horizontal line

Phase Amplitude Coupling – Toolbox Development

This document contains notes from when Yael was originally refactoring Fleming’s code for Devorah’s Rett project, and the questions/inconsistencies that came up, the toolbox now on git has resolved all these issues, but this doc is meant to serve as record in case anyone is curious all the nuances of how we got here!

**Written: Feb 28, 2023, Last updated: Aug 11, 2023**

| **Problem** | **Status** | **Notes/Person who completed it** |
| --- | --- | --- |
| When originally developing the code, we had to confirming whether our PAC analysis matched Flemings and/or Michaels | Complete | [Yael Braverman](mailto:yael.braverman@enders.tch.harvard.edu), our preprocessing matched Fleming’s.  Changes made to PAC analysis  **Number participants flipped during permutation:** Max number of participants that get artificially flipped to another group is N/2 (consistent with Michael) instead of N (what Fleming did).  **Implementing permutation with replacement:** Michael and Fleming’s code ran the permutation analyses without replacement. That is to say that on each iteration we are pulling a random number of participants to flip, but over the course of 1000 iterations we may be repeating these patterns of random combinations.  For example, with a sample of 10 participants there are 252 combinations, so if we loop 1000 times we may be biasing towards certain re-shuffling patterns, which may be impacting the ultimate cluster threshold.  Fix: Now the permutation analyses will compute the number of possible combinations given the sample size and will continue to iterate until it has reached the max number of combinations or 1000, whichever comes first. It keeps a record of which participants were flipped on each round, checks each time that it’s completing a “unique” flip, otherwise it will increase the max\_comb by one and skip that flip and try another.  **Definition of cluster sizes:** The pipeline implements both approaches now, one called   1. Fleming    1. After running a t-test between the MI\_raws and MI\_surrs for the current diagnostic group, you get a matrix low freq X high freq X channels of 1s and 0s (1 where h=1 0 where h=0).    2. Cluster size defined as sum of 1s that are next to each other 2. Michael - Had sample size of 98 (collapsed across ages for permutation analysis)    1. After running a t-test between the MI\_raws and MI\_surrs for the current diagnostic group, you get:       1. a) A matrix low freq X high freq X channels of 1s and 0s (1 where h=1 0 where h=0). (same as above)       2. b) A matrix of low freq X high freq X channels of t-statistics for each of the t-tests run          1. Michael then subtracted the minimum t-statistic required to achieve p <.05 from matrix **b**    2. Clusters positions are defined by 1s that are next to each other in matrix **a**    3. Cluster sizes are computed by summing the t-scores in matrix **b** for the clusters defined by matrix **a** |

