HINT User Guide

Contents

1 HINT Overview

1.1 Panel 1 - preparing the data

The first panel of HINT is for inputting the data, preprocessing the data, selecting ICs of interest, and preparing an initial guess for the EM algorithm. The panel is split into three sub-panels, as can be seen in Figure 1. We describe the function of each of the sub-panels below.

1.1.1 1. Setup panel

The setup panel consists of the following three items, only the first of which is required:

Analysis Folder This is the folder where all results from the analysis will be stored. Select the folder usin the "browse" button or input it manually into the box.

Prefix If a prefix is input, all output files from the analysis will begin with the prefix. This can be helpful if multiple analyses are to be stored in the same folder.

Create session log If this box is ticked, a log will be written to the analysis folder detailing the steps completed.

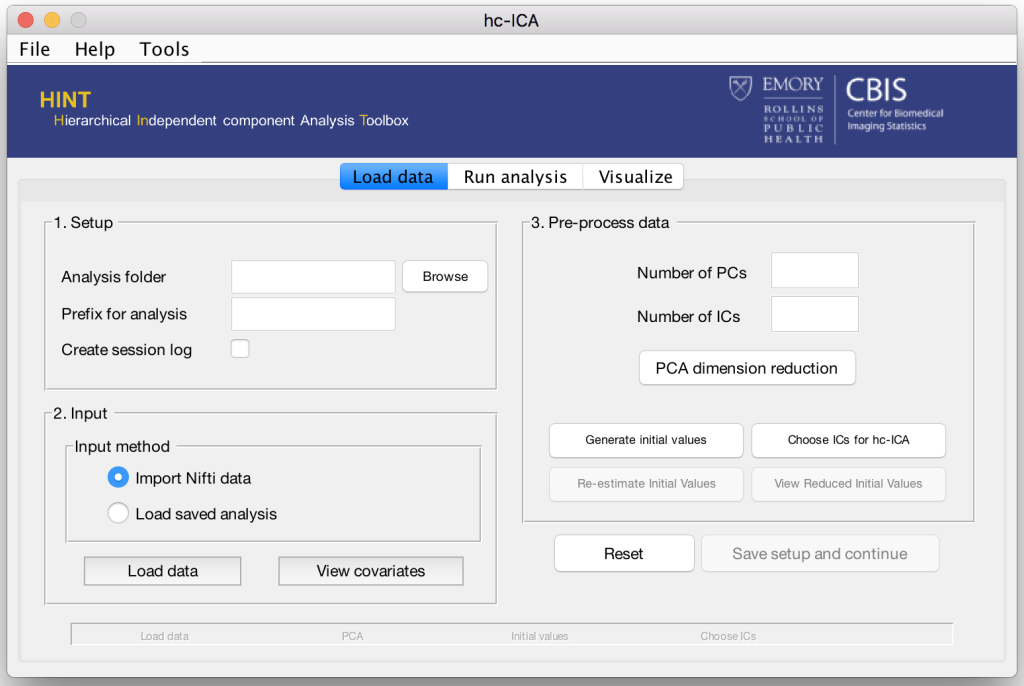


Figure 1: The first panel of HINT. Here the user can input the data, specify interactions among the covariates, pre-preocess, generate an initial guess, and select independent components of interest for the analysis.

1.1.2 2. Input panel

This input panel allows the user to either (A) input the raw data for a new analysis or (B) load the progress of a previous analysis. The user can then specify interactions among the covariates.

1.1.3 Input data for a new analysis

To input data for a new analysis, complete the following steps:

1. Make sure "import Nifti data" is selected and click "Load Data." The window shown in Figure 2 will appear.
2. Input the Nifti files. After clicking "Browse Nii Files" a box will appear asking if the data are in the same folder. This option is provided because some preprocessing pipelines will place each subject's data in their own individual folder. If the data are all in the same folder, select "yes," navigate to the folder, and select all Nifti files of interest. If the data are not in the same folder, select "no," provide a general path to the data, with values that change replaced by astericks (see the examples section for more detail).
3. Input the mask file as a nii file using the "Browse Nii File" box under "Load mask."
4. Load the design matrix as a csv file using the "browse .csv file" button. IMPORTANT: The first column of the csv file identifies the subjects and should consist of the filename from each subject. See the examples section for more details.
5. Click "Load Data."

At this point, the HINT toolbox will reorder the subject Nifti files to be in the order specified in the design matrix. Click on "View covariates" to verify that this process was successful. This window (Figure 3) can also be used to specify interactions among the covariates.

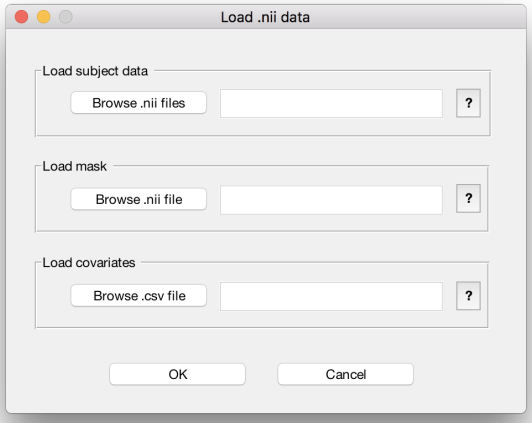


Figure 2: The data input window.

