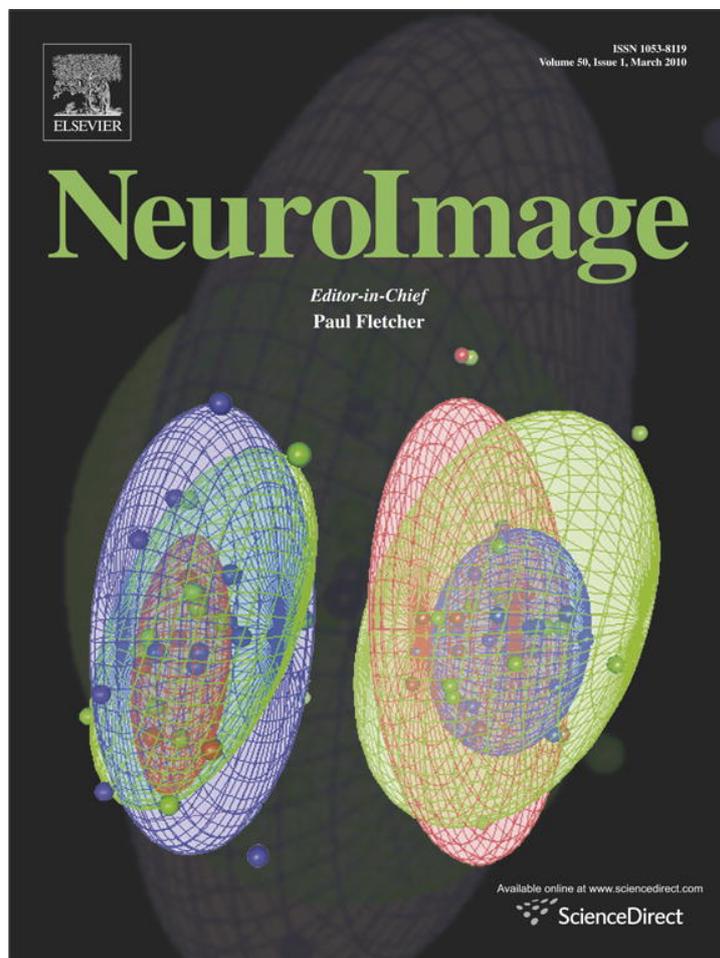


Provided for non-commercial research and education use.
Not for reproduction, distribution or commercial use.



This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

<http://www.elsevier.com/copyright>



Contents lists available at ScienceDirect

NeuroImage

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

Technical Note

Hierarchical control of false discovery rate for phase locking measures of EEG synchrony

Archana K. Singh*, Steven Phillips

Neuroscience Research Institute, National Institute of Advanced Industrial Science and Technology (AIST), Tsukuba Central 2, 1-1-1 Umezono, Tsukuba, Ibaraki, 305-8568, Japan

ARTICLE INFO

Article history:

Received 29 July 2009

Revised 25 October 2009

Accepted 7 December 2009

Available online 16 December 2009

Keywords:

Hierarchical false discover rate

Phase locking values

Synchrony

Electroencephalography

Visual search

ABSTRACT

Computing phase-locking values (PLVs) between EEG signals is becoming a popular measure for quantifying functional connectivity, because it affords a more detailed picture of the synchrony relationships between channels at different times and frequencies. However, the accompanying increase in data dimensionality incurs a serious multiple testing problem for determining PLV significance. Standard methods for controlling Type I error, which treat all hypotheses as belonging to a single family, can fail to detect any significant discoveries. Instead, we propose a novel application of a hierarchical FDR method, which subsumes multiple families, for detecting significant PLV effects. For simulations and experimental data, we show that the proposed hierarchical FDR method is most powerful. This method revealed significant synchrony effects in the expected regions at an acceptable error rate of 5%, where other methods, including standard FDR correction failed to reveal any significant effects.

© 2009 Elsevier Inc. All rights reserved.

Introduction

Investigating changes in functional connectivity associated with task variables raises two closely related issues: (1) determining a suitable measure of functional connectivity; and (2) assessing its reliability. For EEG, computing phase-locking values (PLVs) has become a popular measure for quantifying functional connectivity in terms of synchronization between brain regions (Lachaux et al., 1999). PLVs are computed by wavelet decomposition, providing instantaneous measures of phase differences between two signals at any desired frequency. This method has an advantage over methods based on correlation, covariance, or spectral coherence, in that PLV measures are directly applicable to non-stationary signals and treat phase and amplitude independently (see Lachaux et al., 1999).

Although PLV analysis affords a more detailed picture of the synchrony relationships between channels at different times and frequencies, the accompanying increase in data dimensionality incurs a serious multiple testing problem for assessing the significance of PLV effects. As the test statistics among neighboring frequency and time bands, and electrode pairs are likely to be dependent, standard Bonferroni correction is often too severe. Moreover, researchers are also interested in the direction of these effects, so Type III (direction) error along with Type I error may also be inflated.

There are several approaches to controlling false-positives arising from mass univariate statistical maps of neuroimaging data. Among

the Family-Wise Error Rate (FWER) controlling procedures, Random Field Theory (RFT, Worsley et al., 1992) and resampling-based maxT correction (Nichols and Holmes, 2002; Pantazis et al., 2005; Singh et al., 2008) are the most established methods. RFT method has been applied to the analysis of voxelwise EEG synchrony (see, e.g., Mizuhara and Yamaguchi, 2007), however it requires that EEG be done in conjunction with fMRI. Resampling-based methods have been used in testing EEG synchrony, but not for multiple testing correction (e.g., Lachaux et al., 1999). Another approach, pursued here, is to employ the widely used False Discovery Rate (FDR) control, which controls the proportion of false-positives to total number of detections (Benjamini and Hochberg, 1995). This approach differs from Bonferroni and other FWER methods, which control the probability of there being even a single false-positive. FDR is of more practical benefit to researchers who are willing to accept a small number of false-positives as the cost of increased power.

FDR is now a standard approach for controlling Type I errors in statistical maps obtained from fMRI, EEG/MEG, and fNIRS (Genovese et al., 2002; Chumbley and Friston, 2009; Hemmelmann et al., 2005; Singh and Dan, 2006). But, applications of FDR to maps of EEG synchrony are almost nonexistent. Hemmelmann et al. (2005) compared variations of FDR, FWER and other methods for coherence analysis, and concluded that FDR and FWER were also applicable to high dimensional EEG. For our purposes, though, FDR failed to detect significant PLV effects unless the significance threshold was relaxed to a far less conservative level (e.g., $q = 0.2$). In the absence of suitable methods, researchers have resorted to region of interest analysis, informed by prior studies in the literature, or visual inspection for maxima in the measures of interest (e.g., Phillips and Takeda, 2009).

* Corresponding author.

E-mail addresses: archana@ni.aist.go.jp, sine.arc@gmail.com (A.K. Singh).URL: <http://staff.aist.go.jp/archana.singh> (A.K. Singh).

The alternative has been to conduct descriptive analyses using uncorrected significance tests (Bhattacharya and Petsche, 2005; Fingelkurts et al., 2003; Razoumnikova, 2000; Rodriguez et al., 1999; von Stein et al., 1999; Weiss and Rappelsberger, 2000).

