Received: 5 May 2022 | Accepted: 5 June 2022

DOI: 10.1002/hbm.25995

TECHNICAL REPORT

WILEY

Transfreq: A Python package for computing the theta-to-alpha transition frequency from resting state electroencephalographic data

Elisabetta Vallarino[1] | Sara Sommariva[1,2] | Francesco Famà[3,4] | Michele Piana[1,4] | Flavio Nobili[3,4] | Dario Arnaldi[3,4]

[1]Dipartimento di Matematica (DIMA), Università degli Studi di Genova, Genoa, Italy

[2]CNR-SPIN, Genoa, Italy

[3]Dipartimento di Neuroscienze, Riabilitazione, Oftalmologia, Genetica e Scienze Materno-Infantili (DINOGMI), Università degli Studi di Genova, Genoa, Italy

[4]IRCCS Ospedale Policlinico San Martino, Genoa, Italy

Correspondence
Elisabetta Vallarino, Dipartimento di Matematica (DIMA), Università degli Studi di Genova, Via Dodecaneso 35, 16147 Genoa, Italy.
Email: vallarino@dima.unige.it

Abstract

A classic approach to estimate individual theta-to-alpha transition frequency (TF) requires two electroencephalographic (EEG) recordings, one acquired in a resting state condition and one showing alpha desynchronisation due, for example, to task execution. This translates into long recording sessions that may be cumbersome in studies involving patients. Moreover, an incomplete desynchronisation of the alpha rhythm may compromise TF estimates. Here we present *transfreq*, a publicly available Python library that allows TF computation from resting state data by clustering the spectral profiles associated to the EEG channels based on their content in alpha and theta bands. A detailed overview of *transfreq* core algorithm and software architecture is provided. Its effectiveness and robustness across different experimental setups are demonstrated on a publicly available EEG data set and on in-house recordings, including scenarios where the classic approach fails to estimate TF. We conclude with a proof of concept of the predictive power of *transfreq* TF as a clinical marker. Specifically, we present a scenario where *transfreq* TF shows a stronger correlation with the mini mental state examination score than other widely used EEG features, including individual alpha peak and median/mean frequency. The documentation of *transfreq* and the codes for reproducing the analysis of the article with the open-source data set are available online at https://elisabettavallarino.github.io/transfreq/. Motivated by the results showed in this article, we believe our method will provide a robust tool for discovering markers of neurodegenerative diseases.

KEYWORDS

clustering, machine learning, neurodegenerative diseases, power spectrum, quantitative EEG, transition frequency

Abbreviations: EEG, electroencephalography; IAP, individual α peak; MDF, median frequency; MEF, mean frequency; MMSE, mini mental score examination; PS, power spectrum; qEEG, quantitative EEG; TF, transition frequency.

1 | INTRODUCTION

The analysis of resting state EEG power spectra is a reliable and cheap tool for studying both normal aging (Babiloni et al., 2006; Soininen et al., 1982) and neurodegenerative brain diseases (Klassen et al., 2011; Malek et al., 2017; Moretti et al., 2004). For example, there is evidence that the EEG power in the alpha band and in the slow-wave frequency bands (e.g., theta and delta) shows a direct and an inverse correlation with cognitive performances, respectively. This result has been exploited to support the discrimination of patients affected by the most common neurodegenerative brain diseases from healthy controls (Jaramillo-Jimenez et al., 2021; Klimesch, 1999; Özbek et al., 2021). However, such harmonic behaviours often present significant individual differences (Donoghue et al., 2020) and, moreover, alpha and theta bands, whose power expresses opposite pathophysiological meanings, are contiguous. Therefore, at the individual level the risk is consistent that part of the alpha power band is included in the range of the theta power (i.e., 4–8 Hz), thus implying a wrong interpretation of its (patho)physiological meaning. Establishing the theta-to-alpha transition frequency (TF) at an individual level is therefore of paramount importance in order to avoid misinterpretation of quantitative EEG (qEEG) data. The availability of a computational tool for the determination of TF represents a crucial prerequisite for a meaningful usability of frequency-band power analysis for both research and clinical purposes.

A classic definition of TF was described by Klimesch and colleagues more than 20 years ago (Klimesch, 1999). Their approach, henceforth referred as Klimesch's method, relies on the fact that event-related desynchronisation induces a decrease of the alpha power and an increase of the theta power of the event-related power spectrum, with respect to the power spectrum measured during resting state (Klimesch et al., 1997). It immediately follows that theta-to-alpha TF can be determined by comparison between the task-related and the resting state power spectra. However, two main drawbacks affect this method limiting its actual use in clinical settings: (i) it needs the acquisition of two data sets, that is, a resting state and an event-related time series, and (ii) the task utilised for event-related recording must induce changes in the power spectrum significant enough to allow the identification of variations in the alpha and theta power. To overcome such limitations, a more practical and commonly employed definition of TF consists in finding the lowest value of the (resting state) EEG power spectra between the δ and α peak (Babiloni et al., 2016; Moretti et al., 2004; Moretti et al., 2007). However, due to the noisy nature of EEG time-series, such definition may often result ambiguous as multiple local minima may exist (Poza et al., 2007). Other methods for defining TF have been introduced, based, for example, on the distance from the individual alpha peak (Lansbergen et al., 2011), demonstrating the lack of consensus on a robust method for computing TF, which may also partially explain the limited number of studies exploiting this feature.

The present study aims at fulfilling this lack by introducing *transfreq*, a publicly available Python package implementing a novel algorithm for the automated computation of TF from theta to alpha band that works even when just one resting state EEG recording is available for each subject. This computational approach relies on the determination of appropriate features associated to the power spectrum measured at each channel, and on the application of an unsupervised algorithm that automatically identifies two clusters of EEG sensors associated to the alpha and theta bands, respectively. In *transfreq* we implemented four different strategies characterised by different sensor-level features and different clustering algorithms (Saxena et al., 2017). The workflow of these approaches is illustrated in the case of a test-bed example and validated on both an open-source data set and time series recorded during an experiment performed in our lab. To quantitatively evaluate the results of *transfreq*, we considered only the EEG recordings where Klimesch's method estimated a feasible value of TF and we used such a value as ground truth: for most of the analysed data the value of TF returned by *transfreq* was less than 1 Hz from TF estimated by Klimesch's method. Additionally, we show some typical scenarios in which Klimesch's method fails in capturing the correct TF while *transfreq* still returns plausible estimates.

A detailed documentation of *transfreq* and the codes for reproducing the analysis of the article with the open-source data set are available online at https://elisabettavallarino.github.io/transfreq/.

2 | METHODS AND MATERIALS

2.1 | Klimesch's method

A classic approach to compute theta-to-alpha TF is that proposed by Klimesch and colleagues (Klimesch, 1999) and schematically depicted in Figure 1(a). In detail, Kilmesch's method requires two EEG recordings as input, one acquired during a resting state condition and one acquired while the subject is performing a task. For both recordings and for each of the N EEG sensors, the power spectrum (Bendat & Piersol, 2011; Vallarino et al., 2020) of the corresponding time series is computed and normalised by dividing by the norm over all frequencies, that is, we computed

$$\widetilde{P}_i^{\text{task}}(f) = \frac{P_i^{\text{task}}(f)}{\sum_f P_i^{\text{task}}(f)} \text{ and } \widetilde{P}_i^{\text{rest}}(f) = \frac{P_i^{\text{rest}}(f)}{\sum_f P_i^{\text{rest}}(f)}, \quad (1)$$

where $P_i^{\text{task}}(f)$ and $P_i^{\text{rest}}(f)$ are the power spectra at frequency f of the signal recorded by the ith sensor during the task and the resting state conditions, respectively. Then, the mean over all the EEG channels of the normalised power spectra in (1) is computed to obtain two spectral profiles, namely

$$S_{\text{task}}(f) = \frac{1}{N}\sum_{i=1}^{N} \widetilde{P}_i^{\text{task}}(f) \text{ and } S_{\text{rest}}(f) = \frac{1}{N}\sum_{i=1}^{N} \widetilde{P}_i^{\text{rest}}(f)\,. \quad (2)$$

Klimesch's method relies on the fact that S_{rest} usually presents a peak in the alpha band while, due to task-related alpha desynchronisation, S_{task} presents a lower intensity in the alpha band and a higher

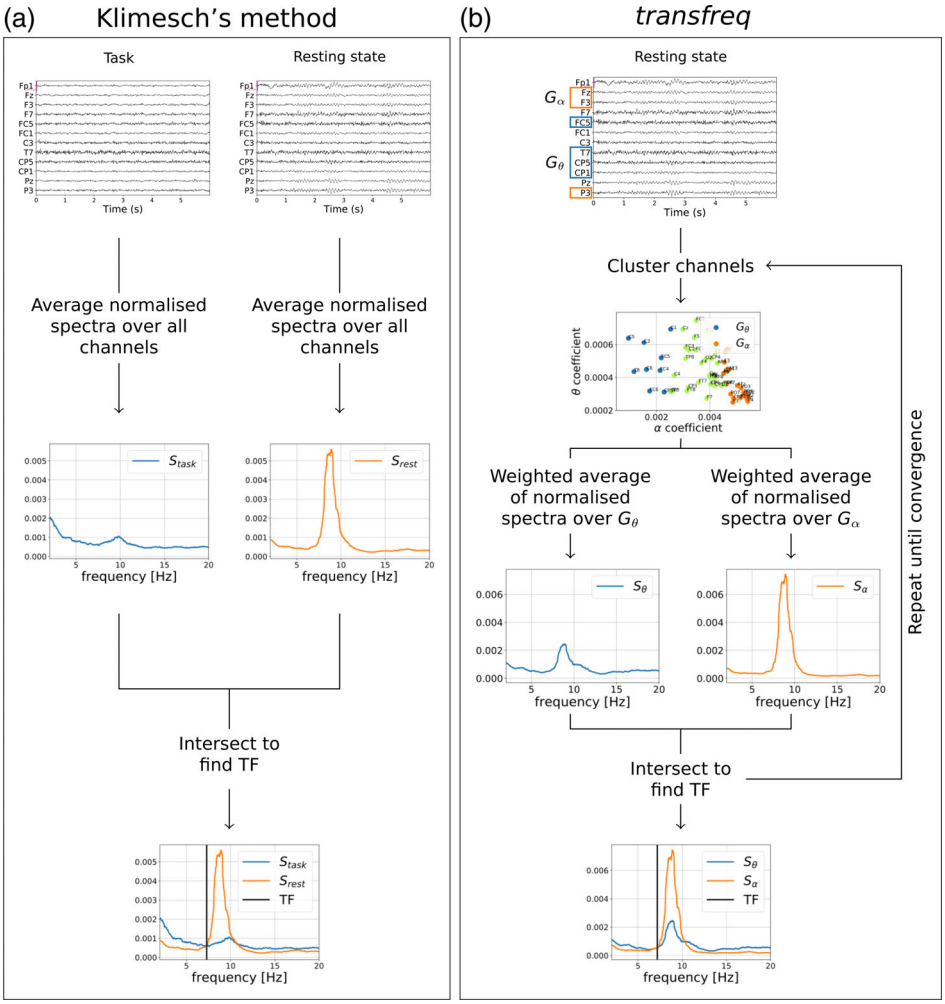


FIGURE 1 Comparison between the pipelines of Kilmesch's method (a) and of *transfreq* (b)

intensity in the theta band with respect to S_{rest} (Klimesch, 1996; Klimesch et al., 1998; Schacter, 1977). TF is thus defined as the highest frequency before the individual alpha peak (IAP) at which S_{task} and S_{rest} intersect. Here, the IAP is defined as the frequency in the range (Jaramillo-Jimenez et al., 2021; Poza et al., 2007) Hz at which S_{rest} peaks (Babiloni et al., 2004).

