The Open EEGLAB portal

Arnaud Delorme, Amitava Majumdar, Subhashini Sivagnanam, Ramon Martinez-Cancino, Kenneth Yoshimoto, Scott Makeig

Abstract— **The EEGLAB signal processing environment is a widely used open source software environment for processing electroencephalographic (EEG) data. The Neuroscience Gateway (*nsgportal.org*) is a software portal allowing users to readily run a variety of neuroimaging software on high performance computing (HPC) resources. We have expanded the current Neuroscience Gateway (NSG) services to enable researchers to freely run EEGLAB processing scripts and pipelines on their EEG or related data via the Neuroscience Gateway. This Open EEGLAB Portal is open to all for use in nonprofit projects and allows researchers to submit unimodal or multimodal EEG data for parallel processing using standard or custom EEGLAB processing pipelines. A detailed user tutorial is available (*sccn.ucsd.edu/wiki/EEGLAB_on_NSG*). As a proof of concept, we apply an EEGLAB pipeline to freely available 128-channel EEG data from 1,097 participants in the Child Mind Institute Healthy Brain Network project (*childmind.org/center/healthy-brain-network*).**

I. INTRODUCTION

Although scalp electroencephalography (EEG) has now been used to observe brain activity supporting human thought and action for over 80 years, its utility has often been perceived as relatively limited. Despite its strong advantages over other brain imaging methods in speed, directness, cost, and portability, EEG research and development have in recent decades been overtaken in use and interest by newer, spatially more detailed, but much slower, less direct, and also significantly more expensive brain imaging modalities, most notably fMRI. However, at least three factors -- 1) the disadvantages of measuring brain activity indirectly by measuring relatively slow brain metabolic processes, 2) availability of new methods enabling cortical source-resolved imaging of EEG data, and 3) the ongoing development of highly mobile, low cost EEG systems allowing data collection under a wide range of natural conditions -- are prompting renewed interest in employing non-invasive, EEG and multimodal brain imaging for a range of purposes.These include discovery of and subsequent cost-effective screening for biomarkers of many neurological and psychiatric conditions [1, 2]

The application of and research into source-level EEG analysis and application of complex statistics to high-density

EEG and to multimodal brain imaging data including EEG is currently limited by the relatively high computational demands posed by large and complex data sets. Further, these demands are ever growing as ever more sophisticated methods of EEG source imaging and brain dynamic modeling are introduced. The EEGLAB community has pioneered the use of highly-efficient GPU processing for ICA decomposition [3]. However, even such solutions are not commonly available to EEG researchers who also may not have the technical expertise available to install and use them. Although publicly available HPC resources exist at national academic supercomputer centers, access by neuroscientists to HPC resources is limited by both administrative and technical barriers including the steep learning curve required to understand HPC hardware, software, and policy environments, to install and run applications, and the need to write peer-reviewed proposals every year for HPC compute time.

The Neuroscience Gateway (NSG) [4-8], an NSF-funded project supporting the computational neuroscience community, greatly reduces these barriers to accessing and mastering use of HPC resources. NSG was first built to serve computational neuroscientists who need access to HPC resources for simulations of large scale neuronal networks and data analyses that involve complex modeling and computing steps including high-dimensional optimization and parameter space exploration. Currently, through a simple web-based portal (*nsgportal.org*) NSG provides a streamlined environment for uploading such models, specifying HPC job parameters, querying running job status, receiving job completion notices, and storing and retrieving output data. On its back end, NSG transparently distributes user jobs to appropriate HPC resources. NSG can be accessed via the NSG web portal or programmatically via a RESTful service programming interface. NSG currently provides neuronal network simulation tool implementations for NEURON, GENESIS, PyNN, BRIAN, MOOSE, and NEST and, more recently, for two widely used tools for analysis and visualization of structural and functional brain imaging data, FreeSurfer (freesurfer.org) and MRtrix. The NSG project also supports a custom pipeline [9] for the Human Connectome research community and BluePyOpt (released by the EU Human Brain Project) an open source framework for data-driven model parameter optimization. Since the NSG portal came online in early 2013, it has been well received by the computational neuroscience and functional brain imaging communities, with currently more than 750 registered research users. NSG leverages UCSD Comet and UT Austin Stampede2 supercomputers, both part of the Extreme Science and Engineering Discovery Environment (XSEDE) (*xsede.org*); NSG also utilizes the XSEDE Jetstream cloud resource at the University of Indiana. A yearly HPC compute time allocation has been funded each year by NSF at increasing levels of support. For the 2018 Federal fiscal year alone, NSG is providing

neuroscience researchers over 10 million core hours of HPC computation (the equivalent of 1,140 cores running continuously through the year).