Standard (direct) FDR treats all hypotheses as belonging to a single family. For multi-dimensional data like EEG, or research questions with multiple facets, it is often more appropriate to group hypotheses into different families depending on the dimension or type of question. This problem arises in microarray analysis, which has been the primary motivation for recent advances in FDR methods. But, it also arises in EEG research, where the distribution of neural events are typically not uniform and are likely to be dependent on time and frequency. For example, synchrony within different frequency bands has been implicated in subserving different cognitive functions, such as top-down versus bottom-up driven shifts in visual attention (Buschman and Miller, 2007). So, the standard FDR (like Bonferroni), which treats all the tests as single family may fail to account for the underlying dependence structure, and tends to be overly conservative.

Recent advances in FDR methods accommodate multi-resolutional testing (Yekutieli et al., 2006; Reiner et al., 2007; Yekutieli, 2008). The advantage of this approach is similar to multi-stage testing as the large proportions of hypotheses that are unlikely to reveal significant effects are excluded from subsequent stages. Yekutieli et al. (2006) and Yekutieli (2008) proposed a hierarchical framework for FDR to estimate an overall FDR bound for the entire study while controlling FDR for each dimension separately at level q . If the number of discoveries for the entire study exceeds the number of dimensions, then the overall FDR bound for the entire study also converges to q . In this study, we apply hierarchical FDR to the analysis of PLVs computed from EEG data. We use this method in combination with an improvised FDR procedure that also controls for directional (Type III) errors (Benjamini and Yekutieli, 2005). The power and robustness of this method are evaluated using simulation, and verified with existing experimental data (Phillips and Takeda, 2009). Our main result is that hFDR reveals significant synchrony effects in the expected regions at a overall false discovery rate of 5%, while standard FDR failed to detect any significant effects at this level.

Materials and methods

Here, we summarize the methods used to control Type I and Type III errors associated with tests of PLV significance. Since our focus is on FDR control, details for computing PLV (Lachaux et al., 1999) and with specific application to visual search (Phillips and Takeda, 2009) are not repeated here. FDR is defined as the expected proportion of false-positives (V) among the detected-positives (R)¹ (see Table 1),

$$FDR = \begin{cases} E(V/R), & \text{if } R > 0 \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

and FWER as the probability of there being at least one false-positive,

$$FWER = P(V \geq 1). \quad (2)$$

Linear step-up FDR (FDRBH)

A linear step-up procedure (FDRBH, Benjamini and Hochberg, 1995) uses ordered p -values $P_1 < P_2 < \dots < P_m$ corresponding to hypotheses H_1, H_2, \dots, H_m . Control at a specified level q is achieved by rejecting hypotheses hypothesis H_1, \dots, H_k , where $k = \max(i : P_i \leq \frac{i}{m} q)$. If such a k does not exist, all hypotheses are accepted. This procedure controls FDR at the level q under the assumption of

Table 1

Variables associated with the number of true-negatives (TN), false-negatives (FN), false-positives (FP) and true-positives (TP), for multiple testing of m null hypotheses.

	Declared non-significant	Declared significant	Total
True H_0	U (TN)	V (FP)	m_0
False H_0	T (FN)	S (TP)	$m - m_0$
Total	$m-R$	R	m

independence or a certain form of positive dependence (Benjamini and Hochberg, 1995; Benjamini and Yekutieli, 2001); the later has been assumed to be the case with the test statistics obtained from most practical neuroimaging data (Genovese et al., 2002).

Mixed directional FDR (mixed FDR)

Testing multiple null hypotheses against two-sided alternatives potentially results in directional (Type III) errors that are in addition to Type I errors. For example, in simultaneous testing of null hypotheses $H_{0i} : \theta_i = 0$, against the alternatives, $H_{1i+} : \theta_i > 0$ or $H_{1i-} : \theta_i < 0$, for $i = 1, \dots, m$, researchers may falsely accept H_{1i+} (or, H_{1i-}) when in fact $\theta_i < 0$ (or, $\theta_i > 0$). Mixed directional FDR (mixed FDR) procedures control the expected proportion of both Type I and Type III errors (Benjamini and Yekutieli, 2005). The procedure for the assumption of independence of test statistic is as follows: apply FDRBH at level q to $H_{0i} : \theta_i = 0$ with two-sided p -values, P_i . Let R denote the number of discoveries made. Declare $\theta_i > 0$ (or, $\theta_i < 0$) if $P_i < \frac{Rq}{m}$ and $\theta_i > 0$ (or, $\theta_i < 0$).

The assumption of positive dependence as mentioned above is difficult to verify with two-sided test statistics. In this case, the authors have suggested separately testing each of the m one-sided hypotheses, as follows: apply FDRBH at level $q/2$ to test the m null hypotheses, $H_{0i+} : \theta_i \leq 0$ using their corresponding one-sided p -values, P_i , and null hypotheses, $H_{0i-} : \theta_i \geq 0$, with p -values, $1 - P_i$. Reject $R = R^+ \cup R^-$, where R^+ and R^- denote the rejected nulls from H_{0i+} and H_{0i-} , respectively. This alternative method for controlling Type III error is less powerful than the previous one, but it ensures that FDR is controlled separately for both positive and negative differences, even when the two-sided statistics are not positively dependent.

Hierarchical FDR (hFDR)

The previous methods assume that all hypotheses belong to a single family. Greater efficiency can be obtained by organizing hypotheses into a family-subfamily tree hierarchy, where each (sub)family is associated with a single hypothesis (hFDR Yekutieli et al., 2006; Yekutieli, 2008). We use our PLV example to illustrate the hierarchical approach. Each PLV test statistic is associated with an electrode pair in a given frequency and time dimension. Suppose the tests are grouped into M frequency and N time families based on the frequency and time band associated with the test. The M frequency families may constitute the first level of the hierarchy. In this case, there are N time subfamilies at the second level for each frequency family, and within each time subfamily are the test statistics, one for each electrode pair, at the third (lowest) level. Alternatively, time may be regarded as the first level, in which case, there are M frequency subfamilies at the second level for each time family. For each (sub)family there is an associated hypothesis. In our PLV example, a hypothesis associated with a first level frequency family corresponds to a significant effect at that frequency band. A hypothesis associated with a second level time family corresponds to there being a significant effect at that frequency and time band, and so on. If a null hypothesis is accepted then all descendants for the associated family are no longer considered for further analysis. Pruning the hypothesis tree in this way prevents control of FDR becoming overly conservative by

¹ FDR can be expressed equivalently as $FDR = \Pr(R > 0)E(V/R | R > 0) + \Pr(R = 0)E(V/R | R = 0) = E(V/\max(R, 1))$.

ignoring groups of hypotheses that are unlikely to be relevant. This process continues by recursively checking each child hypothesis of a parent whose null hypothesis was rejected, and terminates when no children are left to be tested.

Depending on the research question, hFDR procedure can be employed to control FDR for identifying discoveries at all levels (full-tree analysis), at a specific level (level-restricted analysis), or at outer nodes only (end-node analysis). In this article, our focus is on the illustration of the full-tree hFDR application, which can also be used in level-restricted analysis.