2.2 | Minimum method

In many practical scenarios, it is often the case that only data recorded at rest are available and thus Klimesch's method cannot be applied. In this case, a commonly applied method, henceforth referred to as minimum method, consists in defining TF as the frequency lower then IAP where the averaged normalized power spectra S_{rest} has a minimum (Moretti et al., 2004). Here S_{rest} and IAP are computed as described in the previous section.

2.3 | Transfreq algorithm

In this article we introduce transfreq, a method to automatically compute TF from theta to alpha band when only resting state EEG data are available. *Transfreq* relies on a rationale similar to that of Klimesch's method. Namely, TF is defined as the intersection between two spectral profiles differing in their content within the alpha and theta bands. However, with respect to Klimesch's method, such profiles are computed by exploiting the fact that alpha and theta activities are not uniformly expressed across the different EEG channels. In fact, some channels present high alpha activity (typically, channels above the occipital lobe), whereas others show lower alpha and higher theta activities (typically, channels corresponding to temporal and frontal brain areas) (Klimesch, 1996; Nunez et al., 2001). Consequently, two groups of channels can be identified: the first group includes channels characterized by a preponderant alpha activity (this group plays a role analogous to the one of EEG data measured at rest in Klimesch's method); the second group includes channels showing preponderant theta activity and limited alpha activity (this second group plays a role analogous to the one of the task-evoked EEG recordings in Klimesch's method).

The *transfreq* pipeline is schematically illustrated by Figure 1(b) and Algorithm 1. In detail, for each EEG channel the normalised power spectrum is computed as in Equation (1), that is,

$$\widetilde{P}_i(f) = \frac{P_i(f)}{\sum_f P_i(f)}, \quad \forall i \in \{1,\ldots,N\}. \tag{3}$$

TF is determined through the following iterative procedure.

i. Set an initial value for the alpha and theta frequency-bands. Specifically, the alpha frequency-band is identified as a 2 Hz range centred on the IAP, which is defined as the frequency where the power spectrum averaged over all sensors peaks; the theta frequency-band is set equal to (Jaramillo-Jimenez et al., 2021; Klassen et al., 2011) Hz, or to [IAP − 3, IAP − 1] Hz if the previous interval overlaps with the alpha frequency-band.
ii. Compute, for each channel, the alpha and theta coefficients by averaging the normalised power spectrum $\widetilde{P}_i$ over the corresponding frequency band.
iii. Apply a clustering algorithm to identify two groups of channels based on the alpha and theta coefficients. The channels in the first group, denoted as G_θ, will be characterised by low alpha and high theta activities, while the channels in the second group, G_α, will be characterised by high alpha and low theta activities. Two spectral profiles are thus obtained through a weighted average of the power spectra over the two groups, that is,

$$S_\theta(f) = \frac{1}{\sum_{i \in G_\theta} w_i^\theta} \sum_{i \in G_\theta} w_i^\theta \, \widetilde{P}_i(f) \text{ and } S_\alpha(f) = \frac{1}{\sum_{i \in G_\alpha} w_i^\alpha} \sum_{i \in G_\alpha} w_i^\alpha \, \widetilde{P}_i(f), \quad (4)$$

where w_i^θ and w_i^α are the theta and alpha coefficients for channel i, respectively.

iv. Define a first estimate of TF as the highest frequency before the IAP at which S_θ and S_α intersect.
v. Use the value of TF computed in (iv) to define new, more accurate, alpha and theta frequency bands, set equal to [max{IAP−1, TF}, IAP + 1] and [TF − 3, TF − 1], respectively. Such a choice guarantees the intervals to be fully characterised by alpha and theta activation. Indeed, we chose narrower bands with respect to the classic 4 Hz ranges defined in the literature (Bazanova & Vernon, 2014; Klimesch, 1999) and we impose at least a 1 Hz separation between the intervals.

Steps (ii)–(v) are iterated until a desired level of accuracy is reached, quantified as the difference between two consecutive estimates of TF. The desired level of accuracy is set equal to the highest value between 0.1 Hz and the frequency resolution Δ_f. The rationale behind this choice is that 0.1 Hz is an acceptable error when computing TF. However, if the frequency resolution is lower (i.e., Δ_f > 0.1 Hz), setting the desired level of accuracy to 0.1 Hz would be the same as setting it to 0, which is a too strong requirement; therefore in such cases the level of accuracy is set equal to the frequency resolution.

We point out that the effectiveness of *transfreq* depends on the clustering procedure used to define the two groups of channels G_θ and G_α. In *transfreq* we have implemented four different algorithms, described in the next subsections.

Algorithm *transfreq* core algorithm

Input: Resting state EEG data recorded by N sensors
Compute and normalise sensors' power spectra as in (3)
Initialise theta and alpha frequency bands
$\varepsilon := |\mathrm{TF}_{new} - \mathrm{TF}_{old}| = +\infty$
while $\varepsilon \geq toll$ **do**
  Compute alpha and theta coefficients, w_i^α, w_i^θ, $i = 1,\ldots,N$
  Define channel groups, G_θ and G_α, through a clustering method
  Update TF
  Update ε
  Update theta and alpha frequency bands

2.3.1 | Clustering method 1: 1D thresholding

The first clustering method implemented in *transfreq* is based on the ratio between the alpha and theta coefficients computed for each channel. In fact, channels with a low value of such alpha-to-theta ratio are characterised by low alpha and high theta activities, whereas channels with a high value are characterised by high alpha and low theta activities. The first group of channels, G_θ, is thus defined by the four channels showing the lowest values of the alpha-to-theta ratio, while the second group, G_α, is defined by the four channels showing the highest values of the same ratio. A visual representation of this approach on a representative data set can be seen in Figure 2(a). In *transfreq*, the number of channels in each group has been set equal to 4 after computing and visually inspecting the results for different values of such a parameter. In fact, the overall behaviour of the algorithm was similar across the different tested values.

2.3.2 | Clustering method 2: 1D mean-shift

One drawback of the previous approach is the need to heuristically set the number of channels within the two groups, G_α and G_θ. To overcome such a limitation, we implemented a second clustering approach where the Mean Shift algorithm (Comaniciu & Meer, 2002) is used to cluster the EEG sensors with respect to the ratio between the alpha and theta coefficients computed, for each channel, as described in the previous sub-section. To this end we used the MeanShift function available within the Python package Scikit Learn (Pedregosa et al., 2011) that also automatically determines the number of clusters. G_θ is then defined equal to the cluster containing the channel with the lowest value of the alpha-to-theta ratio, while G_α is set equal to the cluster containing the channel with the highest value of the same ratio. A visual representation of this approach on a representative data set can be seen in Figure 2(b).

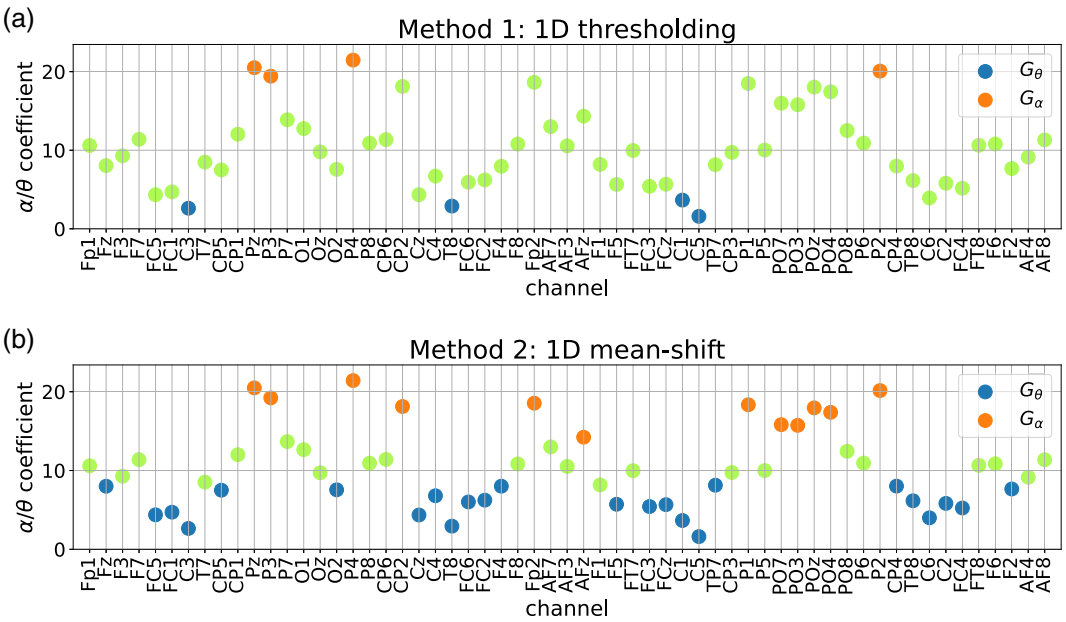
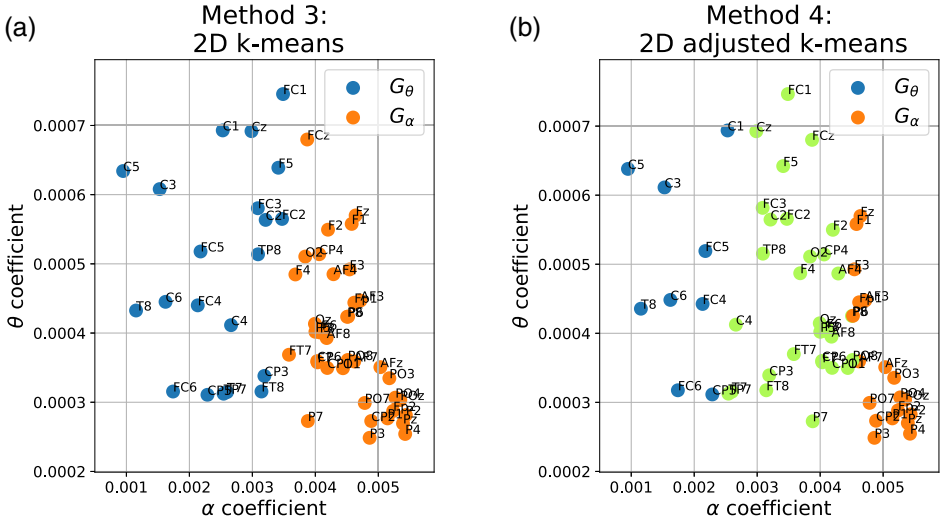


