Distinct Criticality of Phase and Amplitude Dynamics in the Resting Brain

Converging research suggests that the resting brain operates at the cusp of dynamic instability signified by scale-free temporal correlations. We asked if the scaling properties of these correlations differ between amplitude and phase fluctuations, which may reflect different aspects of cortical functioning. Using source-reconstructed magnetoencephalographic signals, we found power-law scaling for the collective amplitude and for phase synchronization, both capturing whole-brain activity. The temporal changes of the amplitude comprise slow, persistent memory processes, whereas phase synchronization exhibits less temporally structured and more complex correlations, indicating a fast and flexible coding. This distinct temporal scaling supports the idea of different roles of amplitude and phase in cortical functioning.

4. Phase-Amplitude Criticality

4.1 Introduction

It has been proposed that the brain is in or near a critical state — its dynamics may be positioned at the border between spatiotemporal order and disorder, reminiscent of non-equilibrium phase transitions in thermodynamic systems [144, 17, 19, 18]. The concept of brain criticality is attractive because critical systems display optimal performance on several characteristics such as information transfer [98, 17], wide dynamic range [131, 200], information capacity [102, 199], and long-term stability [52, 198]. Criticality relates closely to self-organization [9, 52, 198], which is considered crucial to cortical functioning [54, 28, 102, 181, 73].

A hallmark of critical behavior is the presence of power laws [211, 18]. Power laws symbolize scale-free behavior, adopting the same form on all time scales: they are self-similar. Consider the case of a scale-free auto-correlation function $AC$. The corresponding power law obeys the form $AC(s \cdot \tau) = s^{2H} \cdot AC(\tau)$, i.e. if time $\tau$ is rescaled to $s \cdot \tau$, then the shape of $AC$ is preserved and only rescaled by a factor $s^{2H}$. The scaling exponent $H$ is referred to as the Hurst exponent [119] and qualifies the underlying correlation structure: $H = 0.5$ corresponds to an uncorrelated, random process whereas $H > 0.5$ indicates persistent, long-range correlations.

There is accumulating evidence for the presence of power laws in brain activity [25, 107]. Neural spikes come in avalanches that display scale-free distributions [17]. Spectral distributions of encephalographic signals have $1/f$-structures [87, 71, 110, 86, 141] and their auto-correlation structures also show power-law behavior [144, 143]. In the spatial domain, scale-free distributions have been observed in functional as well as neuroanatomical connectivity patterns [74, 40, 83].

Previous work has focused only on spatially local measures of brain activity [87, 144, 17, 143, 25, 107] or considered pairs of nodes in networks [133, 30] rather than analyzing global brain activity. In complex systems, the global activity can be very informative about the generating dynamical structure [102], in particular when studying critical behavior [192]. Here, we adopted these concepts to investigate criticality in the brain. By using source-reconstructed magneto-encephalographic (MEG) signals we sought to disambiguate between the scaling characteristics of amplitude and phase fluctuations because they may resemble different aspects of cortical functioning.

4.2 Methods

4.2.1 MEG data & outcome variables

Magneto-encephalographic (MEG) signals of ten subjects were recorded and sampled at 1 kHz in eyes-closed resting state for approximately five minutes. After down
sampling to 250 Hz, signals were beamformed onto a ninety-node brain parcellation, yielding ninety time series \( y_k(t) \) per subject. Data were previously published by Cabral and coworkers [45].

Signals \( y_k(t) \) were filtered with a second-order IIR-bandpass filter in the alpha band (8-12 Hz) and (upper) beta band (20-30 Hz). With the Hilbert transform we constructed the analytic signal and defined phase \( \phi_k(t, f) \) and amplitude \( a_k(t, f) \) as functions of time \( t \); \( f \) indexes either the alpha or the beta frequency band.