The FDR bound on a hypothesis tree is defined recursively as the sum of the expected proportion of the number of false discoveries to total discoveries for each family. Yekutieli (2008) derived an approximate bound using the following equation

$$\text{bound} = q\delta \frac{N_d + N_f}{N_d + 1} \quad (3)$$

where q is the threshold for controlling FDR in each single family, N_d is the number of observed discoveries, N_f is the number of families tested, and δ is a multiplicative constant, typically about 1, when the number of tests is not high. N_d (and N_f) may either represent all the discoveries (and families tested) in a full-tree analysis, or only the discoveries (and families tested) at the specified level k , in a level- k restricted analysis. For several hundreds of tests, δ with an upper bound of 1.4 is needed.

The value of hFDR bound depends on the distribution of the data, the FDR tree constructed, and the set of FDR tree discoveries that are of interest. If number of discoveries greatly exceeds the number of families tested ($N_d \gg N_f$), so that the multiplier $\frac{N_d + N_f}{N_d + 1} \approx 1$, then hFDR bound $\approx q$ in both full-tree and level-restricted analyses. Otherwise, hFDR bound may exceed q . See appendix for more on approximation of hFDR bounds (Appendix A). Any reasonable value of hFDR bound, even if it is higher than q , can be presented to support the hFDR inference. Alternatively, the q level used in hierarchical method can be made more stringent by substituting it with q^* to ensure that hFDR bound does not exceed q :

$$q^* = q \frac{N_d + 1}{N_d + N_f} \quad (4)$$

Simulation

Simulations were conducted to assess the viability of hFDR for PLV analysis, and to compare this method with standard FDR. We associated PLV differences (between conditions) with 4 frequency

bands, 12 time bands, 10 participants, and 171 electrode pairs (i.e., all unique pairwise combinations of 19 electrodes) constituting a $4 \times 12 \times 10 \times 171$ array, and q was set at 0.05.

As already mentioned, there is more than one way to construct a hierarchy. Ideally, families that are likely to be rejected should be placed higher in the hierarchy so that their member hypotheses are removed from further analysis. So, for example, if the true-positives are clustered within a particular frequency band, but distributed throughout all time bands, then a hierarchy with frequency at the first level and time at the second level is likely to be more efficient than a hierarchy with time at the first level and frequency at the second level. For similar reasons, within a level of the hierarchy, efficiency is likely to be greater when true-positives cluster within a single family, rather than partially distributed over many.

To assess the effect of dimension order and partitioning on hFDR sensitivity, we constructed a series of synchrony patterns and hypothesis trees. The synchrony patterns were subsets of the 48 ($=4 \times 12$) possible frequency–time bands (windows). The first four patterns were designed to assess the sensitivity of hFDR to dimension order and partitioning by defining windows of significance that were either within a single family, or spanned multiple families (see Fig. 1A to D). The fifth pattern was designed to test the effect of identifying multiple synchronies distributed throughout the families (see Fig. 1E). This situation may arise when a cognitive process invokes multiple phases of synchrony at different frequencies, and the experimenter does not have strong prior reasons for asserting particular regions of interest. For frequency–time windows containing significant effects (shaded regions), 12 out of 171 pairs were defined as truly significant. Synchronized pairs were assigned non-zero PLV differences, randomly generated from a normal distribution with parameters chosen from a real PLV dataset. Other pairs were assigned values from a standard normal distribution with mean zero and standard deviation one.

Hypothesis trees were constructed by systematically varying the association between (frequency, time) dimension and the (first, second) level of the tree, and the “width” of each family (i.e. the number of contiguous windows covered by a family). Trees are identified by a two-letter scheme, where the first/second letter corresponds to the first/second level of the tree and subscripts indicate the number of families at the associated level per parent family. So, for example, a $T_{12}F_4$ tree indicates 12 time families at the first level each having 4 frequency families at the second level. Families were arranged contiguously to cover the entire dimension. So a T_6F_2 tree means that there were six time families each covering

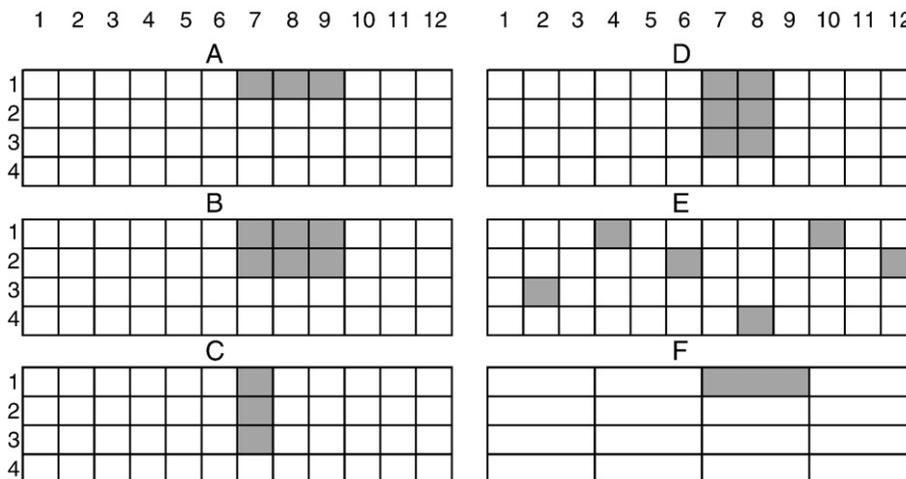


Fig. 1. Patterns of frequency–time windows containing true-positive effects (gray regions).

two time bands consecutively, and within each of those time bands there were two frequency families covering two frequency bands consecutively. The trees were F_4T_{12} , F_4T_6 , F_4T_4 , F_4T_2 , F_2T_{12} , F_2T_6 , F_2T_4 , F_2T_2 and F_1T_{12} ; plus all previous trees with order reversed (e.g., $T_{12}F_4$, etc.), totalling 18 trees. In every case, the third level of the tree consisted of 171 hypotheses (corresponding to the significance of the effect of condition on PLV at each electrode pair across the 10 participants) for each second level family.

FDR procedures (using mixed FDR for independence assumption) were applied to p -values, generated from single-sample two-tailed t -tests on PLV samples. For each simulation the total number of tests was 8028. Reported numbers of false and detected discoveries, and hFDR bounds are averages over 1000 runs. hFDR bounds were computed for full-tree and level-3 restricted hFDR analyses.

Experimental data analysis

To confirm the effectiveness of hFDR, we reanalyzed EEG data acquired from a visual search experiment (Phillips and Takeda, 2009). The purpose of the experiment was to test the hypothesis that top-down driven control of visual attention in humans is accompanied by frontal-parietal synchrony in the lower gamma-band. The design of the experiment followed an early study on monkeys using implanted electrodes (Buschman and Miller, 2007). Top-down signals were induced using distractors that share a feature (e.g., colour, or orientation) with the target, yielding a steep search slope (search time increasing with display set size)—inefficient search. Bottom-up signals were induced using distractors with no feature in common with the target, yielding a flat slope (search time independent of set size)—efficient search. Monkeys showed significantly greater synchrony between frontal and parietal electrodes in the lower gamma-band during inefficient than efficient search. For the human study, a corresponding increase in phase-locking as the measure of synchrony between frontal and parietal scalp electrodes was also observed for the same contrast and frequency band (Phillips and Takeda, 2009). In this case, the significance of the difference between the inefficient and efficient conditions was established by analysis of variance of PLVs obtained from 10 subjects averaged over 768 trials for the pairs and time-frequency regions of interest.