FIGURE 2 Performance illustration of the 1D clustering approaches thresholding (a) and mean-shift (b). Both panels show the value of the ratio between alpha and theta coefficients as function of the EEG sensors. Channels that belong to G_θ and G_α are represented as blue and orange dots, respectively. In *transfreq*, the remaining channels (green dots) are excluded from the subsequent analysis

FIGURE 3 Performance illustration of the 2D clustering approaches k-means (a) and adjusted k-means (b). Both panels show the value of the theta coefficients on the y-axis and that of the alpha coefficients on the *x*-axis. Channels that belong to G_θ and G_α are represented as blue and orange dots, respectively. In *transfreq*, the remaining channels (green dots) are excluded from the subsequent analysis



2.3.3 | Clustering method 3: 2D k-means

Both approaches described in the previous sub-sections rely on 1-dimensional clustering techniques that use the ratio between the alpha and theta coefficients as feature. In the third approach implemented in *transfreq* we exploited the k-means algorithm (Lloyd, 1982) to cluster the EEG sensors by using the alpha and theta coefficients as two distinct features. To this end, we used the KMeans function within the Python package Scikit Learn (Pedregosa et al., 2011). The number of clusters to generate is set equal to 2. Then G_α is defined as the cluster whose centroid shows the highest value of the alpha coefficient, while the other cluster defines G_θ. As illustrated in Figure 3(a), channels belonging to G_α (orange dots) typically present a higher alpha coefficient and a lower theta coefficient than the other ones (blue dots).

2.3.4 | Clustering method 4: 2D adjusted k-means

The fourth clustering approach implemented in *transfreq* takes as input the two sensors groups, G_α and G_θ, computed using the k-means algorithm as described in the previous sub-section. However, the two groups are now adjusted so that only sensors showing the highest intercluster difference in terms of the alpha and theta coefficient values are retained. To this end, as illustrated in Figure 3(b), we removed from G_α and G_θ all points laying between the two lines that

TABLE 1 Functions implemented within *transfreq*

Module 1: Operative functions		
Name	**Description**	**Input**
compute_transfreq	Computation of TF	rest PS
compute_transfreq_manual	Computation of TF (customised clusters)	rest PS; G_α; G_θ
compute_transfreq_klimesch	Computation of TF (Klimesch's method)	rest PS; task PS
compute_transfreq_minimum	Computation of TF (minimum method)	rest PS
Module 2: Visualisation functions		
Name	**Description**	**Input**
plot_psds	Normalised PS power spectrum	rest/task PS
plot_coefficients	α and θ coefficients or their ratio	rest/task PS
plot_clusters	Computed clusters	tfbox
plot_channels	G_α and G_θ on scalp	tfbox; channel locations
plot_transfreq	TF on top of S_α and S_θ	rest PS; tfbox
plot_transfreq_klimesch	TF on top of S_{rest} and S_{task}	rest and task PS; TF value
plot_transfreq_minimum	TF on top of S_{rest}	rest PS; TF value

Note: The table provides the name of each function (first column), a short description of their purpose (second column), and the required input variables (third column). Here, rest PS and task PS stand for resting state and task-related EEG power spectrum, respectively; tfbox is a dedicated dictionary output of the operative functions. For some of the functions, an additional set of optional arguments may be passed by the user, such as predefined alpha and theta frequency-band, or the clustering approach to be used for defining G_α and G_θ. The full list of these additional parameters may be found in the package documentation.

pass through the centroids and are perpendicular to the segment connecting the two centroids.

2.4 | Software architecture

The approach described in the previous section is implemented in the publicly available Python library *transfreq* (https://elisabettavallarino.github.io/transfreq/). As shown in Table 1, *transfreq* comprises two modules: a set of three operative functions, that allow the estimation of TF either with Klimesch's method or with our approach, and a set of six functions to visualise the results.

2.4.1 | Module 1: Operative functions

All the operative functions require in input the power spectra of the recorded EEG data. These power spectra have to be provided as matrices of size $N \times F$, where N is the number of EEG sensors and F is the number of frequencies in which the power spectra are evaluated.

The function *compute_transfreq* implements the iterative procedure described in Algorithm 1. Customised estimation of the transition frequency may be obtained through the function *compute_transfreq_manual* by providing two predefined groups of channels G_α and G_θ. In this case, TF is computed by looking at the intersections between the corresponding spectral profiles S_α and S_θ. Both functions return a dedicated dictionary, called *tfbox* in Table 1, that contains: (i) the results of the clustering procedure, together with the alpha and theta coefficients, w_i^α and w_i^θ, associated to each one of the sensors; (ii) the name of the employed algorithm; (iii) the estimated value of TF.

In order to provide an exhaustive toolbox for computing the theta-to-alpha TF we also implemented a function for the computation of TF with Klimesch's method and a function for the computation of TF with the minimum method. Such functions are named *compute_transfreq_Klimesch* and *compute_transfreq_minimum*, respectively, and only return the estimated value of TF.

2.4.2 | Module 2: Visualisation functions

As shown in Table 1, *transfreq* offers the users two functions to visualise features of the data provided in input, namely the normalised EEG power spectrum (function *plot_psds*) and the corresponding alpha and theta coefficients (function *plot_coefficients*).

Three other functions allow the user to visualise the results from each step of our approach, that is, (i) the alpha and theta coefficients grouped according to the results of the clustering procedure (function *plot_clusters*); (ii) the corresponding channels group G_α and G_θ located on top of topographical maps (function *plot_channels*); (iii) the final estimated value of TF on top of the spectral profiles S_α and S_θ (function *plot_transfreq*). The function *plot_channels* makes use of the Python package visbrain (Combrisson et al., 2019), and, in particular, we modified its function *TopoObj* to optimise it to our visualisation purpose.

Eventually, the functions *plot_transfreq_klimesch* and *plot_transfreq_minimum* are dedicated to plot the value of TF estimated using the Klimesch's method and the minimum method, respectively.

2.5 | Data

We validated *transfreq* by using two EEG data sets. The first one is an open-source data set, while the second one is an in-house data set we recorded in our lab. We used two different data sets to test the robustness of *transfreq* across data recorded in different experimental conditions.

2.5.1 | Open-source data set

This data set contains EEG data available at OpenNeuro (Christopher et al., 2021) at the accession number ds003490 (data set DOI doi:10.18112/openneuro.ds003490.v1.1.0). Data comprise both resting state and stimulus auditory oddball EEG recordings, sampled at 500 Hz, from 25 Parkinson's patients and 25 matched controls. For Parkinson's patients two sessions are available, while for healthy controls one session is available. More information about this data set can be found in the paper by Cavanagh and colleagues (Cavanagh et al., 2018). For each subject and for each session we selected 2 minutes of recording under stimulation, and 1 minute resting state eyes-closed recording measured through 59 EEG channels whose locations are depicted in Figure S1A, Supporting Information.

2.5.2 | In-house data set

This data set included 80 traces acquired in our centre during a previous multicenter study, namely the Innovative Medicines Initiative PharmaCog project: a European Alzheimer's Disease Neuroimaging Initiative (ADNI) study (Galluzzi et al., 2016). This study aimed at investigating multiple markers in a population with amnesic mild cognitive impairment (MCI), by following subjects for 3 years or until conversion to dementia. EEG was repeatedly acquired every 6 months; thus the 80 traces refer to 16 subjects undergoing EEG from one to 7 times. The 16 subjects (8 males, 8 females, age range 55–82 years, mean: 70 ± 6 years; mini-mental state examination [MMSE] score range at first evaluation: 23–30, mean: 26.5 ± 2.13) included 11 subjects who converted to Alzheimer dementia disease during the follow-up, 2 who convert to frontotemporal dementia, and 3 who remained in an MCI stage or even reverted to a normal condition.

For the analysis we selected two and a half minutes of resting state eyes-closed recording and two and a half minutes of resting state eyes-opened recording, where data showed a desynchronisation of the alpha rhythm (Gómez-Ramírez et al., 2017). Both data were recorded with a sampling frequency of 512 Hz by a 19-channel EEG cap schematically represented in in Figure S1B.

2.6 | Data analysis

The recorded time series from both data sets were first preprocessed using the MNE-Python analysis package (Gramfort et al., 2013). For each subject and for each condition, the EEG recording was filtered between 2 and 50 Hz, while bad segments were manually removed and bad channels were interpolated. Then, data were re-referenced using average reference (Offner, 1950) and independent component analysis (ICA) (Jutten & Herault, 1991) was applied for artefact and noise removal. Remaining bad segments were automatically rejected by using the autoreject Python package (Jas et al., 2017). Finally, the preprocessed EEG recordings were visually inspected by experts and discarded when they did not present a visible alpha peak (first exclusion phase). In this way, in the open-source data set we excluded the first session of four subjects and both sessions of one subject. In the in-house data set all sessions involving four subjects were excluded from the analysis.

Power spectra were computed in the 2–30 Hz range with the multitapers method (Thomson, 1982). With such a method the frequency resolution of the power spectra depends on the time resolution and duration of the EEG recordings. In order to apply Klimesch's method, the spectral profiles under the two conditions (rest and task) need to have the same frequency resolution. To this end the length of both recordings was set equal to the length of the shortest one. Average duration of the EEG recordings from the open source data set was 58 s, while average duration of the EEG recordings from the in-house data set was 134 s. Afterwards, TF was computed using both Klimesch's method and *transfreq*. Finally, the results obtained with Klimesch's method were visually inspected by experts and excluded when the method did not provide reliable results (second exclusion phase). Exclusion criteria comprised cases in which the two spectral profiles did not intersect as well as cases in which the two spectral profiles overlapped. This process led to the exclusion of 19 EEG recordings from the open-source data set and 14 EEG recordings from the in-house data set. Therefore, the analysis to validate *transfreq* was performed on a total of 50 EEG recordings from the open-source data set and 45 from the in-house data set.