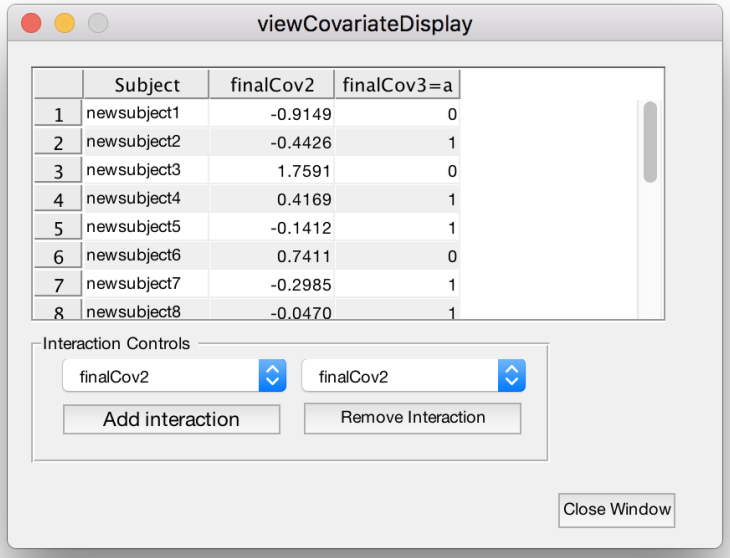


Figure 3: The covariate display window.

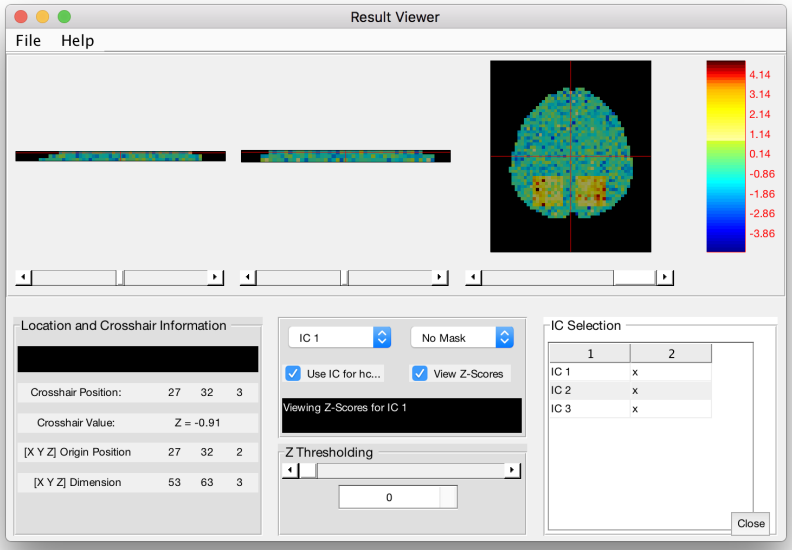


Figure 4: The IC selection window.

1.1.4 3. Pre-process data panel

The pre-process data panel performs PCA dimension reduction of the input data and then prepares an intial guess. The steps in this window are provided below from top to bottom.

Number of PCs Number of principal components used in the PCA decomposition. This number should be larger than the number of ICs.

Number of ICs Number of independent components to be estimated.

PCA dimension reduction This button will perform the PCA decomposition as a final preprocessing step.

Generate intial values This button uses the GIFT toolbox to generate an initial set of guesses for the EM algorithm.

Choose ICs for hc-ICA This button will open up a display window (Figure 4). Here the user can cycle through the estimated ICs for the initial guess. ICs that are not of interest can be deselected; deselected ICs will not be included in the analysis.

Re-estimate intial values This button will be disabled unless ICs were deselected. If ICs have been deselected, this button will use GIFT to re-estimate the intial guess for the EM algorithm after removing the effect of the de-selected ICs.

View Reduced Initial Guess This button will be disabled unless ICs were deselected. After re-estimating an inital guess using the reduced set of ICs, this button will display the new estimates for inspection.

After all preprocessing is complete, select "Save setup and continue" to move to the analysis panel.

1.2 Panel 2 - analysis

The analysis panel (Figure 6) allows the user to select the settings for the EM algorithm and carry out the analysis. There are only a few settings for this panel:

Max Iterations The maximum number of iterations for the EM algorithm. If this number of iterations is reached, the algorithm will terminate even if the stopping criteria has not been met.