For the purpose of comparing methods in the current study, data were confined to 25 electrode pairs for each of the subject, i.e., five frontal electrodes (F7, F3, Fz, F4, F8) by five posterior electrodes (T8, P3, Pz, P4, T6), located according to the International 10–20 system. Frequency was partitioned into lower (22–34 Hz) and upper (36–48 Hz) gamma bands, corresponding to the original studies (Buschman and Miller, 2007; Phillips and Takeda, 2009). Time was partitioned into twelve 50 ms windows for the first 600 ms after stimulus (search display) onset. For each subject, efficient–inefficient contrast values were generated by computing mean PLVs difference between the conditions for these frequency–time windows for each of the 25 electrode pairs. A random effect analysis using one-sample two-tailed t -test on subjects' contrast samples was performed to test multiple hypotheses involved in the following procedures.

For direct FDR, all 600 ($= 2 \times 12 \times 25$) hypotheses were regarded as a single family, and a mixed FDR procedure was applied on the resulting p -values for multiple testing correction.

For hierarchical FDR, we used a 3-level F_2T_{12} tree for applying full-tree and level-3 restricted analyses. The first level testing involved testing a single family containing 2 hypotheses corresponding to each frequency band, $H^0_i: \mu_i = 0$, where μ_i is the mean PLV for i th frequency band. The resulting p -values were thresholded by mixed FDR procedure. The second level testing followed for only the significantly identified frequency band(s). The test family for each significant frequency band included 12 hypotheses for each time band, $H^0_{ij}: \mu_{ij} = 0$, where μ_{ij} is the mean PLV for i th frequency and j th time band, using mixed FDR. The third-level testing followed in a similar way. For

each of the significantly identified frequency–time bands, a family of 25 hypotheses corresponding to the electrode pairs, $H^0_{ijk}: \mu_{ijk} = 0$, were tested, where μ_{ijk} is the mean PLV for i th frequency, j th time band, and k th electrode pair. The full-tree hFDR bound and q -value (q^*) required for this bound to be maintained at 5% were calculated using Eqs. (3) and (4). The level-3 restricted hFDR bound was computed by using the same q -level (q^*) in Eq. (3).

Results

Simulation

Simulation results are summarized in Table 2. For all combinations of hypothesis hierarchy and synchrony pattern, the false discovery rate for hFDR was found to be lower than FDRBH; and for most combinations, the number of pairs of electrodes detected as significant was greater with hFDR. The level-3 restricted bound was lower than full-tree bound for all combinations of hypotheses and synchrony patterns.

For hFDR, the number of discoveries (i.e., families and pairs) and detected pairs varied with hierarchy and synchrony pattern. Dimension order affected the number of detections for a given pattern. With the first pattern, for example, where synchrony was confined to a single frequency band but spanned several time bands, there were generally more detected pairs with frequency than time at the first level (e.g., 21 detections for tree F_4T_{12} versus 6 detections for $T_{12}F_4$). For a particular dimension order, the number of detections also varied with the family width (i.e., number of windows covered by a family). For the F_4T_i trees (i.e., those with four frequency families at the first level), tree F_4T_4 provided the greatest number of detections (24), where the width and location of the time family coincided the synchrony pattern (see Fig. 1F).

Similar effects were observed for the second, third and fourth patterns. For the third pattern, for example, where synchrony was confined to a single time band, but spanned several frequency bands, the converse effect was observed—trees with time as the first level generally detected more pairs, and tree $T_{12}F_4$ provided the greatest number of detections. For some trees (e.g., F_4T_2), however, the number of detections (1) was less than for FDRBH.

For the fifth pattern, containing multiple synchrony windows throughout the frequency and time domains, trees F_2T_{12} and F_2T_4 provided the most detections, identifying 19 and 20 pairs, respectively. In this case, half of the trees provided fewer detections than FDRBH.

Experiment

For the experimental data, hFDR detected 34 full-tree discoveries, including 1 frequency band at the first level, 7 time bands at the second level, and 26 electrodes pairs with significant synchrony (Fig. 2) at the third level, with an estimated FDR bound of .0491 at $q = 0.04$. The value of q was set according to Eq. (4). It revealed significantly greater synchrony for the inefficient than efficient condition in the lower gamma band (22–34 Hz) predominately at 300–500 ms post-stimulus onset. For two time bands, 250–300 ms and 500–550 ms, the frequency–time families were significant, but not at the level of individual electrode pairs. Synchrony in the inefficient condition was not significantly greater than in the efficient condition for the high gamma band (36–48 Hz). hFDR did not reveal significantly greater synchrony for the efficient than inefficient condition in any frequency or time band. The level-3 restricted hFDR bound corresponding to 26 level-3 discoveries, i.e., electrode pairs with significant effects was 0.0488 at $q = 0.04$. FDRBH did not reveal any significant differences in synchronization even at $q = 0.05$. The results from the two versions of mixed FDR

Table 2
Simulation results for synchrony patterns (P) and hierarchy trees (T) direct (FDRBH) and hierarchical (hFDR) procedures.