In the in-house data set each EEG recording was associated with the result of the MMSE performed at time of the recording. To investigate the potential benefit of using TF estimated by *transfreq* as a clinical marker, we considered only the subjects who converted to Alzheimer dementia (7 subjects involved in a total of 21 sessions) and we tested for the presence of a significant correlation between the estimated TF and (i) the individual alpha peak and (ii) the corresponding MMSE score. As a comparison, the same correlation analysis was carried on by considering TF estimated by Klimesch's and minimum method, as well as other widely used individual qEEG features, namely mean frequency (MEF) and median frequency (MDF) (Accornero et al., 2014; Benz et al., 2014; Chotas et al., 1979; Coben et al., 1983; Tonner & Bein, 2006), defined as

$$\mathrm{MEF} = \frac{\sum_f f S_{\mathrm{rest}}(f)}{\sum_f S_{\mathrm{rest}}(f)} \quad \text{and} \quad \mathrm{MDF} = \underset{f}{\operatorname{argmin}} \left\{ \left| \sum_{f' < f} S_{\mathrm{rest}}(f') - \frac{1}{2} \sum_f S_{\mathrm{rest}}(f) \right| \right\}, \tag{5}$$

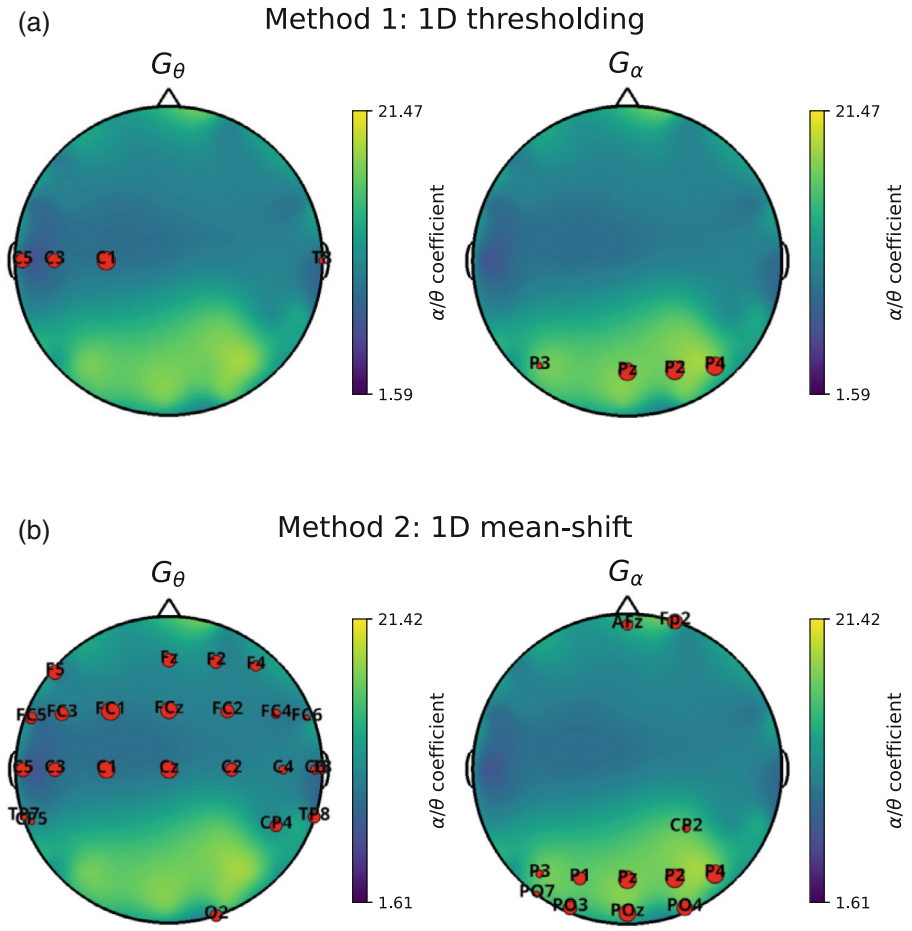


FIGURE 4 Location on the scalp of channels in G_θ (left column) and G_α (right column) for one representative subject from the open-source data set. Sensors have been clustered by using 1D thresholding (upper row) or 1D mean-shift (lower row). In each panel, red dots represent the selected channels and, in the background, the topographical map shows the value of the ratio between alpha and theta coefficients. For the sensors in G_θ, the size of the dots is proportional to the corresponding theta coefficient, w_i^θ, while for those in G_α the size is proportional to the alpha coefficient, w_i^α

$S_{rest}(f)$ being the averaged normalised power spectrum as in Equation (2).

3 | RESULTS

3.1 | Transfreq performances on an illustrative example

We first tested the performances of *transfreq* when applied to an illustrative example picked up from the open-source data set. Figures 4 and 5 show the results provided by the tool when the four different clustering algorithms were applied. For all algorithms, the resulting G_α mainly contained channels that lie over the occipital lobe and showed a higher alpha activity than the channels in G_θ.

When 1D thresholding is used as clustering method, both G_α and G_θ contain a predefined number of sensors (4 in this case). Instead, the other methods automatically estimate the size of G_θ and G_α, and thus the two groups may contain a different number of channels.

While with the clustering method 2D k-means G_α and G_θ span all the EEG channels, the 2D adjusted k-means starts from the two groups defined by using k-means and selects only the channels showing a high inter-cluster difference. Specifically, as illustrated in Figure 5, the channels in G_α (G_θ) showed both a high alpha (theta) activity and a low theta (alpha) activity.

Depending on the selected clustering approach, *transfreq* may return different estimates for TF, as illustrated in Figure 6. With this subject, the value of TF estimated by Klimesch's method was 7.29 Hz, while transfreq returned 7.38 Hz with 1D thresholding, 7.39 Hz with 1D mean-shift, 7.22 Hz with 2D k-means, and 7.19 Hz with 2D adjusted k-means. Finally, the minimum method estimates a value of TF equal to 6.22 Hz.

3.2 | Validation on the open-source data set

As illustrated in Figure 7(a), for most subjects in the open-source data set the difference Δ_{TF} between TF value estimated by *transfreq* and by Klimesch's method was in absolute value below 1 Hz. Specifically, $|\Delta_{TF}|$ was lower than 1 Hz for 82% of the subjects when 1D thresholding was employed for clustering, 76% in the case of 1D mean-shift, 82% for 2D k-means, and 88% for 2D adjusted k-means. Instead, when the minimum method was used the averaged value of $|\Delta_{TF}|$ was around 1.67, and only for 40% of the subject was below 1 Hz (Figure 7(b)). Figure 7(a) also shows that *transfreq* mainly estimated a lower value of TF than Klimesch's method.

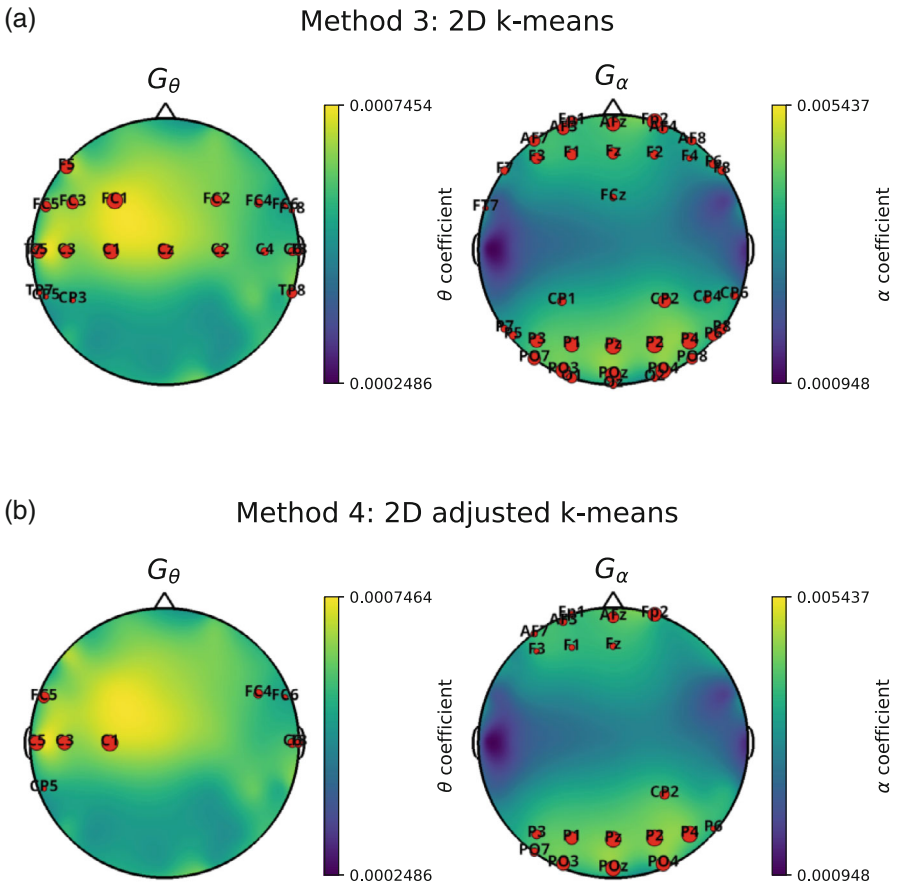


FIGURE 5 Location on the scalp of channels in G_θ (left column) and G_α (right column) for one representative subject from the open-source data set. Sensors have been clustered by using 2D k-means (upper row) or 2D adjusted k-means (lower row). As in Figure 4, the red dots depict the selected channels. In the two panels on the left side, referring to G_θ, the size of the sensors and the background topographical maps represent the theta coefficient, w_i^θ. Instead, the two panels on the right side, referring to G_α, show the alpha coefficients w_i^α

Since the lowest values of $|\Delta_{TF}|$ were obtained by clustering the EEG channels by means of the 2D adjusted k-means method, this is suggested as the default approach within *transfreq*.

3.3 | Improvements of transfreq over Klimesch's method

Klimesch's method relies on an event-related reduction of the alpha activity that may not occur in practical scenarios due, for example, to an incorrect execution of the task. Indeed, as shown in Figure 8, for some of the subjects in the considered data sets the spectral profiles S_{task} and S_{rest} overlapped and thus Klimesch's method failed in computing TF.

On the other hand, some subjects may show an event-related modulation of the alpha frequency (Haegens et al., 2014). As represented in Figure 9, in this case the shift of the alpha peak in S_{task} prevented the use of Klimesch's method because the two spectral profiles S_{task} and S_{rest} did not intersect in the range [0, 10] Hz.