Our aim in the Open EEGLAB Portal project is to leverage unprecedented access to HPC computing power via NSG to overcome barriers that now increasingly keep advanced EEG brain imaging and analysis approaches from being applied outside of the few EEG laboratories that run sophisticated cluster computing facilities. These tools are essential to enable continued development of more precise, computationally intensive data models, for performing statistical analyses across large studies, and for performing (for the first time) meta-analyses across EEG studies.

The Open EEGLAB Portal, built on the NSG framework, now encourages use of more sophisticated and accurate methods in EEG analysis practice, and thus can hasten the transition from lower-precision scalp-channel EEG modeling to higher-precision and more informative source-resolved 3-D electrophysiological imaging and multimodal brain/behavior imaging [10, 11] that takes better advantage of the mobility, low cost, and high temporal resolution of EEG recording. The Open EEGLAB Portal makes feasible wider application of multimodal EEG brain imaging and MoBI approaches. It now enables many more laboratories to test and use more flexible and computationally complex methods for performing time/frequency decomposition, blind source separation [12], and multivariate statistics on individual and group data. As well-annotated EEG data become increasingly available in public archives, meta-analysis of source-resolved EEG data measures across studies will soon become feasible.

II. METHODS

A. Capabilities of the Neuroscience Gateway (NSG)

NSG already has in place a well-established and tested user interface, user account management system, documentation, and ticketing system for bug reports and resolution, a user support mechanism, a training/workshop hosting track record, and policies in place to fairly distribute HPC time to users -- and given that all of these have been in production use for over five years -- the NSG is ideally positioned to give the EEG research community access to HPC resources.
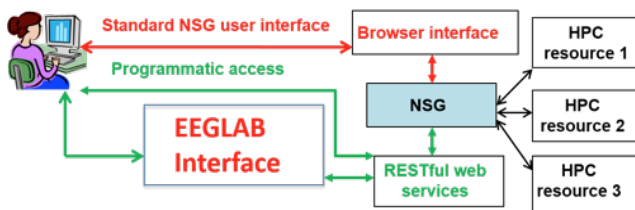

Figure 1. Schematic view of the Open EEG Portal (OEP).

NSG is based on the CIPRES Workbench [13], a flexible software framework for creating science gateways that has been used to create many robust and heavily used science gateways in various domains including phylogenetics, and bioinformatics. Some of these gateways provide many thousands of users access to large numbers of tools running on HPC resources. A recent addition has allowed the development of programmatic access via a software infrastructure [14] using Representational State Transfer ("REST"). The NSG-R (R for REST) uses web services that expose NSG access capabilities via publicly available application programmer interfaces (APIs). We have now developed and made publicly available an EEGLAB implementation of the standard NSG user interface portal, and will soon release an NSG-R API implementing a programmatic access interface.

Figure 1 above shows schematically how EEGLAB users can use the NSG to run EEGLAB analysis pipelines on HPC resources. Below we describe the four work phases to enable EEGLAB users to seamlessly use HPC resources for complex data analyses and meta-analyses running EEGLAB-specific pipelines on the NSG back end.

B. The NSG User Interface for EEGLAB

The first phase OEP now enables EEGLAB users to use the NSG via its web portal: users first acquire an account on the NSG and then interact with NSG through a web browser (Fig. 2). The NSG EEGLAB interface includes an interface for uploading EEGLAB-specific input data, setting EEGLAB-specific processing parameters, and specifying HPC parameters for jobs that are then submitted by NSG to HPC resources on its back end. Via the NSG Tool Module, EEGLAB jobs run on HPC resources readies results for retrieval upon job completion. EEGLAB-specific data may be either raw or preprocessed EEGLAB data in EEGLAB-formatted data files, uploaded to the NSG interface as a single compressed zip file along with an EEGLAB analysis script. HPC resources run EEGLAB scripts written in MATLAB (The Mathworks, Inc.) or in open source Octave (*gnu.org/software/octave*).