# Outstanding Action Items/Questions

| **Items** | **Status** | **Notes** |
| --- | --- | --- |
| Add functionality to have padding and then remove after filtering in PAC | Done | Hi Fleming,  I did a dive into the pactools and do see a "mask" option throughout this script: [pactools/comodulogram.py at master · pactools/pactools · GitHub](https://github.com/pactools/pactools/blob/master/pactools/comodulogram.py) which I'm thinking is what you were referring to.  Do you have your copy of beapp\_calc\_comod.m easily accessible? I'm guessing this is where you would have edited some code to input the mask option when calling the python scripts. Thanks so much! |
| Add segmentation overlap functionality to beapp | Not started | Yael - edit batch\_beapp\_create\_baseline\_segs, beapp\_extract\_segments (to make segments longer by some standard amount , rn hard coded to 2 seconds on each side but can just add to end) |
| Check with | Not started | With our padding, we have much better outcomes (yay!) but this also means we have a large percentage of overlap in our PAC measures - need to figure out how much this would matter from a reviewer perspective |
| Pull the images that match Michael’s paper to see if we are seeing anything similar (Devorah) | In progress |  |
| Run whatson-wheeler test on phase bias values (Yael to implement) | Done | Yael finished writing script - “pac\_strength\_phase\_preference\_stats.R” living here: 'Z:\Public\EEG Analyses + Matlab scripts\PAC\Analyze\_beapp\_PAC\_outputs\post\_analysis\_statistics'  From there - see if we have any matches with Michael’s findings  If we don’t have any, look at how he defined phase bias - Yael implemented Michael’s method as well - checking with April to confirm I did it right but we should have graphs / outputs for both his and fleming’s phase bias measures |
| Check how to pad during segmentation in beapp (Yael) | Done | Look here: [beapp/beapp\_extract\_segments.m at master · lcnbeapp/beapp (github.com)](https://github.com/lcnbeapp/beapp/blob/master/functions/beapp_extract_segments.m)  YB edited the following scripts to add   1. Beapp\_extract\_segments    1. Added tmp\_eeg\_w\_padded which holds the same contents as tmp\_eeg but has 2 seconds of padding on either side of segment 2. Batch\_beapp\_create\_baseline\_segs.m    1. Added EEG\_tmp\_padded variable which inherits the segment rejection info from EEG\_tmp, so segment rejection is just based on the 2 second segments, but then the ultimately saved segments are the overlapping 4 second padded ones 3. Beapp\_msk\_art    1. Commented out the part where it searches for artifacts in data, so that it only masks the first and last 2 seconds of data 4. Changed user inputs to turn on    1. grp\_proc\_info.beapp\_baseline\_msk\_artifact=1; |
| Look at ABCCT dataset in regions (multiple channels) - Yael | Not started |  |
| Confirm the code Fleming used for post beapp analysis | Done | Fleming confirmed post beapp code was PAC\_analysis.m |
| Confirm the user inputs used for PAC (if beapp used) | Done | Yael emailed  From looking at the history of the files - I think up to ICA is on the server - which matches our settings  Confirmed PAC settings, waiting on pre-PAC settings |
| Try running a few ABCCT file on our current settings to see if results change by a lot | Done | Yael ran 5 abcct files on our PAC settings now (just changed segment threshold to 6 to reflect their longer segments)  [ABCCT File Comparison Slides](https://docs.google.com/presentation/d/1mNFZzfWwKt3AGnHRMu2av8vhY3-C1iaPPoftyjxGmOQ/edit?usp=sharing) |
| Try re-running PAC to match the HF/LF resolution in Fleming and Michael’s papers | Done | Devorah in process of re-running all files |

# Verifying Updated Code against original pac\_analysis (by Fleming)

## Checking Contents of PAC\_analysis\_variables.mat - All tests passed

* **Result: All tests passed, all variables equal**
* Variables: MI\_surr, MI\_norm, amp\_dist, MI\_raw
  + New code format: struct with diagnostic groups as fields, old format: separate variable for each
* Method, run sections of new and old code to generate two PAC\_analysis\_variable.mats

**Code:**

%Tests checking contents of new vs old code

assert(isequal(MI\_surr.ASD,ASD\_MIsurr))

assert(isequal(MI\_surr.TD,TD\_MIsurr))

assert(isequal(ASD\_ampdist,amp\_dist\_all.ASD))

assert(isequal(TD\_ampdist,amp\_dist\_all.TD))

assert(isequal(MI\_norm.ASD,ASD\_MInorm))

