Quantifying the impact and profiling functional EEG artifacts

1st Zhenyu Jin
BCI-NE, CSEE
University of Essex
Colchester, UK
frank990226@outlook.com

2nd Fabien Bourban
Mindmaze SA
Lausanne, Switzerland
fabien.bourban@mindmaze.ch

3rd Robert Leeb*
Mindmaze SA
Lausanne, Switzerland
robert.leeb@mindmaze.ch

4th Serafeim Perdikis*
BCI-NE, CSEE
University of Essex
Colchester, UK
0000-0003-2033-2486

***Abstract*—The susceptibility of electroencephalography (EEG) signal to artifacts is considered a major obstacle preventing the deployment of relevant non-invasive neurotechnology. In spite of a large body of literature dedicated to the identification, rejection and removal of artifactual components in EEG, the study of the impact that different artifacts may have on the EEG signal properties has been mostly qualitative and focused on the source (e.g. muscle activity, electromagnetic interference) rather than the function generating them. This work takes advantage of a unique dataset where EEG of 12 participants elicited during the execution of 9 common human activities (e.g., speaking, blinking, etc.) is co-registered with electromyography (EMG), electrooculography (EOG), accelerometer and gyroscope sensors, and baselined to "resting" (artifact-free) intervals to allow an exact, quantified assessment of the impact of artifacts. We examine several metrics capturing different facets of the influence of artifacts on EEG and measure the extent to which a state-of-the-art artifact removal method is able to eliminate them. In addition to an in-depth, quantified profiling of functional EEG artifacts, our work provides valuable information for precisely tuning the hyper-parameters of artifact rejection and removal algorithms and for designing realistic brain-computer interface (BCI) applications.**

***Index Terms*—electroencephalography, artifacts, quantification, functional artifacts, artifact removal**

I. Introduction

Distortion of the temporal dynamics and spectral consistency of electroencephalography (EEG) signals can occur as a result of interference by a variety of sources, either physiological (e.g., electrooculography (EOG) generated by blinking or eye movements, electromyography (EMG) activity of muscles close to the scalp, head and body movements) or external to the body (e.g., electromagnetic activity emitted by nearby devices, or by physical tampering with EEG sensors and cables, etc.). Such perturbations of normal EEG activity are termed "artifacts" and can be more formally defined as electrical potentials that are picked up by EEG sensors but do not reflect the brain's neuronal activity [1].

Owing to its very low amplitude which is often orders of magnitude smaller than that of the interference, EEG signal is known to be extremely vulnerable to artifacts. In most cases, the presence of artifacts can be determined by mere visual inspection of the EEG signal scope. The neuroimaging community has thus been, early on, well aware of the potential risks EEG artifacts may pose to the clinical or other applications of EEG and devised strategies to identify and cope with these [2]–[4]. While the early literature has mostly focused on the description, identification and rejection of artifacts (with a clear focus on EOG sources), the advent of brain-computer interface (BCI), an application that requires alleviating the influence of all possible types of artifact in real-time and in largely uncontrolled environments, has further contributed to the technical progress of this field and shifted attention from artifact detection/rejection to "removal" techniques. The term artifact removal (or correction) encompasses methods concerned with the "cleaning" of artifact-contaminated EEG signal intervals so as to keep them available in the processing pipeline for continuous, uninterrupted BCI, rather than only isolating these intervals and excluding them from further processing (artifact rejection) [5]. A series of surveys capture the gradual (including recent) progress, the increasing algorithmic elaborateness and pervasiveness, and the overall vast amount of work dedicated to this topic up to this day [3], [6]–[14].

Although the study of EEG artifacts has attracted considerable attention, this line of research can be said to suffer certain limitations. To begin with, there seems to be a tendency for qualitative and indirect, rather than strict and quantitative, characterization of artifacts, which further leads to similar shortcomings in the evaluation and application of the corresponding rejection and removal algorithms. In support of this, one may consider that despite the current abundance of powerful artifact rejection and removal approaches, the main defenses against artifacts in experimental designs remain the guidelines to participants towards the avoidance of artifacts in the first place, and the "manual" visual inspection of raw signals by experimenters to remove heavily affected data. Also typical of the current state-of-the-art is the fact that, even when elaborate techniques like blind source separation (BSS) are employed to discern the artifact-inflicted from the sound portions of EEG, the final classification of the extracted components into signal or artifact is still, usually, done manually [15], including in popular toolboxes widely used by the community. Furthermore, even where fully-automatic approaches are pursued, there is no consensus among different studies on the optimal criteria that can be used to reveal

* These authors contributed equally.

the presence and influence of artifacts; similarly, the hyper-parameters (e.g., thresholds applied on the different metrics) are rather arbitrary and not guaranteed to generalize [5], [16]. Another two issues emerging from the recent literature [9], [12]–[14] are the relative lack of universal and holistic EEG artifact studies, with most works focusing exclusively on eye-, muscle- or movement-related artifacts), and the related issue that this research has investigated the differences between potential sources of artifact, but not the "functions" (i.e., the human actions and activities) that create them. Because of this, it is still hard to infer what the real impact of different kinds of artifacts would be, especially in real-world settings.