We used two collective variables to capture whole-brain activity per subject. First, we defined the phase synchronization,

\[
R(t, f) = \frac{1}{90} \left| \sum_{k=1}^{90} e^{i\phi_k(t, f)} \right|
\]

and, second, the mean amplitude,

\[
A(t, f) = \frac{1}{90} \sum_{k=1}^{90} a_k(t, f)
\]

We note that \( R(t, f) \) is the modulo of the complex-valued Kuramoto order parameter [135]. We z-scored \( R(t, f) \) and \( A(t, f) \) to reduce between-subject variability such that we could assess subject-averaged behavior by means of a detrended fluctuation analysis (DFA, [164]), as described below.

To relate our study to the previously established results in RSNs, we also examined the amplitude dynamics in more detail (see also [36]). The expression of RSNs is mainly reflected in the low-frequency content of the amplitudes \( a_k(t, f) \), whose time scale is comparable to those of the blood-oxygenation-level-dependent (BOLD) signal [23]. In order to study these slow amplitude dynamics, we evaluated the longer time scales of the \( a_k(t, f) \) dynamics by extracting its amplitude \( a_k^{(a)}(t, f) \) and phase \( \phi_k^{(a)}(t, f) \). Subsequently, we defined the collective variables associated with the amplitude dynamics:

\[
R^{(a)}(t, f) = \frac{1}{90} \left| \sum_{k=1}^{90} e^{i\phi_k^{(a)}(t, f)} \right|
\]

\[
A^{(a)}(t, f) = \frac{1}{90} \sum_{k=1}^{90} a_k^{(a)}(t, f)
\]

Further analysis was identical to that described above; see Fig. 4.1 for illustration.
Fig. 4.1: Signal \( y_k(t, f) \) after filtering in the alpha band and corresponding Hilbert-amplitude \( a_k(t, f) \) (upper left) and Hilbert-phase \( \phi_k(t, f) \) (lower left). For clarity we decreased the time axis range and rescaled the \( \phi_k(t, f) \) trace in the lower left panel. The amplitude is also displayed (black) in the upper middle panel together with its envelope \( a_k(a)(t, f) \) (red; see main text for the detailed definition). The Hilbert-phase \( \phi_k(t, f) \), equal to \( \phi_k(t, f) \) in the lower left panel, and the Hilbert-phase of \( a_k(t, f) \), \( \phi_k(a)(t, f) \), are displayed in the lower middle panel. Different slopes indicate separate time scales (frequencies). The upper right panel shows amplitudes \( A(t, f) \) and \( A(a)(t, f) \), the lower right one the phase order parameters \( R(t, f) \) and \( R(a)(t, f) \)

4.2.2 Data analysis — DFA

DFA is considered robust against non-stationarity rendering it suitable for analyzing the temporal autocorrelation structure of encephalographic activity, in general, and global amplitude and phase synchronization, in particular. We employed a modified form of DFA including (Bayesian) model selection to verify the presence of power-law behavior [224].

In a nutshell, to quantify the autocorrelation structure of (the cumulative sum of) a signal \( Y(t) \), one divides it into non-overlapping segments \( Y_i(t) \), with \( t = 1, ..., n \) being discrete time steps and \( i = 1, ..., M \) indexing the segments; \( M = \lfloor N/n \rfloor \) is the number of non-overlapping segments of length \( n \). In each segment the linear trend \( Y_i^{\text{trend}}(t) \) is removed providing an estimate of fluctuations in terms of

\[
F_i(n) = \sqrt{\frac{1}{n} \sum_{t=1}^{n} (Y_i(t) - Y_i^{\text{trend}}(t))^2}
\]