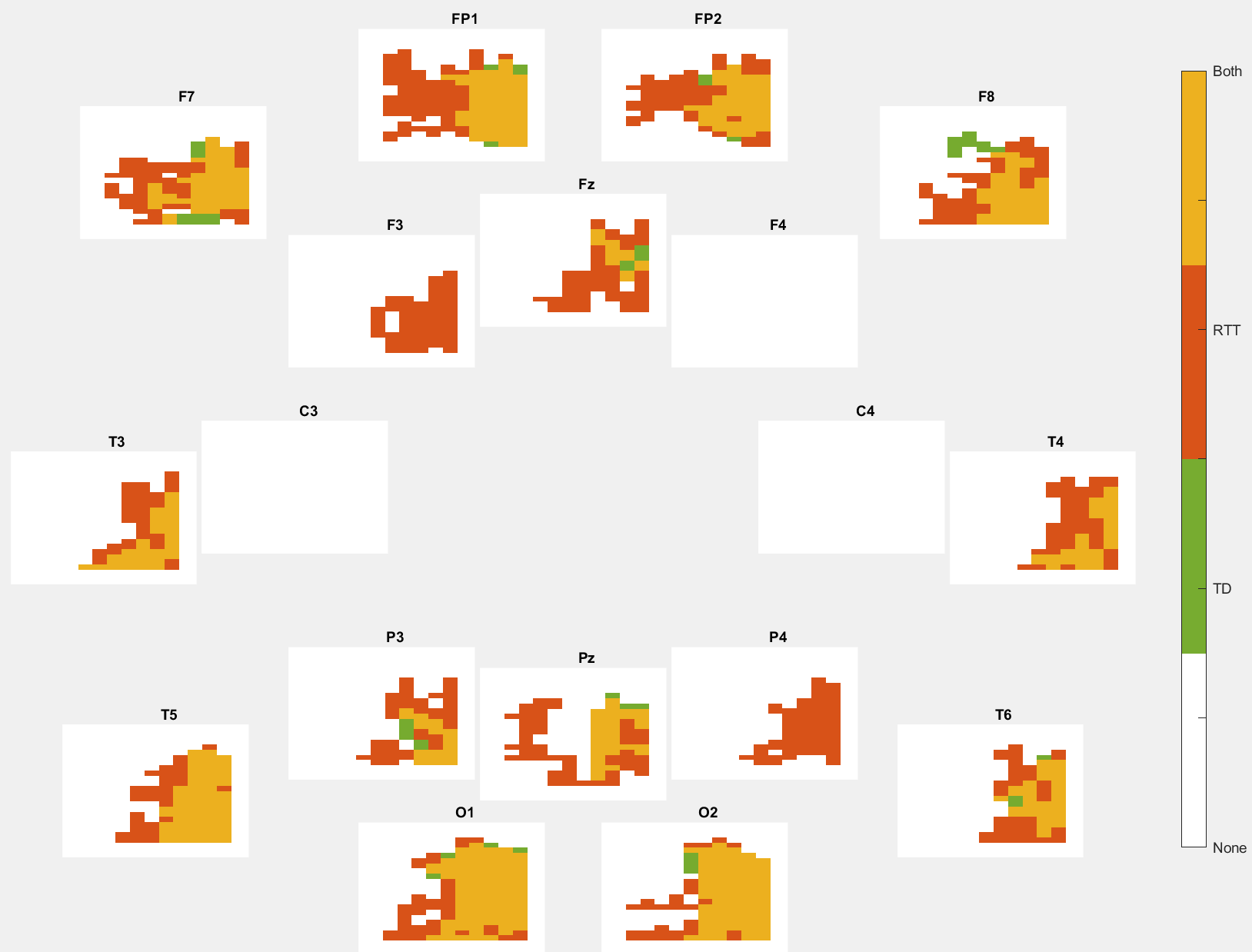
assert(isequal(MI\_norm.TD,TD\_MInorm))

