Project plan – master’s thesis, 21/08-2020

To comply with the program specifications for the master’s education in biomedical engineering at DTU, this document contains a plan for the project period of the master’s thesis along with aims and motivation for this. It also addresses the learning objectives for a master’s thesis at DTU. The project plan is to be handed in within the first month of the project period, and a revised version will be enclosed with the final thesis hand-in.

Name:	Anders Stevnhoved Olsen
Program:	Biomedical engineering
Project period:	03/08-2020 to 03/02-2021
Project title (English):	Quantification of sleep states using dynamical modeling of functional neuroimaging data
Project title (Danish):	Kvantificering af søvnstadier ved brug af dynamisk modellering af data fra funktionelle hjerneskanninger
Supervisor:	Morten Mørup (Ph.D., Professor, DTU Compute)
Co-supervisors:	Sebastian Holst (Ph.D., PostDoc, NRU), Patrick Fisher (Ph.D., Senior researcher, NRU)

Project motivation

Sleep has historically been characterized using polysomnography (PSG), for which guidelines for staging were established already in 1968[1]. Although a viable tool that has aided neurological research in many areas, PSG is burdened by its lack of spatial resolution. Vigilance state characteristics, including sleep graphoelements such as K-complexes and wave frequencies, can be revealed using electroencephalography (EEG), however, the method lacks important spatial information that can help reveal the importance and effects of specific brain areas in vigilance states and the transitions between them.

Using functional magnetic resonance imaging (fMRI), and specifically the ultra-fast sequences magnetic resonance encephalography (MREG)[2] and multiband (MB) that create 10 and 4.6 images per second, respectively, researchers have the opportunity to combine spatial and temporal information in brain studies.

This project will utilize unsupervised dynamical machine learning models to compare the information provided in fMRI-data to the known PSG stage divisions. For this purpose, we have at our disposal a data set of 20 subjects that have participated in three sessions of simultaneous EEG-fMRI recordings. The first session was a baseline scan late afternoon after a normal night of approximately 8 hours of sleep. The two subsequent sessions were acquired after a sleep deprivation period of 32 hours and the administration of either Carvedilol (an adrenergic antagonist) or Placebo. Each session consisted of 10-15 minutes of wakefulness and 30-50 minutes of sleep opportunity and the sequences used for recording were a mix of MREG and MB. The fMRI and the

[1] *The visual scoring of sleep in adults* (Silber et al., 2007)
[2] *Single shot whole brain imaging using spherical stack of spirals trajectories* (Assländer et al., 2013)

PSG have been preprocessed, and the PSG has been scored by two experts according to standard guidelines.[3]

Stevner et al. (2019)[4] implemented Hidden Markov Models (HMM) on a large data set of sleep EEG-fMRI, attempting to group the fMRI-data into a discrete set of states, subsequently comparing these HMM-states to the expert-labeled PSG stages. They selected a number of HMM-states (19) and analyzed the specificity and sensitivity of these states to the PSG stages. The study then examined state transition probabilities and grouped states into modules similar to the conservative sleep stage division, including N1-N3 sleep and wakefulness, in the end presenting spatial activation distributions of these modules.

This study will attempt to extend the knowledge provided in Stevner et al. (2019). In this quest, the starting point will be to implement the HMM from the article on our data, keeping as many variables constant as possible. From here, the project will look at the importance of having a high temporal resolution, the advantages of having switching dynamics, and what the results show at the group analysis level. This is summarized in the following research questions.

Project aims

- With data from ultra-fast fMRI sequences of wakefulness and sleep, can we reveal vigilance states using dynamical modeling, and what do we gain from having a high temporal resolution?

This research question may be regarded as an approach to the Stevner et al. (2019) study. Several experimental conditions are different between the two studies, chief among which the MR-sequences used in data recording. The data in Stevner et al. (2019) stems from a BOLD sequence acquired with a repetition time (TR) of 2.4s whereas we in our study have MREG (TR=0.1s) and MB (TR=0.215s) sequences. Both these are considered to be comparable to BOLD, but a thorough investigation of the literature in this matter must be carried out to ensure this.