The aforementioned limitations mostly stem from the unavailability of suitable datasets (more specifically, the lack of ground truth [17]) and metrics to study EEG artifacts. Profiling of artifacts and the evaluation of the corresponding rejection and removal methods are most often done on EEG segments known to be contaminated with one or more (usually, but not always, of known type) artifacts [11], [14]. Having only such information at hand, there is no way to know which components of the signal correspond to brain and which to foreign activity. As a result, conventional, established metrics for assessing the impact of noise in signal processing, like Signal-to-Noise ratio (SNR), cannot be strictly measured. In certain cases, artifact-free intervals are also not available for comparative analysis. Hence, analyses is forced to be either, as already mentioned, manual/qualitative, or indirect (e.g., increase of classification accuracy of a certain BCI after artifact removal). Alternatively, researchers resort to artificial datasets [5], [18], which however can considerably deviate from realistic scenarios. It is interesting that even in studies where some sort of ground truth is collected (e.g., parallel sensor readings correlating with the artifact source), these may be only used for the artifact rejection/removal algorithm and not for evaluation, or only in expensive [17], artificial or–as shown here–suboptimal ways [9], [18], [19], which highlights the additional need for better metrics, able to optimally exploit such "imperfect" ground truth.

This work attempts to address to some extent such deficiencies in the study of EEG artifacts, by introducing a dataset specifically designed for the assessment of "functional" artifacts, by highlighting the importance of using information-theoretic metrics to optimally exploit indirect ground truth, and by thereafter offering exact quantification of the impact of various functional artifacts on EEG. In more detail, we present and analyze a unique dataset where 12 subjects repetitively executed 9 different activities (e.g., speaking, blinking, etc.) interleaved with resting/idling intervals, while acquiring synchronously sampled EEG, EMG, EOG, accelerometer and gyroscope sensor readings. Our study's contribution is multifaceted. First, we profile and precisely characterize several common functional artifacts, answering several open questions pertaining to the overall impact of specific artifacts and to the spatial and spectral locations most likely to be affected. Second, we show that mutual information is superior to the coefficients of correlation and determination when exploiting indirect ground truth in the form of co-registered artifact-related signals. Third, we examine several measures that have been proposed in the literature for the identification of artifacts and evaluate their fitness, thus also better informing the design and parameterization of artifact rejection/removal methods and of BCI application paradigms. Lastly, we quantitatively evaluate the effectiveness of FORCe [5], a popular, fully-automatic artifact removal method, strictly assessing the state-of-the-art in this area.

II. Methods

A. Participants

Twelve able-bodied, adult volunteers (1 female, all right-handed, age 33.8 ± 7.4) with no known medical condition took part in the study. All subjects were informed in detail about the purpose of the study and the experimental protocol and signed written informed consent. The study was implemented in accordance with the mandates of the Declaration of Helsinki.

B. Experimental design

As motivated in the introduction, the experimental design targeted the simultaneous collection of EEG, facial EMG, EOG, head accelerometer and gyroscope data while participants are executing a broad range of tasks that are likely to contaminate the EEG signal with artifacts, and are essential part of human actions and activities that frequently occur in generic, everyday-life scenarios (i.e., functional tasks). Execution of these tasks is interleaved with "resting" intervals where subjects are required to idle and avoid any movement or action that can generate artifacts, thus providing the artifact-free baseline necessary to perform a comparative analysis. Table I summarizes the ensemble of 10 tasks considered in this study and any task-wise specific instructions given to the subjects or observation for each case.