Transfreq overcomes such limitations of Klimesch's method, since it utilises just resting state recordings, and relies on the selection of specific channels that actually present the desired features, i.e. channels with a low (high) alpha and a high (low) theta activity for G_θ (G_α). Indeed, as shown in the right panel of Figures 8 and 9 for the 2D adjusted k-means and in the Figures S2 and S3 for the remaining clustering approaches, in both previously described scenarios *transfreq* estimated a reliable value for TF. More in general, a visual inspection of the results revealed that Klimesch's method provided an untrustworthy value of TF for 27% of the EEG sessions of the open-source data set, while with *transfreq* only 6% of the results were unreliable. These percentages have been computed by considering all the EEG recordings available after the first exclusion phase described in section 2.6 and by counting the recordings fulfilling the criteria for the second exclusion phase. In other words, a TF estimates was considered as untrustworthy if S_{rest} and S_{task}, in Klimesch's method, or S_θ and S_α, in *transfreq*, overlapped or did not intersect in a reasonable frequency range.

3.4 | Validation on the in-house data set

Figure 10(a) shows that the results obtained by applying *transfreq* on the in-house data set are similar to those obtained on the open-source one. Specifically, *transfreq* generally returned higher estimates of TF with respect to Klimesch's method. However, the absolute value of the difference between the values estimated with the two methods was below 1 Hz for 67% of the subject when 1D thresholding was applied, 58% with 1D mean-shift, 73% with 2D k-means, and 62% with 2D adjusted k-means. As can be seen from Figure 10(b), the value of the difference between TF estimated with Klimesch's and the minimum method is on average around 1.87, and only for 27% of the subject is below 1 Hz (Figure 7(b)).

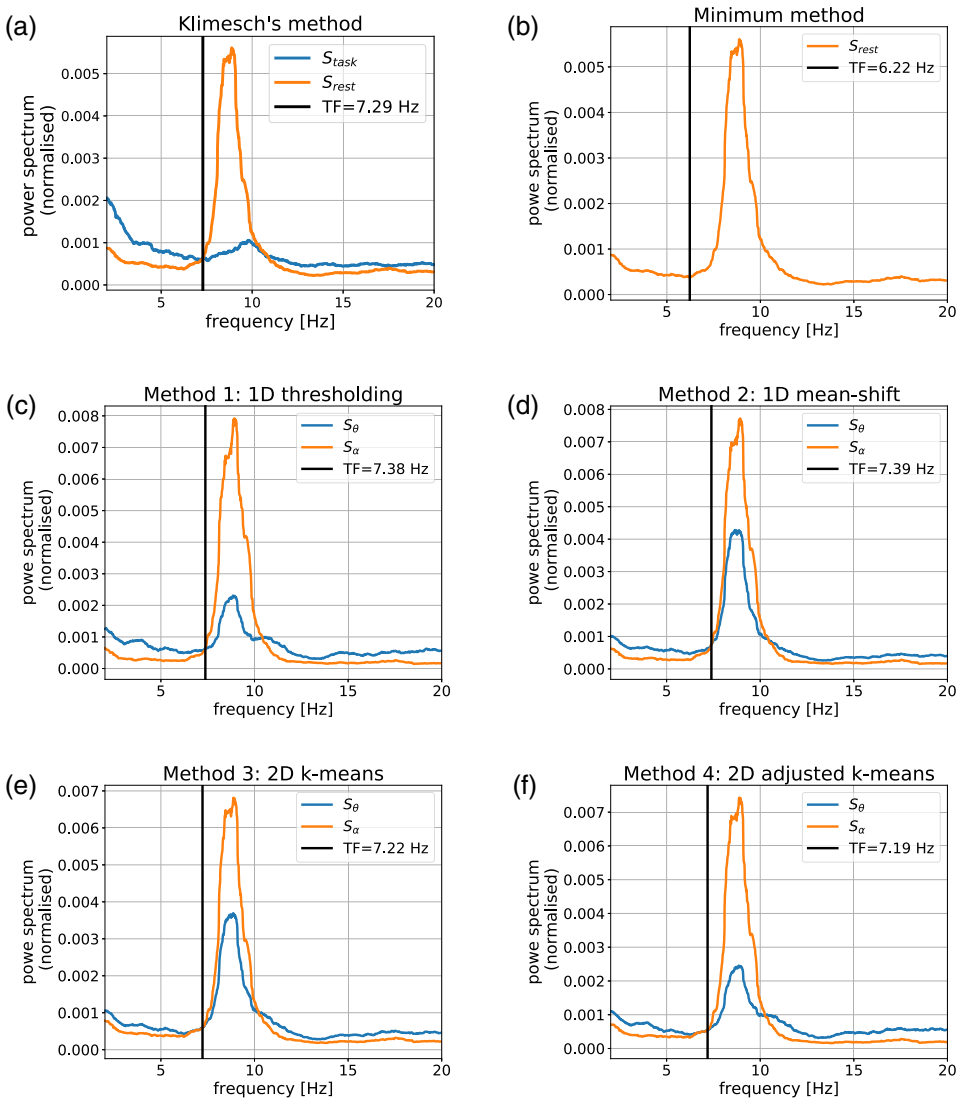


FIGURE 6 TFs estimated with Klimesch's method and with *transfreq* by means of the four clustering methods for one representative subject from the open-source data set. In each panel: The blue line depicts the spectral profile with low alpha and high theta activation, namely S_{task} in Klimesch's method and S_θ in *transfreq*; the orange line shows the spectral profile with high alpha and low theta activation, namely S_{rest} in Klimesch's method and S_α in *transfreq*; the red vertical line indicates the estimated value of TF

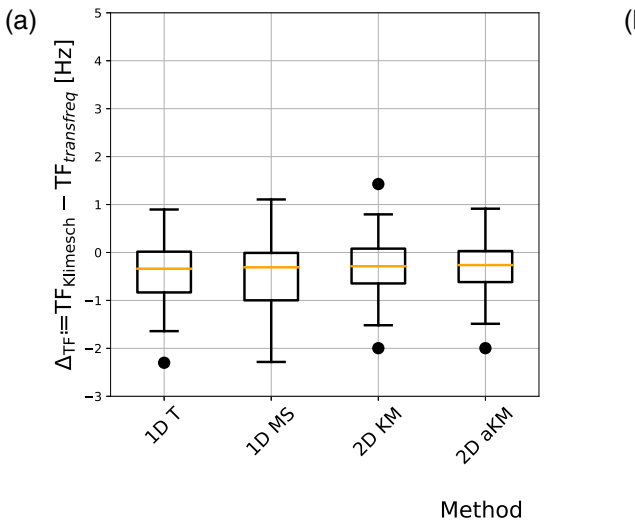


FIGURE 7 (a) Difference between TFs estimated with Klimesch's method ($TF_{Klimesch}$) and with *transfreq* ($TF_{transfreq}$) over the open-source data set. Each boxplot depicts the results obtained when a different clustering approach is used to define the channels group G_θ and G_α, namely: 1D thresholding (method 1D T); 1D mean-shift (method 1D MS); 2D k-means (method 2D KM); and 2D adjusted k-means (method 2D aKM). (b) Difference between TFs estimated with Klimesch's method ($TF_{Klimesch}$) and with minimum method ($TF_{minimum}$) over the open-source data set

3.5 | Proportional bias in estimating TF

We performed a Bland–Altman analysis (Bland & Altman, 1986) to assess proportional bias in the estimates of TF. Figure 11 shows the analysis for the open-source (panel a) and the in-house (panel b) data sets, computed on TF values provided by *transfreq* with adjusted k-means. With the open source data set no proportional bias was present; to confirm this, we computed a regression line and the *p*-value (null hypothesis: slope equal to zero). Differently, the results with the in-house data set showed a statistically significant ($p < .001$) proportional bias. Specifically, Figure 11(b) shows that *transfreq* tends to overestimate TF at higher frequencies (>8 Hz).

3.6 | Computational cost

To investigate whether the computational cost of *transfreq* increases when a 2D clustering approach is used instead of a 1D method, for

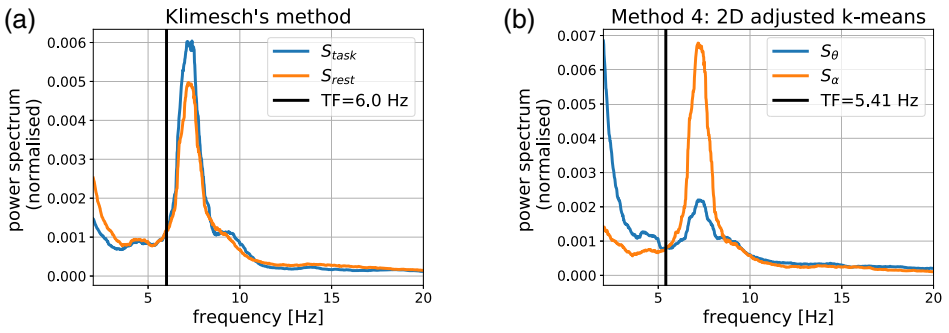


FIGURE 8 Example where Klimesch's method provides unreliable estimate of TF because event-related, S_{task}, and resting state, S_{rest}, spectral profiles overlap. (a) Results obtained with Klimesch's method. (b) Results obtained with *transfreq* by using 2D adjusted k-means to compute the spectral profiles S_θ and S_α

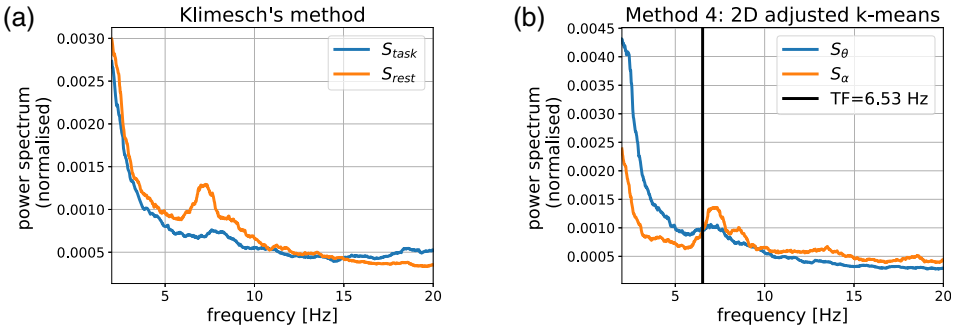


FIGURE 9 Example where Klimesch's method cannot be applied because event-related, S_{task}, and resting state, S_{rest}, spectral profiles do not intersect in a reasonable frequency range. (a) Results obtained with Klimesch's method. An event-related shift of the alpha–peak towards higher frequency can be seen in S_{task}. (b) Results obtained with *transfreq* by using 2D adjusted k-means to compute the spectral profiles S_θ and S_α