The NSG back end engine 1) unzips the input zip file, 2) runs the contained EEGLAB script, 3) recompresses the process folder into an output zip archive, and 4) sends the user an email notice that it is available. Users report this user interface to be useful and easy to use.

C. An EEGLAB REST interface to NSG (NSG-R)

The Open EEGLAB Portal (OEP) will soon expose a set of RESTful web services (now in advanced development) that will provide users the capabilities to access NSG-R via a REST API, in addition to the standard NSG web browser portal interface. The NSG-R web services interface, implemented within the existing NSG CIPRES Workbench framework, will ensure that researchers can readily access HPC resources transparently with low overhead directly from EEGLAB sessions running on Web-connected laptops and desktops (Fig. 3). Researchers will use the client NSG-R web services to automatically initiate EEGLAB processing pipelines on HPC resources, and to download and interact with results within the EEGLAB desktop environment, with minimum user interaction. Analyses will be able to reuse directly their own data files that have been previously uploaded to NSG-R. The RESTful web services capability will allow the seamless functional integration between HPC

resources and EEGLAB and could, we believe, have a broad impact on productivity in advanced EEG brain imaging research.
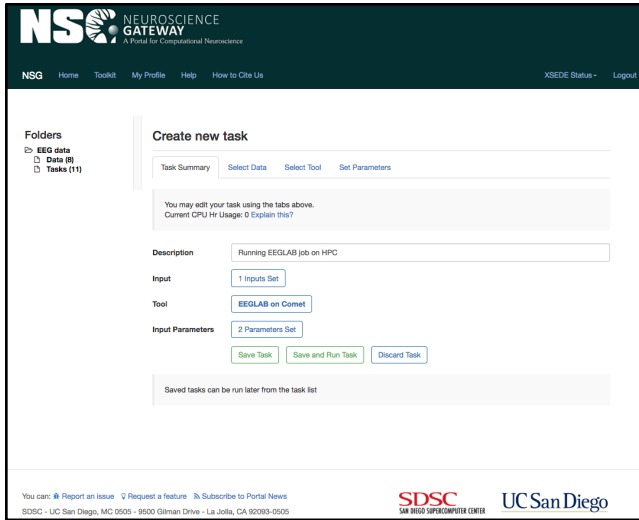

Figure 2. The NSG web interface to EEGLAB.

For this purpose we plan to implement a number EEGLAB pipelines in typical use by advanced EEGLAB users, allowing users to run these pipelines via the OEP plug-in. All the dependencies of various EEGLAB modules that will constitute these pipelines are being implemented in the NSG-R back end on HPC resources and will be provided online (*sccn.ucsd.edu/wiki/EEGLAB_on_NSG*). We also plan to build tutorials and tools allowing more advanced EEGLAB users to easily implement new OEP pipelines.

Figure 3. An EEGLAB interface to NSG-R (under development).

III.   PROOF OF CONCEPT APPLICATION

As a simple example of the Open EEGLAB Portal, we processed EEG data collected and made public by the Child Mind Institute Healthy Brain Network project. We used 1,097 subjects from this release as a proof of concept, here comparing spontaneous EEG dynamics during eyes open and eyes closed rest across developments. The data were uploaded and processed on NSG with custom EEGLAB scripts (EEGLAB v15.0). Data preprocessing was first performed by an EEGLAB script that performed these steps:

1. Download data (2Gb per subject)
2. Import the (128-channel, <10-min, 250-Hz) raw MFF EEG data files (mffmatlabio plugin of EEGLAB) for eyes open/eyes closed recording session.
3. High pass filter the data above 1 Hz (non-causal zero-phase; transition band width, 1 Hz; -6-dB cutoff frequency).
4. Convert data to average-reference and concatenate the two datasets per participant.
5. Clean data of noisy scalp channels and data periods using the EEGLAB *clean_rawdata* plug-in (v 0.34) with default parameters.
6. Recompute the average reference data.
7. Run Infomax ICA decomposition.
8. Apply the ICLabel plug-in (v0.3) [15] to separate brain-based independent component (IC) processes from ICs with high probability of accounting for muscle, eye, channel or line noise, or other activities.
9. Back-project the brain-based ICs to the scalp channels to create artifact-cleaned scalp data.
10. Extract 1-sec data segments during periods in which participants were asked to rest with eyes open or closed; ignore the first 3 secs of each rest period.