P	FDRBH		hFDR												
	NP, FDR (SD)		T	NF	ND	NF3	NP	FDR	(SD)	bound	(SD)	bound3	(SD)		
1	2.69, 0.057 (0.15)		F_4T_{12}	4.14	24.50	2.27	21.35	0.032	(0.03)	0.067	(0.02)	0.054	(0.00)		
			$T_{12}F_4$	2.41	7.35	0.71	5.94	0.028	(0.04)	0.084	(0.02)	0.053	(0.01)		
			F_4T_6	3.30	22.45	1.42	20.15	0.030	(0.03)	0.066	(0.02)	0.053	(0.00)		
			T_6F_4	2.48	12.74	0.75	11.26	0.029	(0.04)	0.077	(0.02)	0.052	(0.01)		
			F_4T_4	2.76	25.49	0.88	23.73	0.031	(0.03)	0.065	(0.02)	0.051	(0.00)		
			T_4F_4	2.75	24.91	0.89	23.16	0.030	(0.03)	0.065	(0.02)	0.052	(0.00)		
			F_4T_2	2.75	20.93	0.87	19.18	0.031	(0.03)	0.066	(0.02)	0.052	(0.00)		
			T_2F_4	2.22	14.50	0.63	13.28	0.031	(0.04)	0.076	(0.02)	0.051	(0.00)		
			F_2T_{12}	2.59	9.21	1.00	7.62	0.030	(0.04)	0.082	(0.02)	0.053	(0.01)		
			$T_{12}F_2$	2.42	6.12	0.72	4.70	0.029	(0.05)	0.085	(0.02)	0.054	(0.01)		
			F_2T_6	2.42	11.31	0.82	9.89	0.032	(0.04)	0.078	(0.02)	0.052	(0.00)		
			T_6F_2	2.46	10.43	0.72	8.97	0.032	(0.04)	0.078	(0.02)	0.053	(0.01)		
			F_2T_4	2.21	14.75	0.62	13.53	0.033	(0.03)	0.076	(0.02)	0.051	(0.00)		
			T_4F_2	2.72	20.52	0.86	18.80	0.033	(0.03)	0.066	(0.02)	0.052	(0.00)		
			F_2T_2	2.17	10.13	0.58	8.95	0.031	(0.04)	0.078	(0.02)	0.052	(0.00)		
			T_2F_2	2.19	10.35	0.59	9.16	0.031	(0.04)	0.078	(0.02)	0.052	(0.00)		
			F_1T_{12}	1.82	2.87	0.40	2.05	0.018	(0.04)	0.092	(0.01)	0.053	(0.01)		
			$T_{12}F_1$	1.70	4.31	0.70	3.61	0.030	(0.05)	0.084	(0.02)	0.056	(0.01)		
		2	11.30, 0.051 (0.08)		F_4T_{12}	7.73	53.03	4.94	46.30	0.032	(0.02)	0.059	(0.01)	0.055	(0.00)
					$T_{12}F_4$	9.33	60.00	5.44	51.67	0.031	(0.02)	0.058	(0.00)	0.055	(0.00)
	F_4T_6			5.88	48.17	3.09	43.29	0.032	(0.03)	0.058	(0.01)	0.053	(0.00)		
	T_6F_4			5.89	52.41	3.14	47.52	0.031	(0.02)	0.056	(0.00)	0.053	(0.00)		
	F_4T_4			4.70	54.83	1.91	51.14	0.031	(0.02)	0.056	(0.01)	0.052	(0.00)		
	T_4F_4			4.23	63.12	2.12	59.89	0.031	(0.02)	0.053	(0.00)	0.052	(0.00)		
	F_4T_2			4.63	45.05	1.84	41.42	0.032	(0.03)	0.057	(0.01)	0.052	(0.00)		
	T_2F_4			4.17	51.21	2.10	48.04	0.033	(0.02)	0.054	(0.00)	0.052	(0.00)		
	F_2T_{12}			5.28	63.83	3.21	59.55	0.031	(0.02)	0.054	(0.00)	0.053	(0.00)		
	$T_{12}F_2$			6.85	59.24	2.96	53.39	0.031	(0.02)	0.056	(0.00)	0.053	(0.00)		
	F_2T_6			4.13	57.78	2.06	54.65	0.032	(0.02)	0.054	(0.00)	0.052	(0.00)		
	T_6F_2			4.55	53.43	1.80	49.88	0.031	(0.02)	0.054	(0.00)	0.052	(0.00)		
	F_2T_4			3.19	61.79	1.12	59.60	0.031	(0.02)	0.053	(0.00)	0.051	(0.00)		
	T_4F_2			3.23	62.07	1.12	59.84	0.031	(0.02)	0.053	(0.00)	0.051	(0.00)		
	F_2T_2			3.17	50.50	1.10	48.33	0.033	(0.02)	0.053	(0.00)	0.051	(0.00)		
	T_2F_2			3.15	50.45	1.08	48.29	0.033	(0.02)	0.053	(0.00)	0.051	(0.00)		
	F_1T_{12}			4.66	44.55	2.72	40.89	0.033	(0.03)	0.058	(0.01)	0.053	(0.00)		
	$T_{12}F_1$			3.89	46.71	2.89	43.82	0.034	(0.03)	0.055	(0.00)	0.053	(0.00)		
3	2.66, 0.061 (0.15)				F_4T_{12}	1.75	3.89	0.37	3.14	0.030	(0.04)	0.091	(0.02)	0.052	(0.01)
					$T_{12}F_4$	5.09	32.73	2.98	28.64	0.030	(0.03)	0.058	(0.00)	0.055	(0.00)
			F_4T_6	1.70	2.92	0.32	2.22	0.022	(0.04)	0.093	(0.01)	0.052	(0.01)		
			T_6F_4	4.09	19.56	2.10	16.47	0.031	(0.04)	0.066	(0.01)	0.056	(0.00)		
			F_4T_4	1.68	2.45	0.30	1.77	0.019	(0.04)	0.093	(0.01)	0.052	(0.01)		
			T_4F_4	3.31	11.84	1.43	9.53	0.029	(0.04)	0.074	(0.02)	0.056	(0.01)		
			F_4T_2	1.71	2.07	0.33	1.36	0.023	(0.05)	0.094	(0.01)	0.055	(0.01)		
			T_2F_4	2.26	4.13	0.63	2.88	0.021	(0.05)	0.088	(0.02)	0.056	(0.01)		
			F_2T_{12}	1.87	7.77	0.45	6.90	0.029	(0.03)	0.085	(0.02)	0.051	(0.00)		
			$T_{12}F_2$	4.00	29.30	1.89	26.29	0.030	(0.03)	0.057	(0.00)	0.053	(0.00)		
			F_2T_6	1.83	6.05	0.41	5.22	0.029	(0.04)	0.087	(0.02)	0.052	(0.01)		
			T_6F_2	3.45	19.90	1.46	17.45	0.032	(0.04)	0.063	(0.01)	0.054	(0.00)		
			F_2T_4	1.84	5.05	0.42	4.21	0.028	(0.04)	0.088	(0.02)	0.052	(0.01)		
			T_4F_2	3.05	13.77	1.17	11.72	0.031	(0.04)	0.070	(0.02)	0.055	(0.01)		
			F_2T_2	1.81	3.47	0.39	2.66	0.033	(0.06)	0.090	(0.02)	0.054	(0.01)		
			T_2F_2	2.28	5.74	0.66	4.46	0.029	(0.05)	0.084	(0.02)	0.055	(0.01)		
			F_1T_{12}	1.92	12.96	0.48	12.04	0.033	(0.03)	0.081	(0.02)	0.051	(0.00)		
			$T_{12}F_1$	2.11	28.91	1.11	27.80	0.032	(0.03)	0.054	(0.00)	0.052	(0.00)		
		4	11.54, 0.051 (0.08)		F_4T_{12}	5.48	31.32	2.93	26.84	0.031	(0.03)	0.068	(0.02)	0.054	(0.00)
					$T_{12}F_4$	9.20	66.67	6.02	58.47	0.031	(0.02)	0.057	(0.00)	0.055	(0.00)
	F_4T_6			4.25	33.69	1.70	30.44	0.030	(0.03)	0.065	(0.02)	0.052	(0.00)		
	T_6F_4			5.24	64.46	3.12	60.21	0.031	(0.02)	0.054	(0.00)	0.053	(0.00)		
	F_4T_4			4.20	30.09	1.65	26.89	0.032	(0.03)	0.066	(0.02)	0.053	(0.00)		
	T_4F_4			5.21	57.61	3.09	53.40	0.032	(0.02)	0.055	(0.00)	0.053	(0.00)		
	F_4T_2			4.09	22.75	1.54	19.66	0.033	(0.04)	0.068	(0.02)	0.053	(0.00)		
	T_2F_4			4.72	38.64	2.66	34.92	0.034	(0.03)	0.057	(0.00)	0.054	(0.00)		
	F_2T_{12}			4.54	41.27	2.31	37.73	0.031	(0.03)	0.061	(0.01)	0.053	(0.00)		
	$T_{12}F_2$			6.91	59.28	3.73	53.37	0.031	(0.02)	0.056	(0.00)	0.053	(0.00)		
	F_2T_6			3.59	42.83	1.36	40.24	0.030	(0.03)	0.059	(0.01)	0.051	(0.00)		
	T_6F_2			4.17	59.20	2.05	56.03	0.031	(0.02)	0.053	(0.00)	0.052	(0.00)		
	F_2T_4			3.54	37.91	1.31	35.36	0.032	(0.03)	0.060	(0.01)	0.052	(0.00)		
	T_4F_2			4.07	51.13	1.96	48.06	0.033	(0.02)	0.054	(0.00)	0.052	(0.00)		
	F_2T_2			3.46	28.23	1.23	25.77	0.033	(0.03)	0.062	(0.01)	0.052	(0.00)		
	T_2F_2			3.74	35.29	1.68	32.55	0.034	(0.03)	0.056	(0.00)	0.053	(0.00)		
	F_1T_{12}			3.95	54.73	2.02	51.78	0.032	(0.02)	0.057	(0.01)	0.052	(0.00)		
	$T_{12}F_1$			3.18	58.50	2.18	56.32	0.032	(0.02)	0.053	(0.00)	0.052	(0.00)		
5	11.49, 0.052 (0.07)				F_4T_{12}	4.28	22.22	2.08	18.94	0.028	(0.03)	0.072	(0.02)	0.054	(0.00)
					$T_{12}F_4$	4.07	16.87	1.59	13.79	0.030	(0.04)	0.075	(0.02)	0.054	(0.01)