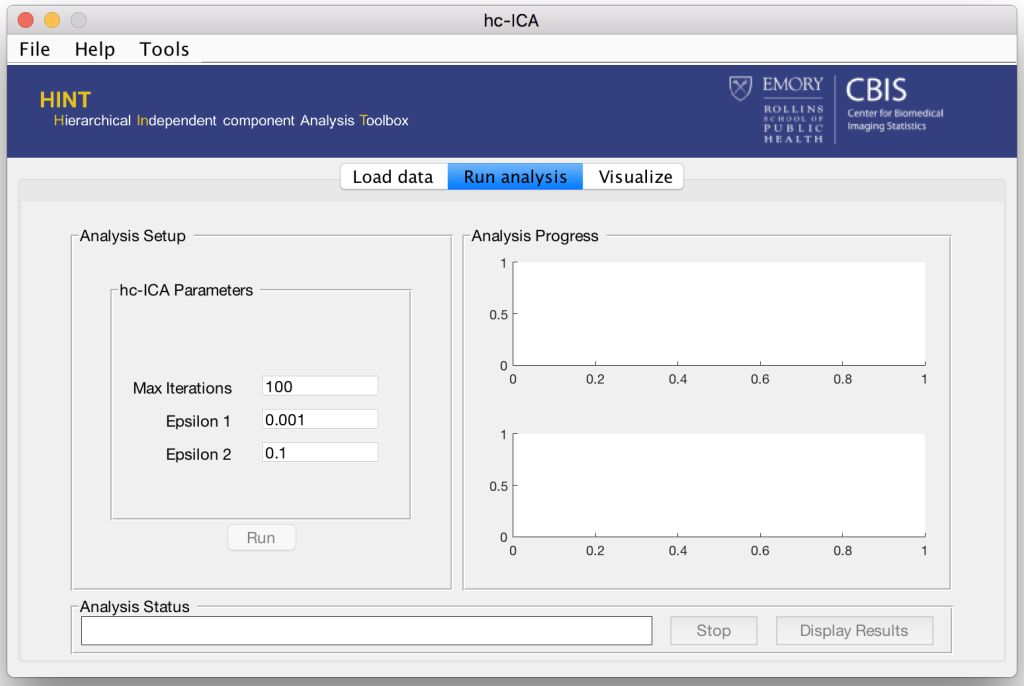


Figure 5: The HINT analysis panel.

Epsilon 1 The normalized change in all parameters other than the regression coefficients required to terminate the EM algorithm.

Epsilon 2 The normalized change in the regression coefficents required to terminate the EM algorithm.

After pressing the "Run" button the EM algorithm will start estimating the parameters. Progress plots of the change in parameter values are provided in the panel, allowing the user to track the progress of the algorithm. If at any time the user desires to stop the EM algorithm, pressing the "Stop" button will cause the algorithm to terminate at the end of the current iteration.

1.3 Panel 3 - visualization

The final panel is for viewing results. If coming to this panel after completing an analysis in the same session, the results with already be loaded and ready for viewing. Otherwise, the results must be loaded. To load the results of a HINT analysis:

1. Click "Browse" and navigate to the foldering containing the "runinfo.mat" file for the analysis of interest. If there are multiple runinfo files in this folder, a selection window will appear asking for the prefix of the analysis of interest.

2. Click "Load Results"

After loading the data, select from four display window types to begin viewing. The available windows are:

Aggregate display maps The overall average map across all subjects for each IC. This window can also be used to create masks.

Sub-population display maps View the estimated IC maps for different combinations of covariate values. This window also allows multiple sub-populations to be compared.

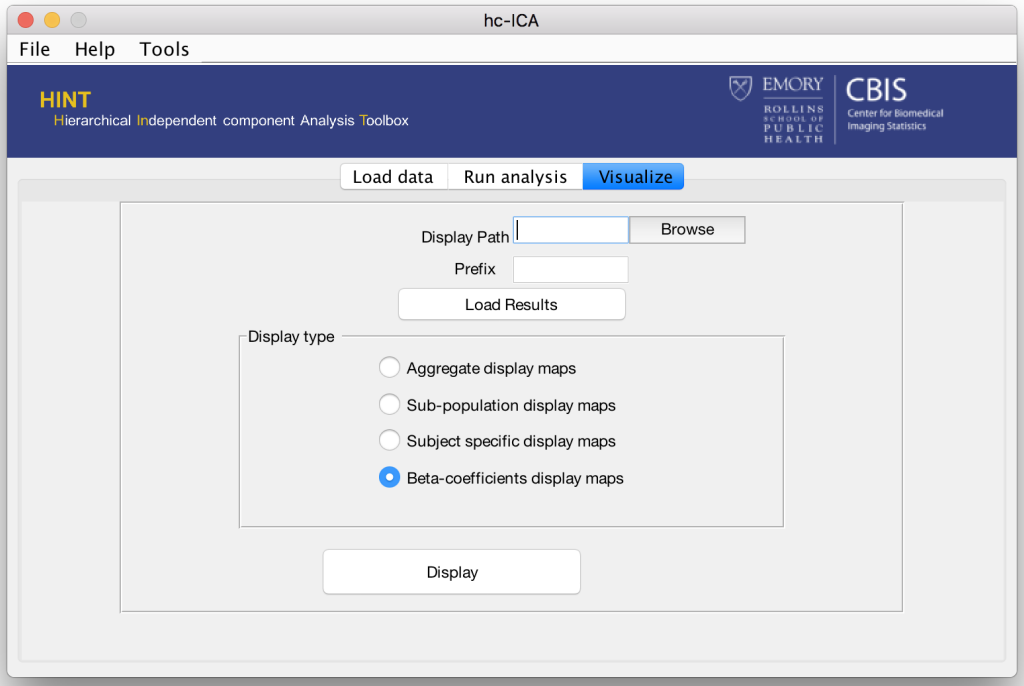


Figure 6: The HINT visualization panel.

Subject specific display maps View the estimated IC map for each subject.