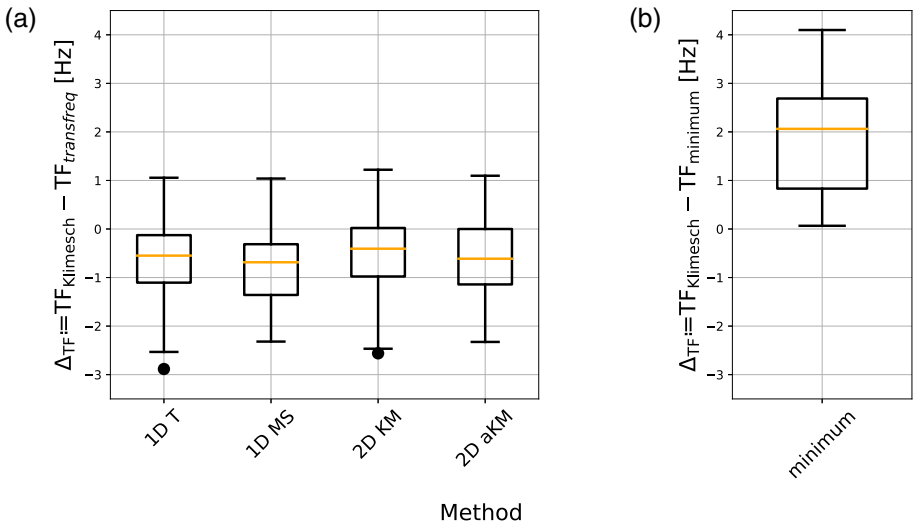


FIGURE 10 (a) Difference between TFs estimated with Klimesch's method ($TF_{Klimesch}$) and with *transfreq* ($TF_{transfreq}$) over the in-house data set. As in Figure 7 each boxplot depicts the results obtained when a different clustering approach is used to define G_θ and G_α, namely: 1D thresholding (method 1D T); 1D mean-shift (method 1D MS); 2D k-means (method 2D KM); and 2D adjusted k-means (method 2D aKM). (b) Difference between TFs estimated with Klimesch's method ($TF_{Klimesch}$) and with minimum method ($TF_{minimum}$) over the in-house data set

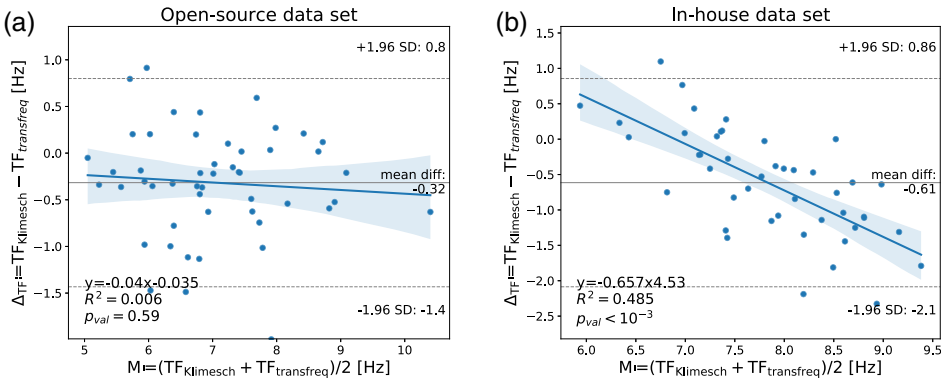


FIGURE 11 Bland–Altman plot between Klimesch's method and *transfreq* with 2D adjusted k-means for the open-source (a) and the in-house (b) data sets. Grey plain and dotted lines show mean bias and corresponding 95% confidence limits, respectively. Proportional bias regression lines are depicted as blue lines, and the corresponding equations are embedded in the lower-left corner of each panel together with the coefficient of determination (R^2) and the *p*-value (*p*) computed testing the null hypothesis that the slope is equal to zero

TABLE 2 Computational cost for all clustering methods proposed within *transfreq* and for both the open-source and in-house data sets, by adding over all recording sessions

	1D thresholding	1D mean-shift	2D k-means	2D adjusted k-means
Open-source data set	0.72	27.38	4.02	3.16
In-house data set	0.82	7.54	3.38	2.19

Note: Time is expressed in seconds.

each implemented method, we computed the overall elapsed time required by *transfreq* to estimate TF from all the data in the two data sets. As can be seen from Table 2, 1D thresholding, being the simplest algorithm, takes only few tenths of seconds. When 2D k-means and 2D adjusted k-means are used the computational cost slightly increases, however in both data sets *transfreq* took less than 5 s to run. A further analysis not shown here demonstrated that the lower computational cost required by the 2D adjusted k-means approach is due to the fact that a lower number of iterations were required for *transfreq* to converge. 1D mean-shift was the most time consuming method reflecting the fact that this is the only method that also estimates the number of clusters used to group the sensors. The differences among the two data sets may be explained by the different number of recording sessions they contain and by the different number of employed EEG sensors.

As a final remark, it is worth mentioning that all methods are fast to run, especially in comparison with the time required by other steps of the analysis, including the preprocessing of EEG time series and the power spectrum computation.

3.7 | Predictive power of *transfreq* TF

As can be seen from Figure 12, the values of TF estimated with *transfreq* equipped with the 2D adjusted k-means clustering approach showed a stronger correlation with the IAP (Sperman's correlation coefficient $\rho = 0.799$, $p < 10^{-4}$) than the values estimated with Klimesch's method ($\rho = 0.633$, $p < .01$). When the minimum method is employed, the values of TF do not show a statistically significant correlation with the IAP ($\rho = 0.266$, $p = .243$) due to the presence of few outliers having an high value of IAP but a low value of TF. Among MEF and MDF, only the latter shows some correlation with the IAP ($\rho = 0.523$, $p < .05$).

These results translate in a stronger predictive power of TF when *transfreq* is used instead of the other classic approaches. Indeed, as shown in Figure 13, the values of TF positively correlate with the MMSE score ($\rho = 0.668$, $p < .001$), while the correlation is weaker when Klimesch's approach is used ($\rho = 0.42$, $p = .058$) and becomes negative with the minimum methods ($\rho = -0.372$, $p = .097$). More interestingly, *transfreq* TF also outperformed some widespread individual spectral features, namely the IAP, the MEF and the MDF, the last two showing a negative or a close-to-zero correlation. The poor performances of MEF and MDF arise from two adversely effects impacting the values of these features, namely the slowing of the IAP which induces a decrease of the values of MEF and MDF for increasing MMSE scores, and the lowering of the α-power which instead induces higher MEF and MDF values for increasing MMSE score.

A further analysis shown in Figures S4 and S5 revealed that among the clustering approaches implemented in *transfreq*, all methods showed a statistically significant correlation with both IAP ($\rho > 0.75$, $p < 10^{-4}$) and MMSE score ($\rho > 0.5$, $p < .05$). More specifically, the best and the worst performances were obtained with the 2D adjusted k-means and the 1D mean-shift approaches, respectively.

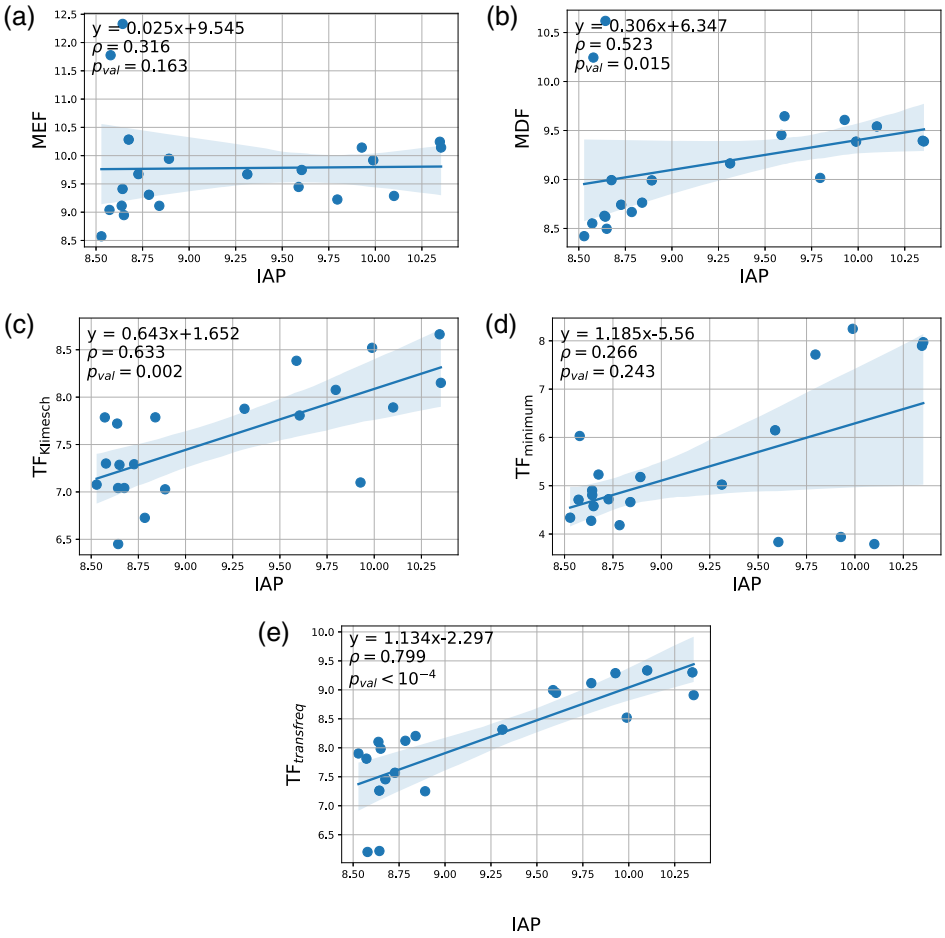


FIGURE 12 Correlation analysis between the individual alpha peak (IAP) and various qEEG features, namely mean frequency (a) median frequency (b) and TF estimated with the different methods, that is, Klimesch's approach (c), the minimum method (d), and *transfreq* with 2D adjusted k-means (e). For this analysis only the subjects who converted to Alzheimer in the in-house data set have been considered. In each panel, the regression line equations are embedded in the top-right corner together with the spearman correlation coefficient (ρ) and its *p*-value (p). The results obtained when the other clustering procedures are used within *transfreq* can be found in Figure S4

4 | DISCUSSION

A classic approach to compute the theta-to-alpha TF is that proposed by Klimesch and colleagues (Klimesch, 1999), which requires the power spectrum of two EEG time series, one recorded while the subject is resting and one while the subject is performing a task. However, in studies involving, for example, patients affected by neurodegenerative diseases, the subject may experiment difficulties in performing the required task and thus the corresponding event-related recording may imply difficult interpretation. On the contrary and similarly to the minimum method (Babiloni et al., 2016; Moretti et al., 2004; Moretti et al., 2007), another classic approach often used in clinical setting, *transfreq* uses only resting state data, which reduces the information at disposal but increases the scenario in which *transfreq* can be applied.

