Caffeine induces age-dependent increases in brain complexity and criticality during sleep

Philipp Thölke[1,2*], Maxine Arcand-Lavigne[1,2], Tarek Lajnef[1,2], Sonia Frenette[2,3], Julie Carrier[2,3], Karim Jerbi[1,2,4,5],

1 Computational and Cognitive Neuroscience Lab (CoCo Lab), Université de Montréal, Montréal, QC, Canada
2 Psychology Department, Université de Montréal, Montréal, QC, Canada
3 Centre for Advanced Research in Sleep Medicine, Université de Montréal, Montréal, QC, Canada
4 MILA (Quebec Artificial Intelligence Institute), Montréal, QC, Canada
5 UNIQUE Center (Quebec Neuro-AI Research Center), Montréal, QC, Canada

* philipp.thoelke@posteo.de

Abstract

Caffeine is the most widely consumed psychoactive stimulant worldwide. Yet important gaps persist in understanding its effects on the brain, especially during sleep. We analyzed sleep EEG in 40 subjects, contrasting 200mg of caffeine against a placebo condition, utilizing inferential statistics and machine learning. We found that caffeine ingestion led to an increase in brain complexity, a widespread flattening of the power spectrum's 1/f-like slope, and a reduction in long-range temporal correlations. Being most prominent during non-REM sleep, these results suggest that caffeine shifts the brain towards a critical regime and more diverse neural dynamics. Interestingly, this was more pronounced in younger adults (20-27 years) compared to middle-aged participants (41-58 years) whose sleep brain dynamics were less affected by caffeine. Interpreting these data in the light of modeling and empirical work on EEG-derived measures of excitation-inhibition balance provides novel insights into the effects caffeine has on the sleeping brain.

Introduction

For many, enjoying the delicate flavors of an espresso coffee ranks high up on the list of life's little pleasures. Unfortunately, although caffeine-containing products have a range of positive effects, such as increased alertness, mental focus and cognitive performance, it also has a disruptive effect on the quality of sleep, which is critical for general health and well-being [1, 2]. On the other hand, caffeine has been shown to exhibit neuroprotective qualities, particularly against Parkinson's disease [3–6], adding to the complexity of its impact on health. Caffeine is a psychoactive stimulant that is consumed by people across all age groups on a daily basis [7] through a wide variety of products such as coffee, tea, soft drinks, energy drinks, chocolate and several pharmaceutical drugs [8]. It is therefore critical to understand how caffeine affects the brain during sleep, and across age.

Caffeine affects the quality of sleep in several ways. It increases sleep latency (i.e. the time it takes you to fall asleep) and it decreases sleep efficiency (the ratio of total

sleep time to the time you spend in bed). The reduction of sleep duration caused by caffeine intake is particularly visible in the amount of time spent in the S2 sleep stage [9–12]. Further, both acute and regular daytime caffeine intake was found to delay REM sleep promotion and lead to a reduction in the quality of awakening [13]. However, a recent study proposed that the caffeine-induced reduction in sleep duration observed in humans could be attributed to a lack of flexibility in wake-up times instead of being a direct effect of caffeine [14]. Nevertheless, lack of sleep and sleep disorders can lead to the deterioration of the proper functioning of sleep-related brain processes, weight gain [15], hypertension [16], cardiovascular diseases, diabetes [17, 18] and increases the risk of depression [19].

As a psychostimulant and an adenosine antagonist, caffeine reduces natural homeostatic sleep pressure by binding to adenosine receptors, thereby inducing a feeling of higher alertness and invigoration [20]. Adenosine, which is crucial for brain homeostasis, has been shown to mediate permeability of the blood-brain barrier (BBB) by binding to A_1 and A_2 adenosine receptors [21, 22]. Activation of these receptors leads to increased permeability of the BBB, allowing transport of macromolecules like dextrans and β-amyloid antibodies into murine brains [21]. Hence, because it is an adenosine antagonist, caffeine is likely to interfere with BBB permeability rhythms during sleep. Despite encouraging progress, the mechanisms by which caffeine alters brain dynamics during sleep remain poorly understood.

Previous research has reported caffeine-induced effects on spectral power of electroencephalography (EEG) signals recorded during sleep. Generally, caffeine was found to decrease power in low frequency oscillations in the delta band and increase power in sigma and beta frequencies [9, 10, 12, 23–26]. Another study conducted in Cynomolgus monkeys, which have been suggested to have sleep patterns similar to those of humans [27, 28], found that caffeine ingestion led to a decrease in delta and theta power ($1 - 8$ Hz) and increased beta and low-gamma power ($20 - 50$ Hz) during sleep [29].

While spectral power analysis provides important insights into the oscillatory properties of brain signals, it does not capture the diversity and full complexity of EEG signals. Complementary insights can be harnessed by assessing brain signal complexity using entropy metrics. High EEG entropy indicates a complex and random pattern of brain activity, while low entropy represents a more ordered and predictable pattern of activity. Typically, high entropy signals are seen in states of wakefulness, while low entropy signals are observed during deep sleep or anesthesia [30]. In fact, entropy was shown to vary consistently across sleep stages with wakefulness exhibiting the highest levels of entropy, followed by REM and NREM sleep [31]. Crucially, brain entropy has also been found to correlate positively with an array of cognitive functions including attention, memory and verbal fluency [32]. Taken together, studies of brain entropy support the view that higher order cognitive processes, as well as states of high vigilance, are mediated by brain dynamics that exhibit heightened complexity. Interestingly, accumulating evidence from a parallel yet closely related stream of research suggests that optimal information processing and behavioral capabilities are supported by critical brain dynamics [33, 34]. Like complexity, brain criticality also drops with reductions in the levels of vigilance or consciousness.

The reported association between brain complexity or criticality on one hand, and cognitive performance or arousal on the other, naturally leads to asking whether caffeine affects the brain by elevating its complexity and criticality. Importantly, age has been related to an increase in entropy and a shift towards the critical point [31, 35–37], it is also tempting to hypothesize that the extent of their putative alteration by caffeine is likely to be age-dependent.

Evidence from one fMRI resting-state study suggests that caffeine induces a global

increase in the entropy of the haemodynamic BOLD response during a resting-state protocol [38]. However, it is unclear whether such entropy increases also occur in direct neurophysiological measures of brain activity, such as EEG. And even more importantly, it is unknown whether caffeine ingestion leads to higher entropy -and possibly higher criticality- during sleep. Finally, whether any such neural alterations are age-dependent is yet another important open question.

The aim of the present study was to address these questions by providing an in-depth assessment of the influence of caffeine on the brain's electrophysiological signals during sleep. In particular, we tested the central hypothesis that caffeine ingestion affects sleep EEG by increasing both its complexity and criticality. Additionally, we also expanded on previous power spectral investigations of the effect of caffeine by disentangling periodic and aperiodic components in the EEG spectra.

By contrasting the effects of caffeine and placebo using inferential statistics and machine learning (ML), we provide the first evidence that caffeine induces a broad increase in EEG brain complexity and criticality during sleep. These effects were more widespread in NREM, compared to REM sleep, and strikingly more prominent in young adults compared to middle-aged individuals.

1 Results

1.1 Caffeine-induced changes in sleep EEG oscillations

Figure 1 shows the effect of caffeine, compared to placebo, on spectral power in five key frequency bands, compensating for changes in the aperiodic component (see Section 3.5). The topographies depict the results of standard statistical testing (T-tests) alongside ML decoding accuracy topographies (SVM and LDA classifiers), and are presented separately for NREM and REM sleep stages.

During NREM sleep, caffeine led to statistically significant reductions in power in the delta, theta and alpha frequency bands (3 left columns in Fig. 1). While this decrease spanned parietal and central channels for the delta and theta bands, the decrease in alpha power was localized between central and prefrontal channels. Interestingly, beta power showed a widespread increase across parietal and frontal channels. Compared to the results of the SVM and LDA classification, the t-test analyses yielded a larger number of channels with statistically significant results. This said, the spatial distributions of the significant results were consistent across all three methods, which speaks to the robustness of the findings across within-sample and out-of-sample analysis methods.

The most prominent caffeine effect observed during REM sleep was a significant reduction of theta power in temporal, parietal and occipital channels. While the t-tests showed statistically significant differences between caffeine and placebo, the decoding accuracy obtained with SVM and LDA, although elevated, did not reach the $p<0.01$ significance threshold level. This said, the LDA-based analysis during REM sleep revealed statistically significant ($p<0.05$) classification using beta power in a frontal cortex. This effect was driven by a reduction in beta power, which contrasts with the strong beta power increase observed during NREM sleep.

1.2 Disentangling contributions from periodic and aperiodic components to caffeine-induced spectral changes

Crucially, and in contrast to previous work, the caffeine-induced power modulations discussed above were obtained using spectral feature computations that correct for the 1/f slope in the power spectrum (see methods). This correction takes into account the

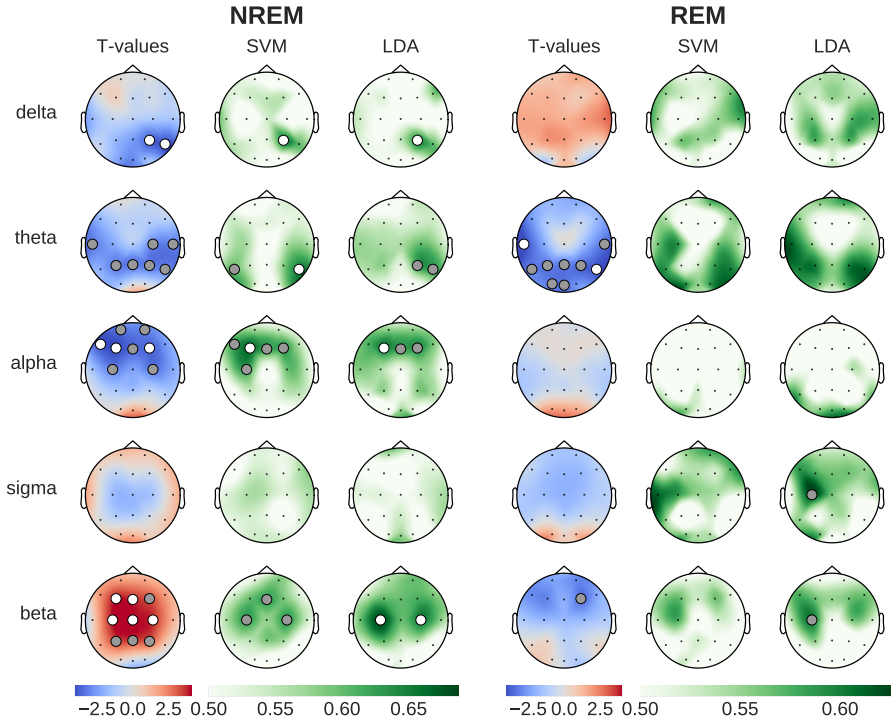


Fig 1. Topographical plots showing the spatial distribution of T-values (caffeine-placebo) and classification decoding accuracy (SVM and LDA) based on EEG power spectral density features across different frequency bands, compensating for changes in the aperiodic component (see Section 3.5). The T-values and the decoding accuracy values were tested for statistical significance using permutation tests. Grey dots represent significance at $p<0.05$ and white dots at $p<0.01$, all corrected for multiple comparisons using the maximum statistics.