TABLE I
EEG ARTIFACT TASKS/CLASSES

	Task	Guidelines and comments
1	Resting	Idling, avoidance of any movement
2	Channel pressing	Press any EEG sensor with finger
3	Blinking	Natural and repetitive
4	Eye movement	Left/right, up/down, circular, random
5	Head movement	Left/right, up/down, circular, random
6	Speaking	Reading out loud presented text
7	Swallowing	Natural, single repetition
8	Jaw clench	Repetitive
9	Frowning	Repetitive
10	Eyebrow raise	Repetitive

All subjects underwent a single experimental session. A session was divided into 8 blocks (each generating a single file) to allow for necessary inter-block breaks taken at the subjects' convenience. A block comprised 36 trials, each dedicated to one of the 9 "active" artifact classes in Table I (i.e., excluding the resting task), so that each specific artifact class is probed 4 times in each block and 32 times in total throughout a subject's session. For eye and head movements, each of the left/right,

up/down, circular or random (i.e., arbitrary and chosen by the subject) movement sub-type is prompted in 2 of the 8 blocks, so that each of these specific eye and head movement sub-types is executed 8 times by each subject in total, and 32 instances are collected for the corresponding overall eye/head movement class, as for all other artifact classes.

Fig. 1a graphically illustrates the timeline of a single trial. Each trial begins with a 2 s long "resting" epoch indicated by a fixation cross. As mentioned, subjects are instructed to "idle" during these epochs, staying as still as possible and avoiding any movement, activity or muscle twitch that could generate artifacts. These epochs collectively provide the aforementioned artifact-free baseline. The resting epoch is followed by a 5 s long artifact epoch during which the participant performs the particular artifact-generating task that the current trial is dedicated to. The artifact class to be executed is cued by an imperative message (e.g. "Blink", "Speak", "Swallow", "Head left/right", etc.). The message stays on for the whole duration of the artifact epoch of each trial. Its disappearance marks the end of the artifact epoch and the beginning of a 2 s long inter-trial epoch concluding the trial and allowing the subject to shortly relax and prepare for the next one. A trial thus lasts 9 s in total. For the Speaking task, a random, short text extract from Shakespeare's plays is also presented during the artifact epoch and subjects are required to read it out loud so as to standardize the task at hand. Participants are further asked to perform all tasks in a natural, ecological fashion, according to their personal style and habits. For artifact tasks that may naturally be very brief or even instantaneous (e.g. Blinking, Eyebrow raise), subjects were instructed to perform several repetitions within the 5 s epoch interval. For the Channel Pressing task, subjects were free to tamper with (e.g., press and/or softly pull) an arbitrary EEG sensor of their choice, which could differ among the different trial repetitions of this task. Overall, this experimental protocol accounts for 576 s of artifact-free baseline and 180 s of each of the 9 artifact tasks per subject. An experimental session would last approximately 90 minutes including preparation and inter-block breaks.

C. Experimental apparatus

During the experiment, subjects were comfortably seated in front of an office desk, wearing a non-commercial mixed-reality device named Elvira, which is developed by Mindmaze SA (Lausanne, Switzerland). Elvira combines an active biosignal acquisition system embedded into a virtual/augmented reality (VR/AR) head-mount display (HMD) for cutting edge, neuro-powered immersive mixed reality and biomedical applications [20]. The experimental paradigm's graphical user interface (GUI) elements were developed in Unity (Fig. 1a) and displayed on Elvira's HMD through cable connection with a local personal computer. EEG and EOG/EMG signals were synchronously sampled together with accelerometer and gyroscope sensor readings at a sampling rate of 500Hz. A total of 31 EEG locations of the international 10-20 placement system were monitored (exact channel layout shown in Fig. 1b) and referenced to both mastoids. The placement of 8 facial EMG/EOG electrodes used is shown as red dots in Fig. 1c. The position of an additional EOG sensor bundled with the EEG electrodes is shown in blue. Six degrees of freedom (DOF) head acceleration and orientation were simultaneously captured through a 3-axis accelerometer and a 3-axis gyroscope embedded in the HMD. We considered that, combined together, these additional, diverse 15 channels adequately describe the dynamics of each of the artifact classes studied here (except for Channel Pressing) to implicitly offer a measure of ground truth for our subsequent analysis.