By comparing with the results obtained with Klimesch's method on two independent data sets, we demonstrated that *transfreq* returns reliable estimates of TF. Indeed, with the best combination of input parameters, the absolute value of the difference between the value of TF estimated with *transfreq* and with Klimesch's method was below 1 Hz for 88% of the analysed data in the open-source data set, and for 73% for our in-house data set (throughout this article Klimesch's method was assumed as ground truth). For the minimum method this percentages reduced to 40 and 27%, respectively. The differences in the performance between the two data sets may be partially due to the noisier nature of the in-house data set and to the different number of EEG sensors in the two experiments. Indeed, in the open-source data set 59 EEG sensors were used, while data of the in-house data set were recorded by means of a 19-channel EEG cap. Moreover, a visual inspection of the estimated values of TF showed that the cases in which the spectral profiles, S_θ and S_α, obtained with *transfreq* intersected ambiguously were considerably less than the cases in which the hypothesis of Klimesch's method on S_{task} and S_{rest} failed.

Among the four approaches implemented in transfreq to accomplish the clustering step, 2D adjusted k-means showed the best performances in the open-source data set while in the in-house data set 2D k-means algorithm performed the best. This is probably due to the fact that these algorithms realise a more accurate selection of the sensors within the two groups G_θ and G_α. Additionally, 1D thresholding always outperformed 1D mean-shift, and it almost matched the results of 2D adjusted k-means in the in-house data. Regarding the computational cost, as can be expected, 1D mean-shift resulted to be the fastest approach, requiring less than 1 s for analysing both the data sets. Instead, 2D k-means and 2D adjusted k-means required between 2 and 4 s depending on the data set, but have the advantage to be fully automated while 1D mean-shift requires the user to set the number of sensors in the two groups G_θ and G_α.

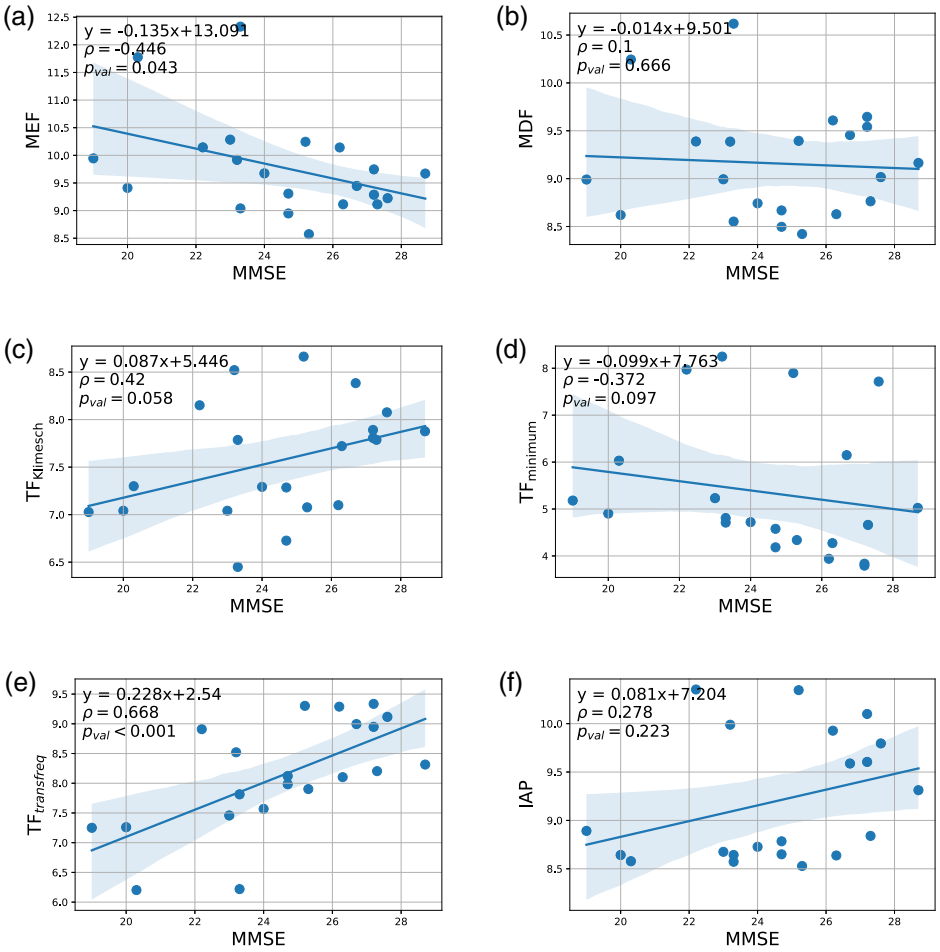


FIGURE 13 Correlation analysis between the mini-mental state examination (MMSE) score and various qEEG features, namely mean frequency (a) median frequency (b), TF estimated with the different methods, i.e., Klimesch's approach (c), the minimum method (d), and *transfreq* with 2D adjusted k-means (e), and the individual alpha peak (IAP). As in Figure 12, for this analysis only the subjects who converted to Alzheimer in the in-house data set have been considered. In each panel, the regression line equations are embedded in the top-right corner together with the Spearman correlation coefficient (ρ) and its p-value (p). The results obtained when the other clustering procedures are used within *transfreq* can be found in Figure S5

Motivated by these results, we suggest to used 2D adjusted k-means as the default method within *transfreq*, however we recommend running also the 2D k-mean and check for consistency between the two approaches when analysing data from a low-density EEG cap.

All four approaches tend to overestimate the value of TF with respect to Klimesch's method, that in turn usually provides higher estimates of TF than the minimum approach. Specifically, the Band-Altman analysis for the in-house data set show that this behaviour seems to be more pronounced for higher values of TF (>8 Hz). This difference between *transfreq* and Klimesch's method is probably related to the fact that only resting state data are used in *transfreq*; as a consequence also channels in G_θ may present a fingerprint of the alpha activity.

While individual qEEG features such as IAP, median and mean frequency have been extensively and successfully used as clinical markers of different neurodegenerative diseases (Benz et al., 2014; Coben et al., 1983; Dierks et al., 1991; Kwak, 2006; Petit et al., 2004), so far very few studies have used TF in clinical scenarios (Babiloni et al., 2006; Babiloni et al., 2016; Moretti et al., 2004; Saad et al., 2018). Additionally, most of these studies use the minimum method whose estimates of TF are usually larger than those of *transfreq*. Therefore, such studies cannot be fully exploited to support the clinical validity of the proposed method and specific experiments need to be designed in order to validate the predictive power of *tranfreq*, especially in those scenarios where it provides different estimates than Klimesch's method. In this work we performed a first experiment in this direction, and we showed that in a group of subjects who converted to Alzheimer dementia, the values of TF estimated with *transfreq* showed a stronger correlation with the MMSE score than those estimated with both Klimesch's and minimum method; *transfreq* TF also outperformed IAP, mean and median frequency. Future effort will be devoted to confirming this result in a wider class of experiments.

The two data sets considered in this article are EEG data. Future studies may be devoted to investigate a possible extension to MEG data.

ACKNOWLEDGMENT

This research study was supported by Fondi per la Ricerca Corrente of the Italian Ministry of Health (MOH), and by grants from the MOH: 5x1000 2018/2019 and NeuroArtP3 (NET-2018-12366666). The study has been developed in teh frame of the 2018–2022 Department of Excellence DINOGMI of the Italin Ministry of University and Research (MIUR). The authors thank the anonymous reviewers for their valuable comments and suggestions.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest that could be perceived as prejudicing the impartiality of the research reported.

DATA AVAILABILITY STATEMENT

The Python library tranfreq can be freely downloaded from the github repository https://github.com/elisabettavallarino/transfreq or through the package manager pip as described in documentation of transfreq (https://elisabettavallarino.github.io/transfreq/). The open-source data set used to test the method is publicly available on OpenNeuro, accession number ds003490, data set DOI doi:10.18112/openneuro.ds003490.v1.1.0. The package documentation also provides the scripts for reproducing the analysis showed in the article on the open-source data set (https://elisabettavallarino.github.io/transfreq/auto_paper/index.html). The raw EEG traces from the in-house data set cannot be made publicly available due to privacy restrictions related to clinical data. However the preprocessed data would be provided upon reasonable request to the corresponding author (vallarino@dima.unige.it).

ORCID

Elisabetta Vallarino https://orcid.org/0000-0002-4247-6192
Sara Sommariva https://orcid.org/0000-0003-1835-9399
Francesco Famà https://orcid.org/0000-0002-9084-9037
Michele Piana https://orcid.org/0000-0003-1700-991X
Flavio Nobili https://orcid.org/0000-0001-9811-0897
Dario Arnaldi https://orcid.org/0000-0001-6823-6069

REFERENCES

Accornero, N., Capozza, M., Pieroni, L., Pro, S., Davì, L., & Mecarelli, O. (2014). EEG mean frequency changes in healthy subjects during prefrontal transcranial direct current stimulation. *Journal of Neurophysiology*, *112*(6), 1367–1375.

Babiloni, C., Binetti, G., Cassarino, A., Dal Forno, G., Del Percio, C., Ferreri, F., Frisoni, G., Galderisi, S., Hirata, K., Lanuzza, B., Miniussi, C., Mucci, A., Nobili, F., Rodriguez, G., Romani, G. L., & Rossini, P. M. (2006). Sources of cortical rhythms in adults during physiological aging: A multicentric EEG study. *Human Brain Mapping*, *27*(2), 162–172.

Babiloni, C., Del Percio, C., Capotosto, P., Noce, G., Infarinato, F., Muratori, C., Marcotulli, C., Bellagamba, G., Righi, E., Soricelli, A., Onorati, P., & Lupattelli, T. (2016). Cortical sources of resting state electroencephalographic rhythms differ in relapsing–remitting and secondary progressive multiple sclerosis. *Clinical Neurophysiology*, *127*(1), 581–590.

Babiloni, C., Miniussi, C., Babiloni, F., Carducci, F., Cincotti, F., Del Percio, C., Sirello, G., Fracassi, C., Nobre, A. C., & Rossini, P. M. (2004). Sub-second "temporal attention" modulates alpha rhythms. A high-resolution EEG study. *Cognitive Brain Research*, *19*(3), 259–268.

Bazanova, O., & Vernon, D. (2014). Interpreting EEG alpha activity. *Neuroscience & Biobehavioral Reviews*, *44*, 94–110.