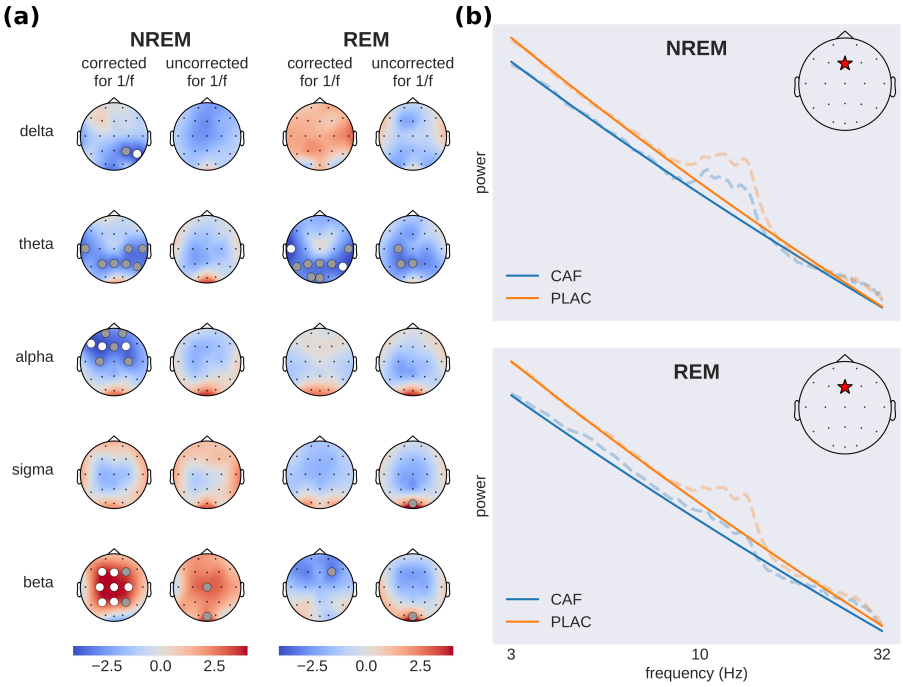


Fig 2. Caffeine-induced power differences in NREM and REM sleep using corrected versus uncorrected spectra **(a)** Topographical plots highlighting the caffeine-induced spectral power modulations before (uncorrected) and after (corrected) retrieval of the aperiodic component (1/f slope) from the power spectrum. As in Fig. 1 the plots represent T-values comparing caffeine and placebo conditions (see Section 3.6) where grey and white dots represent statistical significance at $p<0.05$ and $p<0.01$ respectively (assessed with permutations and maximum statistics correction). **(b)** Illustrative log-log power spectra in NREM (top) and REM (bottom) sleep for caffeine (CAF) and placebo (PLAC) conditions from channel Fz (red star) in a single subject. Dashed lines represent the full power spectrum (i.e. periodic + aperiodic components), solid lines represent the fitted aperiodic component (see methods).

possibility that the aperiodic component of the spectrum (1/f slope) may be different across the two experimental conditions. In which case, not correcting for the 1/f slope may lead to confounding results that amalgamate oscillatory and non-oscillatory signals. To appreciate the impact of this correction on the EEG presented here, we repeated the spectral power comparisons (caffeine vs placebo) shown in Figure 1 but without correcting for the 1/f slope. The results of both approaches are shown in Figure 2a. While the direction of the effect is largely consistent across both approaches (with the exception of delta power during REM sleep), we found the effects to be more pronounced after removing the aperiodic component from the power spectrum (channels with statistically significant T-values are scare without the correction): Several frequency bands do not show any significant effects when looking at the full original power spectrum and only become significant after removal of the aperiodic component.

Figure 2b depicts an illustrative example of power spectra from a single subject (dashed lines) and the fitted aperiodic components (solid lines), which are straight lines in a log-log plot. A flattening of the aperiodic component after the ingestion of caffeine is visible (see Section 1.3).

Taken together, the comparisons between the corrected and uncorrected spectra highlight the utility of removing the aperiodic component when assessing caffeine's impact on neural oscillations during sleep. In addition, the differences we observed between the results obtained with and without this correction strongly suggest that the aperiodic component of the spectrum (i.e. its slope) is affected by caffeine - otherwise removing it would not have modified the results. These changes in the scaling behavior of the EEG spectra can be linked to altered EEG self-similarity, which in turn may be related to shifts in neural criticality. In general, the discrepancies we found between the corrected and uncorrected spectral power results point towards the relevance of considering the aperiodic component of the power spectrum as a candidate EEG feature that is altered by caffeine.

1.3 Caffeine-induced changes in sleep EEG complexity and criticality

To assess the differential impact of caffeine and placebo on EEG complexity and criticality during sleep, we contrasted the two conditions using statistical inference and supervised ML, similar to our assessment of spectral power changes. The results shown in Figure 3 depict the T-values and ML decoding accuracy obtained with multiple metrics in NREM and REM sleep stages. Overall, all EEG complexity measures (SpecEn, SampEn, SpecSampEn and LZc) showed consistent increases under caffeine ($p<0.01$ corrected). In addition, the two metrics we used to probe shifts in critical behavior (i.e. DFA scaling exponent and the slope of the aperiodic activity) both showed consistent reductions, indicative of a shift in critical behavior. During NREM sleep, the caffeine-induced modulations in brain signal complexity and criticality were widespread and prominent. However, during REM sleep, the effects were limited to the occipital regions and remarkably weaker (Fig. 3). More globally, the direction of the effects were consistent among the four measures of complexity, and among the two metrics used to assess critical behavior. Interestingly, the most accurate classification between caffeine and placebo sleep EEG was obtained by using Spectral Sample Entropy as a feature in NREM sleep, with a 75% decoding accuracy, surpassing the best results from using spectral power. It is also worth mentioning that the mean DFA scaling exponent during NREM and REM sleep took values of 1.31 and 1.21 respectively.

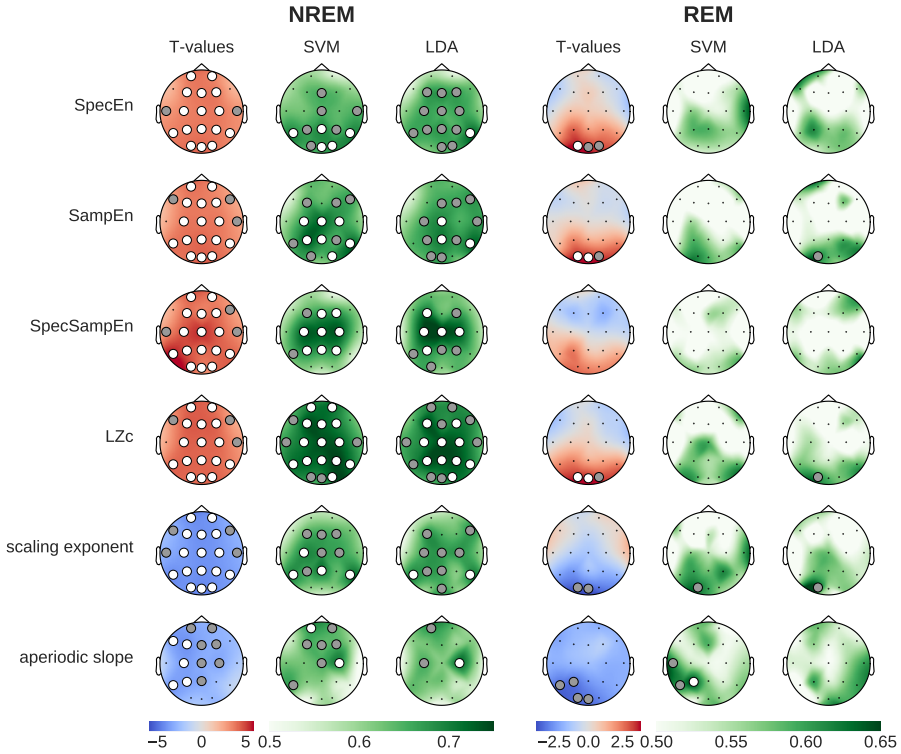


Fig 3. Topographical plots showing the spatial distribution of T-values and decoding accuracy (SVM and LDA) of complexity- and criticality-related measures. T-values were computed using exhaustive permutation tests of the caffeine condition versus placebo. Decoding accuracy was tested for statistical significance using permutation tests. Significance scores were corrected for multiple comparisons using the maximum statistics method. Grey dots represent significance at $p<0.05$ and white dots at $p<0.01$.

1.4 Single-epoch classification

The features used in the classification analyses discussed so far were based on computing mean values at each electrode in each subject. We decided to extend this by asking whether it is possible to train a classifier to discriminate between caffeine and placebo conditions using features computed from single-epochs, without averaging. The findings were largely inline with the results observed with the subject-wise averaged features (Supplementary Figure S1 Fig). This said, the decoding performances were lower because the single-epoch features were obviously noisier than the averaged features. Given the large number of samples, decoding scores in the single-epoch classification were found to be statistically significant ($p<0.01$, corrected) in a broad selection of channels even though passing the statistical decoding chance level only marginally. The overall agreement between the single-epoch and subject-wise ML results speaks to the robustness of the observations as well as to the single-trial sensitivity of the selected features to the effect of caffeine.

1.5 Multi-feature Machine learning

In order to further explore the insights that can be harnessed through a data-driven approach, we built a random forest (RF) classifier which we trained on all extracted features simultaneously (11 features $\times$ 20 channels = 220 total). The spectral features were corrected for the aperiodic component. We chose the RF classifier as it provides straightforward access to individual feature importance scores, allowing us to rank and compare the contributions of the distinct features in discriminating between the caffeine and placebo data. We found that this multi-feature RF classifier achieved substantially higher decoding accuracy in NREM sleep (75.22% $\pm$ 0.15) than in REM sleep (58.63% $\pm$ 0.13) after averaging the scores from 1000 random reinitializations of the classifier. However, more importantly, the feature importance scores from the RF model shown in Figure 4 revealed interesting insights into the nature of the most relevant features as well as their topographical distribution. Remarkably, the feature importance bar plots (colored according to the type of feature) reveals that, among all features, the complexity features are the most useful for the caffeine versus placebo classification during NREM sleep (all the blue features). These were followed by the criticality related metrics (green), and then the corrected spectral power (warm colors). A closer look reveals that LZc and SpecSampEn exhibited high degrees of feature importance during NREM, contributing 37% of the classifier decision across all 11 features. Additionally, SampEn and the DFA scaling exponent saw prominent levels of feature importance. In other words, we found that the complexity and criticality-related measures outperformed the spectral features in terms of their contribution to the classification of caffeine versus placebo samples during NREM.

By contrast, the distribution of features ranked by importance during REM was more heterogeneous and less structured, including high levels of importance in temporal theta power, occipital DFA scaling exponent and SpecEn. The sigma power and the aperiodic slope additionally exhibited reasonably high contributions during REM as well.

Supplementary Table S2 Table provides a list of the ten most important feature-electrode pairs as determined by the RF classifier. These results highlight the prominence and leading contribution of complexity (as measured by LZc and SpecSampEn) over central, parietal and frontal areas.