Upon submitting a job, depending on current HPC load, it might take some time for the job to be allocated to a 24-core node of the Comet supercomputer -- usually from 0 to 30 minutes wait time. MATLAB Parallel Toolbox syntax (replacing "for" loops with "parfor" loops) automatically parallelized the processing on each node with negligible overhead since there is almost no communication between the cores. By this means EEGLAB scripts and functions may be easily parallelized. In our case we used only 8 of the 24 available cores in order to remain within the 128Gb RAM limit of each node. 8 subjects were thus processed in parallel on 8 cores on each node. It took about 40 minutes to process each subject's data including downloading the data - 2Gb per subject, and it took 10 to 11 hours on 8 nodes (64 used cores total) to preprocess the 1,097 subjects and generate 2,194 datasets (one for eyes open and one of eyes closed for each subject).

Group analysis: To explore use of group analysis on this large data set, we ran a second EEGLAB script to perform power spectral decomposition on the overlapping 1-sec data segments on all datasets:

1. Interpolate channels removed during preprocessing (about 3 per subject) using spherical spline interpolation.
2. Taper the 1-sec windows using a hamming window and compute the average log power spectral density using FFT decompositions.
3. Compare spectral power within a 5-by-2 design (age group by eyes-open versus eyes-closed conditions).

Group processing was performed on a single core of Comet and took another 4 hours - no parallelization was used although it would be possible to do so. The graphics output generated by Comet in postscript format were subsequently downloaded and formatted for publication (Fig. 4).

Consistent with other findings, we found scalp alpha power (8-12 Hz) decreases with age during both eyes open and eyes closed conditions [16, 17]. After False Discovery Rate correction for multiple comparisons [18], all scalp channels showed significant effects ($p<0.001$) for both independent variables (age and condition).
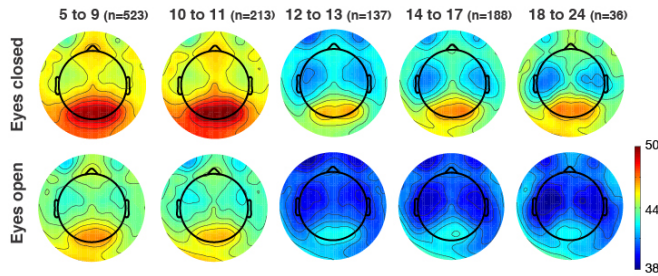

Figure 4. Results of processing 2,194 128-channel EEG datasets from the Child Mind Institute database. Alpha power (color unit, relative power in dB) decreases with age (age groups in years shown in columns) in both eyes open and closed conditions (rows).

The EEG processing here was performed purely as a proof of concept to illustrate the potential of using HPC resources for hypothesis discovery and testing in EEG neuroimaging. Further validation, interpretation, and detailed analysis would be required for standard scientific reporting. The documented analysis scripts used above are available online at (*sccn.ucsd.edu/wiki/EEGLAB_on_NSG*).

IV.   CONCLUSION

Here we have shown that using the simple NSG web interface it is possible to run EEGLAB scripts on national HPC resources by simply uploading the script and data to the NSG interface. Detailed instructions for using the Open EEGLAB Portal are available online (*via the link above*). We believe that the simplicity of this process will help popularize the use of advancing source-resolved and other advanced statistical EEG neuroimaging methods to speed scientific discovery.

V.   ACKNOWLEDGMENT

VI.   REFERENCES

1. Lenartowicz, A., Delorme, A., Walshaw, P.D., Cho, A.L., Bilder, R.M., McGough, J.J., McCracken, J.T., Makeig, S., and Loo, S.K.: 'Electroencephalography correlates of spatial working memory deficits in attention-deficit/hyperactivity disorder: vigilance, encoding, and maintenance.', The Journal of Neuroscience 2014, 34, (4), pp. 1171-1182

2. Rissling, A.J., Miyakoshi, M., Sugar, C.A., Braff, D.L., Makeig, S., and Light, G.A.: 'Cortical substrates and functional correlates of auditory deviance processing deficits in schizophrenia.', NeuroImage. Clinical, 2014, 6, pp. 424-437

3. Raimondo, F., Kamienkowski, J.E., Sigman, M., and Fernandez Slezak, D.: 'CUDAICA: GPU optimization of Infomax-ICA EEG analysis.', Computational Intelligence and Neuroscience, 2012, 2012, pp. 206972

4. Sivagnanam, S., Majumdar, A., Yoshimoto, K., Astakhov, V., Bandrowski, A., Martone, M., and Carnevale, N.T.: 'Introducing The Neuroscience Gateway', in Editor (Ed.)^(Eds.): 'Book Introducing The Neuroscience Gateway' (2013, edn.), pp.