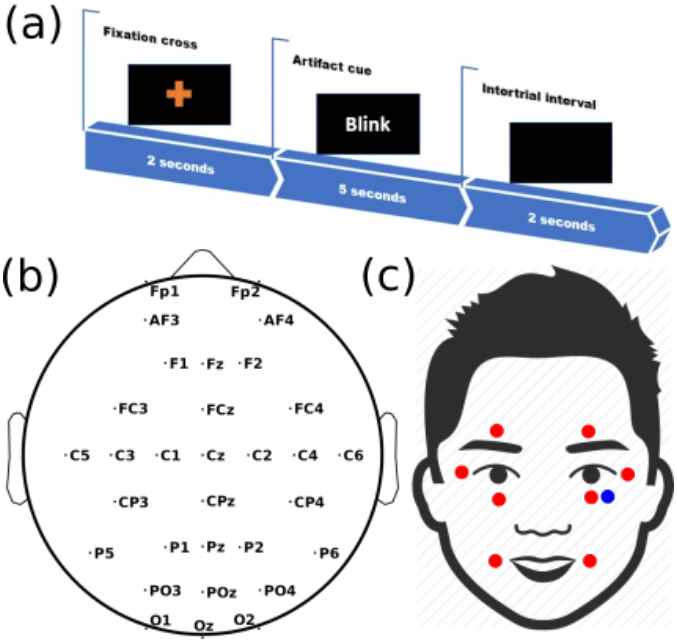

Fig. 1. Experimental protocol and apparatus: (a) Timeline of an exemplary trial of the experimental protocol. (b) EEG channel layout in the international 10-20 system. (c) Placement of facial EOG/EMG sensors.

D. Signal processing

Since our goal is to characterize artifacts and measure their impact rather than reduce it, there is deliberately no use of pre-processing operations on the raw EEG (e.g., spatial filters). The only exception is the application of techniques to detrend the EEG signal and remove the large offset differences observed among different channels, which can subsequently affect the assessment metrics employed although it's completely independent of the presence of artifacts. Specifically, EEG signals are treated with a bandpass finite impulse response (FIR) filter with 0.1 Hz low and 40 Hz high cut-off frequencies applied on the entire block's data (i.e., before epoch extraction). Single epochs of all channels are further treated with mean removal by subtracting the average epoch value from all the epoch's time samples. EEG spectrum is extracted with Welch power spectral density (PSD) (periodogram) [21] using 0.5 s long windows with 20% overlapping. For the band-specific analysis the EEG channels are filtered with a bandpass FIR filter at [1, 4] Hz for the δ, [4, 8] Hz for the θ, [8, 14] Hz for the α, and [18, 30] Hz for the β band, respectively. As