1.6 Age-related effects

Because all the EEG features we explored can show age-related alterations, we decided to investigate the impact of caffeine on sleep EEG, by separating the data into two

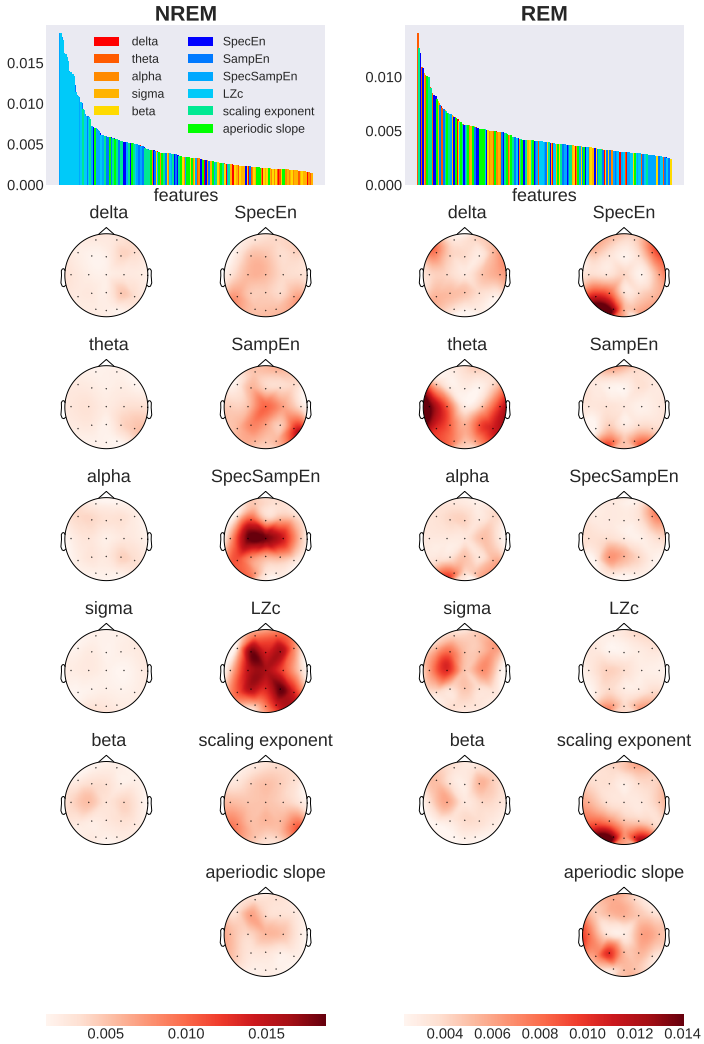


Fig 4. Visualization of feature importance in the random forests trained on a feature space which combines all feature types across all channels (11 features × 20 channels = 220 total). The bar-plots show a ranking of all input dimensions, colored according to the corresponding feature. Warm colors (red to yellow) show spectral power (corrected) in five frequency bands, while cold colors (blue and green) correspond to complexity and criticality measures. The topographical maps depict the spatial distribution of feature importance. 1000 random forests were trained per sleep stage (NREM and REM), here we show the average feature importance across all models from one sleep stage. Higher values indicate higher importance of a feature at a specific channel.

separate age groups (young and middle-aged adults). Figure 5 depicts side-by-side the paired T-test results (caffeine vs. placebo) computed for (i) all subjects combined ("all", 20-58 years old, mean age 35.3 years; n=40, 19 females), (ii) the young adults subgroup (20-27 years old, mean age 22.8 years; n=22, 10 females) and (iii) the middle-aged subgroup (41-58 years old, mean age 50.6 years; n=18, 9 females). Despite the smaller sample sizes within each subgroup, significant effects were observed in both subgroups.

Generally, the feature-wise topographies of caffeine-induced increases and decreases in NREM sleep were consistent across the two subgroups, and they both resembled the results from the full cohort. However, the significant effects were overall more prominent in the young compared to the middle-aged subgroup, which suggests that the effects in the full cohort might be driven by the samples from the young adults. A few exceptions include the decrease in theta power during NREM sleep which, on the contrary, appears to be largely driven by the middle-aged subgroup. As for the REM data, we also found strong similarities between the effect of caffeine in the young adults and in the full population. However, the T-value maps in the middle-aged group were less consistent with those of the younger subjects, showing partly divergent results. In particular, by contrast to the results in the young adults, REM sleep in middle-aged individuals did not show caffeine-induced increases in complexity. Likewise, the effect of caffeine on REM sleep criticality – as assessed by DFA scaling exponent and aperiodic slope – was also absent in the middle-aged group. One tempting explanation that could in theory account for the weaker effects of caffeine – or even absence thereof – in the middle-aged individuals is that their spontaneous brain dynamics have already been modified by age in a manner akin to the acute modifications brought about by caffeine. And therefore, the ingestion of caffeine only has a marginal or insignificant impact in individuals from this age group.

To test this hypothesis, we conducted an analysis where we probed age-related differences in all of the features using data from the placebo nights only. The results of the statistical comparisons between young and middle-aged subjects are shown in Supplementary Figure S2 Fig. Compared to younger adults, the baseline brain activity (placebo nights) of middle-aged individuals exhibits higher complexity and lower DFA scaling exponent and a flatter 1/f slope. Incidentally, these age-related effects are the same as the caffeine-related effects observed in the younger individuals. Interestingly, although these age-related differences were more pronounced in REM sleep, they were also visible during NREM sleep. Taken together, our findings support our intuition that caffeine has a weaker impact on sleep EEG in middle-aged individuals because aging itself is associated with comparable alterations to the EEG baseline.

1.7 Mitigating potential confounds

We ensured that the diminished significance in the middle-aged group, compared to the young group, was not related to the smaller number of subjects (18 vs. 22 in the younger group) by employing bootstrapping (1000 iterations) with the subjects in the younger group. This demonstrated that the results in the young group remain valid when considering only 18 subjects.

Given that caffeine alters total sleep duration and architecture (distribution of sleep stages), it is theoretically conceivable that some of our findings were caused by an unequal amount of epochs between caffeine and placebo. To rule out this possibility we added a control analysis where we used random subsampling to equate the number of epochs used to compute the features across conditions. This was done on a per-subject basis and separately for NREM and REM. This control analysis showed the robustness of the reported results and ruled out that they can entirely be attributed to caffeine-induced changes in sleep duration; we replicated the same statistical effects from Figures 1 and 3, see Supplementary Figure S3 Fig.

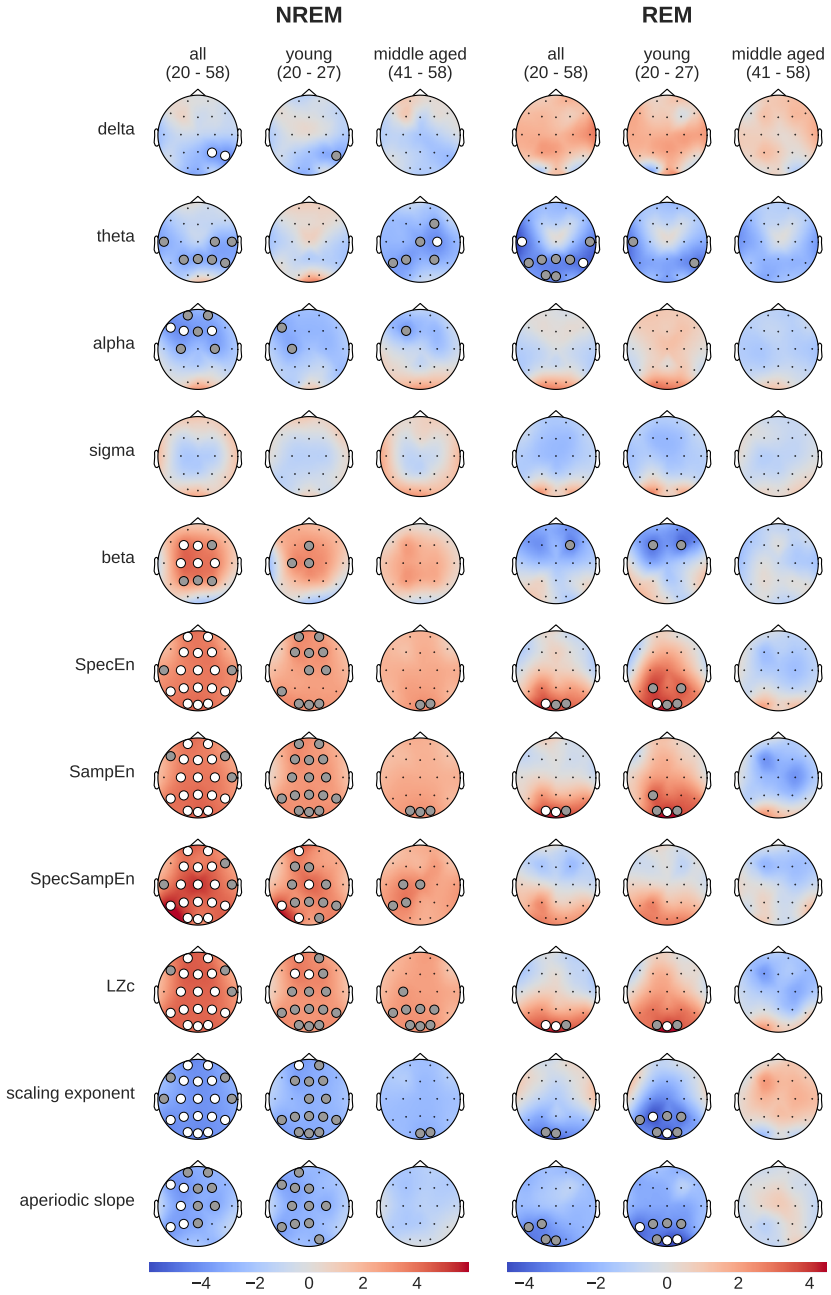


Fig 5. Topographical plots showing the spatial distribution of T-values (caffeine-placebo), compared across age groups. T-values were computed using exhaustive permutation tests of the caffeine condition minus placebo across. Grey dots represent significance at $p<0.05$ and white dots at $p<0.01$.

2 Discussion

The objective of this study was to investigate the effect of caffeine ingestion on the brain during sleep. In particular, we compared the effects of caffeine to that of placebo on neural oscillations, brain complexity and measures of critical behavior. Importantly, we set out to characterize these effects across REM and NREM sleep, and across age. In a nutshell, we found that caffeine induces a broad boost in brain complexity and a prominent shift in criticality. These effects were substantially more widespread in NREM, compared to REM sleep. Remarkably, the brains of the middle-age individuals were strikingly less perturbed by caffeine than those of the younger subjects.

Importantly, throughout this study, we paid particular attention to probing the robustness of our results through rigorous methodological procedures: First, the caffeine and placebo were contrasted using standard inferential statistics (t-tests) as well as two cross-validated ML classification algorithms (SVM and LDA) to check for consistency. Second, the statistical significance of these analyses (including the ML decoding results) were systematically assessed using non-parametric permutation tests and corrected for multiple comparisons. Third, we removed the aperiodic component of the EEG power spectra before analyzing oscillatory power modulations, and we demonstrated how failing to do so may lead to inaccurate interpretations. Fourth, instead of selecting only one metric to probe complexity and only one for criticality, we presented results obtained with 4 distinct complexity metrics and two measures that tap into critical behavior. Our findings were consistent across the multiple metrics for each feature type. Fifth, we tested the robustness and interpretability of our ML framework by adding single-trial decoding as well as multi-feature classification using random forests, and feature importance assessment and visualization. Last but not least, we accounted for an array of possible confounds including age-dependent baseline EEG differences.