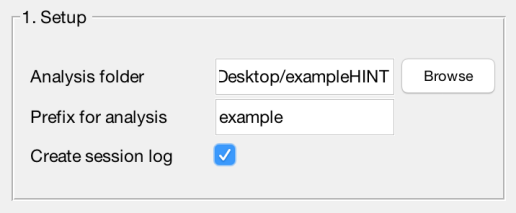
Beta coefficient display maps View the estimated covariate effect maps and test contrasts of the covariates.

Each of these windows is demonstrated in the Example section.

2 Example

Below we provide a walkthrough of the toolbox using simulation data from three slices of the brain. We start by selecting an output folder and a prefix for the analysis:

2.1 Setup



All of our output will be written to this "exampleHINT" folder with the prefix "example."

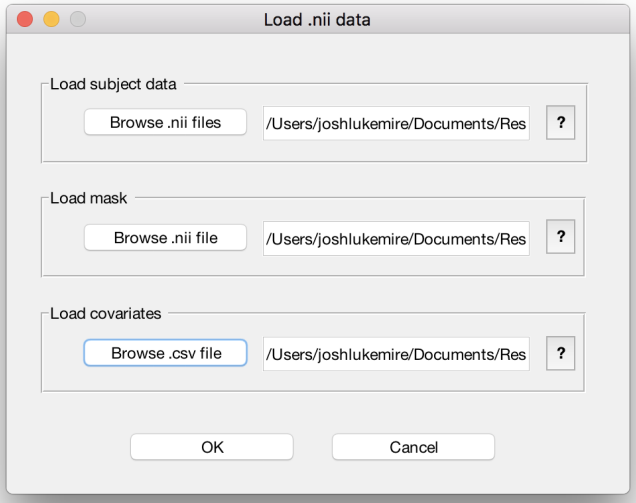
Next, we need to load the data, the design matrix, and our mask. In order to proceed, we need to make sure that our covariates csv file is formatted correctly. The first row of the file should contain "Subject" in the first column and the covariate names in the remaining columns. Each of the following rows will have the subject name in the first column and their covariate settings in the remaining columns. Two important notes:

1. The "subject names" are the filenames for the nii files. For example, if the subject data is stored in "john.nii" then the corresponding row of the covariate file should have "john" in the first column.
2. Discrete covariates (treatment group, disease status, etc.) should be coded using letters, not numbers. Use, for example, "Healthy" and "Control," not 1 and 2. The toolbox will take care of the coding scheme for you.

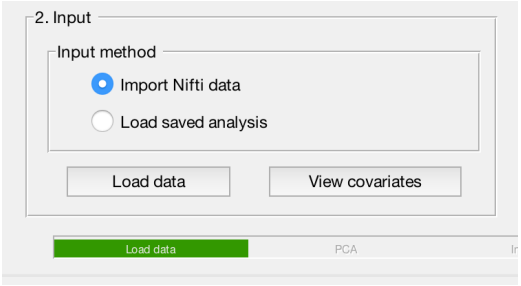
Our covariate file for this example is show below:

	A	B	C
1	Subject	Cont	Dis
2	subj1	-1.6317	a
3	subj2	3.0297	a
4	subj3	-0.99415	b
5	subj4	0.39277	b
6	subj5	-0.82774	b
7	subj6	1.0565	a
8	subj7	-1.5031	b
9	subj8	-0.60871	a
10	subj9	0.52	b

We click on "import Nifti data" then then press "Load data." After filling out each of the boxes our window looks like this:

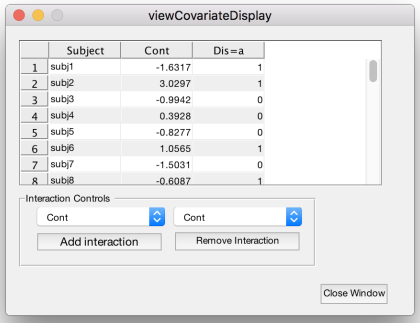


We select "Ok" and are returned to the main GUI window. If the data loaded correctly, the first part of the progress bar at the bottom of the GUI should be filled in, as seen below:



2.2 Checking covariates

Before moving on, we should check that our covariates loaded in correctly. Clock on "View covariates" to open the covariate display window.



Comparing this to our design matrix, we can see that the subjects have been matched properly to the covariates. We also notice that the discrete covariate has been reference cell coded with "b" as the baseline category (notice the column name for the final covariate). Let's add an interaction term. In the "Interaction Controls" box, we select the two covariates of interest from the drop-down boxes and click "add interaction." Now the covariate window looks like this:

	Subject	Cont	Dis=a	(Cont)_x_(...
1	subj1	-1.6317	1	-1.6317
2	subj2	3.0297	1	3.0297
3	subj3	-0.9942	0	0
4	subj4	0.3928	0	0
5	subj5	-0.8277	0	0
6	subj6	1.0565	1	1.0565
7	subj7	-1.5031	0	0
8	subj8	-0.6087	1	-0.6087

