limpert, E., Stahel, W.A., Abbt, M., 2001. Log-normal distributions across the sciences: keys and clues. Bioscience 51, 341–352.
Loewenstein, Y., Kuras, A., Rumpel, S., 2011. Multiplicative dynamics underlie the emergence of the log-normal distribution of spine sizes in the neocortex in vivo. J. Neurosci. 31, 9481–9488.
Mandelbrot, B.B., 1965. Une classe de processus stochastiques homothétiques a soi. Application a la loi climatologique de H. E. Hurst. Comptes Rendus Academic Sciences Paris 260, 3274–3277.
Mandelbrot, B.B., 1983. The Fractal Geometry of Nature. W. H. Freeman and Co., San Francisco.
Markov, N.T., Ercsey-Ravasz, M.M., Ribeiro Gomes, A.R., Lamy, C., Magrou, L., Vezoli, J., Misery, P., Falchier, A., Quilodran, R., Gariel, M.A., Sallet, J., Gamanut, R., Huissoud, C., Clavagnier, S., Giroud, P., Sappey-Marinier, D., Barone, P., Dehay, C., Toroczkai, Z., Knoblauch, K., Van Essen, D.C., Kennedy, H., 2014. A weighted and directed interareal connectivity matrix for macaque cerebral cortex. Cereb. Cortex 24, 17–36.
Matteo, T.D., Aste, T., Dacorogna, M.M., 2005. Long-term memories of developed and emerging markets: using the scaling analysis to characterize their stage of development. J. Bank. Financ. 29, 827–851.
Mizuseki, K., Buzsaki, G., 2013. Preconfigured, skewed distribution of firing rates in the hippocampus and entorhinal cortex. Cell Rep 4, 1010–1021.
Nishida, K., Morishima, Y., Yoshimura, M., Isotani, T., Irisawa, S., Jann, K., Dierks, T., Strik, W., Kinoshita, T., Koenig, T., 2013. EEG microstates associated with salience and frontoparietal networks in frontotemporal dementia, schizophrenia and Alzheimer's disease. Clin. Neurophysiol. 124, 1106–1114.
Obermaier, B., Guger, C., Pfurtscheller, G., 1999. Hidden Markov models used for the offline classification of EEG data. Biomed Tech (Berl) 44, 158–162.
Oh, S.W., Harris, J.A., Ng, L., Winslow, B., Cain, N., Mihalas, S., Wang, Q., Lau, C., Kuan, L., Henry, A.M., Mortrud, M.T., Ouellette, B., Nguyen, T.N., Sorensen, S.A., Slaughterbeck, C.R., Wakeman, W., Li, Y., Feng, D., Ho, A., Nicholas, E., Hirokawa, K.E., Bohn, P., Joines, K.M., Peng, H., Hawrylycz, M.J., Phillips, J.W., Hohmann, J.G., Wohnoutka, P., Gerfen, C.R., Koch, C., Bernard, A., Dang, C., Jones, A.R., Zeng, H., 2014. A mesoscale connectome of the mouse brain. Nature 508, 207–214.
Ouyang, G., Li, J., Liu, X., Li, X., 2013. Dynamic characteristics of absence EEG recordings with multiscale permutation entropy analysis. Epilepsy Res. 104, 246–252.
Park, H., Yun, S., Park, S., Kim, J., Yoo, C.D., 2012. Phoneme classification using constrained variational Gaussian process dynamical system. Advances in Neural Information Processing Systems 25, 2015–2023.
Peng, C.K., Buldyrev, S., Havlin, S., Simons, M., Stanley, H., Goldberger, A., 1994. Mosaic organization of DNA nucleotides. Phys. Rev. E 49, 1685–1689.
Peng, C.-K., Hausdorff, J.M., Goldberger, A.L., 2000. Fractal mechanisms in neural control: human heartbeat and gait dynamics in health and disease. I. In: Walleczek, J. (Ed.), Self-organized Biological Dynamics and Nonlinear Control. Cambridge University Press, Cambridge.
Petersen, A.M., Tenenbaum, J.N., Havlin, S., Stanley, H.E., Perc, M., 2012. Languages cool as they expand: allometric scaling and the decreasing need for new words. Sci. Rep. 2, 943.
Pincus, S.M., 1991. Approximate entropy as a measure of system complexity. Proc. Natl. Acad. Sci. U. S. A. 88, 2297–2301.
Riedi, R.H., Crouse, M.S., Ribeiro, V.J., Baraniuk, R.G., 1999. A multifractal wavelet model with application to network traffic. Information Theory, IEEE Transactions on 45, 992–1018.
Robinson, P., 2003. Time Series with Long Memory. Oxford University Press.
Samorodnitsky, G., 2006. Long range dependence. Foundations and Trends in Stochastic Systems 1, 163–257.
Sarker, M.M.A., 2007. Estimation of the self-similarity parameter in long memeory processes. J Mechanical Engeneering ME 38, 32–37.
Scheffer, M., Bascompte, J., Brock, W.A., Brovkin, V., Carpenter, S.R., Dakos, V., Held, H., van Nes, E.H., Rietkerk, M., Sugihara, G., 2009. Early-warning signals for critical transitions. Nature 461, 53–59.
Van De Ville, D., Britz, J., Michel, C.M., 2010. EEG microstate sequences in healthy humans at rest reveal scale-free dynamics. Proc. Natl. Acad. Sci. U. S. A. 107, 18179–18184.
Varotsos, C., Kirk-Davidoff, D., 2006. Long-memory processes in ozone and temperature variations at the region 60°S–60°N. Atmospheric Chemistry and Physics European Geosciences Union (EGU) 6, 4093–4100.
Wackermann, J., Lehmann, D., Michel, C.M., Strik, W.K., 1993. Adaptive segmentation of spontaneous EEG map series into spatially defined microstates. Int. J. Psychophysiol. 14, 269–283.
Wang, S.S., Shultz, J.R., Burish, M.J., Harrison, K.H., Hof, P.R., Towns, L.C., Wagers, M.W., Wyatt, K.D., 2008. Functional trade-offs in white matter axonal scaling. J. Neurosci. 28, 4047–4056.
Wang, Q., Sporns, O., Bu`rkhalter, A., 2012. Network analysis of corticocortical connections reveals ventral and dorsal processing streams in mouse visual cortex. J. Neurosci. 32, 4386–4399.
Wendt, H., Abry, P., Jaffard, S., 2007. Bootstrap for empirical multifractal analysis. IEEE Signal Process. Mag. 24, 38–48.
Wiedemann, G., Stevens, A., Pauli, P., Dengler, W., 1998. Decreased duration and altered topography of electroencephalographic microstates in patients with panic disorder. Psychiatry Res. 84, 37–48.

Wu, S.D., Wu, C.W., Lin, S.G., Lee, K.Y., Peng, C.K., 2014. Analysis of complex time series using refined composite multiscale entropy. Phys. Lett. A 378, 1369–1374.
Yasumatsu, N., Matsuzaki, M., Miyazaki, T., Noguchi, J., Kasai, H., 2008. Principles of long-term dynamics of dendritic spines. J. Neurosci. 28, 13592–13608.
Yoon, B.J., Vaidyanathan, P.P., 2006. Context-sensitive hidden Markov models for modeling long-range dependencies in symbol sequences. Ieee Transactions on Signal Processing 54, 4169–4184.