In the following, we discuss the key findings from this study in the light of previous work.

An update on the impact of caffeine on neural rhythms during sleep. Our findings on power spectral density, when analyzed across the full power spectrum, align well with previous observations of caffeine's influence (or lack thereof) on EEG power in distinct frequency ranges. The existing literature highlights a decrease in delta and frontal alpha power and an increase in sigma and beta power, as well as in occipital alpha power [9, 10, 39, 40]. Interestingly, when we used 1/f-corrected power spectra (i.e. excluding the arrhythmic component), the results were more enhanced and reached statistical significance over numerous channels. As shown in Figure 2a the caffeine-induced decrease in delta, theta and frontal alpha power during NREM sleep only became statistically significant after removing the aperiodic component from the power spectrum. In addition, discarding the aperiodic component uncovered a strong statistically significant increase in beta power over multiple parietal, central and frontal sites. Interestingly, beta power increases have been linked to increased excitation or reduced inhibition [41–43]. It is therefore possible that the beta increases we found are related to caffeine's antagonistic effects on adenosine receptors. Moreover, while we saw a non-significant decrease in delta power during REM, matching previous reports [23], this effect appeared to be reversed when we used the 1/f-corrected power spectrum. This discrepancy – as well as the other differences we found using corrected versus uncorrected power spectra – can be attributed to the 1/f activity acting as a confounding factor. Crucially, although commonly referred to as 1/f-noise, the aperiodic activity, has been shown to exhibit inter-subject variability that is linked to cognitive ability and age [36]. In sum, the caffeine-induced changes in oscillatory power reported here confirm but also extend the robustness of previous reports by focusing solely on the rhythmic components.

Caffeine ingestion increases EEG entropy during sleep. Prior quantitative

work investigating the effect of caffeine on sleep EEG has predominantly focused on characterizing its effect on brain rhythms. Our findings extend this body of work by uncovering important effects of caffeine on neural complexity and criticality. In fact, by analyzing the feature importance score from a random forest classifier, we found robust evidence that EEG signal entropy played a more important role in distinguishing caffeine from placebo samples, compared to (corrected) spectral power features.

Although our results provide the first evidence that caffeine enhances EEG entropy during sleep, they are coherent with a previous fMRI study showing that caffeine enhances entropy during resting wakefulness [38]. Although widespread, the most significant entropy increases were found in the lateral prefrontal cortex, default mode network, visual cortex and the motor network [38]. Although these fMRI results obtained in the waking state are not directly comparable to our polysomnographic sleep data, they are consistent with the caffeine-related increases we found across all entropy measures (SpecEn, SampEn, SpecSampEn) during NREM and REM sleep. Our findings therefore extend the literature not only by revealing that caffeine elevates the complexity of the brain's electrophysiological activity, but also by detecting such changes during sleep. Finally, our data also suggest that this caffeine-mediated boost in brain entropy is age-dependent.

In a prior investigation assessing the complexity of biological systems, low approximate entropy (ApEn) was associated with the isolation and segregation of a system [44]. Furthermore, experimental work suggests increased segregation of the brain during NREM sleep [45, 46], which was later shown to be reflected in decreased ApEn [47]. These observations provide an interesting framework in which to interpret the caffeine-induced increase in sample entropy (SampEn) which we found during NREM sleep. Given that SampEn is closely related to ApEn, our results may therefore indicate that caffeine ingestion yields reduced isolation. These results also suggest an increase in information processing during NREM, which brain entropy is commonly seen as a proxy for [48].

Caffeine increases Lempel-Ziv complexity and shifts critical dynamics in the sleeping brain. Because LZc measures the randomness and unpredictability of data sequences, it is not surprising to see that caffeine's influence on LZc and the entropy metrics was similar, displaying widespread increases during NREM sleep and a localized increase in occipital channels during REM sleep. Interestingly, LZc has previously been shown to inversely track the distance to criticality in the brain, i.e. an increase in LZc is tied to activity closer to the critical point [33]. LZc has also been found to be inversely related to the aperiodic slope, another marker for critical dynamics [49, 50]. Additionally, the observed increase in entropy and decrease in DFA scaling exponent suggest elevated degrees of activity during NREM sleep [31, 51]. Considering our results in the light of the brain criticality literature suggests that – especially during NREM sleep – the dynamics of the caffeinated brain are shifted closer to the critical point and towards a more active state. Our data also suggest that this effect is more pronounced in younger subjects (20-27 years old) than in middle-aged subjects (41-58 years old), who showed either weaker effects (during NREM) or no effect at all (during REM). It is tempting to associate these findings with the deterioration in sleep quality and its restorative properties, as documented in the literature on caffeine and sleep.

Caffeine-induced neural excitation is associated with a shift towards a critical point. Caffeine works by reducing the activity of adenosine, a neurotransmitter that causes drowsiness. As a result, consuming caffeine enhances alertness, improves information processing, and boosts cognitive performance. Incidentally, according to the brain criticality framework, these same processes are optimized near the "critical point", a state where the system is poised between ordered

and chaotic behavior, resulting in maximum computational efficiency and flexibility [34]. We thus hypothesized that caffeine consumption shifts the brain dynamics closer to a critical regime. But through what mechanism does caffeine generate such a shift in criticality? One way that caffeine could achieve this would be by changing the excitation-inhibition (E:I) ratio. This proposal follows from the theory that E:I ratio can serve as a control parameter driving the brain closer to – or further away – from the critical point. Therefore, any indication of a caffeine-mediated shift in neural E:I ratio towards more excitation would support the interpretation that caffeine moves the brain closer to a critical point. Our data provide the first evidence in support of this hypothesis by revealing that caffeine induces a flattening of the 1/f slope and a drop in the scaling exponent. Indeed, both modeling and experimental studies have shown that reductions in the 1/f slope (or the related scaling exponent and Hurst exponent) reflect an increase in E:I ratio [49, 50, 52–55].

Interestingly, the link we make here between caffeine's antagonistic effect on adenosine receptors and the observed shift in criticality is conceptually consistent with previous research showing that by blocking GABA receptors and reducing inhibitory synaptic transmission, an artificial upward shift in excitation can occur, leading to a supercritical state with larger-than-expected neuronal avalanches [56, 57].

Age-related difference in the impact of caffeine on the sleeping brain. A prior study found increased sensitivity to caffeine among middle-aged participants concerning sleep latency, duration and efficiency, however, did not see any age-related electrophysiological effects [10]. By explicitly testing for age-related differences in our findings, we found that caffeine had a substantially weaker impact on all observed sleep EEG features in middle-aged subjects, compared to the younger participants. We hypothesize that this effect could be related to baseline changes in brain dynamics associated with aging. Broadly speaking, we contend that the brains of the middle-aged group exhibit long-term neurophysiological alterations including changes that resemble acute effects of caffeine. Such changes include naturally occurring age-related reductions in the 1/f slope and loss of A_1 adenosine receptors with age [58]. By contrasting all extracted EEG features in the placebo nights across young adults and middle-aged participants, we found that indeed the sleep EEG in the older group exhibited a flatter aperiodic slope (and lower scaling exponent). These results are in line with the literature on age-related changes in the 1/f-like neural signal [36, 59]. These preexisting age-related reductions in the complexity and power-law characteristics could thus explain why caffeine only induces weak if any further effects in the middle-aged group. The decrease in the number of A_1 adenosine receptors may be related to a decrease in sleep pressure [60] which can be linked to an increased complexity and flatter aperiodic slope. Therefore the effect of caffeine on these features is likely reduced given the the drop in the absolute number of caffeine bindings to A_1 adenosine receptors. Potentially this can further be linked to the reduction in sleep duration and quality that comes with age [61, 62].

Limitations. A few limitations of the present study are worth considering. First, although the statistical comparisons and ML classification revealed significant effects in a substantially larger number of channels (and higher decoding performances) in NREM compared to REM sleep, we need to keep in mind that the different number of epochs available for each stage in each subject can lead to comparatively noisier estimates of the extracted features in REM. This said, during preprocessing, we averaged features across epochs of each sleep stage in each subject, effectively reducing per-sample noise. Second, despite the strong consistency between the effects we found across three entropy metrics - especially between SpecEn and SpecSampEn - we should be mindful that the SpecSampEn is a slight variation of spectral entropy computation that is not commonly used. Therefore, its increased feature importance compared to SpecEn in the

multi-feature random forest analysis (in NREM sleep) is a possibly interesting observation, but it will require further validation and needs to be interpreted with caution. Third, separately examining the two age groups yielded a limited number of subjects, with 22 in the young adults subgroup and 18 in the middle-aged subgroups. Thus special care has to be taken when interpreting significance scores with these relatively low sample sizes. We generally found reduced levels of significance in these subgroups compared to the results with the full population. While this could in theory reflect that significance testing in the subgroups is impaired by the lower number of subjects, we systematically used nonparametric permutation tests and maximum statistics correction. Finally, while some of our interpretations rely on modeling and empirical evidence that the E:I ratio is correlated with the slope of the power spectrum and the DFA scaling exponent, more work is needed to fully establish the mechanistic links between these EEG measures and E:I balance at multiple meso- and macroscopic scales [41].

Conclusion. In this work we investigated the effects of caffeine, the most widely consumed psychoactive drug, on the activity of our brains as we sleep. In particular, we extend previous research on the neural impact of caffeine on sleep by focusing on two promising yet underexplored features of neural dynamics, which are complexity and criticality. This study provides the first evidence that caffeine ingestion leads to a broad increase in EEG complexity, especially during NREM sleep. Crucially, we also discovered that, compared to placebo, caffeine shifts the brain closer towards a state known as a critical regime, where the brain is thought to be most sensitive to inputs, most adaptable, and able to process information most efficiently. Importantly, we propose a mechanistic explanation of the observed shifts by connecting caffeine's effect on adenosine transmission to changes in E:I balance, which are manifest across multiple EEG measures. Finally, our data suggest that the shifts in brain dynamics due to caffeine are more prominent in younger adults than in middle-aged individuals, a difference that could be explained by age-related modifications to the baseline EEG. While our primary focus was on the influence of sleep on brain electrophysiology, we anticipate that future studies will uncover similar outcomes during wakefulness. Considering the widespread consumption of caffeine, gaining a comprehensive understanding of its effects on the brain both during wakefulness and sleep could have far-reaching implications for society and public health.