Bendat, J. S., & Piersol, A. G. (2011). *Random data: Analysis and measurement procedures* (Vol. 729). John Wiley & Sons.

Benz, N., Hatz, F., Bousleiman, H., Ehrensperger, M. M., Gschwandtner, U., Hardmeier, M., Ruegg, S., Schindler, C., Zimmermann, R., Monsch, A. U., & Fuhr, P. (2014). Slowing of EEG background activity in Parkinson's and Alzheimer's disease with early cognitive dysfunction. *Frontiers in Aging Neuroscience*, *6*, 314.

Bland, J. M., & Altman, D. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*, *327*(8476), 307–310.

Cavanagh, J. F., Kumar, P., Mueller, A. A., Richardson, S. P., & Mueen, A. (2018). Diminished EEG habituation to novel events effectively classifies Parkinson's patients. *Clinical Neurophysiology*, *129*(2), 409–418.

Chotas, H., Bourne, J., & Teschan, P. (1979). Heuristic techniques in the quantification of the electroencephalogram in renal failure. *Computers and Biomedical Research*, *12*(4), 299–312.

Christopher, J. M., Krzysztof, J. G., Franklin, F., Ross, B., Yaroslav, O. H., Eric, M., Nell, H., Joe, W.,Oscar, E., Mathias, G., Anita, J., Russell, P. (2021). The OpenNeuro resource for sharing of Neuroscience data. *Elife*, *10*, e71774.

Coben, L. A., Danziger, W. L., & Berg, L. (1983). Frequency analysis of the resting awake EEG in mild senile dementia of Alzheimer type. *Electroencephalography and Clinical Neurophysiology*, *55*(4), 372–380.

Comaniciu, D., & Meer, P. (2002). Mean shift: A robust approach toward feature space analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *24*(5), 603–619.

Combrisson, E., Vallat, R., O'Reilly, C., Jas, M., Pascarella, A., Saive, A.-L., Thiery, T., Meunier, D., Altukhov, D., Lajnef, T., Ruby, P., Guillot, A., & Jerbi, K. (2019). Visbrain: a multi-purpose GPU-accelerated open-source suite for multimodal brain data visualization. *Frontiers in Neuroinformatics*, *13*, 14.

Dierks, T., Perisic, I., Frölich, L., Ihl, R., & Maurer, K. (1991). Topography of the quantitative electroencephalogram in dementia of the Alzheimer type: Relation to severity of dementia. *Psychiatry Research: Neuroimaging*, *40*(3), 181–194.

Donoghue, T., Haller, M., Peterson, E. J., Varma, P., Sebastian, P., Gao, R., Noto, T., Lara, A. H., Wallis, J. D., Knight, R. T., Shestyuk, A., & Voytek, B. (2020). Parameterizing neural power spectra into periodic and aperiodic components. *Nature Neuroscience*, *23*(12), 1655–1665.

Galluzzi, S., Marizzoni, M., Babiloni, C., Albani, D., Antelmi, L., Bagnoli, C., Bartres-Faz, D., Cordone, S., Didic, M., Farotti, L., Fiedler, U., Forloni, G., Girtler, N., Hensch, T., Jovicich, J., Leeuwis, A., Marra, C., Molinuevo, J. L., Nobili, F., … the PharmaCog Consortium. (2016). Clinical and biomarker profiling of prodromal Alzheimer's disease in workpackage 5 of the innovative medicines initiative PharmaCog project: A "European ADNI study". *Journal of Internal Medicine*, *279*(6), 576–591.

Gómez-Ramírez, J., Freedman, S., Mateos, D., Velázquez, J. L. P., & Valiante, T. A. (2017). Exploring the alpha desynchronization hypothesis in resting state networks with intracranial electroencephalography and wiring cost estimates. *Scientific Reports*, *7*(1), 1–11.

Gramfort, A., Luessi, M., Larson, E., Engemann, D. A., Strohmeier, D., Brodbeck, C., Goj, R., Jas, M., Brooks, T., Parkkonen, L., & Hämäläinen, M. (2013). MEG and EEG data analysis with MNE-Python. *Frontiers in Neuroscience*, *7*, 267.

Haegens, S., Cousijn, H., Wallis, G., Harrison, P. J., & Nobre, A. C. (2014). Inter- and intra- individual variability in alpha peak frequency. *NeuroImage*, *92*, 46–55.

Jaramillo-Jimenez, A., Suarez-Revelo, J. X., Ochoa-Gomez, J. F., Arroyave, J. A. C., Bocanegra, Y., Lopera, F., Buriticá, O., Pineda-Salazar, D. A., Gómez, L. M., Quintero, C. A. T., Borda, M. G., Bonanni, L., Ffytche, D. H., Brønnick, K., & Aarsland, D. (2021). Resting-state EEG alpha/theta ratio related to neuropsychological test performance in Parkinson's disease. *Clinical Neurophysiology*, *132*(3), 756–764.

Jas, M., Engemann, D. A., Bekhti, Y., Raimondo, F., & Gramfort, A. (2017). Autoreject: Automated artifact rejection for MEG and EEG data. *NeuroImage*, *159*, 417–429.

Jutten, C., & Herault, J. (1991). Blind separation of sources, part I: An adaptive algorithm based on neuromimetic architecture. *Signal Processing*, *24*(1), 1–10.

Klassen, B., Hentz, J., Shill, H., Driver-Dunckley, E., Evidente, V., Sabbagh, M., Adler, C. H., & Caviness, J. N. (2011). Quantitative EEG

as a predictive biomarker for Parkinson disease dementia. *Neurology*, *77*(2), 118–124.

Klimesch, W. (1996). Memory processes, brain oscillations and EEG synchronization. *International Journal of Psychophysiology*, *24*(1–2), 61–100.

Klimesch, W. (1999). EEG alpha and theta oscillations reflect cognitive and memory performance: A review and analysis. *Brain Research Reviews*, *29*(2–3), 169–195.

Klimesch, W., Doppelmayr, M., Russegger, H., Pachinger, T., & Schwaiger, J. (1998). Induced alpha band power changes in the human EEG and attention. *Neuroscience Letters*, *244*(2), 73–76.

Klimesch, W., Doppelmayr, M., Schimke, H., & Ripper, B. (1997). Theta synchronization and alpha desynchronization in a memory task. *Psychophysiology*, *34*(2), 169–176.

Kwak, Y. T. (2006). Quantitative EEG findings in different stages of Alzheimer's disease. *Journal of Clinical Neurophysiology*, *23*(5), 457–462.

Lansbergen, M. M., Arns, M., van Dongen-Boomsma, M., Spronk, D., & Buitelaar, J. K. (2011). The increase in theta/beta ratio on resting-state EEG in boys with attention-deficit/hyperactivity disorder is mediated by slow alpha peak frequency. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *35*(1), 47–52.

Lloyd, S. (1982). Least squares quantization in PCM. *IEEE Transactions on Information Theory*, *28*(2), 129–137.

Malek, N., Baker, M., Mann, C., & Greene, J. (2017). Electroencephalographic markers in dementia. *Acta Neurologica Scandinavica*, *135*(4), 388–393.

Moretti, D. V., Babiloni, C., Binetti, G., Cassetta, E., Dal Forno, G., Ferreric, F., Ferri, R., Lanuzza, B., Miniussi, C., Nobili, F., Rodriguez, G., Salinari, S., & Rossini, P. M. (2004). Individual analysis of EEG frequency and band power in mild Alzheimer's disease. *Clinical Neurophysiology*, *115*(2), 299–308.

Moretti, D. V., Miniussi, C., Frisoni, G., Zanetti, O., Binetti, G., Geroldi, C., Galluzzi, S., & Rossini, P. M. (2007). Vascular damage and EEG markers in subjects with mild cognitive impairment. *Clinical Neurophysiology*, *118*(8), 1866–1876.

Nunez, P. L., Wingeier, B. M., & Silberstein, R. B. (2001). Spatial-temporal structures of human alpha rhythms: Theory, microcurrent sources, multiscale measurements, and global binding of local networks. *Human Brain Mapping*, *13*(3), 125–164.

Offner, F. F. (1950). The EEG as potential mapping: The value of the average monopolar reference. *Electroencephalography and Clinical Neurophysiology*, *2*(2), 213–214.

Özbek, Y., Fide, E., & Yener, G. G. (2021). Resting-state EEG alpha/theta power ratio discriminates early-onset Alzheimer's disease from healthy controls. *Clinical Neurophysiology*, *132*(9), 2019–2031.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, *12*, 2825–2830.

Petit, D., Gagnon, J. F., Fantini, M. L., Ferini-Strambi, L., & Montplaisir, J. (2004). Sleep and quantitative EEG in neurodegenerative disorders. *Journal of Psychosomatic Research*, *56*(5), 487–496.

Poza, J., Hornero, R., Abásolo, D., Fernández, A., & García, M. (2007). Extraction of spectral based measures from MEG background oscillations in Alzheimer's disease. *Medical Engineering & Physics*, *29*(10), 1073–1083.

Saad, J. F., Kohn, M. R., Clarke, S., Lagopoulos, J., & Hermens, D. F. (2018). Is the theta/beta EEG marker for ADHD inherently flawed? *Journal of Attention Disorders*, *22*(9), 815–826.

Saxena, A., Prasad, M., Gupta, A., Bharill, N., Patel, O. P., Tiwari, A., Er, M. J., Ding, W., & Lin, C. T. (2017). A review of clustering techniques and developments. *Neurocomputing*, *267*, 664–681.

Schacter, D. L. (1977). EEG theta waves and psychological phenomena: A review and analysis. *Biological Psychology*, *5*(1), 47–82.

Soininen, H., Partanen, V., Helkala, E. L., & Riekkinen, P. (1982). EEG findings in senile dementia and normal aging. *Acta Neurologica Scandinavica*, *65*(1), 59–70.

Thomson, D. J. (1982). Spectrum estimation and harmonic analysis. *Proceedings of the IEEE*, *70*(9), 1055–1096.

Tonner, P., & Bein, B. (2006). Classic electroencephalographic parameters: Median frequency, spectral edge frequency, etc. *Best Practice & Research Clinical Anaesthesiology*, *20*(1), 147–159.

Vallarino, E., Sommariva, S., Piana, M., & Sorrentino, A. (2020). On the two-step estimation of the cross-power spectrum for dynamical linear inverse problems. *Inverse Problems*, *36*(4), 045010.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Vallarino, E., Sommariva, S., Famà, F., Piana, M., Nobili, F., & Arnaldi, D. (2022). Transfreq: A Python package for computing the theta-to-alpha transition frequency from resting state electroencephalographic data. *Human Brain Mapping*, *43*(17), 5095–5110. https://doi.org/10.1002/hbm.25995