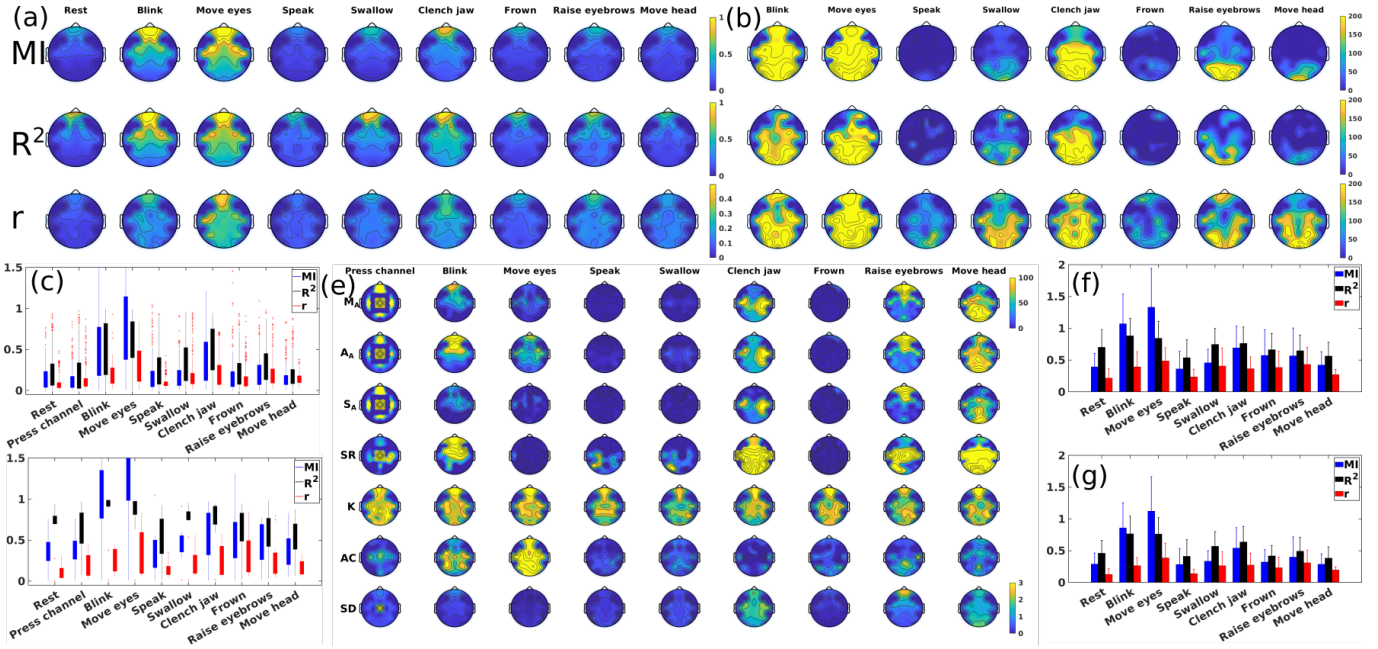

Fig. 2. Artifact impact quantification. (a) Scalp distribution of MI, R^2 and r for resting and all artifact classes and (b) increase (%) over the resting class. (c) Boxplots of MI, R^2 and r for all classes pulling all channels and subjects together and (d) pulling only the subjects' maximum value channel for each artifact. (e) Scalp distribution of increase (%) over the resting class for all artifact classes and EEG-only metrics M_A,A_A,S_A,SR,K,AC,SD. The tampered electrode in Channel Pressing is always exchanged with Cz for intuitive plotting. (f) Average (over subjects) MI, R^2 and r for each artifact class for the (f) maximum channel and the (g) average of 10 maximum channels of each subject and artifact. Note that the Channel Pressing artifact is not calculated for MI,R^2 and r since it is not captured by non-EEG sensors.

an example of the state-of-the-art artifact removal capacity we use the FORCe method elaborately described in [5].

E. Metrics for artifact impact assessment

As discussed, the major contribution of our work is the strict quantification of the influence of EEG artifacts thanks to the parallel acquisition of implicit ground truth in the form of extra, non-EEG sensor data capturing the dynamics of the artifact-inflicting sources. In that respect, we employ the coefficients of correlation (Pearson) r and of determination R^2 (expressing the fraction of explained variance, based here on the standard multiple linear regression estimator), that have been proposed before in studies that did co-register additional sensors for a limited number of artifacts, as well as mutual information MI which, to our best knowledge, has never been used in this context. For estimating the MI, we make use of the method and code in [22]. MI and R^2 are inherently multivariate (i.e., directly take into account all external sensors). For r, we correlate separately the EEG electrodes with each external sensor and use the maximum absolute value of these correlations.

Another contribution we seek to offer is the stricter profiling of artifacts and their benchmarking against an artifact-free baseline in terms of conventional (i.e., based on EEG-only input) metrics that have been used by artifact detection methods, thus potentially better informing their usage and parameterization. In that regard, we include in our analysis a (non-exhaustive) list mainly inspired by the comprehensive work in [5] that comprises the maximum M_A (more precisely: the mean of the 10% highest absolute values), average A_A and standard deviation S_A of the absolute value of the EEG amplitude, the number of spikes per second (spike rate) SR, the kurtosis K of the signal distribution, the sum of the absolute values of the signal's sample autocorrelation function series (for 1000 lags) AC and the spectrum distortion SD (as the average z-score deviation of the power of each band of the artifact class' spectrum from the resting class spectrum). All metrics are computed separately for each of the 31 EEG channels available after concatenating all subject trials corresponding to each artifact class. The grand average spectra used for calculating SD are computed per trial, without concatenation and then averaged within each artifact class.

III. Results

Fig. 2a-d demonstrate how the collected implicit ground truth allows to precisely quantify the impact of artifacts through metrics MI, R^2 and r. The basic prerequisite for any such metric is the existence of a theoretical lower bound corresponding to "completely artifact-free" data, i.e., when EEG signals reflect only the underlying neuronal activity. Ideally, this lower bound should be 0, or at least independent of the experimental conditions (and, thus, reproducible and comparable across all EEG recordings). Experimentally, EEG collected under controlled "resting" conditions should score close to this lower bound. Importantly, this condition must be valid for all potential artifact sources (universality). It is further desirable (but, not necessary) that there also exists an upper bound corresponding to "fully artifact contaminated" case, i.e.,