3 Methods and Materials

3.1 Data acquisition

Sleep EEG (256 Hz) data was recorded from 40 subjects (19 females, 21 males) at the Center for Advanced Research in Sleep Medicine (CARSM), with a channel layout according to the 10-20 international system. Subjects were in good health, aged from 20 to 58 (mean 35.3 $\pm$ 14.3 years). The participants reported moderate caffeine consumption, equivalent to one to three cups of coffee per day. All participants were non-smokers and free of drugs or medicine which could influence the sleep-wake cycle. Subjects also reported no sleep complaints, night work or transmeridian travel in the 3 months before the recording. Participants with a history of psychiatric or neurological illnesses or a body mass index (BMI) above 29 were excluded from the study. Blood sample analysis (complete blood count, serum chemistry including hepatic and renal functions, levels of prolactine, levels of testosterone in men, and levels of oestrogen, follicle-stimulating hormone (FSH) and luteinizing hormone in women) and urinalysis results were examined by a certified physician to exclude significant medical irregularities. To eliminate further irregularities in participants, a polysomnographic

screening night at the laboratory was conducted where a nasal/oral thermistor, electromyographic (EMG) leg electrodes, EEG and electro-oculogram (EOG) were recorded. The presence of sleep disturbances such as sleep apnoeas and hypopnoeas (index per hour > 10), periodic leg movements (index per hour > 10), prolonged sleep latency (> 30 min) or low sleep efficiency (< 85%) resulted in the exclusion of the participant. Peri-menopausal women and women using hormonal contraceptives or receiving hormonal replacement therapy were excluded, premenopausal women reported having regular menstrual cycles (25-32 days) during the year preceding the study, no vasomotor complaints (i.e. hot flashes, night sweats) and showed low FSH levels (< 20 iU L-1) and all postmenopausal women reported an absence of menses during the past year and their FSH levels were > 20 iU L-1. All participants signed an informed consent form which provided detailed information about the nature, purpose and potential risks of the study. The research project was approved by the hospital's ethics committee and subjects received financial compensation for their participation.

3.2 Experimental protocol

After the adaptation and screening night, each participant spent two non-consecutive experimental nights at the sleep laboratory. The time between each night in the laboratory was 6-9 days. On days of recordings, the consumption of foods and beverages containing caffeine was ceased at noon. Instructions were given to maintain a regular sleep-wake pattern within 30 minutes of the habitual sleep-wake cycle, as well as a habitual caffeine consumption one week before the first night to prevent potential withdrawal effects. They were asked to complete the French version of the Pittsburgh Sleep Diary [63] on a daily basis during this time and to abstain from alcohol on experiment days. The participants arrived at the laboratory 6-8 hours before their habitual sleep time and left 1-1.5 hours after habitual wake up time. Bedtime and wake time in the laboratory were determined by averaging each participant's sleep-wake cycle from the sleep diary. The total dose of caffeine administered was 200mg (100mg per capsule) which is considered to be moderate (equivalent to 1-2 cups of coffee) and induces significant changes in the sleep of young subjects [23]. Two-piece telescopic hard capsules were used, allowing the ingestion of caffeine without oral contamination.

3.3 Sleep variables

An extensive analysis of sleep variables and their alteration with caffeine was conducted in previous work in our group with the same data [9, 10]. Briefly, this data exhibits a significant caffeine-induced increase in sleep latency and a decrease in total sleep time, sleep efficiency and time spent in stage 2 sleep [9, 10]. We therefore controlled for the effect of changes in sleep duration in our analysis and found no significant change suggesting our results are robust in this regard. In other words, the difference in the feature estimations was not due to the variable amount of data across conditions but rather to the underlying changes in neural dynamics.

3.4 EEG preprocessing

The data was divided into 20s windows and visually scored by an expert into five sleep stages: S1, S2, S3, S4 and REM according to the Rechtschaffen and Kales manual [64] modified according to scoring based on 20s epochs [10]. From a signal processing standpoint we have no reasons to believe that our results would differ were we to use 30s epochs. After scoring, artifact containing epochs (4s) were eliminated from non-REM data by visual inspection of channels C3 and C4. Due to the relatively low number of epochs in S1, S3 and S4, compared to S2, the epochs from S1, S2, S3 and S4 stages were

combined into a single non-REM (NREM) stage. This is in line with previous research conducted with the same data [9]. After artifact removal, the NREM stage contained 57,594 epochs and REM consisted of 19,341 epochs, which roughly equates to a 3:1 ratio. Before feature extraction, we applied a band-pass filter from 0.5 to 32 Hz to the raw epochs, removing further artifacts from the signal. Features were extracted independently for each channel, resulting in a single value per feature, channel and epoch. To reduce noise and increase comparability between sleep stages, the epoch-wise features were averaged within each subject. For each feature, given that we had 40 individuals, this led to a total of 80 samples: 40 for caffeine and 40 for the placebo condition. Furthermore, the averaged features were z-transformed using the mean and standard deviation from the non-averaged samples across all channels and samples of one feature. The z-transform was applied separately to each subject. This allowed us to ensure strict independence between training and test data sets in the ML analysis.

3.5 Feature extraction

We extracted a selection of relevant hand-crafted EEG features from the neural signals to examine the effect of caffeine on the brain in a data-driven manner. The aim was to differentiate between data from the caffeine or placebo condition using supervised machine learning, trained on these features. The following sections list the specific spectral, complexity and criticality features analysed in this work.

Power spectral density Power spectral density (PSD) was computed on individual 20s EEG epochs using Welch's method [65] with a window length of 4s and overlap of 2s. To isolate changes in periodic (pure oscillatory) activity, we remove the aperiodic 1/f-like component from the power spectrum. The FOOOF algorithm [66] was used to fit a power-law distribution to the power spectrum and extract periodic oscillations in the frequency range from 3 to 32 Hz, limiting the number of periodic peaks to 5. The lower and upper cutoff were determined visually from the full power spectrum. The 1/f-corrected spectrum was then divided into five distinct frequency bands by averaging the power across frequency bins within the given intervals. The final frequency bands were delta (0.5 – 4.0 Hz), theta (4.0 – 8.0 Hz), alpha (8.0 – 12.0 Hz), sigma (12.0 – 16.0 Hz) and beta (16.0 – 32.0 Hz). We extrapolated the power-law distribution to also cover the delta band.

Sample entropy To estimate complexity in the time domain, sample entropy (SampEn) [67] was computed using the preprocessed EEG signal. The definition of SampEn is similar to that of approximate entropy (ApEn), a popular metric for complexity estimation of biomedical times series. ApEn, however, is known to have limitations regarding relative consistency across parameter choices, bias due to counting self-matches and dependence on sample length. SampEn was developed to combat these limitations [68, 69]. It is defined as the negative logarithm of the conditional probability that two matching windows/embeddings of length m still match when increasing their length to $m + 1$. A match between two windows is defined as having a distance smaller than r times the standard deviation of the signal. Distance was calculated using the maximum norm and self-matches were excluded. Mathematically it can be described as $\mathrm{SampEn}(m, r) = -\log(\frac{A}{B})$ with A being the number of matches with window size $m + 1$ and B the number of matches of length m. Therefore SampEn measures how unlikely it is to find a window that matches another part of the signal where the continuation of the window is still a match. We set SampEn's parameters to $m = 2$ and $r = 0.2$ as suggested in the literature [68].

Spectral entropy Complexity of the power spectrum or spectral entropy [70] (SpecEn) can be computed by applying entropy measures to the power spectrum of a signal. Traditionally, spectral entropy uses Shannon entropy [71], which disregards local patterns and only takes the signal's distribution of values into account. When applied

to the power spectrum, Shannon entropy captures the amount of variance in the frequencies, which make up the EEG signal [72, 73]. Additionally, we computed a variant of SpecEn, which we refer to as spectral sample entropy (SpecSampEn): Following the idea of SpecEn, we applied SampEn($m = 2, r = 0.2$) to the full power spectrum. By looking at windows of multiple neighboring frequency bins, SampEn is able to incorporate local patterns in the power spectrum into its complexity estimate. This goes beyond the permutation invariant Shannon entropy and treats the power spectrum more as a coherent signal with meaningful relations between neighboring frequency bins, potentially capturing further aspects of complexity. The power spectrum we used to compute SpecEn and SpecSampEn was estimated as described in Section 3.5 (Power spectral density), however, it was neither corrected for the 1/f-like component nor split into frequency bands.

Lempel-Ziv Complexity Lempel-Ziv complexity (LZc) [74] is a complexity measure designed for binary sequences and text. It counts the number of unique sub-strings in a sequence, thereby measuring how repetitive a signal is. That is, a less complex signal, according to LZc, consists of repetitions of a few different sub-strings while more complex signals are made up of non-repeating segments. As the initially proposed version of LZc is strongly influenced by signal length, we used a normalized variant, which scales LZc by $\frac{\log_b(n)}{n}$ where n is signal length and b the number of unique characters in the signal [75]. In the context of neuroscience, LZc has been useful to track disorders of consciousness (DOC) and was shown to reach a peak at the edge of chaos critical point [33]. We applied a median split to the preprocessed EEG epochs to transform the signal into a binary sequence, allowing it to be processed by the Lempel-Ziv algorithm.

Detrended Fluctuation Analysis To measure the signal's long-term statistical dependence, we performed Detrended Fluctuation Analysis (DFA) [76, 77] on the raw EEG signal. DFA, an estimate of self-affinity, divides the signal into windows and computes standard deviations as a function of window sizes. A power-law with exponent/slope α is then fitted to the resulting graph. The estimated α parameter is a generalization of the Hurst exponent, capable of dealing with non-stationary time series. Similar to the Hurst exponent, a slope less than 0.5 corresponds to an anti-correlated process, a slope of 0.5 indicates a white-noise process and a slope larger 0.5 suggests a correlated signal. In contrast to the Hurst exponent however, DFA can return slopes larger than one in the case of unbounded, non-stationary signals.

Aperiodic slope We used the FOOOF algorithm [66] to estimate the slope of the aperiodic component of the EEG signal, which corresponds to the exponent in the 1/f-like distribution of the signals's power spectrum. FOOOF was fit on the power spectrum between 3 and 32 Hz, limiting the number of periodic peaks to 5. The power spectrum was computed in the same way as described in Section 3.5 (Power spectral density). The aperiodic slope has previously been shown to reflect changes in E/I balance [49, 52, 53], which could link the effect of caffeine, an antagonist of the inhibitory neurotransmitter adenosine, to changes in E/I balance in the brain. For ease of interpretation we report the negative slope (i.e. positive values) in this work. A reduction in slope therefore corresponds to a flatter distribution while increased slope refers to a steeper decline in power.

3.6 Statistical analysis

To evaluate the difference between the caffeine and placebo conditions, two-sided paired permutation-based pseudo T-tests [78] were performed using exhaustive permutations ($n = 10^4$) and corrected for multiple comparisons using the maximum statistics method [79]. The T-tests were carried out on all extracted features (PSD bands,

SpecEn, SampEn, SpecSampEn, PermEn). Statistical significance between the caffeine and placebo condition was evaluated at $p<0.05$ and $p<0.01$. We additionally performed this analysis for young (20-27 years old) and middle-aged (41-58 years old) subgroups separately. To further analyze age as a confounding variable we applied Welch's T-tests to the raw features between the age groups (corrected with the Bonferroni method). This additional step allowed us to put effects seen in only a single age group into context.

3.7 Supervised machine learning analysis

To improve robustness and interpretability of the statistical results, we additionally trained supervised machine learning classifiers to distinguish between samples from the caffeine or placebo condition.