Table 2 (continued)

P	FDRBH	hFDR										
	NP, FDR (SD)	T	NF	ND	NF3	NP	FDR	(SD)	bound	(SD)	bound3	(SD)
		F_4T_6	3.84	15.09	1.64	12.24	0.026	(0.04)	0.075	(0.02)	0.055	(0.00)
		T_6F_4	2.71	7.66	0.86	5.95	0.027	(0.04)	0.084	(0.02)	0.054	(0.01)
		F_4T_4	3.62	11.96	1.42	9.34	0.027	(0.04)	0.077	(0.02)	0.055	(0.01)
		T_4F_4	3.60	11.65	1.40	9.05	0.027	(0.04)	0.078	(0.02)	0.055	(0.01)
		F_4T_2	3.42	8.01	1.22	5.59	0.025	(0.05)	0.081	(0.02)	0.058	(0.01)
		T_2F_4	3.46	7.78	1.19	5.31	0.021	(0.05)	0.082	(0.02)	0.058	(0.01)
		F_2T_{12}	4.68	22.86	2.49	19.18	0.032	(0.04)	0.067	(0.02)	0.055	(0.00)
		$T_{12}F_2$	3.99	14.03	1.51	11.05	0.030	(0.04)	0.076	(0.02)	0.055	(0.01)
		F_2T_6	3.72	11.49	1.54	8.77	0.026	(0.04)	0.076	(0.02)	0.057	(0.01)
		T_6F_2	2.67	6.04	0.82	4.37	0.026	(0.05)	0.085	(0.02)	0.055	(0.01)
		F_2T_4	3.92	23.35	1.73	20.43	0.031	(0.04)	0.065	(0.02)	0.054	(0.00)
		T_4F_2	3.36	15.34	1.16	12.99	0.030	(0.04)	0.074	(0.02)	0.053	(0.01)
		F_2T_2	3.56	13.33	1.37	10.78	0.029	(0.04)	0.071	(0.02)	0.056	(0.01)
		T_2F_2	3.53	11.67	1.25	9.14	0.026	(0.04)	0.075	(0.02)	0.056	(0.01)
		F_1T_{12}	3.35	10.54	1.43	8.19	0.025	(0.05)	0.078	(0.02)	0.057	(0.01)
		$T_{12}F_1$	2.48	10.11	1.48	8.63	0.037	(0.06)	0.075	(0.02)	0.057	(0.01)

Results are indicated for number of families tested and number of discoveries in full-tree analysis (NF and ND), and in level-3 restricted analysis (NF3 and NP), false discovery rate average (FDR), full-tree and level-3 restricted bound (bound and bound3), and their standard deviations (SD).

were almost identical, except that there was one discovery less with the test assuming dependence than assuming independence.

Discussion

The results support our two main points: (1) the standard approach to multiple test correction for phase-locking analysis in EEG is overly conservative and fails to detect significant synchrony effects. (2) The hierarchical FDR method overcomes this problem and reveals significant instances of synchrony that are consistent with true synchrony in the case of simulation, and previous analysis in the case of experimental data.

Simulation and experimental results showed that, generally, hierarchical FDR (hFDR) provided greater sensitivity and lower false discovery rates than direct FDR (FDRBH) (Table 2). Hierarchical FDR affords a multiresolution approach to analysis, which is particularly important when only a few biologically relevant synchronies are expected given the thousands of tests that need to be conducted. The procedure was most effective when unlikely hypotheses were removed early in the hierarchy. For example, in regard to the first pattern, where synchrony was confined to a single frequency band but several time bands, the F_4T_i trees were most effective. Moreover, trees were more effective when the width and location of the family coincided the synchrony window. Again with respect to the first pattern, the F_4T_i trees were more effective than the F_2T_j trees, and within the F_4T_i trees, F_4T_4 (24 detections) was more effective than F_4T_{12} (21), F_4T_6 (20), or F_4T_2

(19). Similar effects were observed for the other patterns and hierarchies (ref. Fig. 1 and Table 2).

Conversely, hFDR is less effective when a large number of hypotheses are not removed early in the analysis. Contrast, for example, the number of detected pairs for F_4T_{12} versus $T_{12}F_4$ on the first pattern, 21 versus 6, and third patterns, 3 versus 29 (respectively). In general, however, the performance of hFDR was never greatly worse than FDRBH (note that hFDR reduces to FDRBH when there is only one family in the hierarchy whose width spans both dimensions). Furthermore, the false discovery rate was generally about half of FDRBH, even though the approximate bound was sometimes higher than 0.05.

The fifth pattern was designed to test hFDR in the case of multiple windows of synchrony distributed throughout the frequency and time dimension (Fig. 1E). When there does not exist a strong reason for asserting a particular hypothesis hierarchy, a simple but effective choice is to use trees that cover each data dimension with a large number of families, which in our study were the F_4T_{12} and $T_{12}F_4$ trees. These trees provided a good number of detections, 19 and 14 pairs, respectively, which was greater than the direct FDR method.