Thus, the comparison between the two studies cannot be regarded as a replication. However, what we can do is to implement the model and subsequent analyses presented in the article, including the temporal HMM-state specificity and sensitivity to PSG stages, graph models of state transitions and spatial activation distributions. Once this analysis pipeline has been established, downsampling our data to the same TR as in Stevner's study may enable us to see potential advantages of having a high temporal resolution in the data.

To compare the two studies, it may be advantageous to use the subset of our data that resembles theirs the most. This could be the baseline session, the placebo sleep-deprivation session, or both while avoiding the Carvedilol activation session. The caveat of using both sessions is that the subjects remain in two different conditions during scanning, namely sleep-deprived and baseline. This may result in different brain activation of both wakefulness, sleep and the transition between these. This may necessitate an extension of the pool of vigilance states to include two types of wakefulness, of which the spatiotemporal characteristics of the second one may resemble non-REM sleep to a degree.

---

[3] *Rules for scoring respiratory events in sleep: Update of the 2007 AASM manual for the scoring of sleep and associated events* (Berry et al., 2012)

[4] *Discovery of key whole-brain transitions and dynamics during human wakefulness and non-REM sleep* (Stevner et al., 2019)

- At the group analysis level, what are differences between the three conditions of which the data originate (Baseline, sleep deprivation + Carvedilol, sleep deprivation + Placebo)

The potential brain activation differences between sleep deprivation and baseline along with the differences between adrenergic inhibition in the locus coeruleus and placebo may be reflected at the group analysis level.

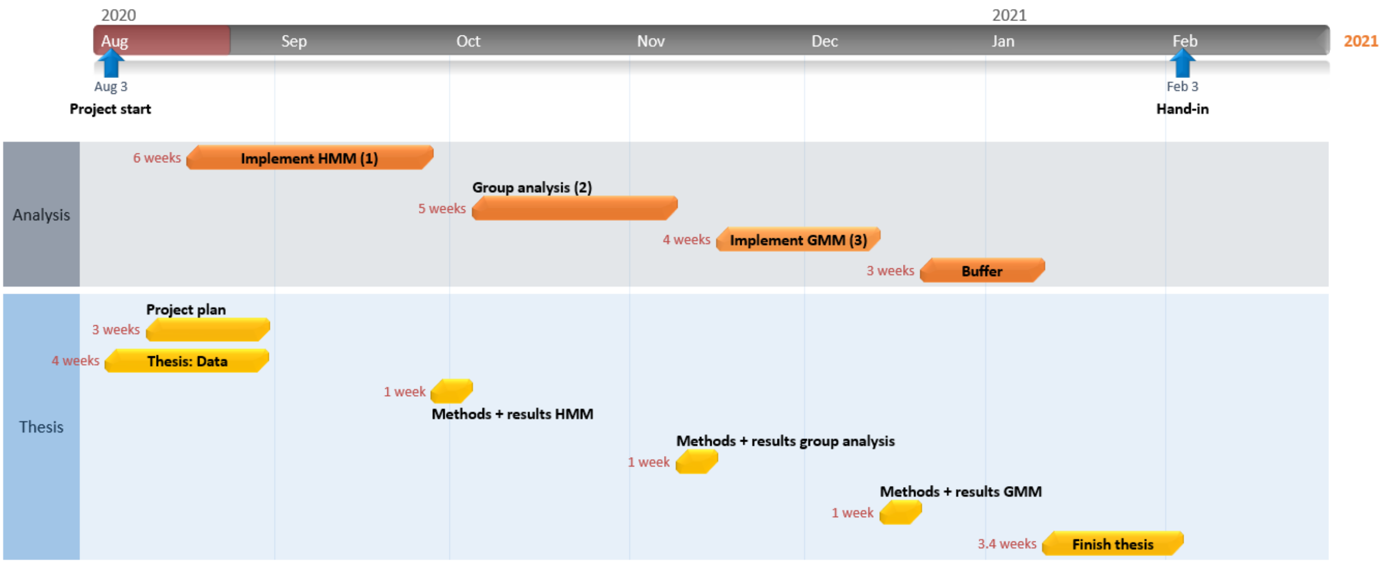
Three state-specific measures are often used when comparing results at this level: dwell time, occupancy, and transition matrices. For this research question, models will be trained using all the pooled data, but condition-specific parameters will be evaluated independently and compared statistically. The analysis pipeline established in the previous aim will be used to provide spatiotemporal characteristics of all three conditions.