Single-feature classification To assess the impact of caffeine on specific features in single channels, different machine learning classifiers were trained on single-feature, binary classification between the placebo and caffeine conditions. Two different classification algorithms were used, namely support vector machine (SVM) [80–83] with an RBF kernel and linear discriminant analysis (LDA) [84, 85]. Models were trained and evaluated separately for the different sleep stages. Permutation tests [86, 87] (n=1000) were applied to the trained model to determine the statistical significance of the classifiers' accuracy scores. The results were corrected for multiple comparisons using the maximum statistics method. During each permutation, the model's average score across folds in a grouped 10-fold cross-validation was used. Due to the low sample size we decided to additionally fit an LDA classifier on single-epoch data (sample sizes NREM: 26,776 caffeine and 30,818 placebo, REM: 9448 caffeine and 9893 placebo), instead of subject-wise averages. Here, we chose to evaluate classifier performance using the balanced accuracy metric (BAcc) due to class imbalance. While accuracy is biased towards the majority class, BAcc was shown not to overestimate classifier performance even in cases of extreme class imbalance [88].

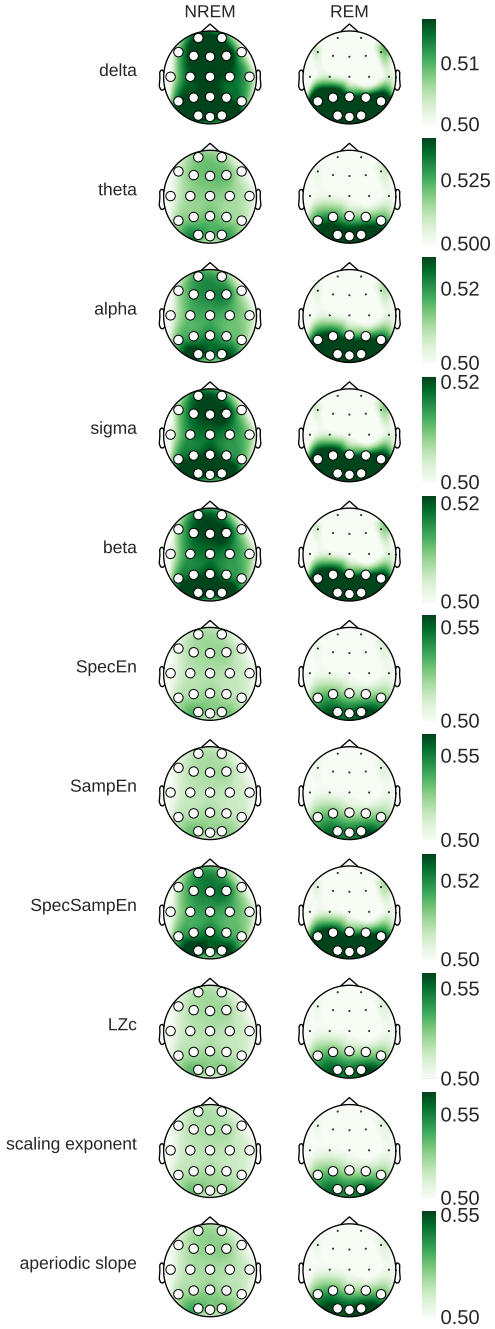
Multi-feature classification In addition to the exploration on a single-channel and single-feature level, we chose to implement a broader multi-feature classifier that takes all features from all channels to build a single model. Spectral power was computed from the power spectrum after removal of the 1/f-like component since our previous analysis showed a better separation between the two conditions when using corrected spectral power. We were able to access the contributions of individual features and channels by examining feature importance. To achieve this we used a random forest, which estimates feature importance by averaging the relative rank of each feature across the decision trees that make up the forest. To measure variance in feature importance, the training process was repeated 1000 times for each sleep stage, applying grid search with grouped 7-fold cross-validation inside a nested cross-validation, leaving out samples from five different subjects in each iteration. See Supplementary Table S1 Table for the selection of hyperparameters optimized during grid search. The overall model score was determined using samples from the left-out testing subjects.

3.8 Open source software and code availability

All analyses were done with Python 3.8, utilizing the software modules scikit-learn [89] for machine learning, MNE-Python [90] for visualization and, together with statsmodels [91], statistical analysis. The SciPy stack [92] was used for efficient data structures, mathematical operations and plotting functionalities. A large share of the code was developed inside the Jupyter Notebook and IPython framework [93]. The machine learning analysis was enabled in part by support provided by the Digital

Research Alliance of Canada (`alliancecan.ca`). All code used for analysis and visualization can be found at `github.com/PhilippThoelke/caffeine-sleep`.

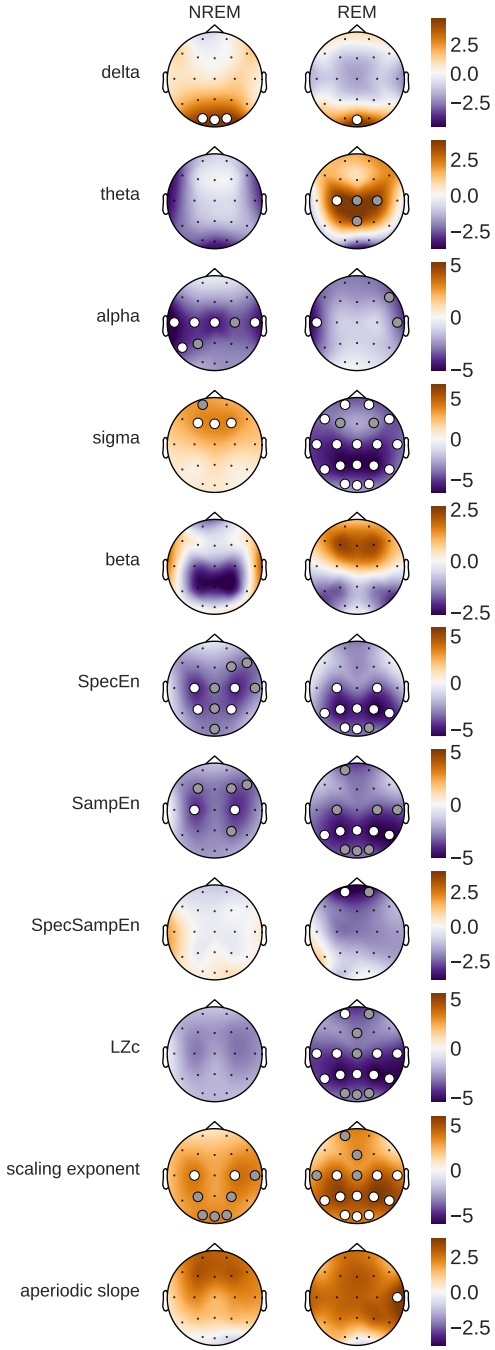
Supporting information



S1 Fig.

Single-epoch Classification. Topographical maps showing LDA classification scores of single-feature, single-channel classification on 20s epochs (instead of averaged features). Due to class imbalance, classification performance is measured using the balanced accuracy (BAcc) metric, resulting in a chance-level score of 0.5. Statistical

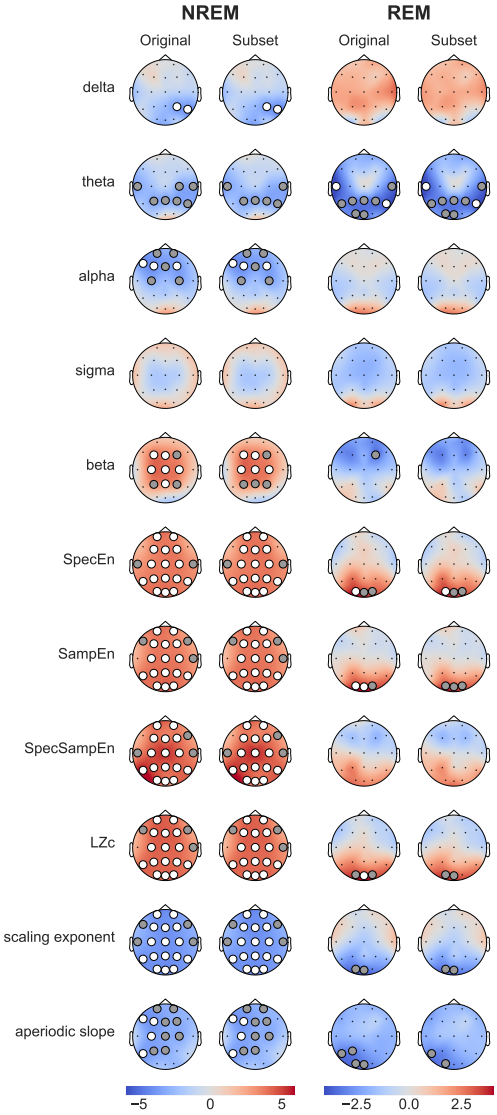
significance was evaluated using permutation tests and corrected for multiple comparisons using the maximum statistics method. Grey dots represent significance at $p<0.05$ and white dots at $p<0.01$.



S2 Fig.

Age-related Effects. T-values from Welch's T-test between the young and middle-aged group. That is, orange corresponds to larger values in the young group while purple indicates larger values in the middle-aged group. We calculated p-values

from the test statistics and corrected for multiple comparisons using the Bonferroni method. Grey dots represent significance at p<0.05 and white dots at p<0.01.



S3 Fig.

Equating Sleep Duration. T-values comparing statistical findings with and without equating the number of epochs between caffeine and placebo conditions. The "Subset" columns show the results of the control analysis for mitigating sleep duration as a potential confound. Data in the "Original" columns is the same as presented in Figures 1 and 3. Grey dots represent significance at p<0.05 and white dots at p<0.01, all corrected for multiple comparisons using the maximum statistics.

S1 Table.

Parameter	Values
#estimators	{75, 100, 150, 200}
max. depth	{inf, 50, 75}
criterion	{gini, entropy}
max. features	{sqrt, log2}
min. samples per leaf	{1, 5, 10}
bootstrap	{true, false}

Grid search parameter selection. Selection of hyperparameters and corresponding value options for the grid-search optimization of the random forest (RF). The values correspond to the hyperparameter options as implemented in scikit-learn.

S2 Table.

NREM			REM		
feature	elec.	import.	feature	elec.	import.
LZc	P4	0.018	theta	T3	0.014
SpecSampEn	C3	0.018	scaling exponent	O1	0.013
SpecSampEn	Cz	0.018	SpecEn	O1	0.012
LZc	F3	0.018	SpecEn	T5	0.011
LZc	P3	0.016	theta	T6	0.011
LZc	Cz	0.016	theta	T4	0.010
LZc	F4	0.015	sigma	C3	0.010
LZc	C3	0.015	aperiodic slope	P3	0.010
LZc	Pz	0.014	scaling exponent	O2	0.010
LZc	Fz	0.014	theta	T5	0.010

Feature importance scores. Ten most important combinations of feature and electrode determined by the random forests' feature importance in descending order. The models were trained and evaluated separately for each sleep stage (NREM and REM).

References

1. McLellan TM, Caldwell JA, Lieberman HR. A review of caffeine's effects on cognitive, physical and occupational performance. Neuroscience and Biobehavioral Reviews. 2016;71:294–312. doi:10.1016/j.neubiorev.2016.09.001.

2. Drake C, Roehrs T, Shambroom J, Roth T. Caffeine effects on sleep taken 0, 3, or 6 hours before going to bed. Journal of clinical sleep medicine: JCSM: official publication of the American Academy of Sleep Medicine. 2013;9(11):1195–1200. doi:10.5664/jcsm.3170.