This definition yields a set of ‘realizations’ of fluctuations \( F_i \) that, in the presence of a power law, scale like \( F_i(n \cdot \tau) = n^{\alpha} \cdot F_i(\tau) \), which is equivalent to \( \log(F_i) = \alpha \cdot \log(n) + \text{const} \). That is, in a log-log representation these fluctuations have a linear relationship with segment size. DFA seeks to identify the scaling exponent \( \alpha \) that provides an estimate for the aforementioned Hurst exponent \( H \).
Instead of computing the mean value of $F_i$ as in conventional DFA [164], we here determined the probability density function $p_n(F_i)$ for every segment length $n$ (see Fig. 4.2). This approach allows for quantifying the appropriateness of a model $f_\theta(\tilde{n})$ for fitting the fluctuation structure $\tilde{F} = \log(F_i)$ as a function of $\tilde{n} = \log(n)$ by means of the log-likelihood function $\ln(\mathcal{L}) = \sum_n \ln(p_n(f_\theta))$, where the tilde indicates a transformation to logarithmic coordinates. Here, the model $f_\theta(\tilde{n})$, parametrized by the set $\theta$, may obey any arbitrary form including the linear one, which corresponds to a power law. We tested this linear relationship against a set of alternative models (Tab. 4.1) using both the Bayesian information criterion and the Akaike information criterion (BIC and AIC$_c$, respectively). The model resulting in the least value of the information criterion was selected as the proper model. Whenever this yielded the linear model $f_1$, we considered power-law behavior to be present and identified the scaling exponent with its slope [224].

Tab. 4.1: The set of candidate models for the selection procedure. The linear model $f_1(x)$ is the form a power law would adopt. The alternative models $f_2 - f_9$ constitute polynomials up to third order. With $f_6(x)$ and $f_9(x)$ we considered two models resembling a (un)stable linear stochastic dynamics.

\[
\begin{align*}
    f_1(x) &= \theta_1 + \theta_2 x \\
    f_2(x) &= \theta_1 + \theta_2 x^2 \\
    f_3(x) &= \theta_1 + \theta_2 x + \theta_3 x^2 \\
    f_4(x) &= \theta_1 + \theta_2 x^3 \\
    f_5(x) &= \theta_1 + \theta_2 x + \theta_3 x^3 \\
    f_6(x) &= \theta_1 + \theta_2 x^2 + \theta_3 x^3 \\
    f_7(x) &= \theta_1 + \theta_2 x + \theta_3 x^2 + \theta_4 x^3 \\
    f_8(x) &= \theta_1 + \theta_2 x + \theta_3 x^2 + \theta_4 x^3 \\
    f_9(x) &= \theta_1 + \frac{1}{\ln(10)} \ln \left( \theta_2 \left( 1 - e^{-\theta_3 e^{\ln(10) x}} \right) \right)
\end{align*}
\]

Since we were interested in the subject-averaged scaling exponents, we determined $p_n$ for every subject individually and averaged over subjects to obtain $\overline{p}_n$. These averaged probability density functions were used both for model selection and to determine the scaling exponent $\alpha$. The scaling range was given by $[n_{\min}, n_{\max}] = [1.875 \cdot 10^2, 1.875 \cdot 10^4] \simeq [0.75, 75]$ seconds, i.e. it spanned two decades. In this range we used one hundred equally spaced window sizes (on a logarithmic scale).

### 4.2.3 Statistics — surrogate data

Bootstrapping served to establish statistical significance using three types of surrogate data. Two of them consisted of randomly permuting temporal order whereas the third one only influenced cross-correlation structure. Of all types we constructed 1000 surrogates and significance values were obtained using $p = 1 - \int_{-\infty}^{\alpha} p_{surr}(h) \, dh$, where $p_{surr}(h)$ denotes the surrogate distribution and $\alpha$ the obtained empirical value
Fig. 4.2: Figure illustrating the relation between $F$ (red dots at the bottom) and the individual densities $p_n$ (red lines, Fig. 4.2(a)) and averaged densities $\bar{p}_n$ (red lines, Fig. 4.2(b)) for $A(t)$ in the alpha band; we show $p_n$ only for a few values of $n$. Fig. 4.2(a) depicts the histograms of $F_i(n)$ on basis of which $p_n$ were determined by kernel density estimation. The $F$ values are the expectation values of $p_n$ and $\bar{p}_n$, respectively.

of the scaling exponent.