- What are the advantages of modeling state transitions, and do Gaussian mixture models provide as much information as HMM?

Both HMM and GMM are unsupervised machine learning tools that model the discrete set of latent variables that could have generated the data. The difference between the two is that GMM does not assume any smoothness in trajectories. HMM can model transition dynamics, occupancy and dwell time, whereas GMM can only show occupancy and dwell time. Evaluating the results from both types of models trained on the same data may reveal the potential advantages of HMM against the simpler GMM.

Project timeline

The project timeline is summarized in the Gantt chart below. The three research questions are to be worked on in sequence, each concluded with a week of thesis writing. The plan includes a 3-week buffer period before finishing the thesis.



Learning objectives of the thesis

The following learning objectives (in orange) are sought fulfilled in the final thesis.

A graduate of the MSc programme from DTU:

- can identify and reflect on technical scientific issues and understand the interaction between the various components that make up an issue
  Several potential technical issues have already been addressed in this project plan including the question of what data to use for modeling. Also, it has previously been shown that HMM suffers from generalization issues when trained across multiple subjects. These issues, among others, will be reflected upon in the final thesis.
- can, on the basis of a clear academic profile, apply elements of current research at international level to develop ideas and solve problems
  The comparison between Stevner et al. (2019) and the data in our study will include analyses also published in Nature Communications, a high-standard international journal. Literature reading will be ongoing throughout the project period, and elements from high-impact articles will be considered in this project to develop ideas and analyze results.
- masters technical scientific methodologies, theories and tools, and has the capacity to take a holistic view of and delimit a complex, open issue, see it in a broader academic and societal perspective and, on this basis, propose a variety of possible actions
  Hidden Markov Models and Gaussian Mixture Models are essential modeling tools in this project and will be studied heavily throughout the project. This includes their mathematical origin, potential limitations and the effect of these in our results. It is expected that the implementation of these tools necessitates tweaking of parameters to understand (negative) results and to suggest alternatives. Moreover, it is the hope that this thesis provides scientific results that may create opportunities for new studies. If this is the case, an effort will be taken to propose new ideas.
- can, via analysis and modelling, develop relevant models, systems and processes for solving technological problems
  In this project, the first research question that is sought answered will also form the basis of an analysis pipeline to be used in subsequent sections. The results from these analyses will help gain insight into the models to develop them further.
- can communicate and mediate research-based knowledge both orally and in writing
  The project background and ideas will be presented during an NRU meeting 01/09-2020 and hopefully once more near the end of the project period to present results and/or rehearse the thesis defense. The actual thesis defense will be open to the public, and the thesis will be made publicly available. The thesis itself will contain an overarching description of the problem, the methods used to solve it, the data available including its preprocessing, and a thorough statistical analysis of results. Finally, a careful discussion will reflect on the results reached and suggest ideas for further studies.
- is familiar with and can seek out leading international research within his/her specialist area.
  Upon writing this project plan, several days have been spent searching and reading literature. The first author of the paper Stevner et al. (2019) has been contacted and will be available for help and collaboration throughout the project period. The bibliography of the thesis will reflect the level of literature search providing the background for the project.

- can work independently and reflect on own learning, academic development and specialization
  This project is overseen by three supervisors with meetings no more than weekly. All programming, thesis writing, and everything in between are carried out by the author. This document will be handed in in a revised form along the thesis, where reflection on own learning process is a viable part.
- masters technical problem-solving at a high level through project work, and has the capacity to work with and manage all phases of a project – including preparation of timetables, design, solution and documentation
  This document also contains a project timetable suggestion for time spent on individual subjects, including thesis writing. Each subsection of the project will require literature search and reading, programming, discussion on results, preparing figures for presentation and thesis writing. The time spent on all these parts will be noted, and deadlines will be set in order to keep on schedule.