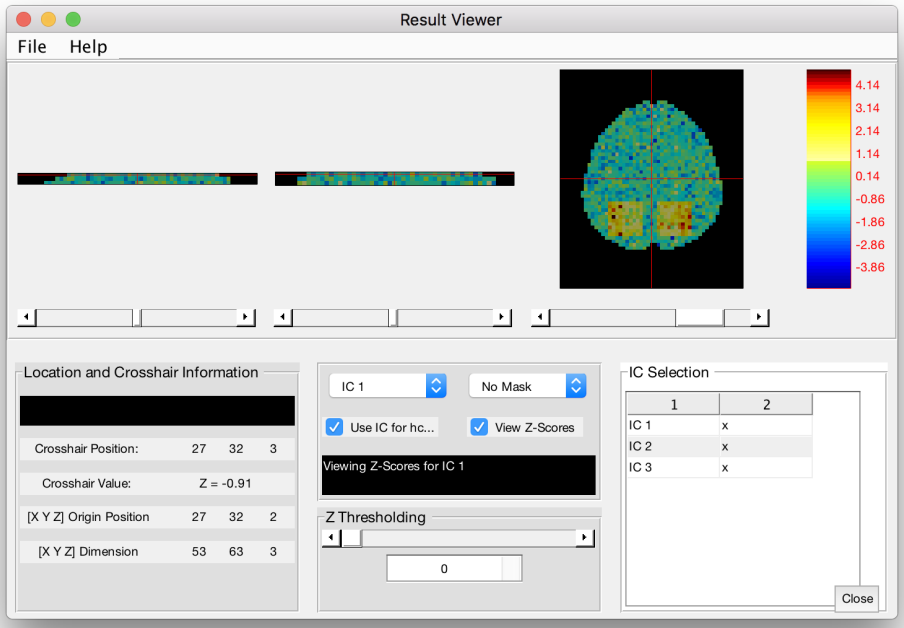
Similarly, interactions can be removed by specifying the two main effects and clicking "remove interaction."

2.3 Preprocessing

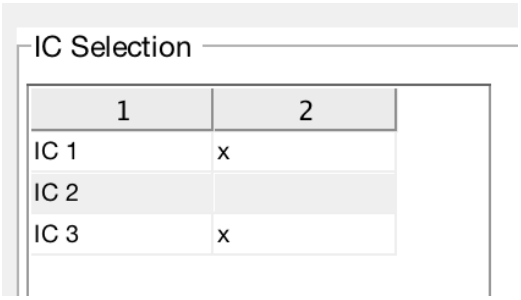
Now that the data has been properly loaded, it is time to preprocess the data and obtain an initial guess. In the "Preprocess Data" panel, we input 6 for the number of PCs and 3 for the number of ICs. The number of PCs should always be greater than the number of ICs. Then, we click on "PCA dimension reduction" to preprocess the data for the hc-ICA analysis. When this step is complete, the second part of the progress bar should turn green.

2.4 Initial Guess and IC Removal

Now that we have completed the preprocessing step, it is time to obtain an initial guess for the EM algorithm. We click on "Generate intial values," which will generate an initial set of guesses using the GIFT toolbox. As soon as the process finishes, a display window opens as seen below:



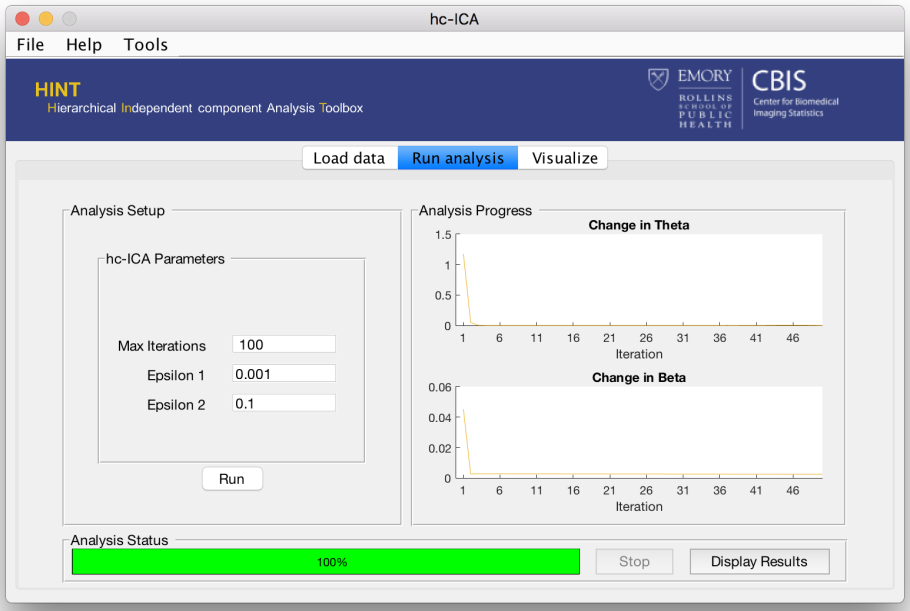
This window allows us to cycle through the different estimates by selecting them from the dropdown menu. For the sake of illustration, we pretend that we are not interested in the second IC. With that IC selected from the dropdown menu, we untick the box labeled "use IC for hc-ICA." Now our IC Selection box looks like this:



This means that the IC will not be used for the final hc-ICA analysis. We click "close" to close the display window. Because we have changed the IC set for the analysis, we need to re-estimate the initial guess, This step would not be required if we had not decided to remove any ICs. We click "Re-estimate intial values" to obtain the new guess. We can then view the new initial guess using the "View reduced intial guess" button if we so desire. At this point, the entire progress bar at the bottom of the screen should be filled out. We can click "Save setup and continue" to move on to the analysis panel. Note that clicking this button will save a "runinfo" file, which can later be used to load this saved analysis and avoid repeating all of the initial work.