3. Frigerio R, Fujishiro H, Maraganore DM, Klos KJ, DelleDonne A, Heckman MG, et al. Comparison of Risk Factor Profiles in Incidental Lewy Body Disease and Parkinson Disease. Archives of Neurology. 2009;66(9):1114–1119. doi:10.1001/archneurol.2009.170.

4. Ren X, Chen JF. Caffeine and Parkinson's Disease: Multiple Benefits and Emerging Mechanisms. Frontiers in Neuroscience. 2020;14.

5. Palacios N, Gao X, McCullough ML, Schwarzschild MA, Shah R, Gapstur S, et al. Caffeine and risk of Parkinson's disease in a large cohort of men and women. Movement Disorders. 2012;27(10):1276–1282. doi:10.1002/mds.25076.

6. Ross GW, Abbott RD, Petrovitch H, Morens DM, Grandinetti A, Tung KH, et al. Association of Coffee and Caffeine Intake With the Risk of Parkinson Disease. JAMA. 2000;283(20):2674–2679. doi:10.1001/jama.283.20.2674.

7. Garriguet D. Beverage consumption of Canadian adults. Health Reports. 2008;19(4):23–29.

8. Mandel HG. Update on caffeine consumption, disposition and action. Food and Chemical Toxicology. 2002;40(9):1231–1234. doi:10.1016/S0278-6915(02)00093-5.

9. Drapeau C, Hamel-Hébert I, Robillard R, Selmaoui B, Filipini D, Carrier J. Challenging sleep in aging: the effects of 200 mg of caffeine during the evening in young and middle-aged moderate caffeine consumers. Journal of Sleep Research. 2006;15(2):133–141. doi:10.1111/j.1365-2869.2006.00518.x.

10. Robillard R, Bouchard M, Cartier A, Nicolau L, Carrier J. Sleep is more sensitive to high doses of caffeine in the middle years of life. Journal of Psychopharmacology. 2015;29(6):688–697. doi:10.1177/0269881115575535.

11. Weibel J, Lin YS, Landolt HP, Kistler J, Rehm S, Rentsch KM, et al. The impact of daily caffeine intake on nighttime sleep in young adult men. Scientific Reports. 2021;11(1):4668. doi:10.1038/s41598-021-84088-x.

12. Clark I, Landolt HP. Coffee, caffeine, and sleep: A systematic review of epidemiological studies and randomized controlled trials. Sleep Medicine Reviews. 2017;31:70–78. doi:10.1016/j.smrv.2016.01.006.

13. Weibel J, Lin YS, Landolt HP, Berthomier C, Brandewinder M, Kistler J, et al. Regular Caffeine Intake Delays REM Sleep Promotion and Attenuates Sleep Quality in Healthy Men. Journal of Biological Rhythms. 2021;36(4):384–394. doi:10.1177/07487304211013995.

14. Aframian K, Yousef Yengej D, Nwaobi S, Raman S, Faas GC, Charles A. Effects of chronic caffeine on patterns of brain blood flow and behavior throughout the sleep–wake cycle in freely behaving mice. PNAS Nexus. 2023;2(9):pgad303. doi:10.1093/pnasnexus/pgad303.

15. Patel SR, Hu FB. Short sleep duration and weight gain: a systematic review. Obesity (Silver Spring, Md). 2008;16(3):643–653. doi:10.1038/oby.2007.118.

16. Gangwisch JE, Heymsfield SB, Boden-Albala B, Buijs RM, Kreier F, Pickering TG, et al. Short sleep duration as a risk factor for hypertension: analyses of the first National Health and Nutrition Examination Survey. Hypertension (Dallas, Tex: 1979). 2006;47(5):833–839. doi:10.1161/01.HYP.0000217362.34748.e0.

17. Cappuccio FP, D'Elia L, Strazzullo P, Miller MA. Quantity and quality of sleep and incidence of type 2 diabetes: a systematic review and meta-analysis. Diabetes Care. 2010;33(2):414–420. doi:10.2337/dc09-1124.

18. Gottlieb DJ, Punjabi NM, Newman AB, Resnick HE, Redline S, Baldwin CM, et al. Association of sleep time with diabetes mellitus and impaired glucose tolerance. Archives of Internal Medicine. 2005;165(8):863–867. doi:10.1001/archinte.165.8.863.

19. Short MA, Gradisar M, Lack LC, Wright HR. The impact of sleep on adolescent depressed mood, alertness and academic performance. Journal of Adolescence. 2013;36(6):1025–1033. doi:10.1016/j.adolescence.2013.08.007.

20. Ribeiro JA, Sebastião AM. Caffeine and adenosine. Journal of Alzheimer's disease: JAD. 2010;20 Suppl 1:S3–15. doi:10.3233/JAD-2010-1379.

21. Carman AJ, Mills JH, Krenz A, Kim DG, Bynoe MS. Adenosine Receptor Signaling Modulates Permeability of the Blood–Brain Barrier. The Journal of Neuroscience. 2011;31(37):13272–13280. doi:10.1523/JNEUROSCI.3337-11.2011.

22. Pan W, Kastin AJ. The Blood-Brain Barrier: Regulatory Roles in Wakefulness and Sleep. The Neuroscientist. 2017;23(2):124–136. doi:10.1177/1073858416639005.

23. Landolt HP, Dijk DJ, Gaus SE, Borbély AA. Caffeine reduces low-frequency delta activity in the human sleep EEG. Neuropsychopharmacology: Official Publication of the American College of Neuropsychopharmacology. 1995;12(3):229–238. doi:10.1016/0893-133X(94)00079-F.

24. Landolt HP, Werth E, Borbély AA, Dijk DJ. Caffeine intake (200 mg) in the morning affects human sleep and EEG power spectra at night. Brain Research. 1995;675(1-2):67–74. doi:10.1016/0006-8993(95)00040-w.

25. Landolt HP, Rétey JV, Tönz K, Gottselig JM, Khatami R, Buckelmüller I, et al. Caffeine attenuates waking and sleep electroencephalographic markers of sleep homeostasis in humans. Neuropsychopharmacology: Official Publication of the American College of Neuropsychopharmacology. 2004;29(10):1933–1939. doi:10.1038/sj.npp.1300526.

26. Reichert CF, Deboer T, Landolt HP. Adenosine, caffeine, and sleep-wake regulation: state of the science and perspectives. Journal of Sleep Research. 2022;31(4):e13597. doi:10.1111/jsr.13597.

27. Porsolt RD. The usefulness of non-human primates in central nervous system safety pharmacology. Journal of Pharmacological and Toxicological Methods. 2013;68(1):23–29. doi:10.1016/j.vascn.2013.03.004.

28. Rachalski A, Authier S, Bassett L, Pouliot M, Tremblay G, Mongrain V. Sleep electroencephalographic characteristics of the Cynomolgus monkey measured by telemetry. Journal of Sleep Research. 2014;23(6):619–627. doi:10.1111/jsr.12189.

29. Authier S, Bassett L, Pouliot M, Rachalski A, Troncy E, Paquette D, et al. Effects of amphetamine, diazepam and caffeine on polysomnography (EEG, EMG, EOG)-derived variables measured using telemetry in Cynomolgus monkeys. Journal of Pharmacological and Toxicological Methods. 2014;70(1):86–93. doi:10.1016/j.vascn.2014.05.003.

30. Bruhn J, Röpcke H, Rehberg B, Bouillon T, Hoeft A. Electroencephalogram approximate entropy correctly classifies the occurrence of burst suppression pattern as increasing anesthetic drug effect. Anesthesiology. 2000;93(4):981–985. doi:10.1097/00000542-200010000-00018.

31. Bruce EN, Bruce MC, Vennelaganti S. Sample Entropy Tracks Changes in EEG Power Spectrum With Sleep State and Aging. Journal of clinical neurophysiology : official publication of the American Electroencephalographic Society. 2009;26(4):257–266. doi:10.1097/WNP.0b013e3181b2f1e3.

32. Manor B, Lipsitz LA. Physiologic Complexity and Aging: Implications for Physical Function and Rehabilitation. Progress in neuro-psychopharmacology & biological psychiatry. 2013;45:287–293. doi:10.1016/j.pnpbp.2012.08.020.

33. Toker D, Pappas I, Lendner JD, Frohlich J, Mateos DM, Muthukumaraswamy S, et al. Consciousness is supported by near-critical slow cortical electrodynamics. Proceedings of the National Academy of Sciences. 2022;119(7):e2024455119. doi:10.1073/pnas.2024455119.

34. O'Byrne J, Jerbi K. How critical is brain criticality? Trends in Neurosciences. 2022;doi:10.1016/j.tins.2022.08.007.

35. Landolt HP, Borbély AA. Age-dependent changes in sleep EEG topography. Clinical Neurophysiology. 2001;112(2):369–377. doi:10.1016/S1388-2457(00)00542-3.

36. Voytek B, Kramer MA, Case J, Lepage KQ, Tempesta ZR, Knight RT, et al. Age-Related Changes in 1/f Neural Electrophysiological Noise. Journal of Neuroscience. 2015;35(38):13257–13265. doi:10.1523/JNEUROSCI.2332-14.2015.

37. Thuwal K, Banerjee A, Roy D. Aperiodic and Periodic Components of Ongoing Oscillatory Brain Dynamics Link Distinct Functional Aspects of Cognition across Adult Lifespan. eNeuro. 2021;8(5):ENEURO.0224–21.2021. doi:10.1523/ENEURO.0224-21.2021.

38. Chang D, Song D, Zhang J, Shang Y, Ge Q, Wang Z. Caffeine Caused a Widespread Increase of Resting Brain Entropy. Scientific Reports. 2018;8(1):1–7. doi:10.1038/s41598-018-21008-6.

39. Carrier J, Paquet J, Fernandez-Bolanos M, Girouard L, Roy J, Selmaoui B, et al. Effects of caffeine on daytime recovery sleep: A double challenge to the sleep–wake cycle in aging. Sleep Medicine. 2009;10(9):1016–1024. doi:10.1016/j.sleep.2009.01.001.

40. Landolt HP, Rétey JV, Adam M. Reduced Neurobehavioral Impairment from Sleep Deprivation in Older Adults: Contribution of Adenosinergic Mechanisms. Frontiers in Neurology. 2012;3. doi:10.3389/fneur.2012.00062.

41. Ahmad J, Ellis C, Leech R, Voytek B, Garces P, Jones E, et al. From mechanisms to markers: novel noninvasive EEG proxy markers of the neural excitation and inhibition system in humans. Translational Psychiatry. 2022;12(1):1–12. doi:10.1038/s41398-022-02218-z.

42. Porjesz B, Almasy L, Edenberg HJ, Wang K, Chorlian DB, Foroud T, et al. Linkage disequilibrium between the beta frequency of the human EEG and a GABAA receptor gene locus. Proceedings of the National Academy of Sciences. 2002;99(6):3729–3733. doi:10.1073/pnas.052716399.

43. Jensen O, Goel P, Kopell N, Pohja M, Hari R, Ermentrout B. On the human sensorimotor-cortex beta rhythm: sources and modeling. NeuroImage. 2005;26(2):347–355. doi:10.1016/j.neuroimage.2005.02.008.

44. Pincus SM. Greater signal regularity may indicate increased system isolation. Mathematical Biosciences. 1994;122(2):161–181. doi:10.1016/0025-5564(94)90056-6.