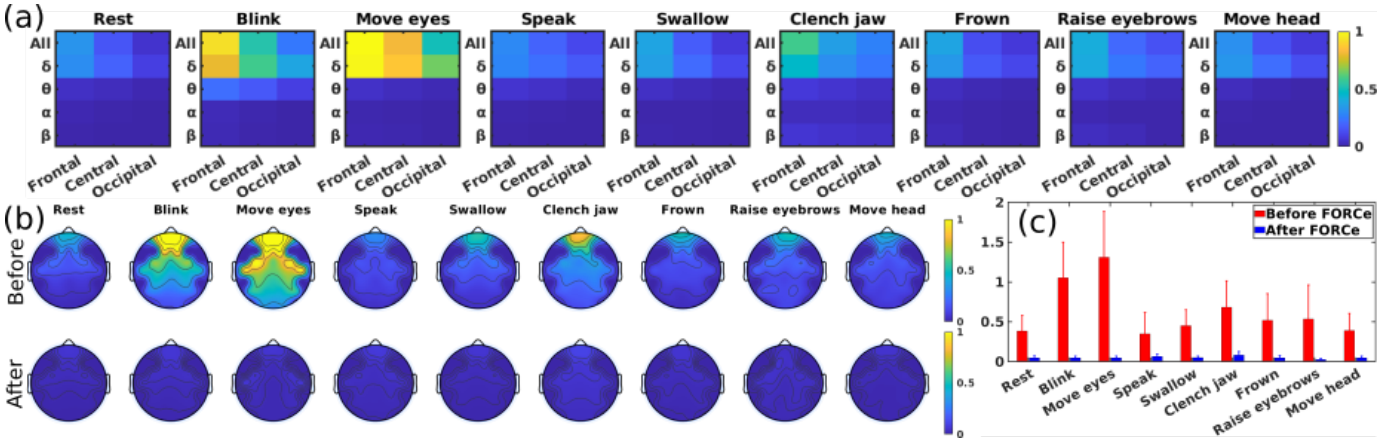

Fig. 3. (a) MI for resting and all artifact classes measured separately over frontal, central and occipital scalp regions and different frequency bands including: full-band ("All"), δ ([1, 4] Hz), θ ([4, 8] Hz), δ ([8, 14] Hz), δ ([18, 30] Hz). (b) Scalp distribution of MI and (c) maximum channel MI for all artifact classes before (first row) and after (second row) the application of FORCe [5].

when the presence of artifacts fully explains the observed EEG which thus do not reflect at all brain activity. Satisfying these conditions permits to not only identify the presence of artifacts, but to exactly arrange them on an ordered "axis of impact" and proceed with exact inferences, e.g., that "artifact X will have double the impact of Y on channel Z". Conventional EEG-only metrics as shown in Fig. 2e do not comply with these conditions. On the contrary, MI, R^2 and r do so, subject to the constraint that the external sensors available can capture all examined artifact sources (this is the reason why Channel Pressing is excluded from this analysis, as it is not captured by any of the EMG, accelerometer and gyroscope sensors available here). These metrics theoretically respect a 0-lower bound (R^2 and r also have a convenient theoretical upper bound of 1). This is experimentally verified in Fig. 2a, as all channels score close to 0 at the Rest class and close to 1 for the frontal channels of the Blink and Eye Movement classes, as anticipated. Deviations from the lower bound at the Rest class may be explained by (ocular or other) artifacts that subjects subconsciously generated despite the instructions given, as well as by unavoidable physiological processes (breathing, heart-beat, etc.) which may slightly influence all sensors. Such deviations can be observed for all three metrics in Fig. 2 a (Rest class), but seem to be more intense for R^2 (frontal channels). Fig. 2c-d shed further light on this comparison, confirming that maximally impacted frontal channels with R^2 deviate substantially from 0 (see Rest in Fig. 2d) even if, over the whole scalp, this effect is diminished (Fig. 2c), apparently due to sensitivity of this metric to EOG sources. This reduces the ability of R^2 to discriminate between Rest and certain artifact classes (Blink, Eye Movements, Frown, etc.). On the other hand, r seems to suffer reduced overall sensitivity (i.e., small differences between Rest and even the strongest artifact classes) and rarely approaches the theoretical maximum (Fig. 2c-d). MI avoids these shortcomings, suggesting that information theoretic metrics may be more suitable for this task compared to statistical measures (r, R^2). MI can also be shown to have a theoretical upper bound, which however is dependent on the entropy of the recorded signals and not very convenient to estimate; this is not particularly detrimental, as in practice the strongest artifacts seem to naturally asymptote to an MI value close to 1.5. Fig. 2b shows that MI also provides the smoothest scalp distributions of impact as % increase over the Rest class.

Fig. 2e stands to show that conventional metrics relying only on EEG are able, in combination, to capture the existence of all functional artifacts studied here. This is very important as, in the general case, EEG neurotechnology may not be able to afford the recording of additional, external signals. However, Figure 2e confirms the non-universality of all of these metrics, as no single measure succeeds in detecting all artifact classes. A possible exception seems to be the Kurtosis K, which, nevertheless, seems to be relatively unstable. By showing the response of each metric to each artifact (as % increase with respect to Rest), this figure can be used to fine-tune the parameters of artifact removal and rejection algorithms employing these criteria.