For the first type of surrogates, we randomly permuted the order parameter time series for each subject. With this we evaluated our DFA and fitting procedure, since the surrogate time series lacked any temporal correlation structure and therefore should result in $\alpha = 0.5$ [147]. With the second type of surrogate we evaluated whether the filtering procedure and Hilbert analysis could have biased the results. For this, we permuted all original time series $y_k(t)$. By permuting $y_k(t)$, all temporal structure was destroyed and therefore this constituted a rather weak null. In the third type of surrogates we performed a random cycling of $y_k(t)$ by shifting the time indices of $y_k(t)$ for each $k$, but keeping their order intact. In this way we retained the original autocorrelation structure of $\phi_k(t, f)$ and $a_k(t, f)$ but destroyed the cross-correlations. Subsequent analyses for all collective variables were identical to those for the original data.

4.3 Results

The presence of power-law scaling was evidenced by clear linear relationships of $A(t, f)$ and $R(t, f)$ fluctuations in log-log scale; see Fig. 4.3. This was confirmed by the AICc and BIC values preferring the linear model in all cases (Tabs 4.2-4.3). In both chosen frequency bands, the brain’s network dynamics, as measured by the collective variables $A(t, f)$ and $R(t, f)$, thus appeared to exhibit scale-free correlations over a very broad range of time scales.
In the alpha band the scaling exponents were 0.88 and 0.62, and in the beta band 0.80 and 0.57, for amplitude $A$ and phase synchronization $R$, respectively. That is, both $A(t,f)$ and $R(t,f)$ showed persistent behavior but amplitude had increased perseverance compared to phase (Fig. 4.3).

We also found long-range temporal correlations in the variables $A^{(a)}(t,f)$ and $R^{(a)}(t,f)$ as shown in Fig. 4.4 and confirmed by the AICc and BIC values by preferring the linear model, except in case of the AICc for $A^{(a)}(t,f)$ in the alpha band (Tabs 4.2-4.3). The difference in scaling behavior between $A^{(a)}(t,f)$ and $R^{(a)}(t,f)$ was similar to that between $A(t,f)$ and $R(t,f)$.

Tab. 4.2: Model selection results for the alpha frequency band using $\text{BIC} = -2 \ln (\mathcal{L}) + K \ln (M)$ and $\text{AICc} = -2 \ln (\mathcal{L}) + 2K + \frac{2K(K-1)}{(M-K-1)}$, with $K$ being the number of the parameters per model. The table shows relative values $\Delta \text{BIC} = \text{BIC} - \min(\text{BIC})$. In all cases the linear model $f_1^o$ resulted in minimal BIC values indicating power-law scaling in all variables. Corresponding $\Delta \text{AICc}$ values are given in brackets.

<table>
<thead>
<tr>
<th>Model</th>
<th>$\Delta \text{BIC}$</th>
<th>$\Delta \text{AICc}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$f_1^o(x)$</td>
<td>0.00 (0.00)</td>
<td>0.00 (0.00)</td>
</tr>
<tr>
<td>$f_2^o(x)$</td>
<td>5.85 (5.85)</td>
<td>3.56 (3.56)</td>
</tr>
<tr>
<td>$f_3^o(x)$</td>
<td>4.60 (2.12)</td>
<td>4.60 (2.12)</td>
</tr>
<tr>
<td>$f_4^o(x)$</td>
<td>16.21 (16.21)</td>
<td>12.09 (12.09)</td>
</tr>
<tr>
<td>$f_5^o(x)$</td>
<td>4.60 (2.12)</td>
<td>4.58 (2.10)</td>
</tr>
<tr>
<td>$f_6^o(x)$</td>
<td>4.89 (2.41)</td>
<td>4.94 (2.46)</td>
</tr>
<tr>
<td>$f_7^o(x)$</td>
<td>9.02 (4.11)</td>
<td>7.01 (2.10)</td>
</tr>
<tr>
<td>$f_8^o(x)$</td>
<td>4.60 (2.12)</td>
<td>4.68 (2.20)</td>
</tr>
<tr>
<td>$f_9^o(x)$</td>
<td>32.05 (29.57)</td>
<td>6.01 (3.53)</td>
</tr>
</tbody>
</table>
Testing against surrogate data confirmed the significance of these correlations. Fig. 4.5 depicts the scaling exponent distributions corresponding to the third type of surrogates. For all variables in both frequency bands the scaling exponents significantly exceeded those of the surrogates \((p < .01)\).