2.5 Running the EM algorithm

The analysis panel is simple to use. We select a maximum number of iterations (100) and two convergence criteria (epsilon 1 and epsilon 2) and then click "Run." We watch the algorithm progress using the two progress plots on the right hand side of the panel. When the analysis is complete, the panel should look similar to this:

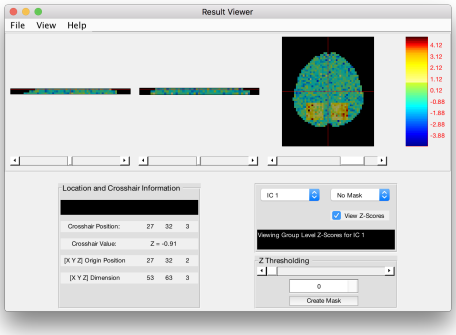


We can now move on to view our results. Clicking "display results" will take us to the final panel.

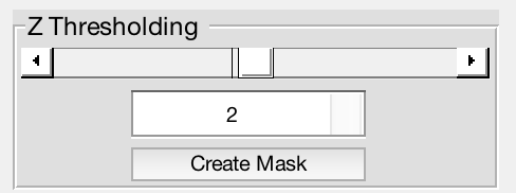
2.6 Viewing the results

2.6.1 Aggregate Maps

In the final panel, we have the chance to view our outcomes. We start by loading the aggregate display maps, which show us the overall average for each IC. These maps will be similar to if we had just run a single group ICA that ignored covariates. The initial view is shown below. We can see two squares in the posterior part of the brain, this is consistent with the way the data were simulated.



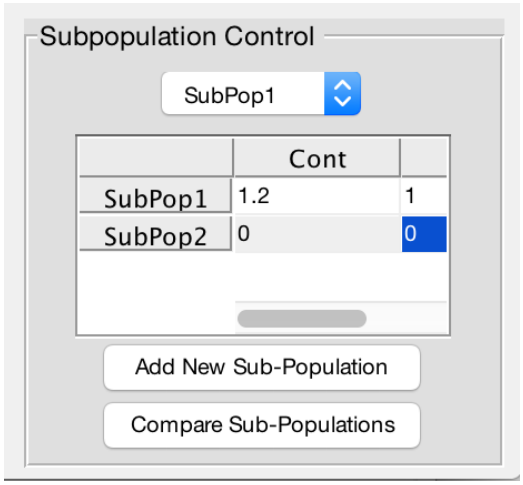
Let's create a mask based on this image. We go down to the Z-thresholding slider and move it to 2.0. Then we click on "create mask" to generate the mask file.



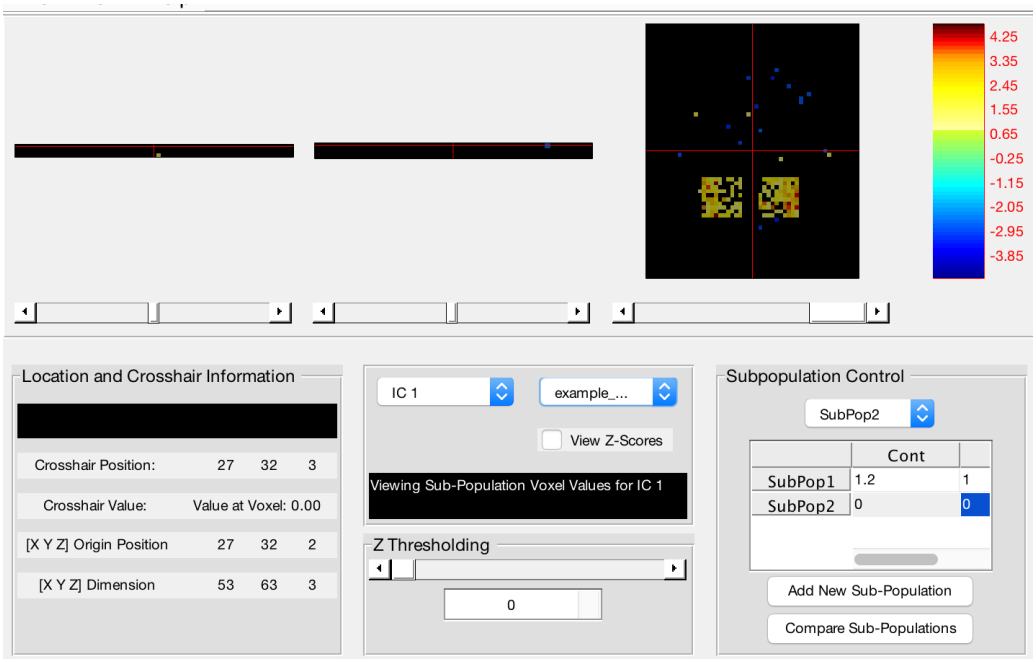
We will see that we can apply this mask to other maps shortly.