Fig. 2f-g compare the impact of the functional artifacts included in this work for profiling purposes. Fig. 2f focuses on local intensity (the most impacted channel is compared), while Fig. 2g takes also into account the spread (average of the 10 most impacted channels). Observed in conjunction with Fig. 2a, the most interesting finding is that many of the functional artifact classes tested, namely, Speaking, Swallowing, Frowning, and Moving the Head, have in fact very small impact on the $[0, 40]$ Hz EEG band (thought to be useful for BCI), negligible when compared with the Rest class (i.e., impact is comparable to the slight artifacts induced even when people are consciously trying to avoid them). On the other hand, the particularly detrimental impact of actions like Blinking, Moving the Eyes, Clenching the Jaw, etc., underlined in the literature, is substantiated by the exact quantification implemented here. It must be also noted that artifact impact does not only depend on the type of artifact, but also on

the individual, as shown by a two-way ANOVA (maximum channel MI as response variable) with significant main effects for factors Artifact Class ($F = 6.57, p < 10^{-7}$) and Subject ($F = 15.4, p < 10^{-13}$). The latter conclusion is reasonable, since different individuals will tend to perform the same task in a different manner, which influences the intensity of generated artifacts.

Further to functional artifact profiling, Fig. 2a-b illustrate in detail the topographic distribution and spread of the impact of different artifacts. The anticipated susceptibility of frontal electrodes in most cases is both apparent and anticipated. What may be less expected and must be highlighted is the considerable spread of Blinking and Eye movement artifacts to occipital regions, as well as the occipital focus of Head Movement artifacts (likely because of the associated neck EMG activity) and Eyebrow Raise (probably because of concomitant head movements). With regard to spectral profiling, Fig. 3 shows that slow EEG rhythms δ and θ account for almost the entirety of artifact impact observed in full-band EEG, irrespective of the artifact type. This is a highly unanticipated result with significant implications for BCI and real-world usage of EEG technology, as it seems that systems employing or studying faster rhythms (e.g., a motor imagery BCI) wont be particularly vulnerable to artifacts in any case.

Lastly, Fig. 3b-c proves that independent component analysis (ICA)-based BSS artifact removal algorithms like FORCe can very effectively eliminate all functional artifacts examined here. It is also demonstrated that artifact-free EEG intervals will indeed approach the theoretical lower bound of 0 MI, as the application of FORCe suggests (Fig. 3c).

IV. Discussion

This study has produced a unique dataset and investigated metrics able to provide exact quantification of the impact of EEG artifacts. We showed that mutual information is superior for this task compared to statistical competitors previously proposed. We have further compared and informed the design and parameterization of EEG-based artifact detection metrics. Profiling of functional artifacts indicates that in most cases the spread and intensity of the influence of such artifacts is minimal or negligible. Importantly, the impact is largely limited to low frequency bands. This, together with the fact that FORCe [5] has successfully alleviated all induced artifacts leads to the conclusion that artifact removal can be broadly considered a solved problem (scientifically), with further efforts directed to technical matters like real-time implementations.

Ours study has been limited by the absence of sensors to (better) quantify electrode tampering, neck activity and power-line noise. Additionally, the computation of the r metric may have been suboptimal and can be improved by investigating suitable combinations of external sensors for each artifact type (e.g., vertical/horizontal EOG bipoles), which, however, must be noted to be a cumbersome or expensive task. Our future work will focus on the benchmarking of popular artifact removal methods with the tools derived here.

References

[1] M. X. Cohen, *Analyzing Neural Time Series Data: Theory and Practice*. The MIT Press, 2014.

[2] S. A. Hillyard and R. Galambos, "Eye movement artifact in the CNV," *Electroencephalography and Clinical Neurophysiology*, vol. 28, no. 2, pp. 173–182, 1970.

[3] R. Croft and R. Barry, "Removal of ocular artifact from the EEG: a review," *Clinical Neurophysiology*, vol. 30, no. 1, pp. 5–19, 2000.

[4] P. L. Nunez and R. Srinivasan, "Fallacies in EEG," in *Electric Fields of the Brain*. Oxford University Press, 2006, pp. 56–98.

[5] I. Daly, R. Scherer, M. Billinger, and G. Müller-Putz, "FORCe: Fully Online and Automated Artifact Removal for Brain-Computer Interfacing," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 23, no. 5, pp. 725–736, 2015.