**Fig. 4.4:** Scaling behavior for \(A^{(a)}(t)\) (red) and \(R^{(a)}(t)\) (black). **Figs 4.4(a), 4.4(b)** show results for the alpha/beta frequency band. Scaling exponents \(\alpha\) are equal to 0.96 \((A^{(a)})\) and 0.78 \((R^{(a)})\) in the alpha band and 0.87 \((A^{(a)})\) and 0.69 \((R^{(a)})\) in the beta band. Shaded areas refer to the 25\(^{th}\) and 75\(^{th}\) percentiles of the subject averaged densities \(p_{\text{avg}}\); see Tabs 4.2-4.3 for the model selection results.

To further highlight the peculiar role of amplitude and phase, we finally contrasted our results with the scaling of fluctuations of the mean MEG activity. For this, we applied our DFA to \(Y(t, f) = \frac{1}{90} \sum_{k=1}^{90} y_k(t, f)\), i.e. we considered not the amplitude and phase but the ‘raw’ MEG signals. This mean activity did not display long-range correlations but rather anti-persistent ones \((\alpha = 0.02, \text{Fig. 4.6})\).

**Tab. 4.3:** Relative values \(\Delta \text{BIC} = \text{BIC} - \min(\text{BIC})\) for all variables in the beta frequency band; cf. Tab. 4.2. In all cases the linear model \(f_{\theta}^1\) resulted in minimal BIC values indicating power-law scaling in all variables. Corresponding \(\Delta \text{AIC}_c\) values are given between brackets.

| \(f_{\theta}^1(x)\)  | 0.00 (0.00) | 0.00 (0.00) | 0.00 (0.00) | 0.00 (0.10) |
| \(f_{\theta}^2(x)\)  | 0.37 (0.37) | 2.72 (2.72) | 0.35 (0.35) | 1.36 (1.36) |
| \(f_{\theta}^3(x)\)  | 4.04 (1.57) | 4.60 (2.12) | 4.17 (1.69) | 4.50 (2.02) |
| \(f_{\theta}^4(x)\)  | 6.29 (6.29) | 11.46 (11.46) | 5.33 (5.33) | 6.83 (6.83) |
| \(f_{\theta}^5(x)\)  | 3.99 (1.51) | 4.60 (2.12) | 4.10 (1.62) | 4.49 (2.01) |
| \(f_{\theta}^6(x)\)  | 4.20 (1.72) | 4.57 (2.09) | 4.33 (1.85) | 4.63 (2.15) |
| \(f_{\theta}^7(x)\)  | 8.15 (3.24) | 9.18 (4.27) | 8.36 (3.44) | 8.89 (3.98) |
| \(f_{\theta}^8(x)\)  | 4.02 (1.54) | 4.62 (2.14) | 4.14 (1.66) | 4.52 (2.04) |
| \(f_{\theta}^9(x)\)  | 55.24 (52.76) | 10.34 (7.86) | 26.24 (23.76) | 12.97 (10.49) |
4.4 Discussion