2.6.2 Sub-populations

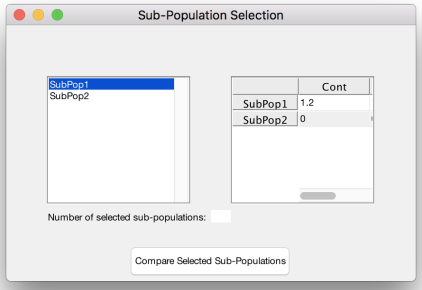
Next let's look at a couple sub-populations. We click on "View" in the toolbar and then select "Sub-population." The display window will change and we will see a new "Subpopulation Control" pane. This will allow us to manage the sub-populations we look at. First let's specify two groups, giving one a setting of 1.2 for the first covariate and 1 for the second covariate (corresponding to being in the "a" group) and giving the other a setting of 0 for the first covariate and 0 for the second covariate ("b" group). We do this by clicking "Add new sub-population" twice and filling out the fields in the table.



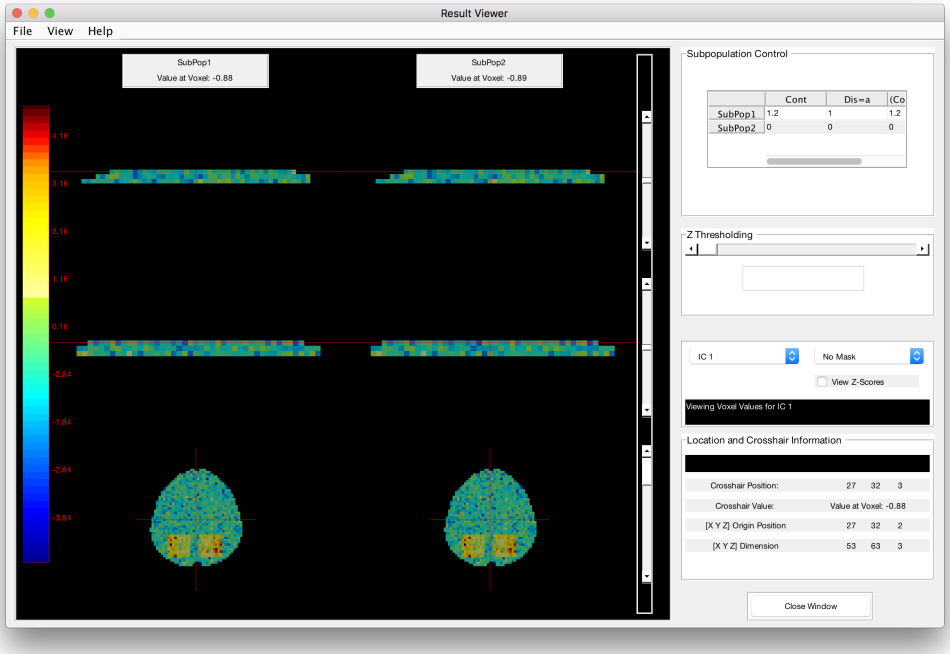
Before we compare the sub-populations side-by-side, we can look at them individually by selecting them in the dropdown menu. We can also apply masks created from the aggregate viewer if we desire. Below is an example of loading the second sub-population and applying the mask we created earlier.



Now we are ready to view the two sub-populations side-by-side. We click on "Compare sub-populations" to open up the selection window.



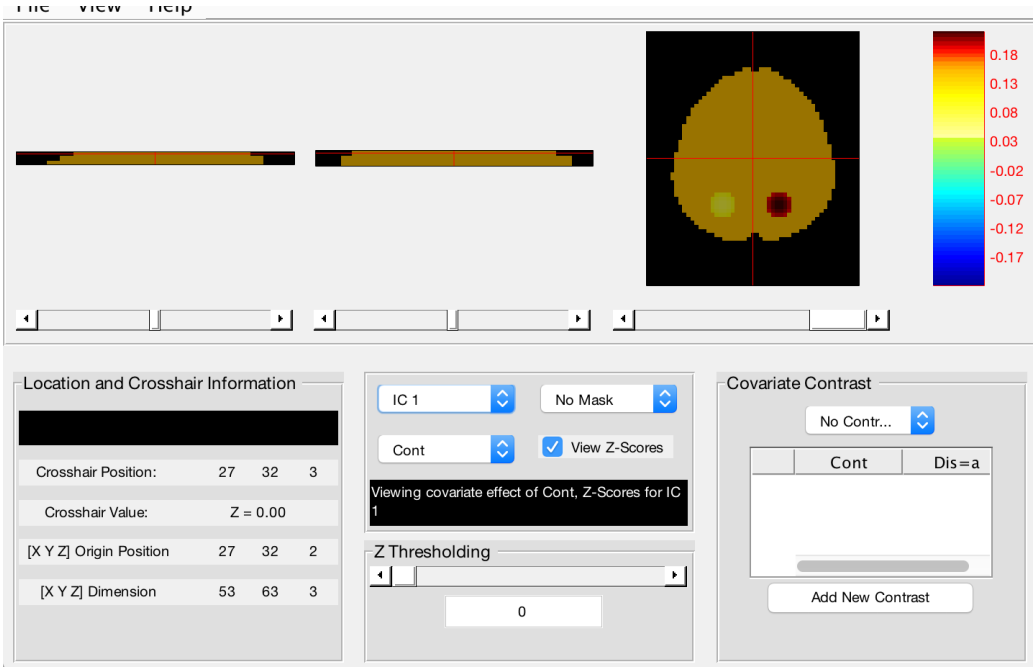
We select "SubPop1" and "SubPop2" in the left hand pane and then click on "Compare selected sub populations" to launch the viewer.