5. Sivagnanam, S., Majumdar, A., Yoshimoto, K., Astakhov, V., Bandrowski, A., Martone, M., and Carnevale, N.T.: 'Early experiences in developing and managing the neuroscience gateway', Concurrency and Computation: Practice and Experience, 2015, 27, (2), pp. 473-488

6. Majumdar, A., Sivagnanam, S., Carnevale, N., Yoshimoto, K., Gleeson, P., Quintana, A., and Silver, R.A.: 'Neuroscience Gateway–Cyberinfrastructure Providing Supercomputing Resources for Large Scale Computational Neuroscience Research', Neuroinformatics, 2016

7. Majumdar, A., Sivagnanam, S., Yoshimoto, K., and Carnevale, T.: 'Understanding the Evolving Cyberinfrastructure Needs of the Neuroscience Community', in Editor (Ed.)^(Eds.): 'Book Understanding the Evolving Cyberinfrastructure Needs of the Neuroscience Community' (ACM, 2016, edn.), pp. 45

8. Sivagnanam, S., Yoshimoto, K., Carnevale, N.T., and Majumdar, A.: 'The Neuroscience Gateway: Enabling Large Scale Modeling and Data Processing in Neuroscience', in Editor (Ed.)^(Eds.): 'Book The Neuroscience Gateway: Enabling Large Scale Modeling and Data Processing in Neuroscience' (ACM, 2018, edn.), pp. 52

9. Schirner, M., Rothmeier, S., Jirsa, V.K., McIntosh, A.R., and Ritter, P.: 'An automated pipeline for constructing personalized virtual brains from multimodal neuroimaging data.', NeuroImage, 2015, 117, pp. 343-357

10. Makeig, S., Gramann, K., Jung, T.-P., Sejnowski, T.J., and Poizner, H.: 'Linking brain, mind and behavior.', International Journal of Psychophysiology, 2009, 73, pp. 95-100

11. Ojeda, A., Bigdely-Shamlo, N., and Makeig, S.: 'MoBILAB: an open source toolbox for analysis and visualization of mobile brain/body imaging data.', Frontiers in Human Neuroscience, 2014, 8, pp. 121

12. Makeig, S., Westerfield, M., Jung, T.||., Enghoff, S., Townsend, J., Courchesne, E., and Sejnowski, T.J.: 'Dynamic brain sources of visual evoked responses', Science, 2002, 295, (5555), pp. 690-694

13. Miller, M.A., Pfeiffer, W., and Schwartz, T.: 'The CIPRES science gateway: a community resource for phylogenetic analyses', in Editor (Ed.)^(Eds.): 'Book The CIPRES science gateway: a community resource for phylogenetic analyses' (ACM Press, 2011, edn.), pp. 1-8

14. Miller, M.A., Schwartz, T., Pickett, B.E., He, S., Klem, E.B., Scheuermann, R.H., Passarotti, M., Kaufman, S., and O'Leary, M.A.: 'A RESTful API for Access to Phylogenetic Tools via the CIPRES Science Gateway.', Evolutionary Bioinformatics Online, 2015, 11, pp. 43-48

15. Pion-Tonachini, L., Makeig, S., and Kreutz-Delgado, K.: 'Crowd labeling latent Dirichlet allocation', Knowledge and Information Systems, 2017, 53, (3), pp. 749-765

16. Miskovic, V., Ma, X., Chou, C.-A., Fan, M., Owens, M., Sayama, H., and Gibb, B.E.: 'Developmental changes in spontaneous electrocortical activity and network organization from early to late childhood', Neuroimage, 2015, 118, pp. 237-247

17. Segalowitz, S.J., Santesso, D.L., and Jetha, M.K.: 'Electrophysiological changes during adolescence: a review', Brain and cognition, 2010, 72, (1), pp. 86-100

18. Benjamini, Y., and Yekutieli, D.: 'The control of the false discovery rate in multiple testing under dependency', Annals of statistics, 2001, pp. 1165-1188