45. Massimini M, Ferrarelli F, Huber R, Esser SK, Singh H, Tononi G. Breakdown of Cortical Effective Connectivity During Sleep. Science. 2005;309(5744):2228–2232. doi:10.1126/science.1117256.

46. Massimini M, Ferrarelli F, Esser SK, Riedner BA, Huber R, Murphy M, et al. Triggering sleep slow waves by transcranial magnetic stimulation. Proceedings of the National Academy of Sciences. 2007;104(20):8496–8501. doi:10.1073/pnas.0702495104.

47. Lee G, Fattinger S, Mouthon AL, Noirhomme Q, Huber R. Electroencephalogram approximate entropy influenced by both age and sleep. Frontiers in Neuroinformatics. 2013;7.

48. Keshmiri S. Entropy and the Brain: An Overview. Entropy. 2020;22(9):917. doi:10.3390/e22090917.

49. Medel V, Irani M, Ossandón T, Boncompte G. Complexity and 1/f slope jointly reflect cortical states across different E/I balances; 2020. Available from: `https://www.biorxiv.org/content/10.1101/2020.09.15.298497v2`.

50. Lombardi F, Herrmann HJ, de Arcangelis L. Balance of excitation and inhibition determines 1/f power spectrum in neuronal networks. Chaos: An Interdisciplinary Journal of Nonlinear Science. 2017;27(4):047402. doi:10.1063/1.4979043.

51. Ma Y, Shi W, Peng CK, Yang AC. Nonlinear dynamical analysis of sleep electroencephalography using fractal and entropy approaches. Sleep Medicine Reviews. 2018;37:85–93. doi:10.1016/j.smrv.2017.01.003.

52. Gao R, Peterson EJ, Voytek B. Inferring synaptic excitation/inhibition balance from field potentials. NeuroImage. 2017;158:70–78. doi:10.1016/j.neuroimage.2017.06.078.

53. Martínez-Cañada P, Panzeri S. Spectral Properties of Local Field Potentials and Electroencephalograms as Indices for Changes in Neural Circuit Parameters. In: Mahmud M, Kaiser MS, Vassanelli S, Dai Q, Zhong N, editors. Brain Informatics. Lecture Notes in Computer Science. Cham: Springer International Publishing; 2021. p. 115–123.

54. Trakoshis S, Martínez-Cañada P, Rocchi F, Canella C, You W, Chakrabarti B, et al. Intrinsic excitation-inhibition imbalance affects medial prefrontal cortex differently in autistic men versus women. eLife. 2020;9:e55684. doi:10.7554/eLife.55684.

55. Stadnitski T. Measuring Fractality. Frontiers in Physiology. 2012;3.

56. Beggs JM, Plenz D. Neuronal Avalanches in Neocortical Circuits. Journal of Neuroscience. 2003;23(35):11167–11177. doi:10.1523/JNEUROSCI.23-35-11167.2003.

57. Shew WL, Yang H, Petermann T, Roy R, Plenz D. Neuronal Avalanches Imply Maximum Dynamic Range in Cortical Networks at Criticality. Journal of Neuroscience. 2009;29(49):15595–15600. doi:10.1523/JNEUROSCI.3864-09.2009.

58. Meyer PT, Elmenhorst D, Boy C, Winz O, Matusch A, Zilles K, et al. Effect of aging on cerebral A1 adenosine receptors: A [18F]CPFPX PET study in humans. Neurobiology of Aging. 2007;28(12):1914–1924. doi:10.1016/j.neurobiolaging.2006.08.005.

59. Dave S, Brothers TA, Swaab TY. 1/f neural noise and electrophysiological indices of contextual prediction in aging. Brain Research. 2018;1691:34–43. doi:10.1016/j.brainres.2018.04.007.

60. Mander BA, Winer JR, Walker MP. Sleep and Human Aging. Neuron. 2017;94(1):19–36. doi:10.1016/j.neuron.2017.02.004.

61. Espiritu JRD. Aging-Related Sleep Changes. Clinics in Geriatric Medicine. 2008;24(1):1–14. doi:10.1016/j.cger.2007.08.007.

62. Wolkove N, Elkholy O, Baltzan M, Palayew M. Sleep and aging: 1. Sleep disorders commonly found in older people. CMAJ. 2007;176(9):1299–1304. doi:10.1503/cmaj.060792.

63. Monk TH, Reynolds CF, Kupfer DJ, Buysse DJ, Coble PA, Hayes AJ, et al. The Pittsburgh Sleep Diary. Journal of Sleep Research. 1994;3:111–120.

64. Allan Rechtschaffen AK. A Manual of Standardized Terminology, Techniques and Scoring System for Sleep Stages of Human Subjects. U. S. National Institute of Neurological Diseases and Blindness, Neurological Information Network; 1968. Available from: `http://archive.org/details/RKManual`.

65. Welch P. The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms. IEEE Transactions on Audio and Electroacoustics. 1967;15(2):70–73. doi:10.1109/TAU.1967.1161901.

66. Donoghue T, Haller M, Peterson EJ, Varma P, Sebastian P, Gao R, et al. Parameterizing neural power spectra into periodic and aperiodic components. Nature Neuroscience. 2020;23(12):1655–1665. doi:10.1038/s41593-020-00744-x.

67. Richman JS, Moorman JR. Physiological time-series analysis using approximate entropy and sample entropy. American Journal of Physiology Heart and Circulatory Physiology. 2000;278(6):H2039–2049. doi:10.1152/ajpheart.2000.278.6.H2039.

68. Aboy M, Cuesta-Frau D, Austin D, Mico-Tormos P. Characterization of Sample Entropy in the Context of Biomedical Signal Analysis. In: 2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society; 2007. p. 5942–5945.

69. Yentes JM, Hunt N, Schmid KK, Kaipust JP, McGrath D, Stergiou N. The appropriate use of approximate entropy and sample entropy with short data sets. Annals of biomedical engineering. 2013;41(2):349–365. doi:10.1007/s10439-012-0668-3.

70. Pardey J, Roberts S, Tarassenko L. A review of parametric modelling techniques for EEG analysis. Medical Engineering & Physics. 1996;18(1):2–11. doi:10.1016/1350-4533(95)00024-0.

71. Shannon CE. A Mathematical Theory of Communication. Bell System Technical Journal. 1948;27(3):379–423. doi:10.1002/j.1538-7305.1948.tb01338.x.

72. Inouye T, Shinosaki K, Sakamoto H, Toi S, Ukai S, Iyama A, et al. Quantification of EEG irregularity by use of the entropy of the power spectrum. Electroencephalography and Clinical Neurophysiology. 1991;79(3):204–210. doi:10.1016/0013-4694(91)90138-T.

73. Urigüen JA, García-Zapirain B, Artieda J, Iriarte J, Valencia M. Comparison of background EEG activity of different groups of patients with idiopathic epilepsy using Shannon spectral entropy and cluster-based permutation statistical testing. PLOS ONE. 2017;12(9):e0184044. doi:10.1371/journal.pone.0184044.

74. Lempel A, Ziv J. On the Complexity of Finite Sequences. IEEE Transactions on Information Theory. 1976;22(1):75–81. doi:10.1109/TIT.1976.1055501.

75. Zhang Y, Hao J, Zhou C, Chang K. Normalized Lempel-Ziv complexity and its application in bio-sequence analysis. Journal of Mathematical Chemistry. 2009;46(4):1203–1212. doi:10.1007/s10910-008-9512-2.

76. Peng CK, Buldyrev SV, Havlin S, Simons M, Stanley HE, Goldberger AL. Mosaic organization of DNA nucleotides. Physical Review E. 1994;49(2):1685–1689. doi:10.1103/PhysRevE.49.1685.

77. Hardstone R, Poil SS, Schiavone G, Jansen R, Nikulin V, Mansvelder H, et al. Detrended Fluctuation Analysis: A Scale-Free View on Neuronal Oscillations. Frontiers in Physiology. 2012;3.

78. Nichols TE, Holmes AP. Nonparametric permutation tests for functional neuroimaging: a primer with examples. Human Brain Mapping. 2002;15(1):1–25. doi:10.1002/hbm.1058.

79. Pantazis D, Nichols TE, Baillet S, Leahy RM. A comparison of random field theory and permutation methods for the statistical analysis of MEG data. NeuroImage. 2005;25(2):383–394. doi:10.1016/j.neuroimage.2004.09.040.

80. Boser BE, Guyon IM, Vapnik VN. A training algorithm for optimal margin classifiers. In: Proceedings of the fifth annual workshop on Computational learning theory. COLT '92. Pittsburgh, Pennsylvania, USA: Association for Computing Machinery; 1992. p. 144–152. Available from: `https://doi.org/10.1145/130385.130401`.

81. Cortes C, Vapnik V. Support-vector networks. Machine Learning. 1995;20(3):273–297. doi:10.1007/BF00994018.

82. Burges CJC. A Tutorial on Support Vector Machines for Pattern Recognition. Data Mining and Knowledge Discovery. 1998;2(2):121–167. doi:10.1023/A:1009715923555.

83. Vapnik V. The Nature of Statistical Learning Theory. Springer Science & Business Media; 2013.

84. Fisher RA. The Use of Multiple Measurements in Taxonomic Problems. Annals of Eugenics. 1936;7(2):179–188. doi:10.1111/j.1469-1809.1936.tb02137.x.

85. Mika S, Ratsch G, Weston J, Scholkopf B, Mullers KR. Fisher discriminant analysis with kernels. In: Neural Networks for Signal Processing IX: Proceedings of the 1999 IEEE Signal Processing Society Workshop (Cat. No.98TH8468); 1999. p. 41–48.

86. Ojala M, Garriga GC. Permutation Tests for Studying Classifier Performance. The Journal of Machine Learning Research. 2010;11:1833–1863.

87. Good P. Permutation Tests: A Practical Guide to Resampling Methods for Testing Hypotheses. Springer Science & Business Media; 2013.

88. Thölke P, Mantilla-Ramos YJ, Abdelhedi H, Maschke C, Dehgan A, Harel Y, et al. Class imbalance should not throw you off balance: Choosing the right classifiers and performance metrics for brain decoding with imbalanced data. NeuroImage. 2023;277:120253. doi:10.1016/j.neuroimage.2023.120253.

89. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research. 2011;12(85):2825–2830.

90. Gramfort A, Luessi M, Larson E, Engemann DA, Strohmeier D, Brodbeck C, et al. MEG and EEG data analysis with MNE-Python. Frontiers in Neuroscience. 2013;7. doi:10.3389/fnins.2013.00267.

91. Seabold S, Perktold J. Statsmodels: Econometric and Statistical Modeling with Python. In: 9th Python in Science Conference. Austin, Texas; 2010. p. 92–96. Available from: `https://conference.scipy.org/proceedings/scipy2010/seabold.html`.

92. Virtanen P, Gommers R, Oliphant TE, Haberland M, Reddy T, Cournapeau D, et al. SciPy 1.0–Fundamental Algorithms for Scientific Computing in Python. Nature Methods. 2020;17(3):261–272. doi:10.1038/s41592-019-0686-2.

93. Perez F, Granger BE. IPython: A System for Interactive Scientific Computing. Computing in Science & Engineering. 2007;9(3):21–29. doi:10.1109/MCSE.2007.53.