[6] P. Anderer, S. Roberts, A. Schlögl, G. Gruber, G. Klösch, W. Herrmann, P. Rappelsberger, O. Filz, M. J. Barbanoj, G. Dorffner, and B. Saletu, "Artifact processing in computerized analysis of sleep EEG – a review," *Neuropsychobiology*, vol. 40, no. 3, pp. 150–157, 1999.

[7] M. Fatourechi, A. Bashashati, R. K. Ward, and G. E. Birch, "EMG and EOG artifacts in brain computer interface systems: A survey," *Clinical Neurophysiology*, vol. 118, no. 3, pp. 480–494, 2007.

[8] K. T. Sweeney, D. Kelly, T. E. Ward, and S. F. McLoone, "A review of the state of the art in artifact removal technologies as used in an assisted living domain," in *IET Seminar on Assisted Living 2011*, 2011, pp. 1–6.

[9] S. D. Muthukumaraswamy, "High-frequency brain activity and muscle artifacts in MEG/EEG: a review and recommendations," *Frontiers in Human Neuroscience*, vol. 7, 2013.

[10] M. K. Islam, A. Rastegarnia, and Z. Yang, "Methods for artifact detection and removal from scalp EEG: A review," *Clinical Neurophysiology*, vol. 46, no. 4-5, pp. 287–305, 2016.

[11] X. Jiang, G.-B. Bian, and Z. Tian, "Removal of artifacts from EEG signals: A review," *Sensors*, vol. 19, no. 5, p. 987, Feb. 2019.

[12] X. Chen, X. Xu, A. Liu, S. Lee, X. Chen, X. Zhang, M. J. McKeown, and Z. J. Wang, "Removal of muscle artifacts from the EEG: A review and recommendations," *IEEE Sensors Journal*, vol. 19, no. 14, pp. 5353–5368, Jul. 2019.

[13] S. Kotte and J. R. K. K. Dabbakuti, "Methods for removal of artifacts from EEG signal: A review," *Journal of Physics: Conference Series*, vol. 1706, no. 1, p. 012093, 2020.

[14] D. Gorjan, K. Gramann, K. D. Pauw, and U. Marusic, "Removal of movement-induced EEG artifacts: current state of the art and guidelines," *Journal of Neural Engineering*, vol. 19, no. 1, p. 011004, Feb. 2022.

[15] T.-P. Jung, S. Makeig, C. Humphries, T.-W. Lee, M. J. McKeown, V. Iragui, and T. J. Sejnowski, "Removing electroencephalographic artifacts by blind source separation," *Psychophysiology*, vol. 37, no. 2, pp. 163–178, Mar. 2000.

[16] I. Daly, N. Nicolaou, S. J. Nasuto, and K. Warwick, "Automated artifact removal from the electroencephalogram," *Clinical EEG and Neuroscience*, vol. 44, no. 4, pp. 291–306, 2013.

[17] N. Oosugi, K. Kitajo, N. Hasegawa, Y. Nagasaka, K. Okanoya, and N. Fujii, "A new method for quantifying the performance of EEG blind source separation algorithms by referencing a simultaneously recorded ECoG signal," *Neural Networks*, vol. 93, pp. 1–6, 2017.

[18] M. Crespo-Garcia, M. Atienza, and J. L. Cantero, "Muscle artifact removal from human sleep EEG by using independent component analysis," *Annals of Biomedical Engineering*, vol. 36, no. 3, pp. 467–475, 2008.

[19] M. Mennes, H. Wouters, B. Vanrumste, L. Lagae, and P. Stiers, "Validation of ICA as a tool to remove eye movement artifacts from EEG/ERP," *Psychophysiology*, 2010.

[20] R. Leeb and D. Pérez-Marcos, "Chapter 14 - Brain-computer interfaces and virtual reality for neurorehabilitation," in *Brain-Computer Interfaces*, ser. Handbook of Clinical Neurology, N. F. Ramsey and J. del R. Millán, Eds. Elsevier, 2020, vol. 168, pp. 183–197.

[21] P. D. Welch, "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 Electroacoustics*, vol. AU, pp. 70–73, 1967.

[22] R. A. Ince, B. L. Giordano, C. Kayser, G. A. Rousselet, J. Gross, and P. G. Schyns, "A statistical framework for neuroimaging data analysis based on mutual information estimated via a gaussian copula," *Human Brain Mapping*, vol. 38, no. 3, pp. 1541–1573, 2016.