The full-tree hFDR analysis with experimental data resulted in 34 discoveries and 9 families being tested, and hFDR bound of about 0.049 at $q = 0.04$ (Fig. 2). These patterns of synchrony conform to those reported in Phillips and Takeda (2009). For our specific application, both full-tree and level-3 restricted hFDR are useful. In a level-restricted analysis, hFDR bound is computed using the same

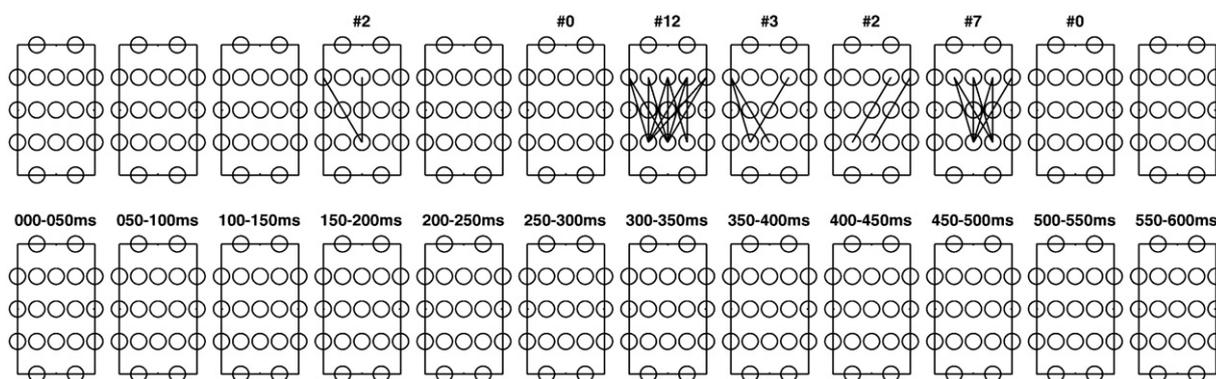


Fig. 2. Synchrony map from hFDR application indicating (number of) electrode pairs showing significantly greater phase-locking for the inefficient than efficient search conditions. The top and bottom rows correspond to lower (22–34 Hz) and higher (36–48 Hz) gamma bands.

procedure as full-tree analysis (Eq. (3)), except the number of discoveries, and families tested, include only those at the specified level. For example, the full-tree analysis shows the numbers of families tested and discoveries (N_d and N_f) made at the third level as 26 and 7, respectively. Substituting these numbers in Eq. (3) and setting $q = 0.04$ and $\delta = 1$, hFDR bound for level-3 inference is 0.0489, which is smaller than full-tree hFDR bound of 0.0491. In our data, as the number of full-tree discoveries include a large proportion of level-3 discoveries, level-3 hFDR bound is smaller than full-tree hFDR bound. This is evident both in simulation as well as experimental data. While the inference in terms of the number of electrode pairs with significant PLV effects is the same in both hFDR methods, there is an important difference between the two. Unlike full-tree hFDR, level-3 restricted hFDR does not provide a multi-decompositional inference in all three dimensions.

The adjustment of q as indicated in Eq. (4) is needed if the researcher needs to control hFDR at the conventional 5% level Type I error bound. Alternatively, the standard 5% threshold for q -value or any other value deemed appropriate for that application can be chosen. For example, in our experimental analysis, substituting q with 0.05 results in hFDR = 0.06 and two more discoveries in the same frequency and time regions (results not shown here). As long as estimated hFDR bound is presented along with the inference, a value slightly higher than conventional 5% should be acceptable.

The results from the two versions of mixed FDR were similar, quite possibly because no significant synchronies were found in the negative direction. For small sample sizes, and the large variances associated with PLV data, the distributional assumptions underlying parametric t -tests may not hold. So, we validated our inference with an additional hFDR analysis using a bootstrap test (with 1000 surrogate data samples) to generate resampling-based p -values (similar to the one used in Lachaux et al., 1999) for the hierarchical framework, which produced similar results with 33 full-tree discoveries at $q = 0.04$. Analysis of PLVs at high frequencies for intervals of 50 ms allows only one or two cycles per interval. However, repeating the analysis using 100 ms time intervals did not change the inference. At lower resolution, the number of discoveries (N_d) and families tested (N_f) also reduced in the same proportion, and still showed significant synchrony in the same regions as detected by the current analysis. For example, a full-tree hFDR analysis using 100 ms time resolution, showed 26 full-tree discoveries, including 19 pairs with significantly greater synchrony in inefficient condition in lower gamma band predominantly at the 300–500 ms post-stimulus onset (results not shown here).

Although EEG data is amenable to a hierarchical approach, the data itself is not naturally hierarchically structured. Simulations showed how the choice of hierarchy may affect hFDR efficiency, although for our application hFDR was more sensitive than FDR in all cases. In practice, where the application domain does not enforce a particular hierarchy, or alternative choices appear equally reasonable, cross-validation or similar method can be used to assess the robustness of the particular hierarchy (model) chosen. That is, the data can be partitioned into selection and evaluation subsets, where multiple hierarchies (models) are employed on the selection set, and the performance of the "best one" is reported for the evaluation set.

Our simulations and experimental data pertained to just 19 electrodes. Yet, even with this number of channels, the problem of multiple comparisons was severe and only addressed by the hierarchical approach. Increasingly, high-density (e.g., 64- to 256-channel) EEG is being employed, where clearly the multiple comparisons problem is exacerbated. We have shown that hFDR offers a practical solution to this problem. Moreover, it is readily applicable to other measures of detecting synchrony, such as mutual information, generalized synchronization (Quiñero et al., 2002), single-trial phase locking (Lachaux et al., 2000), structural synchrony (Fingelkurts et al., 2003), phase resetting (Makinen et al., 2005),

specifically in electrode space when MR images are not available. For a review and comparison of some of these methods, see David et al. (2004); Alba et al. (2007). In the case of PLV, phase and amplitude synchrony effects can be tested separately or together within the same procedure simply by using the level-restricted form of hFDR.

The hierarchical framework also allows one to address more complex questions involving multiple families of hypotheses. This situation occurs with microarray data involving thousands of genes tested for multiple strain effects that result in millions of simultaneous tests, which was the motivating application for this technique (Yekutieli, 2008). The ability to address more complex questions is afforded by the modular nature of the framework where the within-family control procedure is not restricted to a particular version of FDR. In our case, for example, we improvised upon the core approach using the mixed FDR procedure that controls for Type I and Type III error simultaneously.

Finally, an important assumption for hFDR is that the statistical tests between levels are independent. However, hFDR can be employed when there is dependency within levels. In the context of EEG synchrony, a topic of further work is to evaluate the effect of within-level dependency, which occurs for example when there are dependencies between frequency bands, or between time bands, or electrode pairs.

Acknowledgments

This work was supported by a JSPS Postdoctoral Fellowship to the first author. We thank Daniel Yekutieli and Yuji Takeda for their advice, and the reviewers for comments that helped improve the presentation of this work.