%% All tests passed

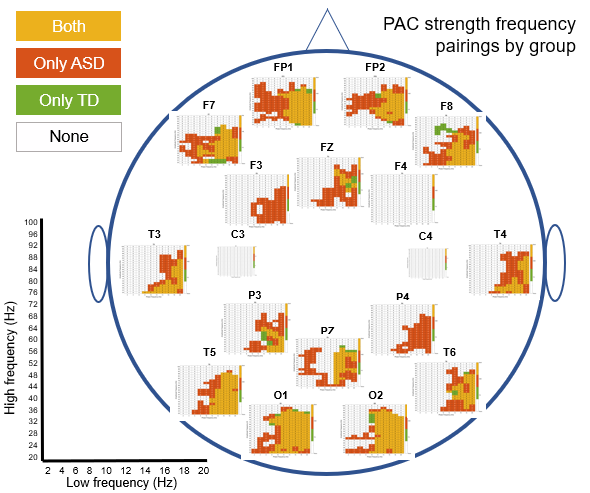
## Checking Permutation Results

* Method:
  + Copy PAC\_analysis\_variables.mat from Z:\Groups\ABCCT\PAC\PAC analysis variables (Which we think is the one used for paper)
  + Convert to new code format (change each variable into struct)
  + Run permutation analyses and compare topoplots visually (since outputs wont be exact same by virtue of the permutation test/swapping)
* **Result: Visuals look the same, cluster threshold slightly different (New run ASD 46, TD 13, Old run ASD 24 TD 10)**

**New Code (RTT should be ASD in colorbar)**



**Old Code - Visual taken from: Z:\Groups\ABCCT\PAC\Images\Group PAC clusters allfreqpairs**



## 

# 

# 

# Agenda for Fleming on 3/13

1. Why did she make 1000 iterations (flipping 1 to n participants instead of 1 n/2 with 200 iterations re: Michael’s approach)
   1. *No reason!*
2. Why were cluster sizes defined by the number of significant lf/hf pairs as opposed to the sum of the t-values in those lf/hf positions for each cluster
   1. *No reason!*
3. Axes limitations on the MI\_norm graphs - why -1.5 to 1.5?
   1. Since we’ve been having issues with low number of sig lf/hf pairs after permutation, we’re trying to determine whether we’re seeing anything meaningful before permutation analyses by looking at the normed comodulograms, and want to confirm if there is an appropriate axis limitation
   2. *Don’t have specific range to visualize whether the MI\_norm graphs had meaningful data, she did compare MI\_norm with MI\_surr and plot a comodulogram of the t-statistics, and masked for t values that surpassed 1.96 which is the significance threshold at p <.05*
4. After using pac\_analysis script, what needs to be done in R / elsewhere for statistics?
   1. She mentioned circular statistics toolbox - watson wheeler R function to compare phase preference: Yael will adapt this script: [rat\_PAC\_PB.R (dropbox.com)](https://www.dropbox.com/home/Mouse%20PAC?preview=rat_PAC_PB.R)
   2. What does she do with the other excel sheets:
      1. **'PAC\_MInorm\_cluster\_averages\_revised\_allfreqpairs.xlsx**
         1. For each region, compute the average MI across 1) all significant clusters where both groups were sig 2) where just diagnostic group was significant
            1. Fleming: pretty sure averaged all those pairs to get pac strength
            2. Could just do t-test, but didn’t end up using for paper
      2. **Cluster\_bin\_proportion\_all\_freq\_pairs.xlsx**
         1. For a given region and angle (e.g bin- (or across all bins) - compute the average amplitude distance across the significant clusters in that region
            1. Right now we do this for bins at 70deg and neg130 but we can change that to be whatever
            2. **Want to look at where circles were different and choose degrees based on that**
            3. **First look at alpha asymmetry before PAC asymmetry**
            4. **Low frequency phase amplitude coupling in fp1/fp2**
      3. **Freqband\_PAC\_metrics\_complete**
         1. For given regions and frequency bands, compute the phase bias, phase values, and pac strength (MI)
            1. Phase values = angle of amp\_dist vector for that lf/hf pair
            2. Phase bias = circmean of the phase values across all the lf/hf combos
         2. Do whatson wheeler test first, if we don’t see anything we can compute phase bias equation michael used (he didn’t use circ mean)
         3. **Write code to read in the phase bias values and do the whatson wheeler test**
      4. **Phase\_stats.xlsx**
         1. For a given region and freq band, check for each participant, which of the 18 phase bins equaled the max of their respective phase bins,
            1. Shows for each participant and across participants, the proportion of when each phase bin contained the maximum amplitude distance value
         2. Fleming didn’t mask for theta-gamma OR all frequencies
         3. Compares phase\_prop\_bins for groups 1 and 2 with a t-test
         4. Fleming said this is stand alone but included in final paper

## 