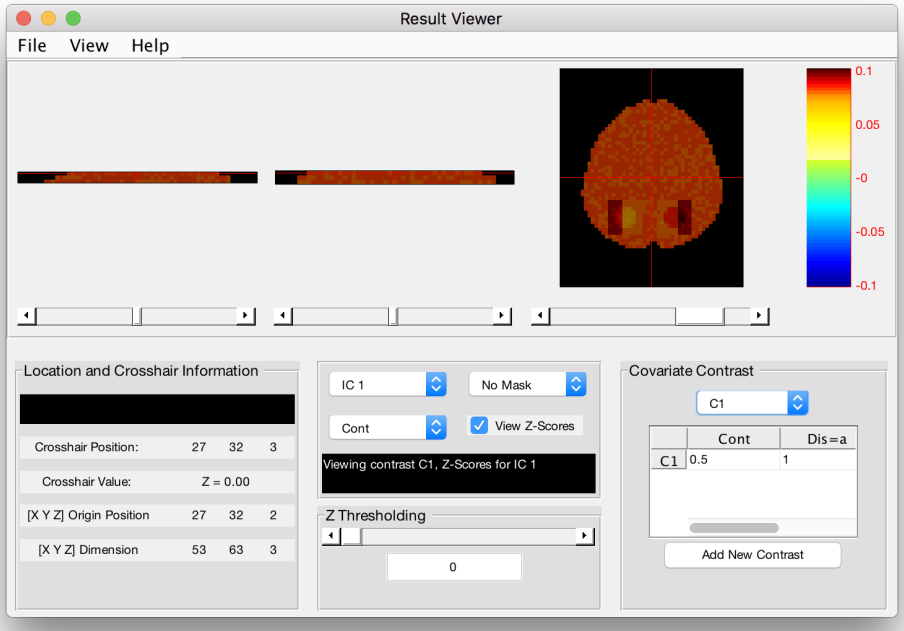
This view works identically to the other viewers, but allows two populations to be viewed simultaneously.

2.6.3 Covariate Maps

Finally, let's look at the estimated covariate effects. We click "View" again and switch to the covariate viewer. Here we have two options. First, we can look at the covariate effects individually; this is the default view. For example, in this image we are looking at the effect of the first covariate on the first IC. There are two circular shapes in the posterior section of the brain slice. This is consistent with how the data were generated.



We can also look at contrasts of the covariates. The pane on the right hand side of the display window allows us to specify contrasts in the same way we specified sub-populations before. Note that any interaction terms will be handled automatically. For example, we can look at the contrast of 0.5 times the continuous factor and a setting of "a" for the discrete factor.



If at any point you are unsure what you are looking at, the black box at the bottom of the control panel displays what the current map represents.

3 Other Functionality

3.1 Running a non-GUI version of the EM algorithm

For a large number of subjects or independent components, it is likely that a cluster will be used. In this case, all of the preprocessing and initial guess estimation should be completed using the GUI, as usual. The EM algorithm can then be run from a script. An example is provided below:

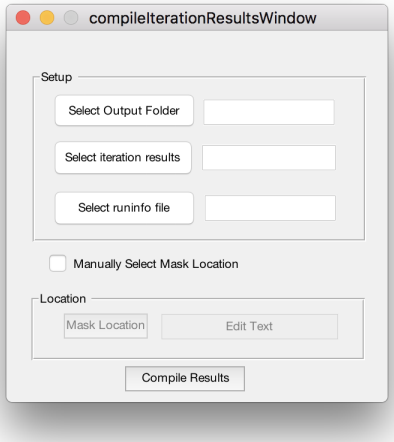
```
% add the path to the toolbox
addpath(genpath('path/to/HINT'))

% Run the estimateFromSavedData function
estimateFromSavedData('/path/to/runinfo/prefix_runinfo.mat', 'approx', 100, 0.001, 0.01)
```

The arguments to the function are (1) The path to the runinfo file, (2) the type of algorithm, this should be left as 'approx' until a future release of the toolbox, (3) the maximum number of iteractions, and (4), (5) the values for epsilon 1 and epsilon 2 respectively.

3.2 Compiling Intermediate Iteration Results

If the algorithm is terminated early, or for some other reason intermediate results are desired, HINT provides a tool for compiling the iteration output from the EM algorithm. Click on "tools" and "compile iteration results." A GUI window will open that looks like this:



To use the window:

1. Select an output folder using "Select Output Folder." The EM results will be stored here
2. Click "Select iteration results" and navigate to the iteration_results.mat file containing the EM output.
3. Click "Select runinfo file" and navigate to the runinfo file for the analysis.
4. If the mask for the analysis has been moved, then check "Manually select mask location" and navigate to the correct mask location, otherwise the path specified in the runinfo file will be used.

After filling out all the fields and pressing "compile results," the HINT toolbox will generate all appropriate output in the specified folder.