Appendix A. Approximation of hierarchical FDR bounds

Yekutieli (2008) used the following equation to approximate the bounds for hierarchical FDR control under the assumption that the test statistics are independently distributed across levels for full-tree and level-restricted hFDR.

$$\text{bound} = q\delta \frac{N_d + N_f m_0}{N_d + 1 m}, \quad (5)$$

where q is the threshold for controlling FDR in each single family, N_d is the number of observed discoveries, N_f is the number of families tested, and δ is a multiplicative constant, typically about 1, when the number of tests is not high. For several hundreds of tests, δ with an upper bound of 1.4 is needed. N_d (and N_f) may either represent all the discoveries (and families tested) in a full-tree analysis, or only the discoveries (and families tested) at the specified level k , in a level- k restricted analysis. $\frac{m_0}{m}$ is the proportion of true null hypotheses to the total number of hypotheses, which is less than 1 if there are any rejections, and the above equation can be rewritten as

$$\text{bound} = q \frac{N_d + N_f}{N_d + 1}. \quad (6)$$

In a full-tree analysis, the bound varies in an interval, $q \leq \text{bound} \leq 2q$. The value $2q$ can be approached only when the proportion of leaf discoveries is too small and $\frac{m_0}{m} = 1$. On the other hand, if number of discoveries greatly exceeds the number of families tested ($N_d \gg N_f$), so that the multiplier $\frac{N_d + N_f}{N_d + 1} \approx 1$, FDR bound converges to the value q . The bound for a level-restricted analysis has no upper limit, but if $N_d \gg N_f$, FDR bound in a level-restricted analysis is also approximately equal to q .

References

- Alba, A., Marroquin, J.L., Pea, J., Harmony, T., Gonzalez-Frankenberger, B., 2007. Exploration of event-induced eeg phase synchronization patterns in cognitive tasks

- using a time–frequency–topography visualization system. *J. Neurosci. Methods* 161 (1), 166–182.
- Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B* 57 (1), 289–300.
- Benjamini, Y., Yekutieli, D., 2001. The control of the false discovery rate in multiple testing under dependency. *Ann. Stat.* 29 (4), 1165–1188.
- Benjamini, Y., Yekutieli, D., 2005. False discovery rate-adjusted multiple confidence intervals for selected parameters. *J. Am. Stat. Assoc.* 100 (469), 71–81.
- Bhattacharya, J., Petsche, H., 2005. Phase synchrony analysis of eeg during music perception reveals changes in functional connectivity due to musical expertise. *Signal Process.* 85 (11), 2161–2177.
- Buschman, T.J., Miller, E.K., 2007. Top-down versus bottom-up control of attention in the prefrontal and posterior parietal cortices. *Science* 315 (5820), 1860–1862.
- Chumbley, J.R., Friston, K.J., 2009. False discovery rate revisited: Fdr and topological inference using Gaussian random fields. *NeuroImage* 44 (1), 62–70.
- David, O., Cosmelli, D., Friston, K.J., 2004. Evaluation of different measures of functional connectivity using a neural mass model. *NeuroImage* 21 (2), 659–673.
- Fingelkurts, A., Fingelkurts, A., Krause, C., Kaplan, A., Borisov, S., Sams, M., 2003. Structural (operational) synchrony of eeg alpha activity during an auditory memory task. *NeuroImage* 20 (1), 529–542.
- Genovese, C.R., Lazar, N.A., Nichols, T., 2002. Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *NeuroImage* 15 (4), 870–878.
- Hemmelmann, C., Horn, M., Ssse, T., Vollandt, R., Weiss, S., 2005. New concepts of multiple tests and their use for evaluating high-dimensional eeg data. *J. Neurosci. Methods* 142 (2), 209–217.
- Lachaux, J., Rodrigue, E., Martinerie, J., Varela, F.J., 1999. Measuring phase synchrony in brain signals. *Hum. Brain Mapp.* 8 (4), 194–208.
- Lachaux, J.-P., Rodriguez, E., Quyen, M.L.V., Lutz, A., Martinerie, J., Varela, F.J., 2000. Studying single-trials of phase synchronous activity in the brain. *Int. J. Bifurc. Chaos Appl. Sci. Eng.* 10 (10), 2429–2439.
- Makinen, V., Tiitinen, H., May, P., 2005. Auditory event-related responses are generated independently of ongoing brain activity. *NeuroImage* 24 (4), 961–968.
- Mizuhara, H., Yamaguchi, Y., 2007. Human cortical circuits for central executive function emerge by theta phase synchronization. *NeuroImage* 36 (1), 232–244.
- Nichols, T.E., Holmes, A.P., 2002. Nonparametric permutation tests for functional neuroimaging: a primer with examples. *Hum. Brain Mapp.* 15 (1), 1–25.
- Pantazis, D., Nichols, T.E., Baillet, S., Leahy, R.M., 2005. A comparison of random field theory and permutation methods for the statistical analysis of meg data. *NeuroImage* 25 (2), 383–394.
- Phillips, S., Takeda, Y., 2009. Greater frontal-parietal synchrony at low gamma-band frequencies for inefficient than efficient visual search in human EEG. *International J. Psychophysiol.* 73 (3), 350–354.
- Quiñero, R., Kraskov, A., Kreuz, T., Grassberger, P., 2002. Performance of different synchronization measures in real data: a case study on electroencephalographic signals. *Phys. Rev. E* 65 (4), 041903.
- Razoumnikova, O.M., 2000. Functional organization of different brain areas during convergent and divergent thinking: an eeg investigation. *Cogn. Brain Res.* 10 (1–2), 11–18.
- Reiner, A., Yekutieli, D., Letwin, N.E., Elmer, G.I., Lee, N.H., Kafkafi, N., Benjamini, Y., 2007. Associating quantitative behavioral traits with gene expression in the brain: searching for diamonds in the hay. *Bioinformatics* 23 (17), 2239–2246.
- Rodriguez, E., George, N., Lachaux, J.P., Martinerie, J., Renault, B., Varela, F., 1999. Perception's shadow: long-distance synchronization of human brain activity. *Nature* 397, 430–433.
- Singh, A.K., Clowney, L., Okamoto, M., Cole, J.B., Dan, I., 2008. Scope of resampling-based tests in fNIRS neuroimaging data analysis. *Stat. Sin.* 18 (4), 1519–1534.
- Singh, A.K., Dan, I., 2006. Exploring the false discovery rate in multichannel NIRS. *NeuroImage* 33 (2), 542–549.
- von Stein, A., Rappelsberger, P., Sarnthein, J., Petsche, H., 1999. Synchronization between temporal and parietal cortex during multimodal object processing in man. *Cereb. Cortex* 9 (2), 137–150.
- Weiss, S., Rappelsberger, P., 2000. Long-range eeg synchronization during word encoding correlates with successful memory performance. *Cogn. Brain Res.* 9 (3), 299–312.
- Worsley, K.J., Evans, A.C., Marrett, S., Neelin, P., 1992. A three-dimensional statistical analysis for CBF activation studies in human brain. *J. Cereb. Blood Flow Metab.* 12 (6), 900–918.
- Yekutieli, D., 2008. Hierarchical false discovery rate controlling methodology. *J. Am. Stat. Assoc.* 103 (481), 309–316.
- Yekutieli, D., Reiner, A., Elmer, G.I., Kafkafi, N., Letwin, N.E., Lee, N.H., Benjamini, Y., 2006. Approaches to multiplicity issues in complex research in microarray analysis. *Stat. Neerl.* 60 (4), 414–437.