We report power-law scaling in both amplitude and phase of collective neural activity on long time scales, which is consistent with the hypothesis that the brain operates in a critical state. Operating in a critical state is not the only way a system can generate power-law scaling. Systems in subcritical states [30] or merely stochastic systems [226, 227] may also display power laws. Biological systems display sub- and supercritical dynamics but they can be tuned into criticality [149, 198, 78]. We favor the interpretation of critical states, also because it is consistent with scale-free autocorrelation structures of single channel EEG activity [144, 143], size and duration of neural avalanches [17, 181, 201] and, in the spatial domain, degree distributions of neuroanatomical and functional connectivity networks [74, 40]. Such scaling laws in neuronal dynamics are also correlated with those found in behavior [162].
Previous studies addressing power-law scaling in neuronal dynamics were solely based on spatially local measures. For instance, power-law scaling has been reported using pair-wise synchronization measures like PLI and $\Delta^2(t, \Delta t)$ by, e.g., Kitzbichler and coworkers [133, 132] and Farmer [79]. In contrast, we evaluated scaling behavior on a global brain scale by using overall amplitude and phase synchrony as collective variables. Analyzing the scaling behavior of these variables enabled us to directly compare phase and amplitude behavior in brain activity.

What does the difference in power-law scaling of amplitudes and phases tell us about information processing in the brain? Despite the fact that amplitude and phase differ in their perseverance, both show long-range correlations over a scale of hundreds of seconds, suggesting that they reflect ‘memory’ of cortical states. A higher scaling exponent reflects a slower decay of auto-correlations and hence a more predictable signal with decreased entropy [49]. Therefore, we speculate that the difference in scaling exponents and their associated complexity reflect that amplitude and phase play different roles in information processing and memory decoding: (low-fidelity) amplitude dynamics decode long-term memory, whereas the (high-fidelity) phase synchrony comprises a more complex and flexible memory coding (in an information theoretic sense).

Volume conduction can be a confounder in analyzing encephalographic recordings [189]. Several methods to mitigate its effects have been proposed, generally relying on removing the instantaneous interactions that signify volume conduction [157, 210, 37]. In consequence these methods can only be applied in a pair-wise fashion, such that they are not applicable when considering the variables $R^{(\cdot)}(t, f)$ and $A^{(\cdot)}(t, f)$. We note, however, that volume conduction does not significantly influence the auto-correlation structure of the signals under study. This finds support by Shriki and coworkers [201] who showed that mere linear mixing cannot ‘transform’ uncorrelated activity to power-law scaling. In fact, if activity displays a power law, linear mixing does not alter this apart from slightly lowering the scaling exponent.

The occurrence of power laws is not only consistent with the ‘criticality hypothesis’ [19]. The macroscopic behavior of self-organizing processes can — in general — be cast into a low-dimensional system when critical [104]; this in fact motivated looking at collective amplitude and phase synchrony. Several modeling studies support the seminal role of self-organization in neural dynamics and often highlight self-organized criticality. For example, synaptic plasticity under the influence of a simple learning rule leads to scale-free networks [54, 28, 73] and power-law distributions of avalanche dynamics [181]. In resting state, the mechanism of local feedback mediated inhibition increases model performance [66] and may also be interpreted as a form of self organization. Furthermore, the slow evolution of the order parameters in self-organizing systems is consistent with the time scale on which RSNs evolve [23, 174, 82, 36].
While previous work has shown that RSNs fluctuate at slow (> 1sec) time scales [23, 36], it has been recently shown that RSNs are also expressed in MEG activity on faster (< 1sec) time scales [10]. Alongside the results in this paper, this represents mounting evidence that RSNs are expressed across a range of time scales, i.e. they are time scale invariant.

In summary we have shown the presence of persistent long-range correlations in the evolution of global brain dynamics, i.e. the auto-correlation function obeys a power-law with scaling exponents exceeding those corresponding to random processes without memory. This adds further support to the hypothesis that the brain is in a (permanently) critical state. The here-reported scaling exponents clearly discriminate amplitude and phase dynamics, suggesting their differential role in whole-brain information processing.